

Populationsgenetik
kupfer- und bronzezeitlicher
Bevölkerungen der
osteuropäischen Steppe

Dissertation
zur Erlangung des Grades
„Doktor der Naturwissenschaften“

Am Fachbereich Biologie
der Johannes Gutenberg-Universität Mainz

Sandra Wilde
geb. in Bad Kreuznach

Mainz 2014

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung: 22.12.2014

D77 (Universität Mainz)

Zusammenfassung

Mit der vorliegenden Arbeit wurden erstmals die prähistorischen Bevölkerungsstrukturen in der osteuropäischen Steppe von der oberthrakischen Tiefebene bis zur Wolga populationsgenetisch untersucht. Mit einem Verfahren aus Multiplex-PCR und *454-sequencing* wurden auf der mitochondrialen DNA neben der Hypervariablen Region I (HVR I) noch 30 kurze Abschnitte in der *coding region* mit 32 kladenspezifischen Polymorphismen, sowie 20 putativ selektierte autosomale SNPs und ein geschlechtsbestimmender Locus analysiert. Von 65 der 180 kupfer- und bronzezeitlichen Individuen wurden Datenanalysen basierend auf HVR I-Sequenzen durchgeführt. (Partielle) Genotypen liegen von 61 Individuen vor. Zu Vergleichszwecken wurden bereits veröffentlichte aDNA-Daten aus Mittel- und Osteuropa, Zentralasien, sowie moderne DNA-Sequenzen herangezogen.

Die genetischen Daten stützen die Annahme, dass frühneolithische Bauern aus Südosteuropa durch demische Diffusion an der Etablierung der Viehwirtschaft in der Steppe beteiligt waren. Die durchweg niedrigen bzw. um Null liegenden F_{ST} -Werte zwischen der frühbronzezeitlichen Jamnaja-Kultur in der Steppe und den aufeinanderfolgenden neolithischen Kulturen Mitteleuropas sprechen für andauernde oder wiederkehrende Kontakte zwischen den beiden Großräumen.

In der der Jamnaja-Kultur nachfolgenden Katakombengrabbkultur steigt der Anteil der in Jäger/Sammler-Populationen im neolithischen Skandinavien und im mesolithischen Nordwestrussland häufigen Haplogruppe U4 von etwa 5 % auf 30 %. Möglicherweise waren Immigranten aus den ostbaltischen Jäger/Sammler-Refugien an der Genese der Katakombengrabbkultur beteiligt. Die niedrigen F_{ST} -Werte zwischen den prähistorischen Steppenpopulationen und der heutigen Bevölkerung Mittel- und Osteuropas weisen auf genetische Kontinuität hin. Dies wird durch die nukleären Genotypenfrequenzen bestätigt. Der moderne europäische Genpool lässt sich nach aktuellem Kenntnisstand auf drei Wurzeln zurückführen: indigene Mesolithiker, frühe Bauern aus dem Nahen Osten und eine nordeurasische Komponente jungpalaeolithischen Ursprungs. Letztere könnte vielleicht über die nordpontische Population in das Erbgut der spätneolithischen Europäer gelangt sein.

Summary

This dissertation presents the first genetic study of prehistoric populations in the Pontic-Caspian steppe from the Upper Thracian Plain to the Volga. Hypervariable region I (HVR I) and 30 short sections of the coding region containing 32 clade-determining polymorphisms on the mitochondrial DNA, as well as 20 putatively naturally selected autosomal SNPs and a sex-determining locus were analysed using a combination of multiplex PCR and 454 sequencing. Data analysis was performed on the HVR I of 65 of the 180 Eneolithic and Bronze Age samples. (Partial) genotypes were generated from 61 individuals. Published ancient DNA data from Central and Eastern Europe and Central Asia, as well as modern DNA sequences were consulted for comparison.

The genetic data support the inference that early Neolithic farmers from Southeast Europe were involved in establishing pastoralism in the steppes by demic diffusion. The consistently low values of the F_{ST} -statistic (the range includes zero) between the Yamnaya Culture of the steppe and a succession of Neolithic cultures in Central Europe indicate continuous or recurrent contacts between the two regions. Between the Yamnaya Culture and its successor, the Catacomb Culture, the incidence of haplogroup U4, which is at high frequency in hunter-gatherer populations of Neolithic Scandinavia and Mesolithic Northwest Russia, rises from approximately 5 % to above 30 %. It is possible that immigrants from Eastern Baltic hunter-gatherer refugia were involved in the genesis of the Catacomb Culture.

The low F_{ST} values between the prehistoric steppe populations and the modern populations of Central and Eastern Europe indicate genetic continuity. This is supported by the nuclear genotype frequencies. According to current knowledge the modern European gene pool can be explained by three roots: indigenous Mesolithic hunter-gatherers, early farmers from the Near East, and an ancient North Eurasian component with an Upper Palaeolithic origin. Maybe the third ancestry component was introduced into the late Neolithic European genome by the North Pontic population.

Inhaltsverzeichnis

1	EINLEITUNG	1
2	PALAEOGENETIK DES MENSCHEN.....	3
3	DAS NEOLITHIKUM IN MITTEL- UND OSTEUROPA	9
4	DER WEST- UND NORDPONTISCHE RAUM VOM NEOLITHIKUM BIS ZUR BRONZEZEIT	18
4.1	DAS UNTERSUCHUNGSGEBIET	18
4.2	DIE ARCHÄOLOGIE DES NORDPONTIKUMS.....	20
4.2.1	Das Neolithikum	21
4.2.2	Das Äneolithikum.....	24
4.2.3	Die Bronzezeit.....	34
5	SELEKTION IN DER HUMANEVOLUTION	40
5.1	AKKLIMATISATION, ADAPTATION & SELEKTION	40
5.2	DIE UNTERSUCHUNG VON SELEKTION.....	44
5.3	SIGNATUREN VON SELEKTION	46
5.4	GENETISCHE MARKER UNTER SELEKTIONSVERDACHT.....	49
6	FRAGEN, ZIELE & STRATEGIEN.....	59
7	METHODISCHE GRUNDLAGEN.....	68
7.1	EIGENSCHAFTEN DEGENERIERTER ALTER DNA.....	68
7.2	<i>NEXT-GENERATION SEQUENCING</i> IN DER PALAEOGENETIK.....	70
7.2.1	Roche 454 Genome Sequencer FLX Sequenzierung	78
7.2.1.1	<i>Library</i>	79
7.2.1.2	<i>Emulsions-PCR</i>	81
7.2.1.3	<i>Pyrosequenzierung</i>	83
7.2.2	Parallelsequenzierung durch <i>Tagging</i> von Proben.....	87
7.2.2.1	<i>Das Tagging nach Meyer et al. 2008</i>	88
7.3	POPULATIONSGENETISCHE DATENANALYSE.....	92
7.3.1	Phylogeographie.....	92
7.3.2	<i>Summary Patterns</i>	93
7.3.3	Explizite Modellierungen	103
8	MATERIAL & METHODEN.....	106
8.1	PROBENMATERIAL	106
8.2	METHODEN.....	111

8.2.1	Maßnahmen zur Kontaminationsvermeidung.....	111
8.2.2	Probenvorbereitung.....	113
8.2.3	DNA-Extraktion	114
8.2.4	DNA-Amplifikation	115
8.2.4.1	<i>Mitochondriale Primersysteme</i>	115
8.2.4.2	<i>Mitochondriale Multiplex-PCR</i>	118
8.2.4.3	<i>Nukleäre Primersysteme</i>	120
8.2.4.4	<i>Nukleäre Multiplex-PCR</i>	123
8.2.4.5	<i>Einzel-PCR</i>	125
8.2.4.6	<i>Agarosegelelektrophorese</i>	126
8.2.4.7	<i>Aufreinigung der PCR-Produkte</i>	127
8.2.5	„Tagging“ der Proben.....	128
8.2.5.1	<i>Barcodes</i>	128
8.2.5.2	<i>Das Tagging-Protokoll</i>	129
8.2.6	DNA-Sequenzierung.....	131
8.2.6.1	<i>Direktsequenzierung</i>	132
8.2.6.2	<i>Aufreinigung der Sequenzierprodukte</i>	132
8.2.6.3	<i>Kapillar-Gelelektrophorese</i>	133
8.2.7	Kontrollen und Replikationen	133
8.2.8	Datenanalyse	134
8.2.8.1	<i>DNA-Sequenzanalyse und Reproduktion der Ergebnisse</i>	134
8.2.8.2	<i>Biostatistische Analysen</i>	137
8.2.8.3	<i>Vergleichsdatensätze</i>	138
9	ERGEBNISSE	141
9.1	METHODISCHE ERGEBNISSE	141
9.1.1	Amplifikationserfolg der mitochondrialen Multiplex-PCR.....	141
9.1.2	Amplifikationserfolg der nukleären Multiplex-PCR	146
9.1.3	Amplifikationserfolg der Einzel-PCR	148
9.1.4	Kontaminationen in den Negativkontrollen	150
9.2	ERGEBNISSE DER PROBENERHALTUNG	152
9.3	ERGEBNISSE DER POPULATIONSGENETISCHEN ANALYSEN	157
9.3.1	Haplotypisierung der Proben.....	157
9.3.2	Genetische Charakterisierung der prähistorischen Populationen.....	162
9.3.3	Genetische Distanzen im nordpontischen Steppenraum	166
9.3.4	Genetische Distanzen zwischen den Kulturen der nordpontischen Steppe und Mitteleuropas	172
9.4	ERGEBNISSE DER ANALYSE DER AUTOSOMALEN DATEN.....	178

9.4.1	Genotypen- und Allelfrequenzen	178
9.4.2	Genetische Distanzen	185
10	DISKUSSION	186
10.1	METHODENDISKUSSION	186
10.1.1	Die Kombination von Multiplex-PCR und 454-Sequenzierung	186
10.1.2	Die Auswahl der genetischen Marker	191
10.1.3	Datenanalyse	192
10.1.4	Probenauswahl	194
10.1.5	Probenerhaltung	195
10.1.6	Authentizität der Daten	196
10.2	DIE ROLLE DER PONTISCH-KASPISCHEN STEPPE IM EUROPÄISCHEN NEOLITHIKUM	199
10.2.1	Der Ursprung der Viehhirten in der pontischen Steppe	200
10.2.2	Die Nachfolge der Linearbandkeramischen Kultur	204
10.2.3	Die Jamnaja und die Trichterbecherkultur	205
10.2.4	Die Genese der Becherkulturen	207
10.2.5	Der Ursprung der Katakombengrabkultur	208
10.2.6	Beitrag der Steppenkulturen zum modernen europäischen Genpool	210
10.2.7	Östliche Kontakte der Steppenpopulationen	212
10.3	DAS ÄNEOLITHIKUM UND DIE BRONZEZEIT IN DER OSTEUEPÄISCHEN STEPPE	213
10.3.1	Das Äneolithikum	213
10.3.2	Die Jamnaja-Kultur	215
10.3.3	Die Katakombengrabkultur	221
10.3.4	Die Poltavka-Kultur	224
10.4	MARKER UNTER SELEKTIONSVERDACHT	224
10.4.1	Pigmentierung & Vitamin D-Haushalt	225
10.4.1.1	<i>Pigmentierungs-Pathway</i>	225
10.4.1.2	<i>TYR</i>	227
10.4.1.3	<i>SLC45A2</i>	228
10.4.1.4	<i>HERC2</i>	230
10.4.1.5	<i>SLC24A5</i>	232
10.4.1.6	<i>Selektionsfaktor Sonneneinstrahlung und Vitamin D-Homöostasis</i>	233
10.4.1.7	<i>CYP3A4</i>	237
10.4.2	Wasser- und Elektrolythaushalt	239
10.4.2.1	<i>CYP3A5</i>	239
10.4.2.2	<i>SLC12A3</i>	241
10.4.3	Temperaturadaptation	243
10.4.3.1	<i>EDAR</i>	244

10.4.3.2	<i>ABCC11</i>	246
10.4.4	Ernährung	249
10.4.4.1	<i>LCT</i>	249
10.4.4.2	<i>TRPV6</i>	254
10.4.4.3	<i>NAT2</i>	258
10.4.4.4	<i>ADH1B</i>	261
10.4.5	Pathogene.....	265
10.4.5.1	<i>CASP12</i>	265
10.4.5.2	<i>ABCB1</i>	269
12	REFERENZEN	277
13	ANHANG.....	320
13.1	LABORMATERIALIEN.....	320
13.2	F_{ST} -WERTE	323
13.3	GENOTYPEN	336

Abbildungsverzeichnis

Abbildung 4.1:	Übersichtskarte des Untersuchungsraums	19
Abbildung 7.1:	Enzymkaskade der photochemischen Reaktion bei der Pyrosequenzierung ...	86
Abbildung 7.2:	Das Prinzip des Taggings nach Meyer et al.	89
Abbildung 8.1:	Geographische Lage der Fundplätze	107
Abbildung 9.1:	Amplifikationserfolg ^S der mitochondrialen Multiplex-PCR	143
Abbildung 9.2:	Boxplot der reads pro locus in der mitochondrialen Multiplex-PCR	145
Abbildung 9.3:	Amplifikationserfolg ^S der nukleären Multiplex-PCR	147
Abbildung 9.4:	Boxplot der reads pro locus in der nukleären Multiplex-PCR	148
Abbildung 9.5:	Amplifikationserfolg der Einzel-PCRs	149
Abbildung 9.6:	Vergleich der mitochondrialen und nukleären Erhaltung	153
Abbildung 9.7:	Haplogruppenverteilung in den prähistorischen Kulturen	162
Abbildung 9.8:	Die U-Linien (assoziiert mit Jägern/Sammlern) in den Steppenkulturen	164
Abbildung 9.9:	Fundplatz-Einheiten als Grundlage für die Bildung von Gruppen	168
Abbildung 9.10:	Multidimensionale Skalierung prähistorischer Populationen Westeurasiens	175
Abbildung 9.11:	Haplotype Sharing zwischen den Steppenpopulationen und diversen Kulturen Westeurasiens	177
Abbildung 9.12:	Vergleich der Allelfrequenzen nach Reproduktionsstatus	179
Abbildung 9.13:	Genotypen- und Allelfrequenzen der kupfer- und bronzezeitlichen Kulturen sowie der modernen Population der Ukraine	184
Abbildung 10.1:	Fundplatz-Einheiten der geographischen Gruppen	217
Abbildung 10.2:	Produktion und Aktivierung von Vitamin D ₃	235
Abbildung 10.3:	Hydrolytische Spaltung von Laktose	250
Abbildung 10.4:	Beispiele für Acetylierungen durch NAT2	259
Abbildung 10.5:	Oxidation von Alkohol	261
Abbildung 13.1:	Karte der Fundplatz-Einheiten	324

Tabellenverzeichnis

Tabelle 5.1:	Übersicht der untersuchten Gene und Polymorphismen unter Selektionsverdacht	56
Tabelle 7.1:	Übersicht der NGS-Plattformen.....	72
Tabelle 8.1:	Probenliste.....	107
Tabelle 8.2:	Mitochondriale Primersysteme	116
Tabelle 8.3:	Konzentrationen der mitochondrialen Primer	118
Tabelle 8.4:	Mitochondriale Multiplex-PCR – Protokolle 1 & 2	119
Tabelle 8.5:	Mitochondriale Multiplex-PCR – Protokoll 3.....	120
Tabelle 8.6:	Nukleäre Primersysteme	121
Tabelle 8.7:	Nukleäre Multiplex-PCR-Protokolle.....	124
Tabelle 8.8:	Einzel-PCR-Protokolle	125
Tabelle 8.9:	Alternative Primersysteme zur Amplifikation der HVR I	126
Tabelle 8.10:	Barcode-Sequenzen.....	128
Tabelle 8.11:	Reaktionsansätze beim Tagging	130
Tabelle 8.12:	Cycle Sequencing-Protokoll	132
Tabelle 8.13:	Mindestanforderung an die Authentifizierung von Genotypen	136
Tabelle 8.14:	HVR I-Vergleichsdatensätze.....	139
Tabelle 9.1:	Amplifikationserfolge ^S der einzelnen Multiplex-PCR-Ansätze.....	144
Tabelle 9.2:	Sequenziererfolge der einzelnen Multiplex-PCR-Ansätze	144
Tabelle 9.3:	Negativkontrollen und Kontaminationen.....	151
Tabelle 9.4:	Kategorisierung der Proben entsprechend ihres Erhaltungszustands	153
Tabelle 9.5:	Nukleärer Reproduktionserfolg in Abhängigkeit vom Erhaltungszustand	155
Tabelle 9.6:	Regionale Unterschiede in der Probenerhaltung	155
Tabelle 9.7:	Temporale Unterschiede in der Probenerhaltung.....	156
Tabelle 9.8:	Mitochondriale Haplotypen der 66 reproduzierten Individuen	159
Tabelle 9.9:	Genetische Diversität und Demographie der prähistorischen Kulturen	165
Tabelle 9.10:	Paarweise F_{ST} -Werte zwischen den Kulturen der Steppe.....	166
Tabelle 9.11:	Fundplatz-Einheiten als Grundlage für die Bildung von Gruppen	167
Tabelle 9.12:	F_{ST} -Werte zwischen dem Äneolithikum und Gruppen der Jamnaja-Kultur	169
Tabelle 9.13:	F_{ST} -Werte zwischen Gruppen der Jamnaja-, der Katakombengrab- und der Poltavka-Kultur	170

Tabelle 9.14:	Jamnaja-Strukturen - westlich und östlich des Südlichen Bug (F_{ST} -Werte)	171
Tabelle 9.15:	Jamnaja-Strukturen - westlich und östlich des Dnjepr (F_{ST} -Werte)	171
Tabelle 9.16:	Jamnaja-Strukturen - Küste versus Landesinneres (F_{ST} -Werte)	171
Tabelle 9.17:	F_{ST} -Werte zwischen den Populationen der Steppe und Westeurasiens.....	173
Tabelle 9.18:	Genotypen- und Allelfrequenzen der autosomalen loci.....	180
Tabelle 9.19:	Genetische Kopplung in ABCB1	185
Tabelle 9.20:	Genetische Distanzen der autosomalen loci.....	185
Tabelle 13.1:	Verwendete Reagenzien und sonstige Chemikalien.....	320
Tabelle 13.2:	Sonstige verwendete Verbrauchsmaterialien	321
Tabelle 13.3:	Verwendete Laborgeräte	322
Tabelle 13.4:	F_{ST} -Werte zwischen den Verbreitungsgebieten neolithischer Kulturen	324
Tabelle 13.5:	F_{ST} -Werte zwischen den Verbreitungsgebieten der Kulturen des frühen Äneolithikums	325
Tabelle 13.6:	F_{ST} -Werte zwischen den Verbreitungsgebieten der Kulturen des mittleren Äneolithikums	325
Tabelle 13.7:	F_{ST} -Werte zwischen den Zwischenstromgebieten	326
Tabelle 13.8:	F_{ST} -Werte zwischen Gruppen entlang der Flüsse	326
Tabelle 13.9:	F_{ST} -Werte zwischen östlichen und westlichen Gruppen.....	327
Tabelle 13.10:	Steppeninterne F_{ST} -Werte	328
Tabelle 13.11:	F_{ST} -Werte zwischen den Kulturen Westeurasiens	333
Tabelle 13.12:	Ergebnisse der Typisierung der autosomalen loci und der Geschlechtsbestimmung.....	336

Verzeichnis der Daten-DVD

Titel der DVD: Wilde – Dissertation

- 1 – Alignments der HVR I-Sequenzen
- 2 – Consensus-Sequenzen der HVR I
und Sequenzen der Vergleichspopulationen
- 3 – Sequenzen zur Bildung der Consensus-Sequenz der HVR I
- 4 – Rohdaten der 454-Sequenzierungen
- 5 – Rohdaten der Direktsequenzierungen
- 6 – Digitale Version dieser Dissertation

1 Einleitung

Durch einen Quantensprung in der DNA-Sequenzierungstechnologie mit der Entwicklung der sogenannten *next-generation sequencing*-Methoden (NGS) kam es im Verlauf des letzten Jahrzehnts zu rasanten Fortschritten in verschiedenen Forschungsbereichen, wie der Medizin, Mikrobiologie, Ökologie, Populationsgenetik, Anthropologie etc. Mit der vorliegenden Arbeit wurde das Ziel verfolgt, alte DNA für diese Technologien verfügbar zu machen, und damit das Feld der Palaeopopulationsgenetik für das NGS zu öffnen. Hierfür wurden klassische Methoden mit NGS kombiniert. Es wurden Protokolle für Multiplex-PCRs entwickelt. Für die Parallelsequenzierung der Proben auf der 454-Plattform wurde die Methode des *sequence taggings* von Meyer *et al.* (2008) für die Anwendung auf alte DNA modifiziert.

Es wurden bereits einige Studien zur *high throughput*-Sequenzierung von DNA aus einzelnen prähistorischen Skelettfunden veröffentlicht (siehe hierzu Kapitel 2). Das interdisziplinäre Projekt „Palaeogenetische Untersuchungen zu wirtschaftlichen Innovationen und sozialer Mobilität in der eurasischen Steppe 3 500 – 300 v. Chr.“ ist eines der ersten, in denen umfangreiche Stichproben aus prähistorischen Populationen mithilfe von NGS sequenziert wurden. Die vorliegende Studie ist ein Teilprojekt dieses Vorhabens. Untersucht wurden kupfer- und bronzezeitliche Populationen aus dem Steppenraum westlich, nördlich und östlich des Schwarzmeeres. Parallel zu den methodischen Aspekten stand vor allem die Gewinnung populationsgenetischer Erkenntnisse im Vordergrund. Im nordpontischen Raum kam es um 3 000 und ab 2 700 v. Chr. jeweils zu deutlichen kulturellen Veränderungen, die der Jamnaja- und der Katakombengrabbkultur zugeschrieben werden. Mit dem Auftauchen von Wagen mit Scheibenrädern in dieser Zeit gab es Hinweise auf eine erhöhte Mobilität, insbesondere in der Jamnaja-Kultur. Anhand der alten DNA-Daten sollten in dieser Arbeit die archäologisch basierten Mobilitäts- und Migrationshypothesen überprüft, sowie Zusammenhänge zwischen Bestattungsriten und anthropologischer Identität der einzelnen Kulturerscheinungen erfasst werden.

Neben diesen Untersuchungen zur steppeninternen Binnenstruktur wurde der nördliche Schwarzmeerraum im Gesamtgefüge des europäischen Neolithikums betrachtet. Für Mitteleuropa gibt es mittlerweile einen umfangreichen aDNA-Gesamtdatensatz aus verschiedenen Veröffentlichungen, anhand dessen die genetische Grundlage der neolithischen Transition zu Ackerbau und Viehzucht untersucht wurde. Archäologischen Befunden zufolge fand in der nordpontischen Steppe von der Mitte des sechsten bis zum Ende des vierten Jahrtausends v. Chr. sukzessiv der Übergang vom Jagen und Sammeln zur Viehwirtschaft statt. In der vorliegenden Studie sollten mögliche Wechselbeziehungen zwischen Mittel- und Osteuropa untersucht und somit der Steppenraum in das sich entwickelnde Bild des Neolithikums integriert werden.

Infrastruktur, Geräte und Verbrauchsmaterialien wurden durch Mittel der Johannes Gutenberg-Universität Mainz bereitgestellt. Die übrige Finanzierung des Forschungsprojektes erfolgte aus Mitteln des Bundesministeriums für Bildung und Forschung (BMBF), Förderschwerpunkt „Wechselwirkungen zwischen Natur- und Geisteswissenschaften“, Fördernummer 01UA0809A. Die parallel von Frau Martina Unterländer verfasste Dissertation mit dem Titel „Populationsgenetik eisenzeitlicher Reiternomaden der Eurasischen Steppe“ (2014) repräsentiert das zweite Teilprojekt des oben erwähnten Verbundvorhabens. In einem dritten Teilprojekt wurde die Domestikation des zweihöckrigen Kamels im südlichen Steppengürtel von Frau Alexandra Trinks im Rahmen einer Masterarbeit mit dem Titel „Populationsgenetische Untersuchung spätbronze- und früheisenzeitlicher Skelettfunde von *Camelus bactrianus* aus West-Sibirien und Süd-Usbekistan“ (2010) untersucht.

Der archäologische Beitrag zur vorliegenden Arbeit wurde vor allem von Prof. Dr. W. S. und Prof. Dr. E. K. von der Freien Universität Berlin geleistet. Weitere Verbundpartner im Gesamtprojekt auf der archäologischen Seite waren Prof. Dr. H. P. (Stiftung Preußischer Kulturbesitz, Berlin) und Prof. Dr. N. B. (Deutsches Archäologisches Institut, Berlin).

2 Palaeogenetik des Menschen

Basierend auf der Diversität rezenter menschlicher Populationen wurden bereits wesentliche Erkenntnisse über unsere Geschichte gewonnen. Allerdings handelt es sich dabei nur um indirekte Beobachtungen zu unserer Vergangenheit. Die zusätzliche Untersuchung alter DNA (aDNA) hingegen liefert einen direkten Einblick in diejenigen Prozesse, die uns Menschen zu dem gemacht haben, was wir heute im Spiegel oder auf der Straße sehen. Die Palaeogenetik ermöglicht uns, vom Zusammenleben von anatomisch modernen Menschen und Neandertalern ein Bild zu zeichnen (Green *et al.* 2010), entfernte Verwandte aus der Gattung *Homo* aufzuspüren (Reich *et al.* 2010), oder den Wandel von Wirtschaftssystemen, wie den Übergang von einer wildbeuterischen zur produzierenden Lebensweise, nachzuzeichnen (Bramanti *et al.* 2009).

Burger formulierte 2007 folgende Definition: „Die Begriffe Palaeogenetik bzw. „ancient DNA“ bezeichnen die spurenanalytische, molekulargenetische Charakterisierung von degradiertem DNA verstorbenen Organismen aus historischem oder naturgeschichtlichem Interesse.“ Ihren eigentlichen Anfang nahm die aDNA-Forschung mit der Veröffentlichung einer 229 Basenpaare langen Nukleotidsequenz des ausgestorbenen Quagga (Higuchi *et al.* 1984). Mit der Einführung der Polymerase-Kettenreaktion (PCR) (Saiki *et al.* 1985, Mullis *et al.* 1986) bekam die Palaeogenetik das notwendige Werkzeug, um sich zu einer „echten“ Wissenschaft zu entwickeln. Vor allem in den Anfangsjahren allerdings wurden eine Reihe von Studien publiziert, bei denen sich im Nachhinein herausstellte, dass es sich bei den produzierten Sequenzen um Artefakte bzw. Kontaminationen handelte (z. B. Pääbo 1985, Cano *et al.* 1993, Woodward *et al.* 1994). Daraufhin wurden einige Standards in Bezug auf Kontaminationsvermeidung und Maßnahmen zur Dekontamination etabliert (siehe u. a. Handt *et al.* 1994, Cooper & Poinar 2000, Pääbo *et al.* 2004).

Gerade mal zwei Jahrzehnte nach ihrer Geburtsstunde erlebte die Palaeogenetik mit der Einführung des *next-generation sequencing* (NGS) einen fundamentalen Wandel, und eine rasante Entwicklung begann. Mit der traditionellen Methode der

Polymerase-Kettenreaktion, gefolgt von einer Sequenzierung nach Sanger *et al.* (1977) konnten aus dem sehr limitierten aDNA-Probenmaterial bisher lediglich einige wenige Genorte untersucht werden. Zu den NGS-Verfahren gehören eine Reihe neu entwickelter Methoden zur *template*-Vorbereitung, Amplifikation und Sequenzierung, die die kostengünstige Produktion unzähliger Daten in kürzester Zeit ermöglichen (siehe Kapitel 7.2). Nur wenige Monate nach der Vorstellung der ersten NGS-Plattform wurde die erste aDNA-Studie veröffentlicht: Poinar *et al.* (2006) hatten 13 Millionen Basenpaare eines Mammutgenoms sequenziert.

Kurze Zeit später folgten neben weiteren Tiergenomen (z. B. Allentoft *et al.* 2009, Miller *et al.* 2009, Willerslev *et al.* 2009, Lindqvist *et al.* 2010) die ersten auf NGS basierenden Studien zum anatomisch modernen Menschen, wie das Genom eines Palaeo-Eskimos der Saqqaq-Kultur (Gilbert *et al.* 2008, Rasmussen *et al.* 2010) und das der Gletschermumie „Ötzi“ (Ermini *et al.* 2008, Keller *et al.* 2012). Dazu kam die medienträchtige Entschlüsselung des Neandertalergenoms (Green *et al.* 2010). Ebenfalls 2010 wurde das Genom eines Individuums unbekannter Art innerhalb der Homininen sequenziert (Reich *et al.* 2010). Es handelte sich um einen Fund in der südsibirischen Denisova-Höhle im Altai-Gebirge. Die gleiche Studie zeigte, dass dieser „Denisova-Mensch“ einen gemeinsamen Vorfahren mit den Neandertalern hat. Wie diese Beispiele zeigen, kommt es in der aDNA-Forschung im Zuge der NGS-Ära zurzeit zu einer Verschiebung von der Untersuchung weniger *loci* in mehreren Individuen zu der Analyse vieler *loci* bzw. ganzer Genome einzelner Individuen. Einen Mittelweg zwischen diesen beiden Strategien bieten sogenannte *target enrichment*- bzw. *DNA capture*-Methoden, mit denen multiple ausgewählte *loci* gezielt angereichert werden, bevor sie auf einer NGS-Plattform sequenziert werden. Damit wurden mitochondriale Genome von Homininen und autosomale DNA-Abschnitte, die im Verlauf der menschlichen Evolution entstandene Mutationen tragen, untersucht (Briggs *et al.* 2009, Burbano *et al.* 2010, Krause *et al.* 2010).

Bereits vor der Verwendung von NGS wurde mithilfe von PCR und Klonierung gezeigt, dass die mitochondriale DNA (mtDNA) der Neandertaler außerhalb der Variationsbreite der heutigen Menschen liegt und auf der Ebene dieses Genoms keine Vermischung nachzuweisen ist (Krings *et al.* 1997, Ovchinnikov *et al.* 2000, Serre *et al.*

2004). Mit dem zunehmenden Datenvolumen aus NGS-basierten Studien wuchsen die Erkenntnisse. So konnte z. B. der Populationssplit zwischen Neandertalern und modernen Menschen in den Zeitrahmen zwischen 270 und 440 000 Jahre vor heute festgesetzt werden (Green *et al.* 2010). Zur Vermischung kam es kaum, sie wird auf maximal 1–4 % geschätzt, und ist heute nur in eurasischen Populationen nachzuweisen (Green *et al.* 2010, Currat & Excoffier 2011).

Insgesamt hat die Wissenschaft in Bezug auf den Ursprung des anatomisch modernen Menschen und seinen Beziehungen zu anderen, extinkten Homininen in den letzten Jahren einen großen Sprung gemacht. Es ist zu erwarten, dass mit der Ausreifung der *capture*- und NGS-Methoden sowie insbesondere der Bioinformatik, und mit zunehmender Zugänglichkeit für immer mehr Arbeitsgruppen, vermehrt Studien und Datensätze mit vielen *loci* von vielen Individuen veröffentlicht werden, um so populationsgenetische Fragestellungen genauer zu beantworten (Veeramah & Hammer 2014). Die größte Herausforderung ist und wird dabei die Sicherstellung der Authentizität dieser Datensätze sein.

Neben der stammesgeschichtlichen Entwicklung des anatomisch modernen Menschen ist auch die rezentere Geschichte von großem Interesse. Weltweit gibt es Projekte, die anhand von mitochondrialer, Y-chromosomaler und nukleärer aDNA die Populationsgenetik menschlicher Gruppen untersuchen, beispielsweise von präkolumbischen Kulturen in Peru (Fehren-Schmitz *et al.* 2010, Fehren-Schmitz 2012), die genetischen Beziehungen zwischen den alten und modernen Völkern der mongolischen Steppe (Keyser-Tracqui *et al.* 2006) oder des antiken Etrusker-Volkes zu den modernen Bewohnern der Toskana (Belle *et al.* 2006). Ein intensiv diskutiertes Thema innerhalb der Populationsgenetik ist die „Neolithische Transition“ in Europa, insbesondere die Art der Verbreitung dieser neuen, produzierenden Lebensweise. Debattiert wurden über Jahrzehnte zwei Grundmodelle: das der Akkulturation der einheimischen Jäger und Sammler durch Wissenstransfer (*cultural diffusion*) und das der Zuwanderung von Bauern aus dem Nahen Osten mit genetischer Vermischung oder gar Ersetzung der Mesolithiker (*demic diffusion*). Genetische, auf der DNA moderner Europäer basierende Studien konnten die Art der Beziehung zwischen Meso- und Neolithikern nicht klären (u. a. Ammerman & Cavalli-Sforza 1984, Sokal *et*

al. 1991, Richards *et al.* 2000, Currat & Excoffier 2005). In einer aDNA-Studie von Bramanti und Kollegen (2009), aufbauend auf Haak *et al.* (Haak *et al.* 2005) wurde Diskontinuität zwischen den beiden Populationen nachgewiesen. Außerdem zeigte sie, dass zur Zusammensetzung des Genpools der modernen Europäer noch eine weitere, bisher unbekannte Komponente hinzugekommen sein muss. Erst vor einigen Wochen wurde eine palaeogenetische Untersuchung veröffentlicht, nach der der Ursprung dieser Komponente im jungpalaeolithischen Nordeurasien liegt (Lazaridis *et al.* 2014). Mittlerweile gibt es eine ganze Reihe von palaeogenetischen Studien, die dazu beitragen, die neolithische Transition weiter aufzuschlüsseln und die Entwicklung zum heutigen Europäer zu klären (z. B. Haak *et al.* 2010, Hervella *et al.* 2012, Skoglund *et al.* 2012, Brotherton *et al.* 2013). Dem Neolithikum Europas ist in dieser Arbeit ein eigenes Kapitel (3, Seite 9 ff) gewidmet.

Der west- und nordpontische Steppenraum (Kapitel 4) fand in der bisherigen palaeogenetischen Forschung kaum Beachtung. Die im Rahmen dieser Arbeit analysierten Proben stammen aus der Zeit nach der neolithischen Transition in Mitteleuropa (ca. 6 500 – 4 000 vor heute). Thematisch geht es auch bei den Steppenpopulationen um die Frage nach dem Übergang zu einer neolithischen Subsistenz, die genetischen Beziehungen aufeinanderfolgender Kulturen und mögliche Verbindungen nach Westen zum Neolithikum Mitteleuropas (siehe auch Kapitel 6 zu den Zielen dieser Arbeit).

Doch nicht nur humane aDNA-Proben dienen der Rekonstruktion der Prähistorie des Menschen, sondern auch genetisches Material von Tieren. Palaeogenetische Erkenntnisse zur Domestikation von Tieren wie Rind, Ziege (Bollongino 2006, Edwards *et al.* 2007, Edwards *et al.* 2011, Scheu 2012), Schwein, Schaf (Larson *et al.* 2007a, Larson *et al.* 2007b, Geörg 2012, Ottoni *et al.* 2013), Pferd (Weinstock *et al.* 2005, Lira *et al.* 2010, Orlando *et al.* 2013) und Hund (Leonard *et al.* 2002, Verginelli *et al.* 2005, Thalmann *et al.* 2013) geben Aufschlüsse über die Lebensweise der zeitgenössischen Menschen. Ähnliches gilt für die Domestikation von Nutzpflanzen, z. B. Mais (Jaenicke-Després *et al.* 2003) oder Berlandier-Gänsefuß¹ (Kistler & Shapiro 2011).

¹ stärkehaltiges Pseudogetreide, verwandt mit Quinoa und Spinat

Pathogene und ihre Verbreitung geben ebenfalls Hinweise auf menschliche Bevölkerungsprozesse wie Migrationen, Erhöhung der Populationsdichte oder den Kontakt mit bestimmten Überträgern. In humanem Skelettmaterial aus dem Mittelalter konnte der Pesterreger *Yersinia pestis* nachgewiesen und seine Verbreitungsrouten nachgezeichnet werden (Haensch *et al.* 2010, Wagner *et al.* 2014). Auch die mittelalterliche Version des Genoms vom *Mycobacterium lepra*, das für die Lepra-Krankheit verantwortlich ist, konnte sequenziert und so Einblick in dessen Herkunft und Transmissionsweise gewonnen werden (Watson & Lockwood 2009, Schuenemann *et al.* 2013). Das *Mycobacterium tuberculosis* ist ein weiterer Erreger, dessen aDNA untersucht wurde (Salo *et al.* 1994, Hershkovitz *et al.* 2008).

Die Koevolution von solchen Pathogenen und humanen Populationen (Stone *et al.* 2009) zeigt sich auch im menschlichen Genom, nämlich in Genen des Immunsystems. Beispielsweise wurde von einer Mutation im *CCR5*-Gen (*chemokine, CC motif, receptor 5*), die mit einer Resistenz gegen das Humane Immundefizienz-Virus (HIV) einhergeht, angenommen, dass ihre Selektion durch *Yersinia pestis* hervorgerufen wurde. Dies wurde durch eine aDNA-Studie widerlegt (Hummel *et al.* 2005). Die Palaeogenetik kann Zeiträume eingrenzen, in denen sich Resistenzen oder das Immunsystem weiterentwickeln (Green *et al.* 2010, Abi-Rached *et al.* 2011, Olalde *et al.* 2014). Neben Krankheiten gibt es noch eine Reihe weiterer Selektionsfaktoren, die die genetische Zusammensetzung von Populationen beeinflussen (näheres hierzu in Kapitel 5). aDNA-Analysen von *loci*, die unter Selektionsverdacht stehen, können, insbesondere mithilfe von Simulationsanalysen, Aufschlüsse über den Zeitpunkt bzw. die Dauer eines Selektionsereignisses, über die Stärke des Einflusses, über bevölkerungsgeschichtliche Entwicklungen oder einfach über phänotypische Eigenschaften alter Populationen geben. Als Beispiele seien hier *FOXP2* (*forkhead box P2*), das Sprachgen (Krause *et al.* 2007), die Selektion von Laktasepersistenz (Burger *et al.* 2007, Itan *et al.* 2009), oder die phänotypischen Merkmale des Palaeo-Eskimos (Rasmussen *et al.* 2010), von „Ötzi“ (Keller *et al.* 2012) oder prähistorischen Populationen der sibirischen Steppe (Bouakaze *et al.* 2009, Keyser *et al.* 2009) genannt.

Bei den in diesem einführenden Kapitel aufgeführten Studien handelt es sich lediglich um einen Ausschnitt aus dem breitgefächerten Themengebiet der aDNA-Forschung. Da die Thematik dieser Arbeit auf Menschen fokussiert ist, sind vor allem Beispiele aufgeführt, die direkt mit dessen Prähistorie verknüpft sind (siehe dazu auch Kirsanow & Burger 2012). Kaum berücksichtigt wurden daher z. B. unzählige Studien zur Stammesgeschichte oder Populationsgenetik diverser, zum Teil extinkter Tiere oder Pflanzen die keinen direkten erkennbaren Einfluss auf die menschliche Entwicklung hatten (abgesehen vielleicht als Fressfeind), aber dennoch die Palaeogenetik prägten. Übersichten hierzu geben z. B. Ramakrishnan & Hadley (2009), Ho & Gilbert (2010) und Shapiro & Hofreiter (2014).

3 Das Neolithikum in Mittel- und Osteuropa

Die Steinzeit lässt sich in drei Abschnitte unterteilen, das Palaeolithikum, das Neolithikum (Lubbock 1865), und eine Übergangsphase dazwischen. Nördlich der Alpen wird diese als Mesolithikum bezeichnet (Westropp 1872), ansonsten ist vom Epipalaeolithikum die Rede. Im Jungpalaeolithikum, vor ca. 45 000 Jahren, wurde Europa erstmals vom anatomisch modernen Menschen aus Afrika heraus besiedelt. Das Letzte Glaziale Maximum vor 24 500 – 19 000 Jahren (Clark *et al.* 2009) machte große Teile Europas unbewohnbar, und die Menschen zogen sich in die wärmeren Regionen im Süden zurück. Das Mesolithikum bzw. Epipalaeolithikum begann ungefähr vor 11 600 Jahren mit einer relativ abrupten Erwärmung, die auch den Beginn des Holozäns markiert. Die Menschen, die Europa nach dem Letzten Glazialen Maximum aus den Refugien rekolonisierten, lebten von der Jagd mit Pfeil, Bogen und Hund, dem Fischfang und dem Sammeln von essbaren Wildpflanzen und -früchten (Terberger 2009). Die Dauer des Mesolithikums war regional unterschiedlich. Das Neolithikum ist durch den Übergang von dieser aneignenden zu einer produzierenden Wirtschaftsweise charakterisiert, was zu enormen gesellschaftlichen, technischen und landschaftlichen Veränderungen führte.

Dieser von Vere Gordon Childe (1936) als „Neolithische Revolution“ bezeichnete Prozess begann im zehnten Jahrtausend v. Chr. in Vorderasien, in einem nicht exakt abzugrenzenden Gebiet, das großräumig den „Fruchtbaren Halbmond“ umfasst. Das sichel- bis halbkreisförmige Gebiet reicht von der östlichen Mittelmeerküste mit der Levante und Südostanatolien im Westen bis zur Nordspitze des Persischen Golfs im Osten. Das Zentrum verläuft entlang und nördlich der Flüsse Euphrat und Tigris (Nordmesopotamien). Nordöstlich wird es durch das Taurus- und das Zagrosgebirge, im Süden durch die syrische Wüste begrenzt (Breasted 1914: 56 f, Özdoğan 2005, Zeder 2008). Erste Hinweise auf beginnende Domestikationsprozesse stammen aus dieser Region vor etwa 11 500 Jahren (Zeder 2008, 2011). Der epipalaeolithischen Kultur Natufien in der Levante wird dabei eine Art Vorläuferrolle zugesprochen. Die Menschen waren teilweise sesshaft und der Verzehr von Wildgetreide spielte eine große Rolle (Bar-Yosef 1998). Über die Gründe für den Wandel von der Lebensweise

als mobile oder semimobile Jäger und Sammler zu sesshaften Bauern wird spekuliert. Wahrscheinlich lag der Auslöser in einer Kombination verschiedenster Faktoren. Die in der Warmzeit guten und ertragreichen Umweltbedingungen in verschiedenen Regionen des Fruchtbaren Halbmondes erlaubten die Sesshaftwerdung von Wildbeutergruppen. Dies wiederum begünstigte das Wachstum der Bevölkerung, da mehr Kinder geboren und aufgezogen werden konnten als auf der Wanderschaft. Möglicherweise ergab sich aus Populationsdruck die Notwendigkeit zum Anbau von Wildpflanzen, woraus sich nach und nach „echte Landwirtschaft“ entwickelte. Auch könnte die Jüngere Dryas, eine zwischenzeitliche Kaltphase mit trockenem Klima zu schlechteren Bedingungen für die als Nahrungsgrundlage dienenden Wildpflanzen geführt haben. Womöglich wurden ihre Keime ausgesät, um einem Mangel entgegenzuwirken. Als weitere Faktoren bei der Entwicklung von Wildbeuter- zu Bauerngesellschaften könnten soziale Aspekte und Strukturen eine Rolle gespielt haben. Einen Überblick über die verschiedenen Modelle gibt Weisdorf (2005). Keramik bzw. Töpferei spielte in dieser frühen Phase noch keine Rolle, weshalb man vom prä- oder vorkeramischen Neolithikum spricht.

Ab etwa 7 000 v. Chr. lässt sich die Ausbreitung der neolithischen Lebensweise in Süd- und Westanatolien nachweisen, Mitte des siebten Jahrtausends v. Chr. in Thessalien und Ostthrakien (Özdoğan 2007, Reingruber 2011). Ab hier lassen sich zwei Verbreitungsrichtungen ausmachen, die mediterrane und die kontinentale Route. Erstere führte entlang des Mittelmeeres bis zur Iberischen Halbinsel, die zwischen ca. 5 500 und 5 000 v. Chr. neolithisiert wurde (Guilaine & Manen 2007, Jeunesse & Willigen 2010). Dieser Weg ist durch die Impressa- und Cardial-Keramik repräsentiert. Die vor allem mit der Linearbandkeramik assoziierte kontinentale Route führte durch Südosteuropa bis in die Ungarische Tiefebene (ca. 6 000 v. Chr.), erreichte die weiten Lössgebiete Mitteleuropas um etwa 5 500 v. Chr. und schließlich die Britischen Inseln ab 4 300 v. Chr. (Müller 2009b). Die Ausbreitung verlief nicht kontinuierlich, sondern schubweise. Die norddeutsche Lössgrenze wurde 4 100 v. Chr. erreicht. Auf den Britischen Inseln, im skandinavischen Raum sowie in den baltischen und russischen Regionen verlor die Verbreitung der neuen Wirtschaftsweise ihre Dynamik und der Übergang vollzog sich nur langsam (Müller 2009b).

Der kontinentale Expansionsweg, der unmittelbar am Untersuchungsgebiet dieser Arbeit vorbeiführt, soll etwas genauer beleuchtet werden. Vom ägäischen Raum breitete sich das Neolithikum zügig auf der Balkanhalbinsel aus. Wenige Funde aus einer frühen Phase, die durch monochrome Keramik charakterisiert ist, repräsentieren erste Vorstöße einiger Bauerngruppen nach Westen. Dann folgte eine schnelle, massive Expansionwelle, mit der Folge, dass zum Ende des siebten Jahrtausends v. Chr. auf fast der gesamten Balkanhalbinsel die neolithische Lebensweise etabliert war. Sie ist mit rot überzogener, polierter Keramik assoziiert (Özdoğan 2011).

Die frühneolithischen Kulturerscheinungen auf dem Balkan und im Karpatenbecken lassen sich in zwei große Kulturkomplexe unterteilen: Der balkanisch-anatolische Karanovo I- und der balkanisch-karpatische Starčevo-Körös-Criș-Komplex. Der Karanovo-Komplex ist nach einem der größten Siedlungshügel Europas benannt, der bis ins zweite Jahrtausend v. Chr. durchweg bewohnt war – abgesehen von einer Unterbrechung gegen Ende der Kupferzeit, als es zu einer plötzlichen Abkühlung des Klimas kam. Manche Archäologen vermuten, dass das Verlassen der Siedlungen in dieser Region die Folge kriegerischer Handlungen (aus der Steppe) war (siehe Anthony 2007: Kap. 11). Der zweite Komplex besteht aus drei regionalen Varianten: die Starčevo-Gruppe in Transdanubien und etwas weiter in östlicher Richtung, die Körös-Gruppe in der Ungarischen Tiefebene (Alföld) und die Criș-Gruppe im heutigen Rumänien (Gronenborn 1999, Müller 2009b). Bei den beiden letzteren handelt es sich um regionale Varianten einer Kultur, die nach dem gleichen Fluss (in unterschiedlichen Sprachen) benannt sind (Anthony 2007: 139). Die Subsistenz der Träger der Starčevo-Körös-Criș-Kultur beruhte vor allem auf Viehwirtschaft, unterstützt durch Ackerbau. Wahrscheinlich spielte das Jagen und Sammeln aufgrund von geographischen und ökologischen Bedingungen in der Bevölkerung der Körös/Criș-Kultur noch eine größere Rolle als in der Starčevo-Gruppe (Gronenborn 1999).

Die Linearbandkeramik (LBK) ist ab etwa 5 700 v. Chr. in Transdanubien in der Region des Balaton (Plattensee) nachgewiesen. Sie entstand unter starkem Einfluss des Starčevo-Körös-Criș-Komplexes, insbesondere Starčevo, bzw. ging wahrscheinlich aus diesem hervor (Quitta 1960b, a, Gronenborn 1999, Pavúk 2004). Pavúk (2004) ist der Meinung, dass auch prä-neolithischen Kulturen bei der LBK-Genese eingeflossen

seien. Im Gebiet des Starčevo-Körös-Criş-Komplexes entstand außerdem die Vinča-Kultur, mit der es zu einer zunehmenden kulturellen Komplexität (ansteigende Siedlungsdichte und -größen, Kupferfunde etc.) im südosteuropäischen Neolithikum kam. In der Großen Ungarischen Tiefebene entwickelte sich, wahrscheinlich durch Kontakte zwischen der Körös-Kultur und späten Mesolithikern, die Alföld-LBK (Gronenborn 1999).

Das Gesellschaftswesen der LBK ist durch Langhäuser in dauerhaften Siedlungen, Gräberfelder, mit Bändern verzierte Keramik und eine produzierende Lebensweise gekennzeichnet. Die entsprechenden Methoden und Technologien (z. B. Tierhaltung, die Bearbeitung der Felder oder die Werkzeugherstellung) wurden während der geographischen Ausbreitung der LBK ständig weiterentwickelt (Müller 2009b). Innerhalb von etwa zweihundert Jahren breitete sich diese Kultur und damit auch das Neolithikum entlang der Donau in Richtung Rhein, nach Norden und nach Osten über große Teile Europas aus. Das Gebiet der LBK reichte von den Niederlanden, Belgien und dem Pariser Becken in Nordfrankreich bis nach Ungarn und den Norden der Ukraine, wo sie in der Waldsteppe an das Gebiet der Bug-Dnjestr-Kultur grenzte. Westlich des Rheins trafen Linearbandkeramiker mit Kulturgruppen wie La-Hoguette, Limburger und Villeneuve-Saint-Germain zusammen, die den äußeren Rand der mediterranen Ausbreitungsroute des Neolithikums bildeten (Müller 2009b, Jeunesse & Willigen 2010). Im Süden bildete das Alpenvorland den Abschluss. Nördlich der Lösszone² blieb Nordeuropa mesolithisch, vor allem repräsentiert durch die Ertebølle- und die Swifterbant-Kultur. Auch wenn ihre Subsistenz weiterhin wildbeuterisch blieb, begannen die Träger dieser Kulturen ab etwa 5 000 v. Chr. mit der Herstellung von Keramik. Außerdem sind im Fundgut Handelsbeziehungen zu den bäuerlichen Gruppen festzustellen (Price 2000, Müller 2009b). Das Erscheinungsbild der LBK ist im Großen und Ganzen sehr homogen, mit wenigen zeitlichen und regionalen Unterschieden (Gronenborn 1999, Price *et al.* 2001).

Ab etwa 4 500 v. Chr. ist die LBK als großräumige Kultureinheit Europas nicht mehr fassbar. Man unterscheidet mehrere regionale Nachfolgekulturen. Im Westen wurde

² Löss ist ein optimales Sediment für Ackerbau.

sie durch die Rössener Kultur (bzw. den Rössen-Großgartach-Komplex), im Osten durch die Stichbandkeramische und südlich davon durch die Lengyel-Kultur abgelöst. In ihrem östlichen Ausläufergebiet bildete sich im westpontischen Raum der Cucuteni-Tripolje-Komplex. In den beiden letztgenannten Regionen, dem balkanisch-karpatischen Raum, entwickelte sich dann bereits die Kupfermetallurgie (Müller 2009b).

Gegen Ende des fünften Jahrtausends v. Chr. setzte auch die Neolithisierung Nordeuropas und der Britischen Inseln ein. Im nördlichen Mitteleuropa und Skandinavien ist mit der Trichterbecherkultur (TBK) erstmals Ackerbau und Viehzucht nachgewiesen. Im Osten reichte die Ausbreitung einiger Gruppen mit Polen und der Ukraine an das Gebiet der Cucuteni-Tripolje-Kultur heran (Müller 2009b). Vorgreifend auf die populationsgenetischen Analysen dieser Arbeit sei hier erwähnt, dass sich die südliche Gruppe der TBK in Mitteldeutschland in mehrere aufeinanderfolgende Untergruppen unterteilen lässt: die Baalberger, Salzmünder und Bernburger Kulturen (Brandt *et al.* 2013). Mit der TBK gab es erstmals Hügelgräber in Mitteleuropa. Um etwa 4 000 v. Chr. verbreitete sich im Westen die Anlage von Megalithbauten (Müller 2009b). In den südosteuropäischen Kulturen des vierten vorchristlichen Jahrtausends, vor allem in der Cernavodă- und der frühen Badener Kultur, sehen die Archäologen wichtige Organe für die Verbreitung von Innovationen in Europa. Zu diesen gehören Rad, Wagen und Kupferverarbeitung (Müller 2009b). Die Trichterbecherkultur dominierte den mittel- bis nordeuropäischen Raum. Es gab aber weiterhin regionale Gruppen mit aneignender (mesolithischer) Lebensweise, wie z. B. Funde in der Blätterhöhle in Westfalen (Bollongino *et al.* 2013) oder die Grübchenkeramische Kultur in Nordeuropa und Skandinavien (Malmström *et al.* 2009) zeigen.

Zwischen 3 300 und 3 100 v. Chr. entwickelte sich nach und nach die Kugelamphorenkultur im Gebiet der östlichen Trichterbecherkultur zwischen Elbe und Dnjepr (Mallory & Adams 1997: 226 f, Szmyt 2001). Es handelte sich dabei um eine Kultur von Viehhirten, die wahrscheinlich aus dem Substrat der TBK entstand und zeitlich mit ihr überlappte. Ihre Mobilität, die Ähnlichkeit ihrer Keramik mit der der Majkop-Kultur im Nordkaukasus und der Unteren Mikhailovka-Kultur in der nordpontischen Steppe wird als Zeichen einer etwaigen indoeuropäischen Expansion

kontrovers diskutiert (Mallory & Adams 1997: 597 f). Die Kultur dauerte regional unterschiedlich lange, aber bis ungefähr 2 300 v. Chr. an (Szmyt 2001).

Der Beginn des späten Neolithikums datiert mit der aufkommenden Schnurkeramischen Kultur in die erste Hälfte des dritten Jahrtausends v. Chr. Kurz darauf folgte die Glockenbecherkultur. Diese beiden kulturellen Gruppen prägten Europa: die Glockenbecherkultur den Westen und die Schnurkeramik den Nordosten, wobei sich die Gebiete in Mitteleuropa überlappen. Letztere entwickelte sich wahrscheinlich auf mehreren Grundlagen, unter anderem der Trichterbecher- und der Kugelamphorenkultur. Es wird auch eine Beteiligung der nordpontischen Steppenkulturen an der Genese der Kultur diskutiert. Lange wurde davon ausgegangen, dass es sich bei den Schnurkeramikern um Nomaden aus dem Osten handelte. Mittlerweile wird die Kultur als eine autochthone Entwicklung betrachtet (Mallory & Adams 1997: 127 f, Häusler 2003: u. a. 9, 17, 20, 69-74, Müller 2009a). Vermutlich waren die Träger der Schnurkeramischen Kultur sehr mobil und betrieben extensive Viehwirtschaft. Ein besonderes Merkmal beider Kulturen ist die geschlechtsdifferenzierte Bestattung. In der Schnurkeramischen Kultur liegen Männer rechtseitig, Frauen linksseitig gehockt. In der Glockenbecherkultur ist es umgekehrt. Das mit der LBK in Mitteleuropa etablierte Dorfleben wurde von Einzelgehöften und kleinen Weilern abgelöst (Müller 2009b). Mit dem Verschwinden der Schnurkeramik und der Glockenbecherkultur gegen Ende des dritten Jahrtausends v. Chr. endete das europäische Neolithikum und die Ära der bronzezeitlichen Kulturen begann.

Zwar kann anhand des archäologischen Fundguts die Entwicklung und Ausbreitung der neolithischen Transition verfolgt werden, doch die Frage nach der Art und Weise der geographischen Expansion nach und durch Europa wird damit nicht beantwortet. Hierfür gibt es zwei grundlegende Ausgangsmodelle: Nach dem Akkulturationsmodell (*cultural diffusion*) entwickelte sich das Neolithikum aus der lokalen mesolithischen Kultur heraus. Ackerbau, Viehzucht und die zugehörigen Technologien wurden von einer lokalen Gruppe an die nächste weitergegeben, ohne dass es zu einer substantiellen Wanderung von Menschen kam. Dem gegenüber steht das Modell der demischen Diffusion (*demic diffusion*). Demnach entspricht die Ausbreitung des Neolithikums einem Populationswachstum mit räumlicher -expansion der

Bauerngesellschaften. Im ersten Fall wären die heutigen Europäer die Nachkommen von Jägern und Sammlern, die die Farmwirtschaft gelernt haben, im zweiten die der Bauern, die Europa im Neolithikum kolonisierten. Dazwischen liegt ein weites Spektrum an weiteren Modellen (Zvelebil 2000, Lichter 2005).

Zunächst wurde versucht, diese populationsgenetische Fragestellung mithilfe moderner genetischer Daten zu beantworten. Cavalli-Sforza *et al.* (1993) führten eine Hauptkomponentenanalyse mit Allelfrequenzen von 95 klassischen Markern in Europa durch und stellten bei der Übertragung auf eine synthetische Landkarte in der ersten Hauptkomponente einen Südost-Nordwest-Gradienten fest. Diesen interpretierten die Autoren als Signatur der Migration (*wave of advance*) neolithischer Bauern. Ihre Deutung ist jedoch umstritten, da dieses und auch weitere beobachtete Muster gleichermaßen auf andere demographische Ereignisse, mathematische Artefakte oder *allelic surfing*³ zurückzuführen sein könnten (Novembre & Stephens 2008, Francois *et al.* 2010). Moderne Daten zeigen, dass die genetischen Distanzen innerhalb Europas mit den geographischen korrelieren, und ein Plot der genetischen Variation der Landkarte Europas äußerst ähnlich ist (Novembre *et al.* 2008). Doch welche Anteile die drei großen demographischen Ereignisse – Erstbesiedelung, Rekolonisierung nach dem Letzten Glazialen Maximum und Neolithische Transition – an diesem Muster haben, kann allein mit modernen Daten nicht geklärt werden.

Durch Untersuchungen alter DNA, z. B. von Haak *et al.* (2005) und Bramanti *et al.* (2009) konnte das genetische *makeup* mitteleuropäischer Mesolithiker und Linearbandkeramiker charakterisiert werden. Ein Datensatz von 16 archäologisch als jungpaläo- und mesolithische Jäger und Sammler identifizierten mitteleuropäischen Individuen weist ausschließlich Typen der mtDNA-Haplogruppe U auf (Bramanti *et al.* 2009, Bollongino *et al.* 2013, Fu *et al.* 2013). In der mittlerweile auf 102 Individuen angewachsenen Sammlung von LBK-Sequenzen besteht das mitochondriale Haplogruppenspektrum aus N1a, T2, K, J HV, V, W, X und H (Haak *et al.* 2005, Haak

³ An der Front einer räumlichen Expansion kommt es aufgrund der Bewegungen der unterschiedlichen Kleingruppen zu mehrfachen lokalen *bottlenecks* mit anschließender Expansion, so dass einzelne Mutationen unter Umständen auf der Expansionswelle „reiten“. Da in diesem Fall die genetische Variation zwischen den Ausbreitungssektoren surfender Allele am größten ist, würde die erste Hauptkomponente rechtwinklig zur Expansionsrichtung verlaufen.

2006, Haak *et al.* 2010, Brandt *et al.* 2013, Brotherton *et al.* 2013). Die Diskontinuität beider Kulturen ist bereits anhand der Haplogruppen erkennbar und wurde auch auf der Nukelotidebene durch koaleszenzbasierte Simulationen bestätigt (Bramanti *et al.* 2009). Bollongino *et al.* (2013) zeigten am Beispiel des Fundplatzes Blätterhöhle, dass Nachkommen der Mesolithiker noch über 2 000 Jahre nach der Ankunft der Bauerngesellschaften ihren wildbeuterischen Lebensstil beibehielten. In wieweit dieser Fundplatz das allgemeine Bild widerspiegelt steht noch in Frage, doch auch in Skandinavien wurde dieses Phänomen beobachtet (Malmström *et al.* 2009).

Haak *et al.* (2010) fanden in ihrer ebenfalls auf mtDNA basierenden Studie eine hohe Affinität der LBK zu modernen Populationen des Nahen Ostens und Anatoliens. Insgesamt stützen die Ergebnisse der verschiedenen palaeogenetischen Studien ein Modell demischer Diffusion aus Südosten nach Mitteleuropa. Sie zeigen aber auch, dass die Linearbandkeramiker allein nicht die direkten Vorfahren der modernen Europäer sein können. Auch Vermischungsmodelle zwischen ausschließlich Mesolithikern und LBK erklären den modernen (maternalen) Genpool nicht. Demnach müssen weitere einflussreiche demographische Ereignisse im Verlauf des mittleren oder späten Neolithikums die europäischen Vorfahren genetisch beeinflusst haben (Bramanti *et al.* 2009, Haak *et al.* 2010, Brandt *et al.* 2013). Erst kürzlich wurde eine Studie von Lazarides *et al.* (2014) veröffentlicht, die anhand moderner und prähistorischer Genomdaten zeigten, dass sich der moderne europäische Genpool aus mindestens drei ancestralen Populationen zusammensetzt: den frühen Bauern aus dem Nahen Osten, den indigenen westeuropäischen Jägern/Sammlern und einer nordeurasischen Population jungpalaeolithischen Ursprungs, die auch zum Genpool amerikanischer Ureinwohner beitrug (Raghavan *et al.* 2014). Interessanterweise konnte eine nordeurasische Abstammung weder in west- und mitteleuropäischen Mesolithikern noch frühneolithischen Bauern nachgewiesen werden, dafür aber in skandinavischen Jäger/Sammlern. Der nordeurasische Anteil in den heutigen Europäern ist vermutlich im Verlauf des Neolithikums hinzugekommen.

In einer umfangreichen Studie an neun archäologischen Kulturen, die einen chronologischen Längsschnitt durch das mitteleuropäische Neolithikum in der Mittel-Elbe-Saale-Region (Sachsen-Anhalt) repräsentieren, identifizierten Brandt *et al.*

(2013) vier Verschiebungen in der genetischen Zusammensetzung, die sie mit populationsgeschichtlichen Ereignissen in Verbindung setzten: 1) die Ausbreitung der LBK, 2) die Neolithisierung Skandinaviens, 3) ein genetischer Einstrom aus Osten bei der Genese der Schnurkeramischen Kultur, 4) ein genetischer Zufluss aus Südwesteuropa (eventuell von der Iberischen Halbinsel) bei der Genese der Glockenbecherkultur. Bezüglich des vorliegenden Projekts ist das dritte vorgeschlagene Ereignis von besonderem Interesse. Könnten die nordpontischen Steppenkulturen der Ausgangspunkt der genetischen Einflüsse aus dem Osten sein?

4 Der west- und nordpontische Raum vom Neolithikum bis zur Bronzezeit

4.1 Das Untersuchungsgebiet

Im Fokus dieser Studie stehen archäologische Kulturen des westlichen Teils des eurasischen Steppengürtels vom Äneolithikum bis zur mittleren Bronzezeit, etwa 6 500 – 2 000 v. Chr. Der Untersuchungsraum reicht von der westlichen Schwarzmeerküste über den nordpontischen Raum bis nördlich des Kaspischen Meeres in das Gebiet der unteren und mittleren Wolga. Größtenteils ist die Schwarzmeerküste die natürliche südliche Begrenzung. Im heutigen Bulgarien reicht das Gebiet bis in die Oberthrakische Tiefebene. Östlich des Asowschen Meeres bildet der Nordkaukasus die Südgrenze (Abbildung 4.1).

Von West nach Ost teilen die großen Zuflüsse des Schwarzen und des Kaspischen Meers den Steppenraum in mehrere Gebiete, nach denen die Kulturen zum Teil benannt wurden. Die bedeutendsten sind dabei im Westen die Donau und ihr Nebenfluss, der Pruth, der entlang der heutigen rumänisch-moldawischen Grenze verläuft. Ungefähr parallel zum Pruth, teilweise entlang der moldawisch-ukrainischen Grenze, fließt der Dnjestr, der bei Odessa in das nordwestliche Schwarzmeer mündet. In der südwestlichen Ukraine fließt der Südliche Bug. Er teilt sich ein ausgeprägtes Mündungsgebiet mit dem Dnjepr. Dieser entspringt in Nordwestrussland, fließt durch Weißrussland, bevor er die heutige Ukraine in zwei Hälften teilt. Bis Kiew fließt er in südlicher Richtung, dann nach Südosten, bis er in den Steppengürtel eindringt. Dort findet sich das sogenannte Dnjepr-Knie. Ab hier fließt der Fluss etwa 90 Kilometer nach Süden, dann nach Westen, und mündet schließlich in südwestlicher Richtung ins Schwarzmeer. In dem nach Süden ausgerichteten Teilabschnitt befanden sich früher über eine Distanz von 75 km die Dnjepr-Stromschnellen, die den Fluss dort unbefahrbar machten (Kubijovyč 2001b). Weiter östlich folgt der Siwerskyj Donez (im Folgenden als Donez bezeichnet), der westliche Nebenfluss des Don. Der Don folgt einem ähnlich kurvigen Verlauf wie der Dnjepr und mündet in die Spitze des

Asowschen Meeres. Ungefähr spiegelbildlich dazu ist der Verlauf der Wolga. Sie und der Don laufen zunächst aufeinander zu, bis sie am Wolga-Knie lediglich 40 km voneinander getrennt sind. Dann fließt die Wolga in östlicher und südöstlicher Richtung zum Kaspischen Meer. Noch weiter östlich bildet der Ural den Grenzfluss zwischen dem europäischen und asiatischen Eurasien.

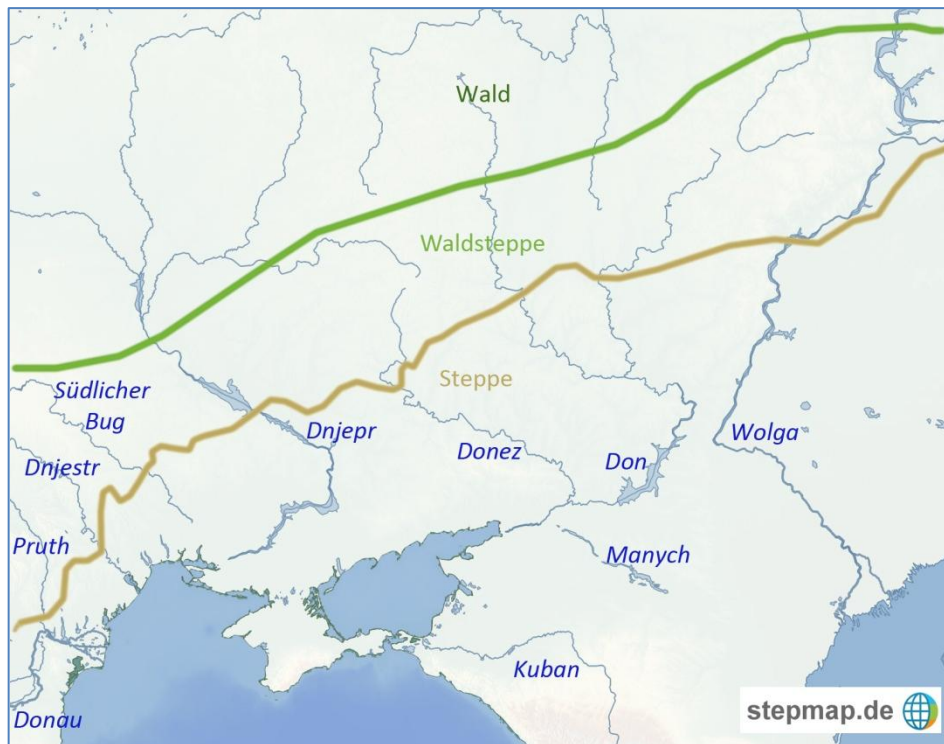


Abbildung 4.1: Übersichtskarte des Untersuchungsraums

Ungefäher Verlauf der Grenze zwischen Steppe und Waldsteppe (beige) und zwischen Waldsteppe und Waldzone (grün) mit Angabe der wichtigsten Flüsse (blau) (gezeichnet nach Kubijovyč 2001a, Stebelsky 2001, Wechler 2001).

Der nordpontische Raum ist von Süden nach Norden in drei natürliche Streifen untergliedert, der Steppenzonen, gefolgt von der Waldsteppe und dem Waldgürtel im Norden (Parzinger 1998a, Kubijovyč 2001a, Stebelsky 2001, Wechler 2001: 2 f). Die reine Steppe erstreckt sich bis etwas nördlich des Dnjepr-Knies entlang der Küste. Im Westen reichen die Ausläufer bis zum Karpatenbecken und den Gebieten um die untere Donau. In östlicher Richtung wird der Steppengürtel stetig breiter. Nördlich des Kaspischen Meeres verläuft die Steppen-Waldsteppengrenze etwas unterhalb der

Wolga-Schleife. Der Streifen der Waldsteppe verläuft annähernd parallel zur reinen Steppe, mit einer Breite von etwa 250 bis 300 km. Am Dnjepr beginnt der Waldgürtel in etwa auf der Höhe von Kiew. Der ungefähre Verlauf der Streifen ist in Abbildung 4.1 eingezeichnet. Die untersuchten Proben stammen fast ausschließlich aus dem Steppengürtel, einige wenige aus Übergangsbereich und dem südlichen Bereich der Waldsteppe.

Das Untersuchungsgebiet ist relativ flach und leicht wellig. Das Klima ist in der Steppenzone arid, mit Trocken- und Dürreperioden. Die Gebiete direkt nördlich der Krim und nordöstlich des Kaspischen Meeres, in der Kaspischen Senke, sind sogar wüstenähnlich. Auf der Krim und im Nordkaukasus beeinflussen die Gebirge das Klima. Abhängig von der Höhenlage reicht es von steppenartig und arid bis bewaldet und humid. Die Waldsteppenzone ist etwas kühler und erfährt nur im Hoch- bis Spätsommer eine kurze Trockenzeit. Hier verläuft der Übergang von semi-aridem zu semi-humidem Klima. Die gesamte Waldsteppen- und Steppenzone zeichnet sich aufgrund von pleistozänen Lössablagerungen durch fruchtbare Böden aus (siehe Parzinger 1998a, Wechler 2001: Kap. 2).

4.2 Die Archäologie des Nordpontikums

Das Bild der Prähistorie im nord- und nordwestpontischen Raum wurde jahrzehntelang von der 1956 von Maria Gimbutas erstmals aufgestellten Kurganhypothese geprägt. Demnach waren die Kurganvölker der Steppe berittene Nomaden, die auf kriegerische Weise die friedlichen, bäuerlichen Kulturen Südosteuropas verdrängten. Man spricht hierbei auch von drei bzw. vier Kurganwellen – also extensiven Migrationen, die vom frühen Äneolithikum bis zur Bronzezeit zur Ausbreitung dieser Völker führten (Gimbutas 1994). Archäologen wie Danilenko und Merpert hielten an der Grundidee einer Invasion aus dem Osten fest, entwickelten jedoch differenzierte Modelle zum Äneolithikum der Steppe. Beschreibungen und Gegenüberstellungen der einflussreichsten Modelle gibt Rassamakin in einem

umfassenden Aufsatz (1999: 59-67). Mittlerweile wurden alternative Ansichten von einem friedlichen und gegenseitigen Austausch zu den Beziehungen zwischen den Steppenvölkern und den bäuerlichen Nachbarn entwickelt. Während in früheren Modellen noch von einem kriegerischen Reiternomadentum die Rede war, wird die Bedeutung des Pferdes als Nutztier inzwischen in Frage gestellt.

Bis heute gibt es noch keine einheitliche Klassifizierung der nordpontischen Kulturerscheinungen. Sowohl die Namensgebung als auch die Charakterisierung von Lebensweisen, Ausbreitungen und Beziehungen variieren. Insbesondere in der relativen und der absoluten Chronologie gibt es unterschiedliche und teils widersprüchliche Ansichten. So sehen manche Autoren die Jamnaja als äneolithische Kultur an, andere ordnen sie bronzezeitlich ein (Rassamakin 1994, Parzinger 2006: 240). Ihr Auftauchen wird zwischen 3 600 (Mallory & Adams 1997: 651) und 2 600 v. Chr. (Masson 1998) datiert.

Die folgenden Ausführungen zu Äneolithikum und Bronzezeit orientieren sich am Modell von Rassamakin (1994, 1999). Wie alle anderen Modelle ist auch das von Rassamakin nicht unumstritten (z. B. Lichardus & Lichardus-Itten 1998). Es wird jedoch von vielen Archäologen anerkannt (E. K., persönliche Kommunikation) und sollte genügen, um einen groben Überblick über die Entwicklungen im nordpontischen Raum zu gewinnen. Ein Großteil der in dieser Arbeit untersuchten Proben stammen aus der frühen bis mittleren Bronzezeit (N = 121), einige aus dem Äneolithikum (N = 32). Dennoch soll hier auch das Neolithikum betrachtet werden, da sich Spuren aus dieser Zeit womöglich im Genom der bronzezeitlichen Steppenbewohner finden.

4.2.1 Das Neolithikum

Der Grundstein für die Entwicklungen im nordpontischen Raum wurde bereits in der dem Untersuchungszeitraum vorangehenden Neolithikum gelegt. Der Beginn des frühen Neolithikums wird mit dem Auftauchen erster Keramiken datiert. Die ältesten Tonwaren fanden sich in der Samara-Region an der mittleren Wolga, etwa 7 000 – 6 500 v. Chr. In der Steppe begann das keramische Neolithikum etwa 6 200 – 5 800

v. Chr. (Anthony 2007: 149). Während das Neolithikum in Europa primär durch die produktive Wirtschaftsweise definiert wird, gibt es für den hier betrachteten Untersuchungsraum keine ausreichenden Hinweise, um das akeramische Neolithikum eindeutig zu erfassen und vom Mesolithikum abzugrenzen.

Man unterscheidet Vertreter fünf verschiedener Kulturen, die den nordpontischen Raum während des Neolithikums besiedelten. Die älteste ist die Bug-Dnjestr-Kultur. Sie ist ab dem späten siebten oder frühen sechsten vorchristlichen Jahrtausend erfasst und war hauptsächlich in der Waldsteppe, in den Flusstälern zwischen dem Pruth und dem Südlichen Bug, verbreitet. Das Gebiet zwischen dem Asowschen Meer und dem Dnjepr-Knie gehörte der Sursk-Dnjepr-Kultur (Parzinger 1998a, 2006: 158). Später begann etwas nördlicher in der Steppe und Waldsteppe die Dnjepr-Donetz-Kultur, deren regionale Gruppen ein weites Gebiet zwischen den beiden Flüssen einnahmen. Sie entstand vermutlich in Folge von Kontakten zwischen ansässigen Mesolithikern, und Trägern der bereits keramisch neolithisierten Bug-Dnjestr- und Sursk-Dnjepr-Kulturen (Dergačev 1998). Entlang des Dons siedelten am Oberlauf im Wald und in der Waldsteppe die Stämme der Mittel-Don-Kultur, im Steppengebiet um den unteren Don die der Rakušečnyj Jar-Kultur. In der Region des nördlichen Kaukasus, an der unteren Wolga bis zur Mündung des Urals war die Seroglazov-Kultur verbreitet. Die Subsistenzwirtschaft in all diesen Kulturen beruhte zunächst auf Jagen, Sammeln und Fischfang (Wechler 2001: Kap. 13, Parzinger 2006: 158-161).

Der erste Impuls zu einem Wandel in der Wirtschaftsweise im Nordpontikum kam vermutlich aus dem Westen. Mit der Criș-Kultur erreichten ca. 5 800 – 5 700 v. Chr. erstmals Bauern vom Donautal mit Rindern und Schafen den nordwestpontischen Raum (Anthony 2007: 138-146). Sie gehörten dem frühneolithischen Starčevo-Körös-Criș-Komplex an, der entscheidenden Einfluss auf die Entstehung der Linearbandkeramischen Kultur, deren Träger die produzierende Lebensweise nach Mitteleuropa brachten, hatte (siehe hierzu Kapitel 3). Die Gruppen der Bug-Dnjestr-Kultur waren, wahrscheinlich aufgrund ihrer Kontakte zu den Bauern der Criș-Kultur, die ersten nordpontischen Stämme, die Viehhaltung und Ackerbau in ihr Wirtschaftssystem integrierten, wenn auch zunächst nur komplementär. Die Hauptversorgung basierte weiterhin auf einer aneignenden Wirtschaftsweise.

Nachdem die Criș-Kultur verschwand, wurden die westlichen Kontakte zur LBK anscheinend fortgeführt, wie sich in importierten und in nachgeahmten Fundstücken zeigt. Der Kontaktbereich beschränkte sich auf das Gebiet westlich des Südlichen Bugs (Wechler 1998). Erst in ihrer späten Phase, als die Bauern der frühen Tripolje-Kultur als Nachfolger der LBK in den östlichen Karpaten siedelte, gingen die Bug-Dnjestr-Gruppen zu einer überwiegend produzierenden Lebensweise über (Wechler 1998, 2001: 221). Bis ca. 5 200 v. Chr. bildete das Dnjestr-Tal eine Art kulturelle Grenze (Anthony 2007: 154). In den weiter östlich verbreiteten Kulturen gab es bis zu diesem Zeitpunkt keine gesicherten Hinweise auf Wirtschaftshaustiere oder Getreidekultivierung. Lediglich im Asowschen Gebiet gibt es zwei umstrittene Ausnahmen sehr früher Haustierfunde. Falls sich diese bestätigen ließen, würde dies auf Einflüsse des transkaukasischen Frühneolithikums hindeuten (Wechler 1998). Erst ab etwa 5 200 v. Chr. breitete sich nach und nach das Hirtentum in der pontisch-kaspischen Steppe aus. Die bäuerliche Tripolje-Kultur drang nach Osten bis zu den Tälern des Südlichen Bugs vor. Ungefähr zu dieser Zeit verschwand die Bug-Dnjestr-Kultur. Auch die Steppenbewohner am Dnjepr begannen Rinder und Schafe zu halten. Gegen Ende der Dnjepr-Donetz-Kultur erreichte der Haustieranteil bis zu 80 % (Parzinger 2006: 161). Getreideanbau spielte aber weiterhin keine oder kaum eine Rolle (Wechler 1998). Über das Neolithikum auf der Halbinsel Krim gibt es nur wenige Daten. Wahrscheinlich begann es eher spät, während des nordpontischen Äneolithikums (Wechler 2001: 237, Motuzaitė-Matuzevičiute *et al.* 2013). Um etwa 5 000 v. Chr. gab es auch im Nordkaukasus Viehherden. In der Wolga-Region blieb das Volk der Seroglazov-Kultur bis zu ihrem Verschwinden in ihrer Lebensweise mesolithisch (Parzinger 2006: 159). Domestizierte Tiere erreichten die Wolga-Ural-Region erst gegen Ende des Neolithikums, etwa 4 700 – 4 600 v. Chr., während es bis 4 200 v. Chr. so gut wie keine Hinweise auf Getreidekultivierung gibt (Anthony 2007: 174). Im Großen und Ganzen deutet alles darauf hin, dass sich die Entwicklung des Neolithikums in den einzelnen Regionen des nordpontischen Raums jeweils auf Grundlage des lokalen Mesolithikums vollzog, mit regional unterschiedlichen Einflüssen aus Europa, Transkaukasien und Mittelasien (Wechler 2001: Kap. 13).

Die bisherigen Erkenntnisse zum Neolithikum in der Steppe beruhen ausschließlich auf archäologischem Fundgut. Der genetische Hintergrund dieser Entwicklungen, wie z. B. die Frage nach der biologischen Identität der Menschen, die begannen, in der Steppe Vieh zu halten, ist noch nicht erforscht. Es könnte sich um eingewanderte Bauern handeln oder um indigene Bevölkerungsgruppen, die von ihren Nachbarn die Neuerungen übernommen haben, also rein durch einen kulturellen Austausch.

4.2.2 Das Äneolithikum

Rassamakin (1999) datiert den Beginn des frühen Äneolithikums auf etwa 4 550 v. Chr., Manzura (2005) etwa 200 Jahre früher. Während die in den nördlichen Waldsteppen- und Wald-Gebieten angesiedelten Gruppen der Dnjepr-Donetz- und die Mittel-Don-Kultur sich weiter entwickelten, wurden die neolithischen Waldsteppen- und Steppenulturen Sursk, Rakušečnyj Jar und Seroglazov von kupferverarbeitenden Kulturen abgelöst. Diese wurden zunächst unter dem Begriff Srednij Stog II zusammengefasst. Von den meisten Autoren wird dabei die verwandte Khvalynsk-Kultur separiert, die der Seroglazov-Kultur an der unteren Wolga nachfolgte (Parzinger 2006: 162, 165). Mit zunehmenden Ausgrabungen wurden später die Gruppen Untere Mikhailovka zwischen dem unteren Dnjepr und dem Asowschen Meer und Novodanilovka zwischen dem Südlichen Bug und dem Asowschen Meer aus der Srednij Stog II-Erscheinung ausgegliedert. Im Verlauf der Forschungsgeschichte tauchten noch diverse andere Gruppierungsmodelle auf. Rassamakin (1994, 1999) kritisiert die Unklarheiten und Widersprüche in der Nomenklatur und der Charakterisierung der äneolithischen Gruppen. Ihm zufolge wird der Srednij Stog II-Begriff den vielen lokalen Differenzierungen in der Steppe nicht gerecht. Er stellt die Srednij Stog II als Kultur in Frage und identifiziert in ihrem Verbreitungsgebiet vier Gruppen, die er, basierend auf Keramik und Bestattungsritus, als eigenständige Kulturen klassifiziert: Skelanska, Stogovska, Kvitanska und Dereivka⁴. In Anlehnung an

⁴ In seiner Abhandlung (1999) ersetzt Rassamakin die in seinem Artikel von 1994 verwendeten Bezeichnungen *Skelanska*, *Kvitanska* und *Nizhnemikhailovka* durch *Skelya*, *Kvityana* und *Lower Mikhailovka Culture*. Die *Stogovska Culture* wird nicht mehr explizit benannt, sondern es ist lediglich von *Sredny Stog II*-typischen Bestattungen die Rede.

das Modell von Danilenko können diese Kulturen als Basis für die Entstehung der Jamnaja-Kultur angesehen werden (Rassamakin 1994: 44). Ihre zeitliche Einordnung erfolgt vor allem durch ihre Synchronisation mit den gut dokumentierten Phasen der Tripolje-Kultur in Kombination mit Radiokarbondatierungen. Wie auch für das Neolithikum stellt sich die Frage, ob es sich bei diesen Gruppierungen tatsächlich um eigenständige Populationen handelte oder eine überregional genetisch vermischte Bevölkerung mit regionalen, kulturellen Besonderheiten.

Das frühe Äneolithikum

Die von ca. 4 500 – 4 100 v. Chr. datierte Skelanska-Gruppe repräsentiert die älteste frühäneolithische Steppenkultur. Sie entstand in der Region zwischen Dnjepr und Don und ersetzte die großen Mariupol⁵-typischen flachen Friedhöfe des Neolithikums (Rassamakin 1999: 75-83, 97). Ihre westlichen Ausläufer reichten bis in das Gebiet zwischen Dnjestr und Donau. Die Gruppen hier werden auch als Suvorovo-Gruppen bezeichnet. Die Bestattungen des am Unteren Dnjepr verbreiteten Novodanilovka-Typs werden von Rassamakin in den Skelanska-Begriff integriert. Die Toten wurden einzeln in ovalen Gruben, zum Teil in Kisten, in gehockter Rückenlage bestattet. Die Orientierung war meist ostwärts. Die Körper wurden mit reichlich Ocker bedeckt und es gab reiche Grabbeigaben, insbesondere Kupfergegenstände. Üblich waren Flachgräber, aber in der Suvorovo- und der Novodanilovka-Gruppe gab es auch erste kleine Aufschüttungen (Häusler 1998).

Nördlich des Kaspischen Meeres entwickelte sich parallel die Khvalynsk- aus der Seroglazov-Kultur. Sie erstreckte sich im Gebiet der unteren Wolga bis zur Waldsteppengrenze. Die Bestattungen erfolgten in Flachgräbern in gehockter Rücken-, seltener in der Seitenlage. Der Ritus weist große Ähnlichkeiten mit dem in der Skelanska-Kultur auf – insbesondere die Grabbeigaben – was für eine Verwandtschaft

⁵ Eine alternative Klassifizierung der hier beschriebenen, östlich des Dnepr bis zur Wolga-Ural-Region verbreiteten Kulturen in der Übergangszeit zum Äneolithikum findet sich unter dem Oberbegriff Mariupol-Gemeinschaft. Unter diesem werden die Aow-Dnjepr-, die Kaspische, die Unterdon- und die Samara-Kultur zusammengefasst (Dergačev 1998). Äneolithische Kulturen werden in der Literatur oft als Post-Mariupol bezeichnet.

der beiden Gruppen spricht. In der Keramik hingegen sind Unterschiede festzustellen (Parzinger 2006: 165).

Historisch bedingt wurde in Bezug auf die Wirtschaftsweise im nordpontischen Raum während des Äneolithikums *a priori* von einer Form mobilen Hirtentums ausgegangen. Diese wurde dann pauschal als saisonal, semi- oder voll-nomadisch bezeichnet, ohne die demographischen Aspekte genauer zu untersuchen. Mittlerweile hat man sich von diesem Konzept gelöst und versucht, das ökonomische System der Steppe basierend auf dem Fundmaterial neu zu rekonstruieren (Rassamakin 1999: 130). Für eine sichere Charakterisierung der Wirtschaftsweise im frühen Äneolithikum gibt es bisher nicht genug eindeutig zuordenbares Fundmaterial. Auf Grundlage der Komposition der Tierfunde in der Siedlung Svobodnoe nördlich des Kaukasus und diversen Einzelfunden in den Gebieten der Skelanska- und der Khvalynsk-Kultur, erlaubt sich Rassamakin (1999: 133) lediglich eine grobe Einschätzung. Er geht von einer sesshaften Lebensweise aus, in der es zwar domestizierte Tiere gab, aber in der die Jagd weiterhin eine bedeutende Rolle spielte. Von den neolithischen Nachbarn wurden nur ausgewählte Innovationen übernommen (Manzura 2005).

Im Nordwesten, in der Waldsteppe zwischen Pruth und Dnjepr, entwickelte sich im frühen Äneolithikum die Cucuteni-Tripolje-Kultur zur größten und am weitesten fortgeschrittenen agrikulturellen Gesellschaft des kupferzeitlichen Europas (Anthony 2007: 279 ff, Parpola 2008). Dies zeigte sich in weiträumigen Siedlungen und einer hohen Bevölkerungsdichte, deren Versorgung durch Ackerbau und Viehzucht gesichert wurde. Im nördlichen Balkangebiet und im Karpatenbecken befanden sich die ersten Kupferabbau- und -verarbeitungscentren. Die Bevölkerung der Cucuteni-Tripolje-Kultur importierte Kupfer und stellte Gebrauchsgegenstände und Ornamente her.

Von dem Gebiet der Cucuteni-Tripolje-Kultur aus gelangten Kupfer und aus Kupfer gefertigte Gegenstände in die Steppe. Analogien in der Keramik sowie importierte Fundobjekte zeugen von Kontakten zwischen den Trägern der Tripolje- und der Skelanska-Kultur (Dergačev 1998, Rassamakin 1999: 100-112). Im Kerngebiet der Skelanska-Kultur tauchten auch die ersten Artefakte auf, die als Prestigeobjekte oder Statussymbole gedient haben mochten (z. B. Tierkopfzepter und Schmuckstücke).

Nach und nach tauchten solche Objekte in der gesamten nordpontischen Steppe auf. Rassamakin spricht der Skelanska-Kultur die Schlüsselrolle in der Verbreitung des „Prestige-Handelssystems“ in der nordpontischen Region zu. Ihre „Elite“ vermittelte sozusagen zwischen der agrikulturellen Welt Südosteuropas und der Steppe bis einschließlich des Wolga-Gebietes und integrierte auch das Kaukasusvorland. Die in den beiden letztgenannten Gebieten gefundenen Gegenstände enthalten Kupfer, das aus den Abbaugebieten des Karpaten-Balkan-Raums stammte. Damit repräsentieren sie die Peripherie der von Chernykh (2008) beschriebenen *Carpatho-Balkan Metallurgical Province (CBMP)*.

Darüber wie das Material über solch weite Strecken übermittelt wurde, kann bisher nur spekuliert werden. Es gibt keine verlässlichen Daten zum Transportwesen in der Steppe. In den westlichen agrikulturellen Kulturen gibt es Hinweise auf Schlitten und Wagen, z. B. in Form von Ton-Modellen. Möglicherweise initiierten die Träger dieser Kulturen das Transportwesen in der Steppe, wobei Ochsen als Zugtiere gedient haben könnten (Rassamakin 1999: 137). Pferde hingegen treten als Prestige-Symbole wie Zeptern oder Statuetten auf. Möglicherweise gab es im frühen Äneolithikum erste Versuche, Pferde zu fangen und zu zähmen, aber von einer Pferdedomestikation oder gar einem Reiternomadentum kann in der Steppe noch nicht gesprochen werden (Rassamakin 1999: 134, 139). Zur Zeit der Skelanska-Kultur gewann die Viehhaltung im Vergleich zum Neolithikum an Bedeutung, da sie sicherlich eine Rolle im Prestigesystem spielte.

Das frühe Äneolithikum kann durch die Entwicklung einer sozialen Stratifikation innerhalb der Steppenkulturen charakterisiert werden. Es scheint, als wurde sie durch die Kontakte zwischen den agrikulturellen Gruppen (inklusive der Kupferabbauregionen) und denen der Skelanska-Kultur initiiert, und durch letztere in der Steppe durch weitläufigen Handel von Prestigeobjekten verbreitet. In der Archäologie zeigt sich dies vor allem in den Veränderungen im Bestattungsritus. Zwar dominierten Flachgräber, aber es gab auch die ersten Kurganaufschüttungen, und bei den Grabbeigaben sind deutliche sozial geprägte Unterschiede festzustellen (Parzinger 2006: 166 f).

Um etwa 4 100 v. Chr. begann, was Rassamakin (1999: 112, 122) als „*Steppe 'Hiatus'*“ bezeichnet, eine ungefähr drei Jahrhunderte andauernde Phase, aus der es keine Funde gibt, und in der die Entwicklungen in der Steppe unterbrochen zu sein scheinen. Falls es in dieser Zeit die für die Skelanska-Kultur vermutete soziale Differenzierung weiterhin gab, so ist sie zumindest im Grabbrauch nicht zu erkennen. In der Stogovska-Kultur, die als Nachfolgekultur oder eher noch als Spätphase der Skelanska-Kultur angesehen wird, sind selten Grabbeigaben zu finden, und wenn, dann handelt es sich um Alltagsgegenstände. Die Stogovska-Kultur existierte wahrscheinlich bis etwa 3 600 v. Chr. (Rassamakin 1994: 37). Was genau während dieser Phase der Fundlücke passierte, ist unklar. Genetische Daten von vor und nach dieser Zeit könnten Hinweise geben, ob es hier möglicherweise zu einem Umbruch in der Bevölkerung kam.

Das mittlere Äneolithikum

Das Ende des Hiatus um ca. 3 800 v. Chr. markierte den Beginn des mittleren Äneolithikums im nordpontischen Raum. Wie auch im frühen Äneolithikum sind die Entwicklungen im nordpontischen Raum von den Beziehungen zu den benachbarten agrikulturellen Gruppen geprägt. Die bisher einheitliche Cucuteni-Tripolje-Kultur teilte sich auf in einen westlichen (Cucuteni) Teil zwischen Dnjestr und Pruth und einen östlichen (Tripolje) Teil in der Bug-Dnjepr-Region mit jeweils eigenständiger Keramik (Rassamakin 1999: 112 f). Traditionell wird das Zerfallen der Cucuteni-Tripolje-Kultur als Folge einer Bedrohung oder sogar Invasion aus der Steppe gesehen (Gimbutas 1956, Dergačev 1998, Anthony 2007). Rassamakin (1999: 114) interpretiert die Daten umgekehrt, nämlich, dass die immer weiter in die Steppe vordringenden Gruppen der Tripolje-Kultur die dortigen Entwicklungen vorantrieb. Als Hinweise hierfür dienen ihm vor allem die weitverbreiteten Tripolje-Importe im Steppenraum. Sollte zwischen den beiden Kulturräumen Genfluss stattgefunden haben, könnte dessen Richtung zur Überprüfung und Klärung dieser unterschiedlichen Sichtweisen beitragen.

Im nordpontischen Raum begannen sich im mittleren Äneolithikum neue Kulturen zu bilden, die ab etwa 3 700 v. Chr. zuverlässig fassbar sind. Die Stämme der

Kvitanska-Kultur waren sowohl in der Steppe als auch der Waldsteppe um den mittleren Dnjepr ansässig. Sie breiteten sich nach Westen bis hin zum Südlichen Bug aus. Möglicherweise erreichten sie sogar die Donau, wo man in Kurganen der Usatovo- und der Unteren Mikhailovka-Kultur auch Kvitanska-Bestattungen identifizierte. Im Osten sind sie bis zum Don nachgewiesen (Rassamakin 1994: 39 f, 1999: 83-87). Ihre Keramik weist auf starke Einflüsse der Stogovska-Kultur hin. Wahrscheinlich gab es schon eine frühere Phase der Kvitanska-Kultur, parallel zur Stogovska. Zum Bestattungswesen gehörten Kurganaufschüttungen mit mehreren schmalen Gräbern. Die für das Neolithikum typische gestreckte Bestattung in Rückenlage, meist in Richtung Osten orientiert, wurde wieder praktiziert. Schmuckstücke und Werkzeuge aus Kupfer wurden beigegeben. Feuer scheint eine bedeutende Rolle im Ritual gespielt zu haben. Für eine Beurteilung der Wirtschaftsweise der Kvitanska-Bevölkerung ist die Datengrundlage zu gering. Die Lage der Fundplätze entlang des Dnjeprs und anderen Flüssen deuten auf eine Subsistenz wie im Neolithikum hin. Die Kurgane, die in der späteren Phase in der Asowschen Steppe und im Westen (weg von den Flüssen) errichtet wurden, geben Hinweise auf eine Haltung von Tieren (Rassamakin 1999: 143).

Im mittleren Äneolithikum war die Dereivka-Kultur, die sich aus den neolithischen Gruppen herausbildete, ausschließlich in der Waldsteppenzone verbreitet. Im Süden koexistierten ihre Stämme mit denen der Kvitanska-Kultur im Gebiet um und zwischen Dnjepr und nördlichen Donez. Aufgrund der unzureichenden Datengrundlage ist nur eine sehr grobe Datierung der Dereivka-Kultur auf ca. 3 700 bis 3 150 v. Chr. möglich (Rassamakin 1994: 40 ff). Dereivka gilt als eine der am ausgiebigsten erforschten Siedlungen, aber auch als eine der problematischsten. Eine Vielzahl an Funden von Pferdeknochen sowie ein als Pferdetrense identifizierter Fund führten zur umstrittenen Annahme einer frühen Pferdedomestikation. Es ist allerdings unklar, ob es sich bei den Knochenfunden um wilde oder domestizierte Pferde handelt. Auch lassen jüngere Objekte im Fundgut an der Authentizität zweifeln (Anthony & Brown 2000, Parzinger 2006: 164). Während z. B. Lichardus und Lichardus-Itten (1998) bereits für das frühe Äneolithikum von Reitpferden und Pferdeherden sprechen, schlussfolgert Rassamakin (1999: 147) aus dem Vorhandensein von Pferdeknochen

lediglich, dass das Pferd gejagt und als Fleischquelle verwendet wurde. Als domestizierte Tiere wurden hauptsächlich Rinder, aber auch Schweine, Schafe und Ziegen gehalten. Hacken und andere Werkzeuge lassen auf eine primitive Landwirtschaft schließen. Über das Grabwesen ist außer dem Vorhandensein von Kurganaufschüttungen kaum etwas bekannt (Parzinger 2006: 164).

Die Untere Mikhailovka-Kultur zwischen 3 700 und 3 000 v. Chr. beschränkt sich auf den südlichen Steppenraum von der Donau bis zum unteren Don. Das Gebiet nördlich und östlich des Asowschen Meeres ist hierbei noch nicht gut genug erforscht, um eine genaue Grenze festzulegen (Rassamakin 1999: 92). Wie auch bei den vorigen Kulturen ist ein früheres Auftreten denkbar, aber nicht sicher nachzuweisen. Im Überlappungsgebiet mit der Kvitanska-Kultur ist im Fundgut eine Vermischung der Eigenschaften der beiden Kulturen zu erkennen. Anders als bei der Kvitanska-Kultur wurden die Toten der Unteren Mikhailovka-Kultur in gehockter Seitenlage in ovalen Gruben bestattet. Die Kurgane wurden von Gräben umringt. Der Charakter der Grabrituale war sehr einheitlich mit lokalen Besonderheiten (Rassamakin 1994: 42). Das Wirtschaftssystem der Unteren Mikhailovka-Bevölkerung beruhte vor allem auf Tierhaltung, wobei Schaf und Ziege dominierten. Getreideeindrücke an Gefäßen weisen auf agrikulturelle Aktivitäten, gefördert von der direkten Nachbarschaft der nach Süden gewanderten Tripolje-Gruppen, hin. Der Wildtieranteil ist in den Fundplätzen eher gering (Rassamakin 1999: 142 f). Rassamakin und Danilenko sehen die Skelanska-, Stogovska-, Kvitanska- und Dereivka-Kulturen als Vorläufer bei der Formation der Jamnaja-Kultur, während die Untere Mikhailovka entlang der Küstenlinie eine separate Entwicklung im Steppenäneolithikum darstellte (Rassamakin 1994: 44). Ob diese als distinkte Population im genetischen Sinne anzusprechen ist, gilt es noch zu untersuchen.

In der Waldsteppe um den mittleren Don entwickelte sich die Repin-Kultur. In ihrer späten Phase breitete sie sich in den Regionen des unteren Dons und der Wolga sowie in Richtung Dnjepr aus. Ähnlichkeiten in der Keramik lassen darauf schließen, dass die Repin-, genau wie die Kvitanska-Kultur, in der Dnjepr-Region unter dem Einfluss der Stogovska-Kultur entstand (Rassamakin 1999: 117). Bestattungen waren in östlich ausgerichteter gehockter Rückenlage. Megalithen gehörten bei der Repin-Kultur zur

Kurgantradition. Bezüglich ihrer Lebensweise beschreibt sie Rassamakin (1999: 151) als Jäger und Viehhalter.

Bei der Konstantinovka-Kultur in der Steppe am unteren Don handelt es sich um eine Fortführung der Skelanska-Tradition, wobei hier Kurgane errichtet wurden. Die Grabbeigaben waren reichlich und enthielten unter anderem Prestige-Objekte, die allerdings nicht aus dem Westen sondern aus der Region der Majkop-Kultur stammten. Die Majkop-Kultur erschien im mittleren Äneolithikum in Norden des Kaukasus. Ihr Einfluss reichte jedoch nicht weiter als bis zum unteren Don. Die Tonwaren der Konstantinovka-Kultur weisen auf Kontakte zu Vertretern der Kvitanska-Kultur hin (Rassamakin 1999: 117, 122). Ihre Lebensweise beruhte sowohl auf der Haltung von hauptsächlich Schafs- und Ziegenherden, als auch auf der Jagd. Auch Fisch gehörte zur Ernährung. Hinweise auf agrikulturelle Aktivitäten der Konstantinovka-Stämme gibt es jedoch trotz ihrer Verbindungen zu denen der Majkop-Kultur keine. Diese hatten ein komplexes Wirtschaftssystem mit Ackerbau und Rindern, die in Ställen gehalten wurden (Rassamakin 1999: 149).

Für die frühäneolithischen Khvalynsk-Kultur in der Wolga-Region gibt es Hinweise, dass sie bis ins mittlere Äneolithikum fortbestand (Rassamakin 1999: 122). Auf der anderen Seite des Untersuchungsraums, im unteren Donau-Tal westlich des Schwarzmeeres, verschwand die Gumelnița-Kultur und die Cernavodă I-Kultur kam auf. Da nun plötzlich Steppenelemente in dieser Region zu finden sind, entstand die Kultur möglicherweise durch die Assimilation einheimischer Gruppen und einwandernden Migranten der Suvorovo-Gruppe (Dergačev 1998, Anthony 2007: 260).

Das späte Äneolithikum

Mit Beginn des späten Äneolithikums um 3 500 v. Chr. zerfiel der Tripolje-Komplex vollständig in kleine, unabhängige sogenannte Post-Tripolje-Gruppen. Südwärts bildete sich am unteren Dnjestr die Usatovo-Gruppe in der Steppe (Häusler 1998). Einflüsse aus der Steppe, wahrscheinlich von der Unteren Mikhailovka-Kultur sind deutlich zu erkennen. So gab es sowohl Flachgräberfelder als auch

Kurganbestattungen. Die Menschen der Usatovo-Kultur bestatteten ihre Toten jedoch nicht in der verbreiteten Rückenhocklage, sondern linksseitlich gehockt. Hauptsächlich in den Kurganen finden sich bereits Artefakte aus arsenhaltiger Bronze. Viehzucht schien in der Usatovo-Kultur eine bedeutende Rolle gespielt zu haben (Parzinger 1998a, 2006: 240). Die Tripolje-Splittergruppe in den Gebieten des mittleren Dnjestr, Pruth und Südlichen Bugs war die Gordenŝti- (Kasperovka-) Gemeinschaft. In der Waldsteppe des mittleren Dnjepr lebte die Vertreter der Sofievka-Gruppe (Rassamakin 1999: 123).

Die Kulturen des mittleren Äneolithikums blieben bestehen, doch ihre Verbreitungsgebiete wurden kleiner. So ist z. B. die Kvitanska-Kultur nur noch in ihrer ursprünglichen Kernzone am Dnjepr zu fassen. In den Kontaktzonen mit den Tripolje-Gruppen zwischen den Gebieten der Kvitanska- und der Unteren Mikhailovka-Kultur sind vermischte Charakteristika zu beobachten. Auch die Dereivka-Kultur weist zunehmende Tripolje-Einflüsse auf (Rassamakin 1999: 122 f).

Rassamakin identifiziert zwei Migrationswellen gegen Ende des Äneolithikums. Die erste zeigt sich durch eine Reihe von Bestattungen in der Steppe und Waldsteppe vom Dnjestr bis zum Samara, die als Životilovka-Volčansk-Gruppe bezeichnet werden. Sie finden sich in kleinen Ansammlungen entlang der Flüsse. Die Toten wurden in bereits vorhandenen Kurganen nachbestattet. Sie wurden in extrem gehockter Seitenlage und westlicher Ausrichtung in rechteckige Gruben, teilweise mit einem Absatz, gelegt. Als Ausgangspunkt für die Migration wird die Gordenŝti-/Kasperovka-Gruppe vermutet, deren Keramik das Fundgut der Životilovka-Volčansk ausmacht (Rassamakin 1994: 45, 1999: 92, 97). Durch die Migration wurde eine Verbindung durch den beinahe gesamten nordpontischen Waldsteppen- und Steppenraum geschaffen. In diesem Zuge wurden auch Elemente der späten Majkop-Kultur des Nordkavkasus in die Steppenzone getragen. Die Kontakte zwischen der Životilovka-Volčansk- und der Majkop-Kultur führten möglicherweise zur Ausbildung der Novosbodnaja-Gruppe innerhalb der Majkop (Rassamakin 1999: 123 ff).

Die zweite von Rassamakin erfasste Migration am Ende des Äneolithikums ist durch die Verbreitung der Repin-Kultur, ausgehend vom mittleren Don nach Norden zum

oberen Don, nach Südwesten in die Dnjepr-Region und nach Süden zum unteren Don und der unteren Wolga, repräsentiert (Rassamakin 1999: 125). Die Konstantinovka-Kultur verschwand, während sich die Gebiete der Kvitanska- und der Dereivka-Kultur verkleinerten. Manche Forscher sehen die Repin-Kultur als eine frühe Phase der Jamnaja-Kultur (z. B. Parzinger 1998a). Rassamakin hingegen sieht sie als eigenständige Kultur, die lokal an der Bildung der Jamnaja-Kultur beteiligt war. Die beiden Migrationshypothesen basieren auf Charakteristiken des archäologischen Fundguts, z. B. Besonderheiten in der Keramik und den Bestattungen. Möglicherweise gibt es genetische Spuren, die die Hypothesen stützen, dass hier tatsächlich Menschen gewandert sind.

Im späten Äneolithikum passten die Gruppen der zerfallenden Tripolje-Kultur ihre Subsistenzwirtschaft den lokalen Bedingungen an. In der Waldsteppe wurde weiterhin Landwirtschaft betrieben und Viehhaltung und Jagd verstärkt. Die Usatovo-Stämme in der Steppe entwickelten eine Ökonomie, die auf Schafzucht, Rinderherden und der Kultivierung einer Reihe von Getreiden basierte. Gejagt wurde kaum noch (Rassamakin 1999: 149). Sie ähnelten damit sehr der benachbarten Unteren Mikhailovka-Kultur. Die Vertreter der östlicheren Kulturen behielten ihre gemischte Wirtschaftsweise bei, vornehmlich Schaf- und Ziegenhaltung, in Kombination mit Jagd. In diese Zeitscheibe fallen auch die frühesten Funde von hölzernen Rädern als Grabbeigaben in der Steppe. Rassamakin (1999: 151) sieht eine Verbindung zwischen der Životilovka-Volčansk-Migration und der Verbreitung von Transport auf Rädern im nordpontischen Raum. Gleichzeitig tauchten sowohl in der Steppe als auch Waldsteppe vermehrt Spinnwirtel aus Ton auf.

Insgesamt charakterisiert Rassamakin das späte Äneolithikum als eine Phase des Zusammenbruchs des sesshaften agrikulturellen Wirtschaftssystems und einer Hinwendung zu mobilen Formen der Viehhaltung.

4.2.3 Die Bronzezeit

Die frühe Bronzezeit

Die Životilovka-Volčansk- und die Repin-Kultur markieren die Übergangsphase vom Äneolithikum in die Bronzezeit. In dieser Zeit, ab etwa 3300 v. Chr., begann sich auch die frühe Jamnaja-Kultur im Steppengürtel herauszubilden. Auch wenn sie sich deutlich von den äneolithischen Kulturen unterscheidet, ist ihr Ursprung laut Rassamakin (1994, 1999: 125-127) innerhalb des nordpontischen Raums zu suchen, und nicht auf einen Populationswechsel durch eine Migration aus Osten, wie teilweise postuliert wurde, zurückzuführen. Sie verbreitete sich zunächst zwischen dem Südlichen Bug und unterer Wolga (Parzinger 1998a). Spätestens mit der entwickelten Jamnaja-Kultur um 3000 v. Chr. wird von der frühen Bronzezeit gesprochen. Die letzten äneolithischen Kulturen wurden mittlerweile abgelöst. Die Träger der Jamnaja-Kultur weiteten ihr Gebiet über die Steppe hinaus aus. Es umfasste nun auch die Waldsteppenzone und die Region nördlich des Kaukasus. Im Westen siedelten sie bis nach Bulgarien und das untere Donaugebiet. Einige Gruppen kamen bis nach Ungarn ins östliche Karpatenbecken (Kalicz 1998, Parzinger 1998a, 2006: 240, Gerling *et al.* 2012). Manche Forscher sehen hier eine Migration, die zur Indogermanisierung Europas beitrug (Gimbutas 1956, 1994, Dergačev 1998, Anthony 2007: 305 f). Hingegen gibt es laut Häusler (1998) keine Hinweise auf einen andauernden Einfluss der Kultur in dieser Region. Das Wolga-Ural-Gebiet wird als Ausgangspunkt der Jamnaja-Kultur angenommen, von wo aus sie sich rapide nach Westen ausbreiteten (Masson 1998). Die am ältesten datierten Jamnaja-Fundstellen liegen aber nicht nur um die Wolga, sondern auch in der gegenüberliegenden Peripherie, an der nordwestlichen Schwarzmeerküste. Die geographisch dazwischen liegenden Datierungen sind jünger (Chernykh 2008). Laut Rassamakin (1999: 125) ist nicht festzustellen, ob die Jamnaja-Kultur sich von einem Ursprungsort ausbreitete. Bisher fehlt noch jeglicher populationsgenetischer Ansatz zur Überprüfung der verschiedenen archäologischen Hypothesen zum Ursprung und der Verbreitung dieser Kultur. Von besonderem Interesse dürfte die Natur der archäologisch aufgezeigten Kontakte zu den westlicheren Kulturen Europas sein.

Charakteristisch für die Jamnaja-Kultur ist der einheitliche Bestattungsritus, der scheinbar schlagartig im gesamten nordpontischen Raum praktiziert wurde. Es handelte sich um eine reine Kurgankultur. Flachgräber, wie es sie in den vorigen Kulturen noch gab, wurden nicht mehr angelegt. Die Kurganarchitektur an sich wurde mit dem Erscheinen der Jamnaja einfacher und uniform. Es gab keine komplexen Strukturen wie Megalithen oder Gräben mehr (Rassamakin 1999: 127). In den Kurganen wurden die Toten in rechteckigen Gruben bestattet, die mit verschiedenen Materialien, z. B. Holz oder Stein, abgedeckt wurden. Die Gruben der Nachbestattungen in bereits verwendeten Kurganen wurden meist kreisförmig um deren Mitte angeordnet, wobei die Ausrichtung der Toten von der Lage des Grabes im Kurgan abhing (Nikolova 2007). Vorherrschend war zunächst die gehockte Rückenlage in der Totenstellung, dann kamen auch seitliche Hockstellungen hinzu. Wie auch schon teilweise im Äneolithikum praktiziert, wurden die Toten und die Grubengräber oft mit gemahlenem Ocker eingefärbt. Daher wird die Jamnaja-Kultur⁶ auch als Gruben- oder Ockergrabkultur bezeichnet. Insgesamt waren die Gräber eher spärlich ausgestattet. Teilweise fehlten Grabbeigaben völlig. Aber es gab auch reich ausgestattete Gräber, unter anderem auch Anlagen mit vierrädrigen Wagen mit Scheibenrädern. Einige Gräber der Jamnaja-Kultur wurden, basierend auf ihren Beigaben, als Handwerkerbestattungen identifiziert. Diese Unterschiede deuten auf eine soziale Stratifizierung innerhalb der Kultur hin (Parzinger 1998a). Messer und andere metallene Werkzeugbeigaben bestanden aus Bronze und Kupfer (Nikolova 2007).

Auch wenn das Bild der Jamnaja-Kultur insgesamt sehr homogen erscheint, so gibt es dennoch regionale Unterschiede. Diese werden vor allem in der Keramik deutlich, die jeweils enge Beziehungen zu den lokalen äneolithischen Vorläufern erkennen lässt. So übernahmen z. B. die Gruppen im nordwestlichen Schwarzmeerraum Elemente der Cernavodă III-Kultur. Im Dnjepr-Gebiet ist die Verwandtschaft zur Dereivka-Kultur erkennbar, wobei Importe der Kugelamphorenkultur auf westliche Einflüsse schließen lassen (Dergačev 1998). In der Kuban-Region zeigte die der Jamnaja-Kultur angehörige Novotitarovka-Gruppe Elemente der späten Majkop-Kultur. In der Wolga-Region

⁶ Im Russischen und Ukrainischen leitet sich der Name der Kultur „Ямна“ [Jamna] von dem dem Begriff „Яма“, Grube ab.

entspricht die Keramik in etwa der der Khvalynsk-Kultur (Parzinger 1998a, Rassamakin 1999: 127). Die Jamnaja-Tradition könnte sich vom Wolga-Gebiet entweder im Zuge eines großräumigen migratorischen Ereignisses oder durch die Übernahme der Riten von den verschiedenen Bevölkerungsgruppen (Masson 1998) verbreitet haben. Die große Anzahl von Kurganen aus der frühen Bronzezeit deutet auf ein Bevölkerungswachstum der Träger der Jamnaja-Kultur hin.

Die Grundlage für das Wirtschaftssystem der Jamnaja-Kultur bildete die zum Ende des Äneolithikums praktizierte mobile Tierhaltung. Während zuvor jedoch Schafe und Ziegen die Hauptnutztiere waren, spezialisierte sich die Jamnaja-Bevölkerung auf das Rind. Schafe, Ziegen spielten im Großen und Ganzen eher eine untergeordnete Rolle, Schweine praktisch gar keine (Kaiser 2010). Ochsen wurden von den Menschen als Zugtiere vor hölzernen Wagen eingesetzt (Rassamakin 1999: 153). Über das Pferd als Nutztier herrscht auch für die frühe Bronzezeit keine Einigkeit unter den Archäologen. Einige können sich die mobile Lebensweise der Träger der Jamnaja-Kultur nicht ohne das Pferd als Reit- und Zugtier vorstellen (z. B. Parzinger 1998a, Anthony 2007: 221 f). Zumindest soll es als Fleischquelle gedient haben (Benecke 1998). Rassamakin (1999: 153 f) hingegen sieht keine eindeutigen Nachweise im archäologischen Fundgut für die Nutzung des Pferdes im Jamnaja-Horizont. Bei den Funden in Mikhailovka, eine der wenigen bekannten Siedlungen, sei nicht sicher, aus welcher Zeitschicht das Knochenmaterial stamme.

Der Großteil der Jamnaja-Kurgane befindet sich in der Nähe von Flüssen. Ein Novum war allerdings die Errichtung kleinerer Kurgane in der offenen Steppe. Auch bei diesen kam es zu Nachbestattungen. Es handelte sich dabei wahrscheinlich um Areale, die von den Jamnaja-Viehhirten als saisonale Weideflächen genutzt wurden (Rassamakin 1999: 152, Shishlina 2001). Diese Art des Subsistenzwesens dürfte zumindest teilweise auf das mit dem Beginn des Subboreals trockener werdende Klima zurückzuführen sein, da Weideflächen weniger ergiebig wurden. Auf der anderen Seite war die für regelmäßige Wechsel der Nutzflächen vorausgesetzte Mobilität durch die Innovation der Wagen mit Rädern gegeben. Hinweise auf Getreideanbau gibt es lediglich westlich des Dnjepr und auch dort eher sporadisch.

Innerhalb der Jamnaja-Kultur lässt sich eine Differenzierung zwischen den Gemeinschaften westlich und östlich des Don ausmachen. Während im Westen einige Siedlungen gefunden wurden (z. B. Mikhailovka), gibt es im Osten keine Hinweise auf ein Siedlungswesen. Die östlichen Jamnaja-Gruppen besaßen mehr Schafe und Ziegen. Anthony (2007: 304) schließt daraus, dass die Mobilität in den östlichen Populationen höher war. Es stellt sich die Frage, ob diese Unterschiede in der Lebensweise auf eine genetisch differenzierte Population zurückgehen oder die Strukturen lediglich auf äußeren Faktoren beruhende Verhaltensmuster darstellen.

Die mittlere Bronzezeit

Die Jamnaja-Kultur ist bis ungefähr 2 300 v. Chr. nachgewiesen. Doch bereits einige Jahrhunderte vorher, frühestens um 2 700 v. Chr. tauchte eine neue Bestattungstradition im Steppengebiet auf (Kaiser 2003: 65-68, 2007, Chernykh 2008). Dabei handelte es sich um Grabkammern mit seitlich anschließendem vertikalen Eingangsschacht, sogenannte Katakombengräber, die der Kultur ihren Namen geben. Beim Übergang vom Schacht in die Kammer findet sich häufig eine Stufe nach unten. Meist wurden sie in bereits angelegte Kurgane der Jamnaja-Kultur eingelassen. Wie auch bei der Jamnaja-Kultur folgte die Anordnung der Gräber nach bestimmten Mustern (Kaiser 2007). Interessant ist hier sicherlich, ob ein Zusammenhang zwischen den Neuerungen im Bestattungsritus und der biologisch-anthropologischen Identität besteht. Es gilt zu klären, ob die Menschen der Jamnaja-Kultur die neue Sitte eigenständig etablierten, oder ob vielleicht Einflüsse von außerhalb, in Form von Einwanderern, diese Entwicklung vorantrieben.

Die frühesten Gräber finden sich zwischen Don und Donez, dann auch nördlich des Kaukasus bis nach Kalmückien. Erst in ihrer entwickelten Phase, ab etwa 2 450 v. Chr., breitete sich die Katakombengrabkultur auch westlich des Dnjepr bis zum Südlichen Bug aus (Chernykh 2008). Insgesamt war das Verbreitungsgebiet der Katakombengrabkultur kleiner als das der Jamnaja-Kultur. Es war auf den Steppengürtel mit vereinzelt Funden im Süden der Waldsteppe beschränkt. In der Waldsteppe bestanden die Jamnaja-Traditionen größtenteils noch eine Weile fort.

Noch weiter nördlich verbreiteten sich zu dieser Zeit Schnurkeramische Gruppen, insbesondere die Dnjepr-Don-Kultur (Parzinger 1998a). In der Wolga-Region setzte sich die Tradition der Grubengräber in der Poltavka-Kultur fort. Sie wird als eine Weiterentwicklung der Jamnaja-Kultur angesehen (Mallory & Adams 1997: 439 f, Häusler 1998).

Zwischen der frühen und entwickelten Katakombengrabkultur sind unterschiedliche Tendenzen im Bestattungsritus zu erkennen (Kaiser 2007). In der frühen Phase wurden die Toten meist in gehockter Stellung, vor allem in Seitenlage, abgelegt. In den Fällen, in denen Grabbeigaben gefunden wurden, waren die frühen Gräber reichhaltiger ausgestattet. Die Grabkammern waren meistens etwa eckig, und der Zugang konnte durch Steinblöcke versperrt sein. Die Katakomben der frühen Phase liegen häufig tiefer als die späteren Gräber.

Während die frühe Katakombengrabkultur sehr einheitlich erscheint, lassen sich in der entwickelten Phase zwei großräumige Gruppen unterscheiden. In der Gruppe um den Dnjepr und westlich davon, auch die Ingul-Gruppe (nach einem Nebenfluss des Südlichen Bug benannt) genannt, wurden die Toten in gestreckter Rückenlage in ovalen bis runden Kammern bestattet; in der Region um den Donez und den unteren Don hingegen gehockt in eckigen Kammern. Nördlich des Asowschen Meeres, zwischen Dnjepr und Donez, befand sich ein Übergangsgebiet, in dem beide Formen der Bestattung vorkamen (Kaiser 2003: 55). Im Dnjepr-Gebiet wurden teilweise die Schädel nach Entfernen der Weichteile mit einer tonartigen Masse ausgefüllt. Wie auch in der frühen Phase fehlen Grabbeigaben bei vielen Bestattungen. Und wenn, dann finden sich bei der entwickelten Kultur selten mehr als zwei Gefäße oder Geräte im Grab. Westlich des Dnjepr waren amphorenförmige Gefäße aus ungebranntem Ton eine typische Beigabe (Kaiser 2007). Es gibt allerdings auch Ausnahmen, z. B. Katakombengräber mit Wagenfunden oder Handwerkerutensilien (Parzinger 1998a). Auch die Träger der Katakombengrabkultur verwendeten Ocker in ihren Bestattungen. Ähnlich der Jamnaja-Kultur gab es neben den prinzipiellen Unterschieden zwischen West und Ost regionale Formen, vor allem in der Keramik (Parzinger 2006: 350 f). Erneut stellt sich die Frage, ob auch auf der biologischen Ebene temporale und

regionale Subpopulationen bzw. Strukturen in der Metapopulation Katakombengrabkultur zu fassen sind.

Die Subsistenzwirtschaft änderte sich im Zuge des neuen Grabbrauchs scheinbar wenig. Die Träger der Katakombengrabkultur führten wie die der Jamnaja-Kultur eine halbnomadische Lebensweise als Viehzüchter und entwickelten dieses Wirtschaftssystem weiter (Shishlina 2001). Die Fundstätten deuten auf einen temporären bzw. saisonalen Charakter der meisten Siedlungen hin. Sie befanden sich vor allem an Flussufern, aber auch in der offenen Steppe. Ein Teil der Bevölkerung lebte möglicherweise in längerfristigen Siedlungen und betrieben Landwirtschaft (Parzinger 1998a). Bei den Haustieren dominierte das Rind mit über 70 %, gefolgt von Schafen und Ziegen. Schweine machten einen sehr geringen Anteil aus. Das Pferd wurde möglicherweise nun auch als Reittier eingesetzt (Benecke 1998, Kaiser 2010).

Die Katakombengrabkultur wird von einigen Autoren noch als frühbronzezeitlich eingeordnet, andere verknüpfen sie mit der mittleren Bronzezeit (z. B. Parzinger 1998a). Sie bestand bis etwa 1950 v. Chr. Abgelöst wurde sie durch die Kultur mit Mehrwulstkeramik, die sich über die Grenzen der Katakombengrabkultur hinaus bis in den südlichen Waldbereich ausbreitete, wo bis dato die Bevölkerung der Mittleren Dnjepr-Kultur ansässig war. Der Mehrwulstkeramik folgten in der späten Bronzezeit im nordwestlichen Schwarzmeerraum die Sabatinovka- und am Dnjepr und östlich davon die Srubna-Kultur (Parzinger 1998a).

5 Selektion in der Humanevolution

5.1 Akklimatisation, Adaptation & Selektion

Betrachtet man die heutige Weltbevölkerung, fällt sofort die große phänotypische Vielfalt auf. Auf der genetischen Ebene hingegen ist die Diversität in der noch jungen Spezies Mensch gering. Die modernen Menschen sind untereinander zu etwa 99 % identisch. Unterteilt man sie in die geographischen Großgruppen Subsahara-Afrika, Amerika, Ozeanien, Ostasien und Eurasien, so liegen etwa 80 – 90 % der genetischen Variabilität innerhalb und lediglich 10 – 15 % zwischen diesen Gruppen (z. B. Lewontin 1972, Barbujani *et al.* 1997, Jorde *et al.* 2000, Romualdi *et al.* 2002, Excoffier & Hamilton 2003, Elhaik 2012). Die Unterschiede zwischen den menschlichen Populationen lassen sich auf ein komplexes Zusammenspiel von demographischen und selektiven Prozessen in der rezenten Humanevolution zurückführen.

Diese Evolution ist geprägt von der Ausbreitung des modernen Menschen zunächst in Afrika, und anschließend in die verschiedenen Kontinente. Es folgten weitere Migrationen, wie z. B. der Rückzug aus Mitteleuropa und dessen Rekolonisierung nach dem Letzten Glazialen Maximum. Im Verlauf solcher Migrationen bzw. spatialen Expansionen kommt es zu einer Serie von Gründereffekten (*founder effects*), wenn Gruppen von Individuen aus einer ursprünglichen Bevölkerung migrieren und neue Populationen bilden. In diesen kleinen Populationen ist die genetische Variation reduziert. Da in den Randgebieten einer räumlichen Ausbreitung die Bevölkerungsdichte gering ist, haben niederfrequente oder neue Mutationen eine größere Chance im Zuge des anschließenden Populationswachstums an Häufigkeit zuzunehmen, als welche, die in einer großen, dichten Bevölkerung entstehen. Dies kann schnell zu zufälligen Veränderungen in der Allelkomposition führen, insbesondere wenn einzelne Mutationen sozusagen auf der Expansionswelle „reiten“ (*allelic surfing*). Sie erreichen ihre höchsten Frequenzen weit weg von ihrem Ursprungsort (Edmonds *et al.* 2004, Klopstein *et al.* 2006). Verschiedene Allele werden so in unterschiedliche Regionen getragen. Demographische Entwicklungen und zum Teil damit gekoppelte zufällige Faktoren wie genetische Drift, Mutationen,

Anzahl der Nachkommen etc. erklären einen Großteil der Unterschiede, die zwischen den verschiedenen menschlichen Populationen beobachtet werden.

Gleichzeitig spielen auch äußere Einflüsse eine wichtige Rolle in der Humanevolution. Zusätzlich zu dem bis vor 12 000 Jahren global unbeständigen Klima (Jobling *et al.* 2004: 283 f), wurden die Menschen bei ihrem Vordringen in neue Gebiete mit den unterschiedlichsten lokalen Umweltbedingungen konfrontiert. Um sich dauerhaft etablieren zu können, mussten sie sich an diese anpassen. Auf der individuellen Ebene findet zum einen Akklimatisation, d. h. eine reversible physiologische Anpassung des Organismus an eine akute Situation, statt. Sie erfolgt innerhalb der genetisch vorgegebenen Reaktionsnorm; beispielsweise die kurz- oder langfristige Akklimatisierung (auch Habituation) an die niedrigen Temperaturen und geringe Sauerstoffdichte in größeren Höhenlagen, unter anderem durch erhöhte Atmung und Herzschlag. Zum anderen kann die genetische Konstitution eines Individuums zufällig in Bezug auf die äußeren Umstände vorteilhaft sein. Im letzteren Fall bezeichnet man das vorteilhafte Merkmal als *Adaptation*⁷ und spricht auf der genetischen Ebene von natürlicher positiver Selektion. Der Merkmalsträger hat unter den vorherrschenden Bedingungen eine bessere Fitness, d. h. eine höhere Überlebens- und Fortpflanzungschance als Individuen mit einer anderen, weniger vorteilhaften genetischen Konstitution. Hierdurch steigt die Wahrscheinlichkeit, dass das vorteilhafte Merkmal in die nächste Generation weitergetragen wird. Klassische Beispiele für *Adaptation* sind die von Bergmann und Allen an endothermen Tieren beobachteten Zusammenhänge zwischen Temperatur und Morphologie. Die sogenannten ökogeographischen Regeln besagen, dass in Populationen, die in kälteren Regionen leben, die Tiere tendenziell größer bzw. kräftiger sind als in wärmeren Regionen (Bergmann 1847) und kleinere Extremitäten und Körperanhänge haben (Allen 1877). Im Prinzip wird durch ein kleineres Verhältnis von Körperoberfläche zu Körpervolumen der Wärmeverlust in kalten Gebieten verringert. Diese Regeln scheinen bedingt auch für den Menschen zu gelten. Klassische morphologische Studien an Menschen zeigen, dass sich Populationen in kälteren

⁷ Hierbei ist der evolutionsbiologische Adaptationsbegriff gemeint, nicht zu verwechseln mit dem in der Sinnesphysiologie, der eine reversible Anpassung der Reaktion auf unterschiedliche Reizintensitäten beschreibt.

Regionen tendenziell durch einen kompakten Körperbau auszeichnen (Ruff 1994, Ruff 2002, Holliday & Hilton 2010 und Referenzen darin). Die ostafrikanischen Massai hingegen sind zwar groß, aber gleichzeitig auch sehr schlank mit langen Gliedmaßen. Sie werden oft als Beispiel für die ökogeographische Regel von Allen (1877) beim Menschen herangezogen.

Die Populationsgenetik betrachtet die den phänotypischen Merkmalen zugrunde liegenden Polymorphismen im Erbgut. Dabei ist sie nicht an der genetischen Konstitution und Fitness einzelner Individuen interessiert, sondern an den Auswirkungen von Selektion auf Populationsebene. Eine gerichtete Selektion kann auf eine neu entstandene Mutation bzw. den damit assoziierten Phänotyp wirken, oder auf eine bereits in der Population vorhandene, zuvor neutrale Variation, die durch Veränderungen z. B. von Umweltbedingungen nun vorteilhaft oder nachteilig wird. Als Folge nimmt die Häufigkeit des betroffenen Allels in der Population zu bzw. ab. Positive Selektion führt „im Idealfall“ letztendlich zur Fixierung der vorteilhaften Variante. Ein Beispiel ist die Selektion und Fixierung von zwei Mutationen im Sprachgen *FOXP2* im Verlauf der menschlichen Evolution (Enard *et al.* 2002). Da Veränderungen in der DNA meist neutral oder schädlich sind, spielt negative Selektion eine wesentlich bedeutendere Rolle. Durch sie wird ein nachteiliges Allel aus der Population entfernt. Dies betrifft insbesondere Mutationen in lebenswichtigen Genen. Ausgleichende Selektion hingegen fördert den Erhalt eines Polymorphismus und trägt damit zur Vielfalt in einer Population bei. Als Mechanismen dienen dabei eine zwischenzeitlich gerichtete Selektion in Abhängigkeit von den gegenwärtigen Allelfrequenzen (*frequency-dependant selection*) und der sogenannte Heterozygotenvorteil, bei dem Träger eines heterozygoten Genotyps eine höhere Fitness haben als Homozygote. Die ausgleichende Selektion spielt vor allem in der Immunabwehr eine wichtige Rolle, z. B. in der Erhaltung der Diversität des *major histocompatibility complex (MHC) locus* (Hughes & Nei 1988, 1989, Carrington *et al.* 1999, Cooke & Hill 2001, Prugnolle *et al.* 2005, Buhler & Sanchez-Mazas 2011 etc.). Die Gene dieses Clusters gehören zu den polymorphsten im menschlichen Genom. Die Genprodukte sind Zelloberflächenmoleküle, die eigene und fremde Antigene erkennen und diese den Zellen des Immunsystems präsentieren. Heterozygotenvorteil und

frequenzabhängige Selektion gewährleisten eine hohe Diversität, so dass ein breites Spektrum von Pathogenen erkannt und somit bekämpft werden kann (Sommer 2005).

Neben Pathogenen gibt es noch weitere Selektionsfaktoren, die Druck auf menschliche Populationen ausüben. Das wohl auffälligste Beispiel eines angepassten Merkmals ist die heute zu beobachtende Korrelation der Hautfarbe mit der geographischen Region. In Zonen mit hoher Sonnenintensität schützt eine dunkle Hautpigmentierung vor dem Eindringen der schädlichen UV-Strahlen, während in Gebieten wie z. B. Nordeuropa die Haut einen möglichst hohen Anteil der schwachen Sonnenstrahlen durchlassen sollte, um dem Körper die Produktion des lebenswichtigen Vitamin D zu ermöglichen. Der Phänotyp Pigmentierung wird durch mindestens 16 Gene gesteuert (Valenzuela *et al.* 2010). Menschen in der Prähistorie, die aufgrund von zufälligen Mutationen in diesen Genen eine helle Pigmentierung aufwiesen, hatten in Mittel- und Nordeuropa im Vergleich zu anderen Menschen vermutlich einen stabileren Vitamin D-Haushalt und dadurch eine bessere Gesundheit und höhere Überlebens- und Fortpflanzungschancen. In diesem Fall führte der Selektionsfaktor Sonnenstrahlung aufgrund der Polygenie zu Veränderungen von Allelfrequenzen in verschiedenen Genen (näheres hierzu unter den Punkten 5.4 und 10.4.1). Ein weiteres Beispiel für positive Selektion steht im Zusammenhang mit der oben angesprochenen Akklimatisation an einen zeitweiligen Aufenthalt in Höhenlagen. In tibetischen Populationen wurden vor kurzem Gene identifiziert (u. a. *EPAS1* (*endothelial pas domain protein 1*), *EGLN1* (*EGL9, C. elegans, homolog of, 1*)), die das Ziel von Selektion für eine Adaptation an die Höhe zu sein scheinen (Beall *et al.* 2010, Wuren *et al.* 2014).

5.2 Die Untersuchung von Selektion

Es gibt zwei prinzipielle, sich ergänzende Herangehensweisen in der Selektionsforschung zur Identifizierung von selektierten Markern: a) die Untersuchung einzelner Kandidatengene und b) die Durchführung von Genomscans. Bei beiden Ansätzen kommen für die eigentlichen Selektionsnachweise im Prinzip die gleichen mathematisch-statistischen Testverfahren zum Einsatz. Sie beruhen auf den Spuren, die Selektion in der DNA-Sequenz hinterlässt.

Bis vor wenigen Jahren lag der Schwerpunkt auf der Untersuchung *a priori* ausgewählter Kandidatengene. Es handelt sich dabei um Gene, die aufgrund ihrer Funktion oder ihrer Lage im Verdacht stehen, mit einem beobachteten Phänotyp assoziiert zu sein. Eines der bekanntesten Beispiele für diesen Ansatz ist die Analyse der Duffy-Blutgruppen. Sie basieren auf Polymorphismen im *DARC*-Gen (*Duffy antigen/receptor for chemokines*), welches die Duffy-Antigene, Membranrezeptoren in den Zellmembranen von Erythrocyten, kodiert. Neben zwei Spleißvarianten (*FY*A*, *FY*B*) tritt eine Mutation in der Promotorregion des Gens auf, die dessen Transkription verhindert (Tournamille *et al.* 1995a, Tournamille *et al.* 1995b). Bei homozygot Duffy-negativen (*FY*O*) Individuen wurde in mehreren Studien eine absolute Resistenz gegen den Malariaerreger *Plasmodium vivax* nachgewiesen (Miller *et al.* 1976, Livingstone 1984 und Referenzen darin). Das *FY*O*-Allel ist in großen Teilen Afrikas fixiert oder kurz vor der Fixierung, während es in der restlichen Welt äußerst selten vorkommt (Howes *et al.* 2011). Anlässlich dieser auffälligen geographischen Differenzierung und der nachgewiesenen Malariaresistenz führten Hamblin *et al.* (2000, 2002) statistische Analysen durch, mit denen die Hypothese einer positiven Selektion dieses *locus* erfolgreich bestätigt werden konnte.

Aufgrund der rasanten technologischen Entwicklungen liegt der Hauptfokus der Selektionsforschung mittlerweile auf genomweiten Analysen. 2001 wurde der erste aus der DNA mehrerer anonymisierter Individuen erarbeitete Entwurf eines menschlichen Genoms veröffentlicht (International Human Genome Sequencing Consortium 2001). Eine zweite Sequenz wurde zeitgleich durch ein kommerzielles Unternehmen erstellt (Venter *et al.* 2001). Einer Publikation des ersten mit der

Sanger-Methode sequenzierten individuellen Genoms (Levy *et al.* 2007) folgten fast vollständige Genome mithilfe von *next-generation sequencing*-Verfahren (u. a. Bentley *et al.* 2008, Wheeler *et al.* 2008).

Der Schwerpunkt der Genomforschung wurde auf die Genotypisierung von SNPs gelegt. Zwischen 1999 und 2001 wurden durch die Stiftung THE SNP CONSORTIUM LTD. über 1,4 Millionen SNPs im menschlichen Genom identifiziert (The International SNP Map Working Group 2001). Die Initiative zur Erstellung einer frei zugänglichen Datenbank zur menschlichen Variation wurde mit dem 2002 ins Leben gerufene INTERNATIONAL HAPMAP PROJECT fortgeführt. In zwei Phasen wurde in Kollaboration mit PERLEGEN SCIENCES eine Haplotypenkarte des Genoms mit über 3,1 Millionen SNPs angefertigt (Hinds *et al.* 2005, The International HapMap Consortium 2005, 2007). Die Proben stammten aus vier Populationen: US-amerikanische Bürger nordwesteuropäischer Herkunft (CEU), Yoruba aus Ibadan, Nigeria (YRI), Han Chinesen aus Beijing (HCB) und Japaner aus Tokio (JPT). Inzwischen wurde der Datensatz in einer dritten Phase erweitert, um weitere Populationen zu integrieren und auch seltene SNPs ($MAF \leq 5\%$)⁸ erfassen zu können (The International HapMap 3 Consortium 2010). Seither wachsen die Sequenzdatenbanken und Kataloge zur genetischen Variation stetig. Mit dem 1000 GENOMES PROJECT wurde die umfangreichste Erfassung des menschlichen Genoms in Angriff genommen, mit dem Ziel, 2 500 menschliche Genome zu sequenzieren (The 1000 Genomes Project Consortium 2010).

Die Verfügbarkeit solcher Variationskarten bzw. -kataloge ermöglichte einen neuen Ansatz bei der Identifizierung von unter Selektion stehender *loci*. Anstatt biologisch basierte Selektionshypothesen für einzelne Kandidatengene durch Nachweis der entsprechenden Signaturen zu überprüfen, werden mittlerweile die SNP-Kataloge oder die Sequenzdaten gesamter Genome gezielt nach diesen Selektionssignalen durchsucht. Einen Überblick zu diesen Signalen gibt der nachfolgende Abschnitt (5.3). Rein auf der Ebene der DNA werden mögliche Ziele von Selektion identifiziert, ohne Vorkenntnisse bezüglich Genidentitäten oder Phänotypen. Allerdings kann dabei Selektion von Demographie nicht immer sicher unterschieden werden, da

⁸ *Minor allele frequency*: Häufigkeit des selteneren Allels eines Polymorphismus in einer Population.

demographische Ereignisse im Genom ähnliche Spuren wie ein Selektionsevent hinterlassen können. Hier sei insbesondere das bereits erwähnte *allelic surfing* im Zuge einer räumlichen Populationsexpansion genannt. Es führt zu einer genetischen Signatur, wie sie auch durch natürliche Selektion hervorgerufen werden kann (Currat *et al.* 2006, Nielsen *et al.* 2007, Hofer *et al.* 2009). Früher vermutete man, dass demographische Faktoren alle neutralen *loci* gleichmäßig beeinflussen, und folglich alle *loci*, die drastische Unterschiede in den Allelfrequenzen zwischen geographischen Regionen aufweisen, unter Selektion stünden. Mittlerweile geht man davon aus, dass Phänomene wie das *allelic surfing* einen Teil dieser Unterschiede erklären (Hofer *et al.* 2009). Insgesamt ist das Zusammenspiel der Faktoren, die die Menschheitsgeschichte beeinflussen äußerst komplex. Die Effekte von Expansionen, über die Zeit verschieden gerichteten Migrationen, *bottlenecks*, kulturellen Entwicklungen und sich wandelnden Selektionsdrücken auf den menschlichen Genpool können einander verstärken oder sich auch gegenseitig aufheben oder verzerren (Balaesque *et al.* 2007, Coop *et al.* 2009, Hofer *et al.* 2009, Novembre & Di Rienzo 2009). Daher wird es auch in der Zukunft eine Herausforderung bleiben, das Geflecht aus Demographie und Selektion zu entwirren und ihren jeweiligen Anteil an der menschlichen Entwicklung zu bestimmen.

5.3 Signaturen von Selektion

Selektionsereignisse werden mithilfe sogenannter Neutralitätstests nachgewiesen. Hierbei werden ausgewählte Charakteristiken der beobachteten Diversität mit den Erwartungen unter neutraler Evolution verglichen. Die Neutralitätstheorie (Kimura 1968) dient als Nullhypothese. Nach ihr sind ein Großteil der evolutiven Veränderungen neutral und unterliegen vor allem der genetischen Drift. Selektionssignale werden als lokale Abweichungen von dem allgemeinen Muster der genetischen Variation identifiziert. Dabei gilt es zu differenzieren, ob die festgestellten

Signaturen auf ein Selektionsereignis oder auf demographische Prozesse zurückzuführen sind.

Im Falle einer neu entstandenen, positiv selektierten Mutation nimmt nicht nur deren Häufigkeit in der Population zu, sondern auch die benachbarter, auf dem gleichen Chromosom liegender Varianten. Diese gekoppelten Allele unterliegen damit dem sogenannten *hitchhiking*-Effekt. Als Folge nimmt die Variation in dieser Genomregion ab, es kommt zu einem *selective sweep* (Smith & Haigh 1974). Ist der Selektionsdruck stark genug, wird die neue Mutation schnell fixiert und der Haplotyp, der die vorteilhafte Mutation trägt, kann unter Umständen zum einzigen in der Population werden. Solange die selektierte Position in der Population polymorph ist, spricht man von einem *incomplete sweep*; kommt es zu einer Fixierung, handelt es sich um einen *complete sweep*. Der Grad der Variationsreduktion in der Region des *sweeps* (im Vergleich zu ungekoppelten, neutral evolvierenden *loci*) und die Größe des betroffenen Bereichs geben Anhaltspunkte über Stärke und Zeitraum der Selektion. Der *long range haplotype (LRH) test* (Sabeti *et al.* 2002) beispielsweise identifiziert Regionen, in denen die genetische Kopplung (*linkage disequilibrium*) weiter reicht, als bei der gegebenen Häufigkeit des Allels (bzw. Haplotypen) in der Population unter neutralen Bedingungen zu erwarten wäre. Auch die *iHS*-Statistik (*integrated haplotype score*; Voight *et al.* 2006) erfasst diese Spur der langen Haplotypen.

Im Verlauf der Zeit ändert sich die Signatur der Selektion. Das Signal der langen Haplotypen bleibt nur über eine Spanne bis zu 30 000 Jahren erhalten. In diesem Zeitraum haben Rekombinationsereignisse die durch den *sweep* entstandenen langen Haplotypen soweit degradiert, dass diese Signatur nicht mehr nachweisbar ist (Sabeti *et al.* 2006). Nach und nach wird auch die verlorene Diversität in der *sweep*-Region durch neue Mutationen wieder hergestellt. Dieser Prozess spiegelt sich in einem überdurchschnittlich hohen Anteil seltener Allele wider. Er kann mehrere hunderttausend Jahre andauern und z. B. durch Tajima's *D* nachgewiesen werden (Tajima 1989, siehe auch Abschnitt 7.3.2). Allerdings kann dieses Phänomen einer Verschiebung des Allelfrequenzspektrums in Richtung seltener Allele auch durch eine Populationsexpansion hervorgerufen werden. Auf der anderen Seite nehmen junge Mutationen, die mit dem selektierten Allel gekoppelt sind, schneller an Häufigkeit zu,

als unter neutralen Bedingungen zu erwarten wäre. Das Selektionssignal einer überdurchschnittlichen Anzahl hochfrequenter abgeleiteter Allele wird z. B. durch Fay & Wu's *H test* (Fay & Wu 2000) aufgedeckt.

Neben dem beschriebenen klassischen, auch als *hard sweep* bezeichneten Modell, gibt es zwei weitere Szenarien, die man als *soft sweeps* klassifiziert (Hermisson & Pennings 2005). Zu einem *soft sweep* kommt es durch die positive Selektion eines bereits in der Population verbreiteten Allels (*standing variation*), das zuvor neutral oder leicht negativ selektiert war. Im zweiten Szenario wirkt ein Selektionsdruck auf mehrere unabhängig voneinander vorteilhafte Mutationen an einem *locus*, die auf unterschiedlichen Haplotypen liegen und daraufhin alle an Häufigkeit in der Population zunehmen. Bei *soft sweeps* sind die oben beschriebenen Signale schwächer ausgeprägt und daher schwerer zu detektieren. Es gibt mehrere favorisierte Haplotypen und die Ausbreitung der Signatur auf dem Chromosom ist geringer als bei einem *hard sweep*.

Bei negativer Selektion kann die Eliminierung der unvorteilhaften Mutation aus der Population auch zur Entfernung gekoppelter nicht-schädlicher Varianten führen. Dieser Variabilitätsverlust wird als *background selection* bezeichnet (Charlesworth *et al.* 1993). Über einen längeren Zeitraum andauernde ausgleichende Selektion hingegen führt zu einer erhöhten Diversität in der umgebenden Region, mit einer Anhäufung von langlebigen Allelen mit mittleren Frequenzen (Charlesworth 2006, DeGiorgio *et al.* 2014).

Liegt ein Selektionsereignis länger als 250 000 Jahre zurück, ist von der Signatur eines *selective sweep* nichts mehr erhalten (Sabeti *et al.* 2006). Bis zu mehrere Millionen Jahre alte Ereignisse können unter bestimmten Voraussetzungen auf der interspezifischen Ebene nachgewiesen werden. Auf Proteine wirkende, positive Selektion kann sich durch ein erhöhtes Vorkommen nicht-synonymer Mutationen auszeichnen. Dieses Signal lässt sich durch den Vergleich homologer Sequenzen verschiedener Spezies aufspüren. Zu den Selektionstests, die auf interspezifischen Vergleichen basieren, gehören z. B. der HKA- (Hudson *et al.* 1987) und der McDonald-Kreitman-Test ((McDonald & Kreitman 1991).

Verschiedene Übersichtsartikel geben einen Überblick zu den Signaturen von Selektion und den gängigen Tests und Statistiken (z. B. Biswas & Akey 2006, Harris & Meyer 2006, Sabeti *et al.* 2006, Oleksyk *et al.* 2010).

5.4 Genetische Marker unter Selektionsverdacht

Einige Beispiele für Selektion wurden bereits genannt. An dieser Stelle sollen nun die *loci* kurz vorgestellt werden, die im Rahmen dieser Dissertation untersucht wurden. Eine tabellarische Übersicht findet sich am Ende dieses Kapitels (Tabelle 5.1). Detaillierte Beschreibungen zu Genfunktionen, SNPs und Selektionshypothesen werden im Rahmen der Diskussion der entsprechenden Analysen in Kapitel 10.4 gegeben.

Die Auswahl der *loci* fand nach intensiven Literaturrecherchen in 2007 und 2008 unter folgenden Gesichtspunkten statt: Methodisch bedingt sollte es sich bei den Polymorphismen um einzelne Basensubstitutionen handeln, da diese mittels 454-Sequenzierung oder Direktsequenzierung am besten erfasst werden können. Außerdem musste eine gemeinsame Amplifikation aller *loci* in einem Multiplex-PCR-Ansatz möglich sein. Während im menschlichen Genom unzählige SNPs ein Selektionssignal aufweisen, ist bei vielen nichts über die funktionelle Bedeutung des Genortes im Allgemeinen, geschweige denn über die des Polymorphismus bekannt. Es wurden lediglich Kandidaten ausgewählt, bei denen die Funktion(en) des Gens und die Auswirkung(en) der Variation zumindest ansatzweise erforscht sind, und zu denen bereits Selektionshypothesen aufgestellt wurden. Um arbeitsgruppeninterne Überschneidungen mit anderen Projekten zu vermeiden, wurde auf die Untersuchung einiger mit Infektionsanfälligkeiten assoziierter Marker verzichtet (z. B. das *HFE*-Gen oder *G6PD* (*glucose-6-phosphate dehydrogenase*)). Da die Forschungsgebiete des Projekts von M. U. und dieser Arbeit in Eurasien liegen, wurden auch ausschließlich in Afrika selektierte *loci* nicht verwendet.

Bei den auf den folgenden Seiten angegebenen rs-Nummern handelt es sich um die Referenznummern, unter denen die einzelnen SNPs in den genetischen Datenbanken archiviert sind (NCBI - dbSNP). Die Abkürzungen der Aminosäuren entsprechen der IUPAC-Nomenklatur (IUPAC-IUB Joint Commission on Biochemical Nomenclature 1993).

Im Laufe der Humanevolution hat sich die Hautfarbe mehrmals verändert. Mit dem Verlust der vollen Körperbehaarung wurde die ursprünglich blasse Haut dunkler, um den Organismus vor schädigender UV-Strahlung zu schützen. Mit der Ausbreitung des modernen Menschen über den gesamten Globus setzte sich in Gebieten mit schwacher Sonneneinstrahlung wieder eine hellere Pigmentierung durch (Jablonski & Chaplin 2000, Norton *et al.* 2007, Jablonski 2012). Wie bereits erwähnt, sind eine Vielzahl von Genen an der Pigmentierung von Haut, Haaren und Augen beteiligt, von denen nur einige genauer charakterisiert sind. Es wurden SNPs in vier Genen untersucht, die direkt oder indirekt an der Synthese des Pigments Melanin beteiligt sind, und die nachweislich die Pigmentierungsintensität beeinflussen. Dabei trägt das jeweils abgeleitete Allel zu einem helleren Haut-, Augen- oder Haartyp bei und ist vermutlich in europäischen Populationen selektiert.

Die Tyrosinase katalysiert die ersten beiden Schritte der Melaninsynthese. Der SNP rs1042602 (Ser192Tyr) im *Tyrosinase*-Gen (*TYR*) beeinflusst die Aktivität des Enzyms (Hudjashov *et al.* 2013).

Bei *SLC45A2* (*solute carrier family 45, member 2*) handelt es sich wahrscheinlich um einen Zucker/Protonen-Transporter in Melanocyten, aber über die mögliche Funktion wird noch spekuliert. Er könnte am Transport von Pigmentierungsenzymen zu den Melanosomen oder der Regulation des Volumens oder des pH-Wertes dieser Organellen beteiligt sein (Newton *et al.* 2001, Costin *et al.* 2003). Eine zu einem Aminosäureaustausch (Leu374Phe) führende Transversion (rs16891982) wirkt vermutlich auf das Milieu in den Melanosomen ein (Graf *et al.* 2005).

Auch bei *SLC24A5* (*solute carrier family 24, member 5*) handelt es sich um ein Transporterprotein, ein Natrium-Calcium-Austauscher. Der Polymorphismus rs1426654 (Ala111Thr) beeinflusst vermutlich die Calcium-Konzentration in den

Melanosomen und so deren Reifung. Möglicherweise ist das Protein indirekt an der Regulierung des pH-Werts in den Organeln beteiligt (Lamason *et al.* 2005, Scherer & Kumar 2010).

Als viertes wurde der in *HERC2* (*hect domain and RCC1-like domain 2*) lokalisierte, aber einem regulatorischen Element von *OCA2* (*oculocutaneous albinism II*) zugehörige SNP rs12913832 in der prähistorischen Population genotypisiert. Die Augenfarbe (blau oder nicht blau) wird durch diesen SNP determiniert, indem das abgeleitete Allel zu einer verminderten Expression von *OCA2* führt (Eiberg *et al.* 2008, Sturm *et al.* 2008).

Ein weiteres Gen, für das der Vitamin D-Haushalt die Selektionshypothese darstellt, ist *CYP3A4* (*cytochrome P450, subfamily IIIA, polypeptide 4*). Es gehört zu einer Familie von Enzymen, die am Metabolismus von ungefähr der Hälfte aller üblich verwendeten Pharmazeutika beteiligt sind (Schirmer *et al.* 2006). *CYP3A4* trägt unter anderem zum Vitamin D-Haushalt bei, sowohl in der Produktion als auch im Abbau des aktiven Hormons (Wang *et al.* 2013). Das ancestrale Allel des regulatorischen SNPs rs2740574 steht unter Verdacht, in nicht-afrikanischen Populationen unter negativer Selektion zu stehen (Schirmer *et al.* 2006).

Konkurrierend zu dem Polymorphismus in *CYP3A4* wurde innerhalb des *CYP3A*-Genclusters ein SNP in *CYP3A5* (*cytochrome P450, subfamily IIIA, polypeptide 5*) als Ziel eines Selektionsdrucks vorgeschlagen (Thompson *et al.* 2004). Die beiden Gene haben überlappende Aufgabenbereiche, aber in diesem Fall wurde die Prädisposition zu Bluthochdruck als Selektionshypothese formuliert. 6 β -Hydroxycortisol, das Produkt einer Katalyse durch *CYP3A5* in der Niere, stimuliert die Natriumretention (Grogan *et al.* 1990). In tropischen Regionen ist dies vorteilhaft, wenn nicht sogar überlebenswichtig. Aber in kühleren, salzreichen Umgebungen, und insbesondere bei der in modernen Populationen kochsalzreichen Ernährung, trägt zusätzliche Salzretention zu einem hohen, gesundheitsgefährdenden Blutdruck bei. Rs776746 beschreibt eine Mutation an einer Spleißposition, die zur Synthese eines nicht-funktionellen Enzyms führt, welches mit zunehmender Entfernung zum Äquator an Häufigkeit zunimmt (Thompson *et al.* 2004).

Ebenfalls am Elektrolythaushalt beteiligt ist der von *SLC12A3* (*solute carrier family 12, member 3*) kodierte Natrium-Chlorid-Co-Transporter in der Niere, durch den ein Teil des Natriums aus dem Primärharn reabsorbiert wird. Ein Aminosäureaustausch (Gly264Ala), hervorgerufen durch rs1529927, führt zu einer weniger aktiven Variante des Membranproteins, die möglicherweise in Populationen europäischer Abstammung unter Selektion steht. Die verminderte Na⁺-Reabsorption könnte sich zum einen blutdrucksenkend auswirken, zum anderen geht sie indirekt mit einer stärkeren Rückresorption von Calcium einher (Vormfelde *et al.* 2007). Bei Vitamin D-Mangel, wie z. B. in Regionen mit schwacher Sonneneinstrahlung, wird Calcium nicht effektiv aus der Nahrung aufgenommen, so dass es essentiell ist, einem Verlust des Ions über den Urin entgegenzuwirken.

Dem Klimafaktor Kälte tragen (möglicherweise) zwei der untersuchten Gene Rechnung. *EDAR* (*ectodysplasin A receptor*) kodiert einen Rezeptor, der in Signalwegen zur Entwicklung von ektodermalen Strukturen und Anhängen wie Haaren, Zähnen und exokrinen Drüsen (z. B. Schweißdrüsen) eine wichtige Rolle spielt (Clarke 1987, Monreal *et al.* 1999). Ein Aminosäureaustausch (rs3827760, Val370Ala) beeinflusst die Bindungsaktivität des Rezeptors und damit die Signaltransduktion. Das abgeleitete Allel kommt fast ausschließlich in Asien und Amerika vor und ist höchstwahrscheinlich für das im Vergleich zu anderen Populationen dicke, glatte Haar verantwortlich. Die Variante soll in Ostasien unter Selektion stehen oder gestanden haben, um vor einem Verlust von Körperwärme über den Kopf zu schützen (Bryk *et al.* 2008, Fujimoto *et al.* 2008a).

Rs17822931 (Gly180Arg) in *ABCC11* (*ATP-binding cassette, subfamily C, member 11*) erlangte vor allem Bekanntheit durch seine enge Assoziation mit der Konsistenz und Farbe von Ohrenschmalz (Yoshiura *et al.* 2006), aber auch mit Achselgeruch (Martin *et al.* 2010). Die Mutation führt zum Fehlen eines Transporterproteins in apokrinen Drüsen. Dadurch können diverse Komponenten des Cerumen oder des Achselschweißes nicht sezerniert werden. In Ostasien ist die Variante mit trockenem Ohrenschmalz und fehlendem oder minimalem Achselgeruch verbreitet und vermutlich selektiert. Geringerer Wärmeverlust durch vermindertes Schwitzen im

kalten, trockenen Klima Nordostasiens oder auch sexuelle Selektion sind als Hypothesen vorgeschlagen.

Selektion kann auch auf kulturellen Entwicklungen, z. B. Veränderungen in der Ernährung, begründet sein. Eines der bekanntesten unter Selektion evolvierten Gene ist die *Laktase (LCT)*. Das Enzym spaltet Laktose im Dünndarm, um die Aufnahme des Zuckers zu ermöglichen. Mit dem Ende der Stillzeit, also noch im Kindesalter, wird das Gen fast vollständig herunterreguliert. Wahrscheinlich entstand mit der Einführung von Milchwirtschaft im Rahmen der neolithischen Transition ein selektives Umfeld für eine Genvariante, bei der die Laktaseproduktion nicht eingestellt wird (Bersaglieri *et al.* 2004). C-13 910T (rs4988235) scheint der kausative SNP für Laktasepersistenz in Europa zu sein. Ebenfalls mit diesem Phänotyp assoziiert ist G-22 018A (rs182549), wobei dies wahrscheinlich auf genetischer Kopplung zu C-13 910T beruht (Enattah *et al.* 2002).

TRPV6 (*transient receptor potential cation channel, subfamily V (vanilloid), member 6*) ist ein Membrankanal, der hochspezifisch Ca^{2+} transportiert. Im Darm ist er für die Aufnahme von Calcium aus der Nahrung in die Epithelzellen zuständig (Peng *et al.* 1999). Es handelt sich dabei um den ersten, geschwindigkeitsdeterminierenden Schritt in der Calciumabsorption. Ein Haplotyp, der die abgeleiteten Varianten von drei SNPs, rs4987657 (Arg157Cys), rs4987667 (Val378Met) und rs4987682 (Thr681Met), trägt, ist in nicht-afrikanischen Populationen wahrscheinlich aufgrund eines *soft sweeps* beinahe fixiert (Akey *et al.* 2004, Akey *et al.* 2006, Hughes *et al.* 2008). Bezüglich des Selektionsdrucks wird lediglich spekuliert. Nach einer von mehreren Hypothesen spielte vielleicht die Ernährungsumstellung im Rahmen der Einführung der agrikulturn Wirtschaft eine Rolle.

Ebenfalls in Folge der agrikulturn Transition vermutlich unter Selektion stehend ist *NAT2 (N-acetyltransferase 2)*. Das NAT2-Enzym ist am Metabolismus und der Detoxifikation verschiedener Substanzen mit karzinogenen Eigenschaften beteiligt, indem es diese acetyliert. Einer von mehreren SNPs, dessen abgeleitetes Allel zu einer langsamen Acetylierung führt, ist rs1801280 (T341C, Ile114Thr) (Vatsis *et al.* 1991). Diese Variante kommt in modernen Populationen vor allem in agrikulturn, der

anzestrale, „schnelle“ Typ in Jäger/Sammler-Gesellschaften vor. Es wird vermutet, dass es im Verlauf der Ernährungsumstellung im Neolithikum zur Selektion des abgeleiteten langsamen acetylierenden Allels kam (Patin *et al.* 2006, Luca *et al.* 2008, Magalon *et al.* 2008).

Alkoholdehydrogenasen katalysieren den ersten Schritt in der Alkoholoxidation. Eine besonders aktive Variante des Enzyms trägt den durch rs1229984 hervorgerufenen Aminosäureaustausch Arg48His in der Untereinheit ADH1B (alcohol dehydrogenase 1B, class 1, beta polypeptide). Sie kommt vor allem in der asiatischen Bevölkerung vor und geht mit einer Hypersensitivität gegenüber Alkohol einher. Der SNP steht unter Verdacht im Zuge der Domestikation von Reis, welcher teilweise vergoren konsumiert wurde, selektiert worden zu sein (Han *et al.* 2007, Li *et al.* 2008, Peng *et al.* 2010). Daneben wurde ein weiterer SNP (rs3811801) in der Promotorregion identifiziert, der ebenfalls das Ziel des Selektionsdrucks gewesen sein könnte (Han *et al.* 2007).

Das erste von zwei Genen, die in die Kategorie Immunsystem eingeordnet sind, ist *CASP12* (caspase 12, apoptosis-related cysteine protease). Das Gen gehört zu einer Familie von Proteasen, die in einem komplexen Zusammenspiel im Immunsystem aktiv sind. Unter anderem dämpft CASP12 die Antwort auf bakterielle Infektionen, indem es eine inhibitorische Wirkung auf CASP1 ausübt, welches die Reifung von Inflammationsmediatoren katalysiert. Dadurch ist die Gefahr größer, dass ein bakterieller Entzündungsherd nicht erfolgreich kontrolliert wird und lokal eingedämmt bleibt, sondern es zu einer systematischen Entzündung im gesamten Organismus kommt. Man spricht dann von einer Sepsis (Saleh *et al.* 2004, Saleh *et al.* 2006). Unter positiver Selektion steht ein SNP in Exon 4 (rs497116), der ein vorzeitiges Stoppcodon hervorruft und *CASP12* zu einem Pseudogen macht (Wang *et al.* 2006, Xue *et al.* 2006). Phänotypisch führt dies zu einer stärkeren Immunantwort auf pathogene Reize und einer effizienteren Beseitigung von Bakterien. Während das Pseudogen in modernen Populationen außerhalb Afrikas fixiert ist, wurde in Afrikanern eine Korrelation des Genotyps mit dem Auftreten von schwerer Sepsis festgestellt.

ABCB1 (*ATP-binding cassette, subfamily B, member 1*) kodiert einen Membrantransporter, der ein weites Spektrum an cytotoxischen Substanzen aus Zellen herausbefördert. Er ist insbesondere in diversen Barrieren vor sensitiven Geweben aktiv, z. B. in der Blut-Hirn-, der Blut-Hoden-, in einer Art Blut-Herz- und in der Plazenta-Schranke. Es wurden drei Polymorphismen, für die in verschiedenen weltweiten Populationen Selektionssignale beobachtet wurden, untersucht: C1236T (rs1128503; synonym), G2677A/T (rs2032582; Ala893Thr/Ser) und C3435T (rs1045642; synonym) (Tang *et al.* 2004, Wang *et al.* 2007b). Sie beeinflussen die Aktivität des Transporters und bewirken damit unterschiedliche Resistenzen gegen Arzneimittel und andere zellfremde Stoffe. Der Selektionsdruck wird im Bereich der *ABCB1*-Funktion in der Blut-Hirn-Schranke gesucht, wobei eine Reihe alternativer Erklärungen in Frage kommen.

Tabelle 5.1: Übersicht der untersuchten Gene und Polymorphismen unter Selektionsverdacht

In der ersten Spalte sind die Namen der Gene und der Nummern-Code der Amplikons für die einzelnen SNPs (siehe Tabelle 8.6) gelistet. Die zweite Spalte gibt die Archivierungsnummer des Gens in der OMIM-Datenbank (NCBI - OMIM) und seine Lage auf der chromosomalen Ebene an. Die vierte Spalte gibt die Referenznummer des SNPs in der dbSNP-Datenbank (NCBI - dbSNP) und seine Nukleotidposition auf dem Chromosom entsprechend der Nummerierung des *Genome Reference Consortium Human Genome build 38* (GRCh38) an. Bei den Mutationen ist das jeweils vermutlich positiv selektierte, abgeleitete Allel fettgedruckt. ^f (*forward strand*) und ^r (*reverse strand*) zeigen die Orientierung des SNPs auf dem *reference assembly* (nicht dem Gen) an. Die abgeleiteten Allel- (oben) und Heterozygotenfrequenzen (unten) stammen aus 1000 Genomes (*Release 14*, The 1000 Genomes Project Consortium 2010). Die Daten der Amerikaner afrikanischer Abstammung (ASW) wurden aus der afrikanischen Population ausgeschlossen, da sie zu einem unbekanntem Grad vermischt sind. Anzahl der Chromosomen (2N): Europa: 2N = 758, Asien: 2N = 572, Afrika: 2N = 370.

Gen Amplikon	OMIM Chromosom	Genfunktion	rs-Nummer Position (GRCh38)	Mutation Mutationstyp	Phänotyp des abgeleiteten Allels / Selektionshypothese	Allelfrequenz (abgeleitet) Heterozygotenfrequenz			Auswahl an Referenzen (+ <i>et al.</i>)
						Europa	Asien	Afrika	
ABCB1 S01 S02 S03	*171050 7q21.12	Effluxpumpe für cytotoxische Substanzen zum Schutz sensitiver Gewebe	rs1128503	C > T (1236) ^r	Veränderte Expression; erhöhter Schutz vor (neuen) xenobiotischen Substanzen, insbesondere gegenüber Angriffen an der Blut-Hirn-Schranke	0,427	0,656	0,124	Tang 2004 Wang 2007
			87550285	synonym (Gly)		0,517	0,402	0,238	
			rs2032582	G > A/T (2677) ^r		0,431	0,446	0,011	
			87531302	Ala893Thr/Ser		0,504	0,360	0,022	
			rs1045642	C > A/T (3435) ^r		0,529	0,397	0,132	
87509329	synonym (Ile)	0,509	0,444	0,243					
ABCC11 S04	*607040 16q12.1	Membrantransporter, unter anderem in apokrinen Drüsen	rs17822931	C > T ^f	Nicht-funktionelles Protein; trockenes Cerumen, verringertes/r Schwitzen / Schweißgeruch; Kälteadaptation / sexuelle Selektion	0,144	0,900	0,000	Yoshiura 2006 Martin 2009
			48258198	Gly180Arg		0,224	0,157	0,000	
ADH1B S05 S06	*103720 4q23	Oxidation von Ethanol zu Acetaldehyd (1. Schritt des Alkoholabbaus)	rs3811801	C > T ^r	Erhöhte Enzymaktivität; Hypersensitivität gegen Alkohol; Aversion gegen übermäßigen Alkoholkonsum	0,000	0,607	0,000	Han 2007 Li 2008 Peng 2010
			99323162	Regulator		0,000	0,500	0,000	
			rs1229984	G > A ^r		0,018	0,729	0,000	
			99318162	Arg48His		0,037	0,388	0,000	
CASP12 S08	*608633 11q22.3	Dämpfung der Immunantwort auf bakterielle Angriffe durch Inhibition von <i>CASP1</i>	rs497116	C > T ^r	Funktionsverlust; effizientere Bakterienbekämpfung; vermindertes Sepsis-Risiko	1,000	0,998	0,882	Wang 2006 Xue 2006
			104892390	Arg125Stop		0,000	0,003	0,303	

Gen Amplikon	OMIM Chromosom	Genfunktion	rs-Nummer Position (GRCh38)	Mutation Mutationstyp	Phänotyp des abgeleiteten Allels / Selektionshypothese	Allelfrequenz (abgeleitet) Heterozygotenfrequenz			Auswahl an Referenzen (+ <i>et al.</i>)
						Europa	Asien	Afrika	
<i>CYP3A4</i> S09	*124010 7q22.1	Metabolismus vieler endo- und exogener Substanzen (vor allem in der Leber)	rs2740574 99784473	G > A ^r Regulator	Verringerte Expression und Enzym- aktivität; Verlangsamung des Vitamin D- Katabolismus; Schutz vor Rachitis	0,982 0,035	1,000 0,000	0,203 0,297	Schirmer 2006
<i>CYP3A5</i> S10	*605325 7q22.1	Metabolismus vieler endo- und exogener Substanzen (vor allem in der Niere)	rs776746 99672916	A > G ^r Splicing Site	Nicht-funktionelles Protein; verminderte Salzretention; Schutz vor Bluthochdruck, Anpassung an feuchtere Umwelt	0,946 0,086	0,715 0,409	0,141 0,249	Thompson 2006
<i>EDAR</i> S11	*604095 2q12.3	Rezeptor im Signalweg für die Entwicklung von Haaren, Zähnen und exokrinen Drüsen	rs3827760 108897145	T > C ^r Val370Ala	Veränderte Signaltransduktion; dickes, glattes, rundes Haar, vergrößerte Talgdrüsen etc; Schutz vor Wärmeverlust / Austrocknung in trocken-kaltem Klima	0,000 0,000	0,878 0,203	0,000 0,000	Bryk 2008 Fujimoto 2008a Xue 2009
(<i>HERC2</i>) <i>OCA2</i> S12	*611409 15q12-13	Membrantransporter, z. B. Anionen (=> pH-Regulation)	rs12913832 28120472	A > G ^f Regulator	Verminderte <i>OCA2</i> -Expression; keine Eumelanin-Produktion; blaue Augenfarbe; Sonnenbräune; sexuelle Selektion	0,710 0,359	0,002 0,003	0,000 0,000	Eiberg 2008 Wilde 2014
(<i>MCM6</i>) <i>LCT</i> S13 S14	*603202 2q21.3	Hydrolytische Spaltung von Laktose, um deren Aufnahme zu ermöglichen	rs4988235 135851076 rs182549 135859184	C > T ^r Regulator G > A ^r Regulator	Laktasepersistenz; Fähigkeit, dauerhaft Milchzucker zu verdauen; Co-Evolution mit Milchwirtschaft	0,524 0,377 0,712 0,435	0,000 0,000 0,000 0,000	0,000 0,000 0,000 0,000	Enattah 2002 Bersaglieri 2004
<i>NAT2</i> S15	*612182 8p22	De-/Aktivierung von toxischen Verbindungen durch Acetylierung	rs1801280 18400344	T > C ^f Ile114Thr	langsame Acetylierungs-Aktivität; Anpassung an Ernährungsumstellung	0,453 0,462	0,028 0,056	0,314 0,411	Patin 2006
<i>SLC12A3</i> S16	*600968 16q13	Na ⁺ -Cl ⁻ -Symporter in der Niere zur Feinregulation des Elektrolythaushalts	rs1529927 56870675	G > C ^f Gly264Ala	niedrigere Transportaktivität mit positivem Einfluss auf Ca ²⁺ -Absorption; Stärkung des Ca ²⁺ -Haushalts oder Verhinderung von Bluthochdruck	0,047 0,074	0,000 0,000	0,000 0,000	Vormfelde 2007
<i>SLC24A5</i> S17	*609802 15q21.1	K ⁺ -abhängiger Na ⁺ /Ca ²⁺ - Austauscher; Melanosomenentwicklung	rs1426654 48134287	G > A ^f Ala111Thr	Verminderte Transporteraktivität; helle Hautpigmentierung; bessere Vitamin D- Aufnahme	0,996 0,008	0,014 0,028	0,041 0,070	Lamason 2005 Beleza 2012

Gen Amplikon	OMIM Chromosom	Genfunktion	rs-Nummer Position (GRCh38)	Mutation Mutationstyp	Phänotyp des abgeleiteten Allels / Selektionshypothese	Allelfrequenz (abgeleitet) Heterozygotenfrequenz			Auswahl an Referenzen (+ <i>et al.</i>)
						Europa	Asien	Afrika	
SLC45A2 S18	*606202 5p13.2	Transporter von Anionen und Enzymen; Melanosomen- entwicklung	rs16891982	C > G ^f	Einfluss auf osmotisches Potential und pH-Wert von Melanocyten; helle Pigmentierung; bessere Vitamin D- Aufnahme	0,970	0,007	0,000	Soejima 2006 Norton 2007
			33951588	Leu374Phe		0,055	0,014	0,000	
TRPV6 S19 S20 S21	*606680 7q34	Transmembraner Ca ²⁺ - Transporter, für die Calcium- Aufnahme aus der Nahrung im Darm	rs4987657	C > T ^r	verminderte Calciumaufnahme?, Anpassung an Ernährungsumstellung oder neue Pathogene	0,925	0,977	0,492	Akey 2006 Hughes 2008
			142877160	Arg157Cys		0,129	0,045	0,465	
			rs4987667	G > A ^r		0,926	0,977	0,481	
			142875155	Val378Met		0,127	0,045	0,476	
			rs4987682	C > T ^r		0,926	0,977	0,451	
			142871843	Thr681Met		0,127	0,045	0,492	
TYR S22	*606933 11q14.3	Oxidation von Tyrosin und L- DOPA (Schritte 1+2 der Melanogenese)	rs1042602	C > A ^f	Geringere Enzymaktivität; keine Sommersprossen; helle Pigmentierung; bessere Vitamin D-Aufnahme	0,368	0,002	0,000	Shriver 2003 Sulem 2007
			89178528	Ser192Tyr		0,409	0,003	0,000	
AMELX/Y S23	*300391 Xp22.2 *410000 Yp11.2	Protein Amelogenin als Komponente von Zahnschmelz		6bp Deletion in Intron 1 in AMELX	Geschlechtsbestimmung				Sullivan 1993

6 Fragen, Ziele & Strategien

Das Projekt, in dessen Rahmen die vorliegende Dissertation verfasst wurde, lässt sich in drei zusammenhängende Themenschwerpunkte einteilen:

- 1) Entwicklung und Umsetzung einer Hybridstrategie von klassischen und *next-generation sequencing*-Methoden zur Analyse alter DNA
- 2) Populationsgenetische Analysen zur Prähistorie Europas
 - a. Die Rolle der pontisch-kaspischen Steppe im europäischen Neolithikum
 - b. Die Kupfer- und Bronzezeit in der pontisch-kaspischen Steppe
- 3) Erfassung der Allelfrequenzen von unter Selektionsverdacht stehenden SNPs in der nordpontischen Steppe vor dem Hintergrund der demographischen und adaptiven Evolution.

- 1) *Entwicklung und Umsetzung einer Hybridstrategie von klassischen und next-generation sequencing-Methoden zur Analyse alter DNA*

Schwerpunkt 1 deckt die methodische Zielsetzung dieser Arbeit ab. Als klassische Methoden seien hier die Polymerase-Kettenreaktion und die Sequenzierung mit der Strangabbruchsynthese nach Sanger *et al.* (1977) genannt, die speziell auf kurze Fragmente angepasst werden. Die meisten (klassischen) Populationsstudien an alter DNA beschränken sich aus materialtechnischen Gründen auf die Analyse der Hypervariablen Region I (HVR I). Erst in den letzten Jahren kamen vermehrt Studien mit einzelnen oder einigen wenigen autosomalen oder Y-chromosomal *loci* auf. Populationsgenetische aDNA-Studien mithilfe von *NGS*-Methoden gibt es in dem Sinne noch keine. Stattdessen liegt der Fokus auf der Rekonstruktion der Humanevolution mithilfe einiger weniger Hominidengenomen.

In dieser Arbeit wird ein Ansatz verfolgt, der klassische Methoden mit *NGS* verbindet. Die klassische Multiplex-PCR erhöht die Anzahl studierbarer *loci* um ein Vielfaches, wobei die Kontaminationsgefahr durch seit Jahren etablierte Maßnahmen minimiert und relativ kontrollierbar ist. Die anschließende 454-Sequenzierung (das zur

Zeit des Projektbeginns am weitesten entwickelte NGS-Verfahren) erlaubt die „klonale“ Sequenzierung von tausenden von *loci* innerhalb eines *runs*⁹. Ein ähnlicher Ansatz wurde bei der Untersuchung von „Ötzi“ angewendet, als das komplette mitochondriale Genom dieses einzelnen Individuums in Multiplex-PCR-Ansätzen amplifiziert und pyrosequenziert wurde (Ermini *et al.* 2008). Näheres zum NGS, insbesondere dem *454 sequencing*, findet sich in Kapitel 7.2.

Die Vorgehensweise bei der Etablierung dieser Hybridstrategie wurde in drei zeitlich überlappende Phasen gegliedert. Phase 1 begann mit der Entwicklung eines nukleären Multiplex-PCR-Protokolls, zunächst mit rezenter, die anschließende Feinregulierung mit alter DNA. Martina Unterländer erarbeitete parallel ein Protokoll für eine mitochondriale Multiplex-PCR (2014). Im Rahmen einer Magisterarbeit wurden von Silja Dillenberger (2013) kladenspezifische Y-chromosomale *loci* auf die gleiche Weise kombiniert. Eine PCR mit neutralen Markern sollte folgen. Phase 2 mit der Modifizierung bzw. Anpassung des *tagging*-Protokolls von Meyer *et al.* (2008) auf PCR-Produkte alter DNA für die Parallelsequenzierung begann kurze Zeit später und verlief parallel mit der PCR-Optimierung. Der Übergang in Phase 3, die Datenproduktion, war fließend. Nachdem die Multiplex-PCR-Protokolle vom jeweiligen Entwickler an den eigenen Proben erfolgreich getestet waren, sollten sie überhändig werden. Mit Erhalt der Sequenzdaten des ersten *454*-Testlaufs begann die Entwicklung einer Strategie zur Auswertung des großen Datenvolumens.

Insgesamt musste die Strategie während des Prozesses der Methodenentwicklung relativ dynamisch und flexibel sein. Da die Datenproduktion durch die *454*-Sequenzierung schubweise, in größeren zeitlichen Abständen erfolgte, war es erforderlich, einen Mittelweg zwischen gezielten Methodenoptimierungsversuchen und Datenproduktion zu finden.

⁹ Als einen *run* bezeichnet man eine Durchführung einer NGS-Sequenzierung, d. h. der Durchlauf einer Sequenzierplatte.

2) *Populationsgenetische Analysen zur Prähistorie Europas*

Mit Schwerpunkt 2 beginnt der inhaltliche Teil dieser Arbeit. Neben der Etablierung einer PCR-454-Hybridmethode waren die Ziele dieser Arbeit, (a) die Rolle des west- und nordpontischen Steppenraumes im europäischen Neolithikum erstmals molekulargenetisch zu beleuchten, und (b) die Entwicklungen innerhalb der Steppe im 3. Jahrtausend v. Chr. zu untersuchen.

2a) *Die Rolle der pontisch-kaspischen Steppe im europäischen Neolithikum*

In Mitteleuropa wird das Neolithikum mit der Verbreitung von Ackerbau und Viehzucht, vor allem durch die Linearbandkeramiker, gleichgesetzt. Palaeogenetische Studien trugen wesentlich dazu bei, die zugrundeliegenden bevölkerungsbiologischen Prozesse zu charakterisieren (u. a. Haak *et al.* 2005, Bramanti *et al.* 2009, Haak *et al.* 2010, Bollongino *et al.* 2013, Brandt *et al.* 2013). Der osteuropäische Steppenraum liegt in der Peripherie des studierten Areals der kontinentalen Ausbreitungsroute der LBK. Wie sich der Wandel von Jägern und Sammlern zu Viehhütern in der osteuropäischen Steppe vollzog, ist bisher kaum untersucht. Archäologischen Kenntnissen zufolge liegen die Anfänge der Viehhaltung in der Steppe wahrscheinlich in der nordwestpontischen Waldsteppe, als Folge von Kontakten zwischen der indigenen Bug-Dnjestr-Kultur mit benachbarten Gruppen der Criș-Kultur. Der Modus der Ausbreitung der neuen Wirtschaftsweise in der Steppe ist nicht geklärt. Da es aus dem Gebiet der heutigen Ukraine bisher so gut wie keine prähistorischen Sequenzdaten und auch nur eine sehr begrenzte Anzahl veröffentlichter moderner Daten gibt, sind sämtliche archäologischen Befunde und Hypothesen den Untersuchungsraum betreffend ohne genetischen und damit bevölkerungsbiologischen Hintergrund.

Das erste Ziel dieser Arbeit ist demnach, die archäologisch definierten Kulturen der nordpontischen Steppe genetisch zu charakterisieren und damit eine Grundlage zu schaffen, um dieses Gebiet überhaupt in die voranschreitenden Untersuchungen des

europäischen Neolithikums einzubinden zu können. Einige Fragen hierzu sollen mit diesem Promotionsvorhaben beleuchtet werden:

Bereits angedeutet wurde die Frage nach dem Ursprung der Viehhirten des osteuropäischen Steppenraums. Hier soll insbesondere die Hypothese der Kontakte zwischen Criș-Kultur, später LBK, dann Tripolje-Kultur und der Bug-Dnjestr-Kultur überprüft werden, und die Erkenntnisse zur Ausbreitungsrichtung und -weise des Neolithikums entsprechend gestützt oder erweitert werden. Sind die bronzezeitlichen Viehhirten der nordpontischen Steppe die Nachkommen von Jägern und Sammlern, die die Haustierwirtschaft von den benachbarten Bauerngesellschaften übernommen haben, oder die Nachfahren von Bauern, die in die Steppe eingewandert sind und sich niedergelassen haben?

Die nächste Frage beschäftigt sich mit der Kontaktsituation nach dem ersten Zusammentreffen zwischen Steppenbewohnern und Bauern, zum einen intern, innerhalb der Steppe (falls Bauern eingewandert sind), zum anderen in Richtung der mitteleuropäischen Nachfolgekulturen der LBK. Gab es andauernde Kontakte zwischen den Kulturkreisen Mitteleuropa und Osteuropa oder blieb man genetisch voneinander isoliert bzw. „driftete“ wieder auseinander? Sowohl die Kupferimportfunde in der Steppe als auch die Funde in Mitteleuropa stammen aus den balkanisch-karpatischen Abbaugebieten (Häusler 1998, Chernykh 2008), so dass hier zumindest eine indirekte Verbindung zwischen den beiden Gebieten besteht.

Spätestens für die Zeit der Jamnaja-Kultur, ab etwa 3 000 v. Chr., gibt es archäologische Hinweise auf Migrationen in Richtung Westen – mindestens bis nach Ungarn (Kalicz 1998, Parzinger 1998a, 2006: 240, Gerling *et al.* 2012), wo Jamnaja-Bestattungen nachgewiesen sind. Die Schnurkeramische Kultur tauchte kurz nach der Jamnaja-Kultur auf und hatte ein Verbreitungsgebiet von Mitteleuropa bis nach Westrussland. Einige Archäologen sehen in der Jamnaja-Kultur, zusammen mit der Trichterbecher- und der Kugelamphorenkultur, den Ursprung der Schnurkeramik (Mallory & Adams 1997: 127 f, Häusler 2003: 20, Müller 2009a). Ist es möglich, oder ist es auszuschließen, dass die Träger der Jamnaja-Kultur zum Genpool der

Schnurkeramiker beigetragen haben? Sind Kontakte der Jamnaja-Kultur zu südosteuropäischen Kulturgruppen genetisch nachweisbar?

Das Ausbreitungsgebiet der Katakombengrabkultur war kleiner als das der Jamnaja-Kultur zuvor und auf die Steppe beschränkt. Kam es hier zu einer Art Rückzug und Abbruch oder Lockerung etwaiger Beziehungen zu den mitteleuropäischen Populationen?

Der moderne europäische Genpool lässt sich nicht rein durch eine Vermischung indigener mesolithischer Jäger/Sammler und früher Neolithiker (LBK) erklären (Bramanti *et al.* 2009, Haak *et al.* 2010). Wie oben erwähnt sind Jamnaja-Gruppen in Richtung Mitteleuropa migriert, was die Vermutung aufbrachte, dass sie vielleicht die fehlende Komponente in der heutigen europäischen Bevölkerung sein könnte. Gibt es Hinweise für eine Bestätigung dieser Hypothese? Vor kurzem wurde eine Studie veröffentlicht, laut der wahrscheinlich im Verlauf des Neolithikums ein jungpaläolithisches Element nordeurasischen Ursprungs zum europäischen Genpool hinzukam (Lazaridis *et al.* 2014). Möglicherweise spielte die nordpontische Steppe hier eine Rolle?

Da der nordpontische Steppenraum eine Art Bindeglied oder Korridor zwischen Europa und Westsibirien darstellt, stellt sich schließlich noch die Frage, ob genetische Affinitäten zu Kulturen östlich des Untersuchungsraums, oder Einflüsse des westsibirischen Neolithikums im Datensatz der Steppe feststellbar sind. Die Majkop-Kultur nördlich des Kaukasus soll Verbindungen über das Kaukasusgebirge hinweg gehabt haben. Gibt es eine Möglichkeit dies zu überprüfen?

2b) Die Kupfer- und Bronzezeit in der pontisch-kaspischen Steppe

Bei den obigen Fragestellungen wurden die Kulturen der Steppe als ganze Einheiten betrachtet. Im zweiten Teil des Schwerpunkts „Populationsgenetik“ sollen die inneren Strukturen der Steppenbevölkerung etwas detaillierter betrachtet werden. Die allgemeine Charakterisierung der Population wurde bereits im Rahmen der Frage zum Ursprung der Viehhirten angesprochen. Hier soll nun behandelt werden, in wieweit

sich die archäologisch definierten Kulturen genetisch erfassen bzw. unterscheiden lassen.

Im späten Äneolithikum (ca. 3 500–3 000 v. Chr.) besiedelten mehrere archäologisch distinkte Kulturgruppen den nordpontischen Raum (siehe Kapitel 4.2.2). Unterscheiden diese sich auch genetisch? Gibt es Hinweise auf ein Zusammenleben von möglicherweise eingewanderten Bauern und Jägern/Sammlern? Wie ist die genetische Beziehung zu den früheren mesolithischen und neolithischen Bewohnern der Region?

Das Dnjepr-Tal bildete im Neolithikum bis etwa 5 200 v. Chr. eine Art kulturelle, archäologisch definierte Grenze zwischen Gruppen, die bereits Getreide kultivierten und Wirtschaftshautiere hatten im Westen, und den wildbeuterischen Populationen östlich dieser Region. Schließlich verschob sich diese Grenze in östlicher Richtung zum Dnjepr (Anthony 2007: 166, 277). Waren mit diesen kulturellen auch genetische Barrieren verbunden, die sich im Äneolithikum nachweisen lassen?

Zu Beginn der Bronzezeit (ca. 3 000 v. Chr.) kam es zu einem relativ abrupten Wandel im kulturellen Erscheinungsbild. Fand der Übergang zu einem einheitlichen Grabbrauch im gesamten nordpontischen Steppenraum auf Grundlage der lokalen äneolithischen Gruppen statt? Oder basierte er auf einer Migration oder Bevölkerungsexpansion der Jamnaja-Kultur ausgehend von einem oder zwei Ursprungsgebieten in der Wolga/Ural- und Don-Region? Gab es externe Einflüsse auf die Entwicklungen, bzw. einen genetischen Eintrag von Kulturen außerhalb der Steppe?

Zwar erscheint die Jamnaja-Kultur auf den ersten Blick sehr homogen, doch gibt es archäologische Hinweise auf Unterschiede in der Lebensweise innerhalb der Kultur. Während im west- und nordpontischen Gebiet Siedlungsreste gefunden wurden, gibt es östlich des Dons praktisch keine Hinweise auf ein Siedlungswesen. Dafür ist dort ein hoher Anteil an Schafen und Ziegen nachgewiesen, was auf eine höhere Mobilität der östlichen Jamnaja-Gruppen schließen lässt (Anthony 2007: 304). Auch hier stellt sich die Frage, ob sich solche Populationsstrukturen auch in einer genetischen

Distanzierung widerspiegeln. Oder gibt es möglicherweise andere genetische Strukturen innerhalb der Jamnaja-Kultur?

Mitten in der Jamnaja-Phase, ab etwa 2 700 v. Chr. tauchten bereits die ersten Gräber der frühen Phase der Katakombengrabkultur zwischen Don und Donez auf. In der zeitlichen Abfolge lassen sich zwischen der frühen und entwickelten Phase (ab ca. 2 500 v. Chr.) der Kultur Unterschiede im Bestattungswesen ausmachen. Außerdem breitete sich die Kultur anscheinend erst in ihrer entwickelten Phase in Richtung Westen über den Dnjepr hinaus aus, woraufhin sich eine westliche (Ingul-) und eine östliche Gruppe archäologisch unterscheiden lassen (Kaiser 2003: 55). War es eine Jamnaja-Gruppe, die einen neuen Grabbau entwickelte? Kam die Idee oder kamen die Erbauer der Katakomben eventuell von außerhalb der Steppe? Gab es einen genetischen Umbruch mit Beginn der entwickelten Katakombengrabkultur, als die Jamnaja-Kultur abgelöst wurde? Lebten die beiden Gruppen der entwickelten Kultur unter sich oder gab es einen genetischen Austausch über das gesamte Verbreitungsgebiet?

Parallel zur Entwicklung der Katakombengrabkultur ging in der Wolga-Region, nördlich des Kaspischen Meeres die Jamnaja- in die Poltavka-Kultur über. Im Fundgut unterscheiden sich die beiden lediglich in der Keramik (Mallory & Adams 1997: 439 f, Häusler 1998). Unterscheiden sie sich genetisch? Wie war das Verhältnis der Poltavka-zur zeitgleichen Katakombengrabkultur nördlich des Schwarzmeeres?

Die ursprüngliche Strategie für die populationsgenetischen Analysen war eine Betrachtung maternaler, paternaler und neutraler autosomaler Marker. Diese musste aus organisatorischen Gründen auf eine rein maternale Betrachtungsweise beschränkt werden. Dabei lag der Schwerpunkt auf Methoden der deskriptiven Statistik (Diversitäts- und Distanzmaße). Abhängig von den Ergebnissen dieser Analysen können herausgearbeitete Hypothesen oder Modelle dann in weiterführenden, auf diesem Projekt aufbauenden Studien durch computerbasierte Simulationsanalysen getestet werden.

3) *Erfassung der Allelfrequenzen von unter Selektionsverdacht stehenden SNPs in der nordpontischen Steppe vor dem Hintergrund der demographischen und adaptiven Evolution*

Bei den autosomalen Markern handelt es sich um Punktmutationen, sogenannte SNPs (*single nucleotide polymorphisms*), für die zum Teil überzeugende, zum Teil auch weniger überzeugende Selektionsnachweise erbracht und -hypothesen publiziert wurden. Es geht in diesem Teil der Arbeit nicht um eine phänotypische Charakterisierung einzelner Individuen, sondern um die Veränderung von Allelfrequenzen auf der Populationsebene in der zeitlichen Dimension. Dabei lassen sich zwei grundlegende Fragestellungen unterscheiden:

Die Erkenntnisse zu den demographischen Entwicklungen in Ost- und Mitteleuropa basieren auf der Analyse der mitochondrialen DNA. Können diese Befunde durch die nukleären Daten gestützt, ergänzt bzw. erweitert oder gar in Frage gestellt werden?

Bei der zweiten Frage liegt der Fokus auf den in der Literatur beschriebenen Selektionshypothesen. Dieser Teil der Diskussion wird sich mit den physiologischen Funktionen der betroffenen Gene beschäftigen. Die Hauptfrage ist, ob die etwa 5 000 Jahre alte Steppenpopulation von den hypothetischen Selektionsereignissen betroffen gewesen sein könnte. In wieweit können die nukleären Daten Aufschlüsse über Selektion und Demographie bei der Entwicklung der Allelfrequenzen über die Zeit geben?

Aufgrund des Fehlens von neutral evolvierenden nukleären Markern für Vergleichszwecke, beschränkte sich die Analyse der putativ selektierten SNPs auf den Vergleich zwischen Genotypen- und Allelhäufigkeiten zwischen den prähistorischen und den modernen Populationen aus dem Untersuchungsraum. *Loci*, für die signifikante Frequenzunterschiede festgestellt wurden, sind bereits durch Simulationsanalysen bezüglich ihrer Evolution überprüft und die Ergebnisse vor kurzem veröffentlicht worden (Wilde *et al.* 2014).

Im Großen und Ganzen war das Ziel der vorliegenden Arbeit, das noch unvollständige Bild von Europa im Neolithikum durch den Beitrag der genetischen Daten aus der osteuropäischen Steppe zu ergänzen. In diesem Zuge sollte durch die PCR-454-Hybridmethode ein erster Schritt in der Entwicklung eines *next-generation sequencing*-Protokolls getan werden, welches das Potential von aDNA optimal ausschöpft. Die untersuchten mitochondrialen und nukleären Marker bilden dabei einen Grundstock an *loci* bzw. Daten, der in zukünftigen NGS-Arbeiten deutlich erweitert werden kann.

7 Methodische Grundlagen

7.1 Eigenschaften degenerierter alter DNA

Alte DNA (aDNA) kann aus verschiedenen Quellen gewonnen werden. Dazu gehören Knochen, Zähne, Haare, Exkremente etc. Die Untersuchung von aDNA bringt den großen Nutzen des bis dato direktesten Einblicks in vergangene bevölkerungsbiologische Prozesse, über die mit rein modernen Daten lediglich indirekt Rückschlüsse gezogen werden können. Dieser Vorteil ist jedoch auch an einige Herausforderungen gekoppelt. Das Material ist in vielerlei Hinsicht limitiert, wofür vor allem die Degradierung der DNA durch hydrolytische Prozesse verantwortlich ist.

Unmittelbar mit dem Eintritt des Todes beginnt im Rahmen der Autolyse die Auflösung von toten Zellen durch körpereigene Enzyme. Dabei wird die DNA durch hydrolytische Spaltung der Phosphodiesterbindungen durch Nukleasen abgebaut. Hinzu kommt die postmortale Zersetzung des Organismus durch Mikroorganismen. Daher liegt aDNA in Fragmenten von wenigen hundert oder noch kleineren Fragmenten vor (Hofreiter *et al.* 2001, Pääbo *et al.* 2004). Hydrolytische und oxidative Prozesse führen zu Schäden einzelner Nukleotide innerhalb dieser Fragmente. Die Zerstörung der N-glykosidischen Bindung führt zum Basenverlust. Insbesondere sind davon die Purinbasen Adenin und Guanin betroffen (Lindahl 1993). Dadurch kommt es bei der Polymerase-Kettenreaktion zum Abbruch der Strangsynthese.

Desweiteren kommt es zu Veränderungen der Basenstruktur, z. B. Deaminierungen, dem Verlust der Aminogruppe. Am weitaus anfälligsten ist dabei das Cytosin, das dadurch zu Uracil wird. Für die Bearbeitung hat dies zur Folge, dass es bei der Amplifikation zu einer Fehlpaarung kommt, und die DNA-Sequenz an dieser Stelle scheinbar eine C > T-Transition aufweist. Bei der wesentlich selteneren Deaminierung von Adenin zu Hypoxanthin kommt es bei der PCR zur Paarung mit Guanin (Lindahl 1993, Gilbert *et al.* 2003b). Verliert Guanin seine Aminogruppe und wird zu Xanthin, so hat dies nicht unbedingt eine falsche Basenpaarung zur Folge (Stiller *et al.* 2006). Freie Wasserstoffradikale jedoch führen zur Oxidation von Guanin zu 8-Hydroxyguanin,

welches eine scheinbare Transversion zu Thymin hervorruft. Außerdem kann die Ringstruktur der Pyrimidine zerstört werden. Es entstehen Hydantoine, die keine Wasserstoffbrücken mit anderen Basen ausbilden können, was, ebenso wie die Bildung von Thymindimeren bei der PCR zum Strangabbruch führt. Auch chemische Mutagene, die z. B. im Liegesubstrat vorkommen können, oder mit denen ein Leichnam behandelt wurde, führen zu Veränderungen der DNA-Struktur. Um solche postmortalen Schäden aufzudecken und von authentischen Mutationen zu unterscheiden, müssen Proben mehrfach amplifiziert und sequenziert werden. Im Gegensatz zu authentischen Mutationen, die ein Individuum charakterisieren, trägt nicht jedes DNA-Molekül an der gleichen Stelle eine deaminierte Base. Das in dieser Arbeit angewandte Reproduktionsschema ist in Kapitel 8.2.7 erläutert. Der Verlauf der DNA-Degradierung hängt von vielen Faktoren ab, insbesondere dem Liegemilieu (Mikroorganismen, pH-Wert, Temperatur, Feuchtigkeit etc.), der Liegezeit und der Lagerung der Proben nach der Entnahme aus dem Fundort, sowie der Art des DNA-haltigen Materials (Burger *et al.* 1999).

Damit gekoppelt ist ein weiterer limitierender Faktor bezüglich der Verfügbarkeit und Akquisition von Probenmaterial. Zum einen gibt es kulturelle Unterschiede im Bestattungsritus (Bestattungsort, Behandlung des Verstorbenen, soziale Strukturen im Bestattungswesen etc.), zum anderen unterscheiden sich an verschiedenen Fundplätzen die klimatische Bedingungen und das Liegemilieu. Auch die Kooperation zwischen dem Molekulargenetiker mit den Archäologen beeinflusst die Probennahme. Das dem Bearbeiter vorliegende Probenmaterial, und viel mehr noch, die Proben, aus denen erfolgreich DNA extrahiert werden kann, stammen oft nicht aus einer temporal und geographisch homogenen Verteilung, sondern eher aus punktuellen Fundansammlungen. Dies muss vor allem bei der biostatistischen Datenanalyse berücksichtigt werden.

Die Degradierung der DNA hat zur Folge, dass ihre Konzentration im Extrakt oft nur im nano- oder pikomolaren Bereich liegt, d. h. die PCR startet von wenigen *template*-Molekülen aus. Ein Eintrag von exogener DNA führt daher schnell dazu, dass diese anstatt der endogenen amplifiziert und sequenziert wird. Daher sind dekontaminierende Maßnahmen und kontaminationsvermeidendes Arbeiten oberstes

Gebot bei der Extraktion und Amplifikation alter DNA. Kontaminationsquellen sind Organismen im Liegemilieu, Archäologen, Molekulargenetiker, Reagenzien, Arbeitsbereiche, und zeitnah bearbeitete andere Proben. Oberflächenkontakte und sogar Luft dienen als Überträger. Die entsprechenden Maßnahmen und Vorkehrungen sind in Kapitel 8.2.1 und in diversen Publikationen beschrieben (z. B. Handt *et al.* 1994, Cooper & Poinar 2000, Pääbo *et al.* 2004). Wie im folgenden Abschnitt näher erläutert, spielen beim *next-generation sequencing* Mikroorganismen diesbezüglich eine besondere Rolle. Für einen ausführlichen Überblick über die aktuellen Möglichkeiten, Herausforderungen und Strategien beim Arbeiten mit alter DNA siehe Kirsanow & Burger (2012).

7.2 *Next-Generation Sequencing* in der Palaeogenetik

Seit über dreißig Jahren ist die Strangabbruch-Synthese nach Sanger *et al.* (1977) die Standardmethode zur DNA-Sequenzierung. Zwar gab es Weiterentwicklungen in der Effizienz der Enzyme, der Chemie und Technik zur Detektion der Nukleotidsequenz und in der Automation des Verfahrens, doch das Prinzip der Methode blieb das Gleiche und muss hier wohl nicht näher erläutert werden. In den letzten Jahren sind sogenannte *next-generation sequencing-Verfahren* (NGS) auf dem Vormarsch, die langsam die Sanger-Sequenzierung ablösen. Den Grundstein dafür legten Entwicklungen wie z. B. die Pyrosequenzierung (Ronaghi *et al.* 1996, Ronaghi 2001), die oberflächengebundene (*solid-phase*) Amplifikation (Adessi *et al.* 2000), und die Emulsions-PCR (Dressman *et al.* 2003). Mit der erstmaligen Veröffentlichung einer 454-Sequenzierung [454 LIFE SCIENCES, später ROCHE APPLIED SCIENCE] von Margulies *et al.* im Jahr 2005 begann eine rasante Entwicklung von NGS-Systemen, zunächst die Solexa/Illumina- [SOLEXA, INC., später ILLUMINA, INC.] (Bentley *et al.* 2008), dann die SOLiD-Plattform (*sequencing by oligo ligation detection*) [APPLIED BIOSYSTEMS INC., später LIFE TECHNOLOGIES™ CORP.] (Shendure *et al.* 2005, Valouev *et al.* 2008). 2010 kam der

Ion Torrent™ Semiconductor Sequencer [ION TORRENT SYSTEMS, INC., später LIFE TECHNOLOGIES™ CORP.] (Rothberg *et al.* 2011) auf den Markt.

Der prinzipielle Ablauf einer Probenbearbeitung im Rahmen einer NGS-Sequenzierung sieht wie folgt aus: Zunächst wird eine *DNA library* erstellt. Synthetische Oligonukleotide werden kovalent an die Enden der Fragmente im DNA-Extrakt ligiert. Diese universellen Adapter dienen später als Amplifikations- und Sequenzier-Primer. Im Unterschied zu einer herkömmlichen PCR oder Sanger-Sequenzierung muss also nicht die DNA-Probe selbst die Bindestelle für die Hybridisierung eines spezifischen Primers liefern. Die *library* wird auf einer festen Oberfläche, an der zu den Adaptersequenzen komplementäre Primer kovalent gebunden sind, amplifiziert. Somit entstehen räumlich voneinander getrennte Cluster aus „Klonsequenzen“ von jeweils einem einzelnen *DNA library*-Strang. Bei der anschließenden Sequenzierung finden die DNA-Synthese und die Detektion der eingebauten Nukleotide simultan statt (*sequencing by synthesis*), anders als bei der Sanger-Sequenzierung, bei der die enzymatische Reaktion und die Kapillarelektrophorese zwei distinkte Schritte sind. Die Länge der Sequenzen, die sogenannte *read length*, wird durch das zunehmende Hintergrundrauschen, z. B. Phasenverschiebungen bei der Synthese, begrenzt, was bei aDNA aber keine Einschränkung darstellt. Tabelle 7.1 zeigt grob die Unterschiede zwischen den vier NGS-Plattformen. Die Abläufe der 454-Sequenzierung, die in dieser Arbeit angewandt wurde, sind im folgenden Kapitel (7.2.1) detailliert beschrieben.

Im Jahr 2010 kamen die ersten *third-generation sequencing* oder auch *single-molecule next-generation sequencing*-Systeme auf den Markt. Hier ist der Sequenzierung keine PCR vorgeschaltet, so dass das Problem einer ungleichmäßigen Amplifikation umgangen wird. Bei der PacBio-Plattform [Pacific Biosystems] (Eid *et al.* 2009) z. B. sind einzelne Polymerasen an die Oberfläche eines mikroskopischen Objektträgers gebunden. Sie baut markierte Nukleotide in einen DNA-Strang ein, deren Fluorophor an die terminale Phosphatgruppe gebunden ist, anstatt wie bei anderen Methoden an die Nukleobase. Dieses wird mithilfe hochsensitiver Nanotechnologie detektiert, bevor es im nächsten Syntheseschritt abgespalten wird. Die DNA-Synthese kann demnach in Echtzeit verfolgt werden. Der erste *single-*

molecule-Ansatz war die Heliscope-Technologie von HELICOS BIOSCIENCES CORP. bei der die DNA während der Sequenzierung über Oligomere an eine *array*-Oberfläche fixiert ist. Anders als auf der PacBio-Plattform wird der Fluorophor an den Nukleotiden in einem separaten Schritt abgespalten. Der Vertrieb der Heliscope-Sequenzierer wurde 2011 eingestellt (Glenn 2011) und die Firma meldete 2012 Insolvenz an¹⁰. Das Verfahren sei hier erwähnt, da es der Vorreiter des *third generation sequencing* war und in vielen NGS-Reviews als Methode aufgeführt wird (z. B. Mardis 2008, Shendure & Ji 2008, Metzker 2010, Glenn 2011).

Tabelle 7.1: Übersicht der NGS-Plattformen

Plattform	Amplifikation	Synthese	Detektion
454	EmPCR	Polymerase	enzymatische Lichtreaktion nach der Freisetzung von Pyrophosphat beim Nukleotideinbau
Illumina	<i>Bridge</i> PCR	Polymerase	Reversibel geblockte, mit entfernbarem Fluorophor markierte dNTPs
SOLiD	EmPCR	Ligase	markierte 8nt-Oligonukleotide, deren Fluorophor-tragenden Enden abgespalten werden
Ion Torrent	EmPCR	Polymerase	Änderung des pH-Werts / der Spannung durch Freisetzung von Protonen beim Nukleotideinbau
PacBio	keine	Polymerase	in Echtzeit, durch Fluorophor-markierte dNTPs

Mittlerweile gibt es bereits Nachfolgemodelle der NGS-Systeme (zusammen mit neuen Reagenzien), die noch mehr Daten und längere *reads* generieren. Aktuell liegen die Zahlen von *reads/run* (d. h. Anzahl der generierten Sequenzen bei einem Sequenzierlauf mit voll beladener Platte) im Millionenbereich, bei Laufzeiten von wenigen Stunden. Auch gibt es Modelle, die für weniger umfangreiche Projekte wie z. B. die Sequenzierung kleiner Genome oder gezielter Genorte, geeignet sind. Für einen genaueren Einblick in die verschiedenen Systeme, Vergleiche von Stärken und Schwäche, Datenmengen, Kosten etc. seien einige der zahlreichen Übersichtsartikel zu empfehlen: Metzker 2010, Glenn 2011¹¹, Liu *et al.* 2012, Mardis 2013.

¹⁰ http://en.wikipedia.org/wiki/Helicos_Biosciences [Stand: 01.10.2014]

¹¹ Zu Glenn *et al.* 2011 sind unter <http://www.molecularrecologist.com/next-gen-fieldguide-2014/> Aktualisierungen zu den Tabellen (Laufzeiten, Sequenzlängen, Kosten, Fehlerraten, etc.) für das Jahr 2014 finden [Stand: 01.10.2014].

Das *next-generation sequencing* mit seinen vor allem in der Anfangszeit sehr kurzen *read*-Längen schien wie für die Sequenzierung von alter DNA gemacht. Mit einer PCR können lediglich die DNA-Fragmente erfasst werden, an die beide Primer hybridisieren können, demnach müssen sie sowohl die Bedingung der passenden Sequenz, als auch die einer Mindestlänge von etwa 60 – 70 (eher 100) Nukleotiden (nt) erfüllen. Bei der NGS-Sequenzierung eines Neandertalers waren über die Hälfte der Sequenzen kürzer als 70 nt (Green *et al.* 2008). Durch das *shotgun*-Prinzip beim NGS könnte theoretisch das gesamte Extrakt erfasst werden, anstatt z. B. nur die HVR I und vielleicht einige wenige autosomale *loci*. Ein großes Problem ist jedoch die Zusammensetzung von aDNA-Extrakten. Wenn es sich nicht um extrem gut erhaltene Proben, z. B. aus dem Permafrost handelt, macht die endogene DNA lediglich einen geringen Anteil aus, weit unter 10 % (Green *et al.* 2006, Green *et al.* 2010) oder sogar unter 1 % (Briggs *et al.* 2009, Burbano *et al.* 2010). Beim Rest handelt es sich fast ausschließlich um mikrobielle DNA¹².

Diesem Problem entgegen wirken *target enrichment*- bzw. *DNA capture*-Technologien. Dabei werden vor der Sequenzierung genomische Regionen oder eine Vielzahl unabhängiger *loci* selektiv aus der Probe herausgefischt. Einen Überblick über die verschiedenen Strategien geben Mamanova *et al.* (2010). Auch die PCR und die Multiplex-PCR zählen im Prinzip zu den *target enrichment*-Methoden. Sie haben von allen Methoden die höchste Spezifität, allerdings ist die Anzahl der Primersysteme und folglich Genorte stark eingeschränkt. Zu viele Primer in einer Multiplex-PCR führen zu unspezifischer Amplifikation z. B. aufgrund von Interaktionen untereinander, oder für einzelne Systeme ungünstige Hybridisierungstemperatur. Außerdem kommt es immer wieder zu Nichtamplifikationen einzelner Fragmente (*dropouts*). Um diese Probleme zu minimieren, müssen Primer, Reagenzien und die sonstigen Amplifikationsbedingungen wie Temperatur, Zyklenzahl und Dauer der einzelnen Schritte gut aufeinander abgestimmt sein. Eine weitere PCR-Methode ist die *microdroplet-based PCR* (Tewhey *et al.* 2009). Sie funktioniert ähnlich wie eine

¹² Ein Testversuch einer *shotgun library* aus aDNA-Extrakten von Proben aus diesem Projekt und anschließender Sequenzierung auf dem Illumina Genome Analyzer lieferte so wenige endogene Sequenzen, dass sie inhaltlich nicht ausgewertet wurden. Daher wird dieser Versuch in der vorliegenden Arbeit auch nicht weiter besprochen.

Emulsions-PCR (Dressman *et al.* 2003), indem jedes *template* in einem eigenen wässrigen Mikro-Tropfen, umgeben von Öl, amplifiziert wird. Aufgrund der benötigten Ausgangsmenge an DNA von 7,5 µg ist diese Methode bei aDNA-Analysen jedoch nicht anwendbar.

Bereits mit einer Ausgangsmenge von 200 ng DNA kommt die Anreicherung mit *molecular inversion probes* aus (Hardenbol *et al.* 2005, Deng *et al.* 2009). Bei diesen inversen Sonden handelt es sich um Oligonukleotide mit zwei *target*-spezifischen Sequenzabschnitten – ähnlich wie Primer, allerdings für den gleichen Strang – und einer Verbindungssequenz dazwischen. Die beiden äußeren Stücke hybridisieren mit den flankierenden Bereichen der Zielsequenz. Eine Ligase verbindet anschließend die beiden Enden der Sonde durch Synthese des Gegenstranges der DNA dazwischen, so dass ein zirkuläres Molekül entsteht. Die Verbindungssequenz der Sonde dient als universeller Primer für die Amplifikation der gefischten Sequenzen.

Immer mehr an Bedeutung gewinnen die kommerzialisierten *hybrid capture*-Methoden. Man unterscheidet dabei zwischen *on-array* und *in-solution capture*. Dem *capture* vorgeschaltet ist die Erstellung einer genomischen (*shotgun*) *library* der Probe. Ausgewählte Sequenzen werden dann durch Hybridisierung an Sonden herausgefischt. Als Sonden dienen einzelsträngige *target*-spezifische DNA- oder RNA-Stücke, deren Länge abhängig ist vom Hersteller, z. B. 120 nt beim Agilent SureSelect™ Target Enrichment¹³, ein *in-solution capture*, mit dem bis zu 24 Megabytes DNA gefischt werden können. Bei der *in-solution*-Variante sind die Sonden biotinyliert, und die Hybridisierung findet frei in Lösung statt. Anschließend werden die Produkte durch Bindung an mit Streptavidin überzogene *beads* von der restlichen DNA isoliert (Gnirke *et al.* 2009). Beim *on-array capture* sind die Sonden an der Oberfläche eines *microarrays* fixiert (Albert *et al.* 2007). Zwei Studien zu Neandertaler-Sequenzierungen demonstrieren das *hybrid capture* von aDNA. Nach einem *in-solution enrichment* belief sich der Anteil an Sequenzen bzw. *reads* des Neandertalers aus Vindija,

¹³ Dieses Produkt wurde für einen Testversuch zur Anreicherung von mtDNA von drei aDNA-Proben aus dieser Arbeit verwendet. Aufgrund von Kreuzkontaminationen durch gleichzeitig angereicherte, sehr gut erhaltene eisenzeitliche Proben waren die Ergebnisse allerdings nicht verwertbar. Daher wird dieser Versuch in der vorliegenden Arbeit nicht weiter besprochen. Mittlerweile wurde die Methode für aDNA optimiert und wird in den Nachfolgeprojekten zu dieser Arbeit verwendet.

Kroatien, in der Studie von Noonan und Kollegen (2006) auf etwa 30 % der Gesamt-*reads*. Ein *on-array capture* wurde mit dem zu 99,8 % aus mikrobieller DNA bestehenden Extrakt des Individuums aus der El Sidrón Höhle in Spanien durchgeführt (Burbano *et al.* 2010). Der Anteil der *reads* „*on target*“ betrug etwa 37 %.

Insgesamt erhöht das *next-generation sequencing* in Kombination mit *DNA capture* die Ausbeute bei der Analyse degradierter DNA im Vergleich zu konventioneller PCR mit Sanger-Sequenzierung um ein Vielfaches. Eine einmal erstellte aDNA-*library* kann theoretisch wiederholt amplifiziert werden, so dass das Ausgangsmaterial längerfristig zur Verfügung steht. Inwieweit dies praktisch umzusetzen ist, wird zurzeit am Institut für Anthropologie der Universität Mainz getestet. Mit einem *DNA capture* können wesentlich mehr *loci* aus einem Extrakt untersucht werden. Außerdem werden auch kürzere Fragmente erfasst, die bei einer PCR keine Berücksichtigung finden. Bisher wurde davon ausgegangen, dass aDNA unter idealen Bedingungen maximal 100 000 Jahre erhalten bleibt (Lindahl 1993). Mit der Zugänglichkeit kürzerer DNA-Stücke scheint sich diese Zeitspanne auf über 300 000 Jahre zu erhöhen (Dabney *et al.* 2013, Meyer *et al.* 2014). Dennoch sollte man solche Zahlen mit entsprechender Vorsicht betrachten.

Bei allen Vorzügen, die das *NGS* für die Palaeogenetik bietet, bestehen die Problemfelder Kontaminationen und Authentizität weiterhin. Auch die Kosten *versus* Nutzen-Frage muss sich gestellt werden. Zwar ist die Sequenzierung in Bezug auf Kosten pro *read* günstig, doch sollte die Sequenzierstrategie (welche Proben und welche *loci* beantworten die Fragestellung, welche *pooling*-Strategien sind möglich etc.) gut überlegt sein, insbesondere aufgrund der Unberechenbarkeit von degradierter DNA in Bezug auf Amplifikations- und Sequenziererfolg und den hohen Anteilen an mikrobieller DNA in den Extrakten.

Eine der Herausforderungen auf der labortechnischen Seite ist allgemein die zur Verfügung stehende Menge an alter DNA und ihre Quantifizierung. In den Standard-Protokollen zum *pooling* von Proben, zu den Schritten des *target enrichment* und den Vorbereitungen für die Sequenzierung gibt es Mengen-, Volumen-, Konzentrations- und Äquimolaritätsvorgaben, die beim Arbeiten mit aDNA oft nicht einzuhalten sind.

Dies stellt einen weiteren Unsicherheitsfaktor für den Erfolg der arbeits- und kostenaufwendigen Prozedur dar. Auch sind die Verluste an Material, vor allem während der vielen Aufreinigungsschritte nicht zu unterschätzen (Knapp & Hofreiter 2010).

Das größte Kontaminationsrisiko bei einer *NGS*-Sequenzierung besteht bei der Erstellung der *aDNA library* (Green *et al.* 2009). Sie besteht aus einer Reihe von Einzelschritten, und es kommen viele Reagenzien zum Einsatz. Insbesondere die Kontamination der Adapter-Lösung wäre fatal. Sobald die Adapter fest an die DNA-Fragmente gebunden sind, minimiert sich das Risiko des Eintrags kontaminierender Sequenzen auf das *carry-over* von anderen *libraries*, bei denen die gleichen Adapter verwendet wurden. Die Verwendung von Adaptern mit zusätzlichen Erkennungssequenzen (*tags*), z. B. zur Projekt- oder Probenidentifikation ist zu empfehlen. Ein Vorteil des *NGS* in Bezug auf Kontaminationen ist die bereits erwähnte Erfassung von Molekülen in der Größenordnung zwischen ca. 30 und 60 nt, die mittels PCR nicht amplifiziert werden können. Zwar können moderne Kontaminationen auch degradiert sein, jedoch nicht unbedingt in dem Maße wie die prähistorische DNA. Da die extrem kurzen Fragmente einen Großteil der endogenen alten DNA in einem Extrakt ausmachen (Green *et al.* 2009, Krause *et al.* 2010), verschiebt sich bei das Verhältnis der Anzahl erfasster Moleküle zwischen *aDNA* und moderner Kontamination deutlich in Richtung *aDNA*.

Eine der größten Herausforderung im Zeitalter des *NGS* ist die Datenauswertung. Im Rahmen des *NGS*-„Booms“ wurden massenhaft Sequenzdaten produziert, während für diese Datenvolumina geeignete bioinformatische Werkzeuge erst nach und nach entwickelt wurden. Mittlerweile gibt es eine große Auswahl an Programmen, die speziell an die Ansprüche von *NGS* angepasst sind, z. B. in der Sequenzbeschneidung, dem *assembling*¹⁴, dem Abgleich der Sequenzen mit Datenbanken (Blast), die *coverage*¹⁵ betreffende Berechnungen, und die SNP-Identifizierung. Bei alter DNA kommt zusätzlich noch die Hürde der Authentifizierung der Sequenzierungs-

¹⁴ Zusammenführung von Nukleotidsequenzen zu längeren Fragmenten durch Sequenzalignments.

¹⁵ Die *coverage* eines *locus* bzw. einer Nukleotidposition gibt an, wie viele *reads* der *NGS*-Sequenzierung diese Position abdecken. Bei einer *shotgun*-Sequenzierung entspricht die *coverage* der Anzahl der originalen *template*-Sequenzen aus dem Extrakt.

Ergebnisse hinzu. Zur Detektion von Kontaminationen werden bei den traditionellen Verfahren Leerkontrollen auf das Vorhandensein von DNA überprüft. Dies geschieht nach dem Ja/Nein-Prinzip. Gegebenenfalls kann die Sequenzierung einer positiven Leerkontrolle Aufschluss über die Kontaminationsquelle geben oder darüber, ob die im gleichen Versuchsansatz bearbeiteten Proben den Typ aus der Kontrolle aufweisen. Die Authentifizierung erfolgt durch Replikationen der Ergebnisse.

Beim NGS liegt die Hauptlast der Identifikation von Kontaminationen und Sequenzierfehlern auf der bioinformatischen Ebene. Verschiedene Studien suchten, meist anhand der Neandertaler-*libraries*, nach Mustern in der Degradierung alter DNA, um so authentische Sequenzen von Kontaminanten oder auch Sequenzierfehlern abzugrenzen. Es wird hauptsächlich nach Degradierungsmustern in alter DNA gesucht. Wie schon erwähnt, können Unterschiede in den Sequenzlängen ein Unterscheidungskriterium sein, da die kontaminierenden DNA-Fragmente durchschnittlich länger sind; es gibt allerdings auch Gegenbeispiele (Krause *et al.* 2010). Briggs *et al.* (2007) zeigten dass aDNA-Sequenzen häufig direkt nach einem Nukleotid mit Purinbase einen Strangbruch aufweisen, möglicherweise gefördert durch Instabilität des Moleküls nach einer Depurinierung (siehe auch Sawyer *et al.* 2012). Ein weiterer Ansatzpunkt ist das Muster von Sequenzierfehlern aufgrund von Deaminierungen, die fast ausschließlich Cytosin betreffen (Gilbert *et al.* 2003a). Bei NGS-Sequenzierungen treten „scheinbare“ C > T-Substitutionen gehäuft am 5'-Ende, G > A-Transitionen eher am 3'-Ende der Fragmente auf. In beiden Fällen handelt es sich um Deaminierungen von Cytosin, die beim *fragment end polishing* und dem *nick repair* bei der *library*-Konstruktion durch Strangsynthesen am 3'-Ende als G > A-Austausche erscheinen (Briggs *et al.* 2009, Krause *et al.* 2010, Sawyer *et al.* 2012). Die modernen Kontaminationen in diesen beiden Studien wiesen eine gleichmäßige Verteilung der in weitaus geringerem Maße vorkommenden Austausche auf (siehe dazu aber auch García-Garcerà *et al.* 2011). Ein weiteres Merkmal von als Mutationen erscheinenden postmortalen Schäden in kodierenden DNA-Abschnitten ist, dass sie alle Codon-Positionen betreffen. Im Gegensatz dazu sind echte Mutationen vor allem an der dritten Codon-Position zu beobachten (Gilbert *et al.* 2003b). Eine bei Vergleichen zwischen der DNA von Neandertalern und anatomisch modernen

Menschen verwendete Methode zur Quantifizierung von Kontaminationen ist die Identifikation von mtDNA-Positionen, von denen im Vorfeld durch PCR-Analysen bekannt war, dass sie in der heutigen Weltbevölkerung fixiert sind, und der Neandertaler sich davon unterscheidet. Dadurch konnte der Anteil menschlicher Sequenzen für diese *loci* ermittelt werden und die Kontaminationsrate der mtDNA bestimmt werden (Green *et al.* 2010). In dieser Studie sind noch weitere Authentifizierungsmethoden, z. B. die Betrachtung von Transversionen zwischen Neandertaler und dem anatomisch modernem Menschen etc., aufgeführt.

Die beschriebenen Muster geben eine theoretische bzw. statistische Erwartung an das Verhalten alter DNA vor, eine Art Modell. Doch bekanntermaßen ist die Varianz zwischen Proben und Genorten bei der Degradierung enorm. Im Prinzip ist es nicht möglich, die *reads* einer Probe vollständig in authentische und kontaminierende Sequenzen einzuteilen. Es kann lediglich eine Kontaminationsrate auf Basis einer Reihe von *loci* ermittelt werden und anschließend eine Wahrscheinlichkeitsabschätzung über die Authentizität der gesamten Daten gemacht werden, insbesondere bei diploiden Genomen. Auch beim *NGS* wird man auf Ergebnisreproduktionen nicht so einfach verzichten können, oder zumindest muss die *coverage* hoch genug sein, um das Vorhandensein alternativer Sequenzen zu erkennen (Ho & Gilbert 2010). Wie auch bei den traditionellen Analysemethoden wird das Problem der Authentifizierung wesentlich größer, wenn menschliche Populationen untersucht werden, vor allem wenn sie nicht 30 000 Jahre (Krause *et al.* 2010), sondern nur beispielsweise 3 000 Jahre von uns „entfernt“ sind.

7.2.1 Roche 454 Genome Sequencer FLX Sequenzierung

Für die palaeogenetischen Untersuchungen in der vorliegenden Arbeit wurde ein Herangehensweise gewählt, die eine klassische Form des *target enrichments*, die Multiplex-PCR, mit *next-generation sequencing* kombiniert. Die Vertrautheit des Lesers mit den Grundlagen zu PCR und Multiplex-PCR seien vorausgesetzt oder können anderweitig recherchiert werden. Einzelheiten zur Erstellung der Multiplex-PCRs für dieses Projekt sind dem Kapitel zu Material und Methoden (8.2.4) zu

entnehmen. Allgemeines zum *next-generation sequencing* in der Palaeogenetik findet sich im vorhergehenden Abschnitt. Hier soll speziell auf die 454-Sequenzierung und anschließend auf die im Rahmen dieses Projektes etablierte *tagging*-Methode für die parallele Sequenzierung vieler Proben eingegangen werden.

Für die 454-Sequenzierung der im Verlauf dieser Arbeit amplifizierten DNA wurde das Sequenzierunternehmen GATC BIOTECH GMBH in Konstanz beauftragt. Zunächst wurden die Sequenzierungen auf einem 454 Genome Sequencer FLX Gerät der ersten Generation durchgeführt. Dann wurde 2009 die „Titanium“-Chemie eingeführt, welche die *wells* in den Sequenzierplatten lichttechnisch besser isolierte und die *read*-Längen auf über 400 nt erhöhte (Voelkerding *et al.* 2009). Bereits ein Jahr später verwendete die GATC BIOTECH GMBH das ROCHE GS FLX+-System, das sich durch eine *read length* von durchschnittlich 700 Basen (bis zu 1000) auszeichnete und etwa eine Millionen *reads pro run* produzieren kann. Diese „Verbesserungen“ der Sequenzlängen sorgten dafür, dass für die aDNA-PCR-Produkte von der *library*-Konstruktion bis zur 454-Sequenzierung Anpassungen durchgeführt werden mussten, damit es nicht zur Verwerfung von Proben aufgrund zu kurzer Fragmentlängen kam.

Da die entsprechenden Schritte der Probenbearbeitung nicht selbst durchgeführt wurden, soll hier zumindest ein Überblick über die allgemeine Vorgehensweise bei dieser Sequenziermethode gegeben werden.

7.2.1.1 *Library*

Aus extrahierter DNA wird zunächst eine *library* erstellt. Bei moderner genomischer DNA wird diese zunächst in 400 – 800 Basenpaare (bp) lange Stücke geschert (z. B. durch Nebulisierung / Zerstäubung, Ultraschall, Zentrifugation). Bei alter DNA oder kurzen PCR-Produkten kann dieser Schritt übersprungen werden. Fragmentierte DNA-Doppelstränge weisen Einzelstrang-Überhänge auf. Diese werden beim sogenannten „*fragment end polishing*“ durch Exonukleolyse oder Extension des Gegenstranges entfernt. Außerdem werden 3'-Phosphatreste entfernt und 5'-Enden phosphoryliert. Diese entstehen, wenn die Bruchstelle des DNA-Stranges zwischen dem C5-Atom der

Pentose und der Phosphatgruppe liegt. Das *end polishing* ist auch bei PCR-Produkten notwendig, wenn dephosphorylierte Primer oder eine Polymerase, die Überhänge produziert, eingesetzt wurden.

An die glatten Enden der DNA-Fragmente werden im nächsten Schritt die Adapter ligiert. Hierbei handelt es sich um zwei doppelsträngige Oligonukleotide (Adapter A und B). Sie sind dephosphoryliert, um bei der Ligation die Bildung von Adapterconcatemeren zu verhindern. Ausgehend von dem Ende, das mit dem DNA-Fragment verknüpft wird, sind die Adapter wie folgt aufgebaut: Eine vier Basen lange nicht-palindromische Erkennungssequenz dient bei der Sequenzierung der Identifizierung legitimer *reads*. Ihr folgen nacheinander Anlagerungsstellen für PCR- und Sequenzier-Primer. Außen tragen beide Adapter einen vier Basen langen 5'-Überhang. Hierdurch wird eine direktionale Ligation gewährleistet. Adapter B trägt an diesem Überhang eine Biotin-Markierung. Gegebenenfalls können Adapter verwendet werden, die eine kurze Erkennungssequenz tragen, sogenannte *molecular identifier (MID) tags*, so dass mehrere *libraries* zusammen sequenziert werden können.

Bei der Ligation entstehen neben DNA-Fragmenten, die sowohl einen A- als auch einen B-Adapter tragen, auch unerwünschte Produkte, wie z. B. Fragmente mit zwei gleichen oder nur einem Adapter. Um diese weitgehend auszusortieren, werden alle Produkte mit mindestens einem B-Adapter an mit Streptavidin beschichtete *magnetic beads* gebunden. Fragmente, die keine oder nur A-Adapter tragen, sind nicht biotinyliert und werden ausgewaschen. Produkte, die nur den B- aber keine A-Adapter tragen, spielen in späteren Reaktionen keine Rolle, da sie z. T. noch entfernt werden, bzw. die folgende PCR aufgrund der fehlenden Andockstelle für den Primer nicht stattfinden kann.

Da die Adapter an ihren 5'-Enden keine Phosphatgruppen tragen, kann bei der Ligation jeweils nur der Gegenstrang kovalent an das DNA-Fragment gebunden werden. Der sogenannte „*nick*“ (Kerbe; Unterbrechung im Einzelstrang) zwischen 3'-DNA und 5'-Adapter wird nach der *library*-Immobilisierung in einer *fill-in*-Reaktion durch eine *strand displacement polymerase* repariert. Hierbei werden auch die Überhänge komplementär aufgefüllt.

Um eine *single strand library* zu erhalten, wird die immer noch an die Streptavidin-*beads* gebundene DNA denaturiert. Dies geschieht entweder durch eine Behandlung mit Natriumhydroxid oder durch Erhitzen. Im ersten Fall bleiben die Streptavidin-Biotin-Bindungen erhalten, die *library* enthält ausschließlich die abgelösten nicht-biotinylierten Stränge, also die, bei denen der B-Adapter das 3'-Ende bildet. Da diese Methode mit einem extrem hohen Verlust an Material einhergeht, schlugen Maricic und Pääbo (2009) eine Inkubation des Ansatzes bei 90 °C als Alternative vor. Die Verluste sind hier minimal, außerdem kommt es bei Hitze zu einer Dissoziation des Streptavidin-Biotin-Komplexes. Daher lösen sich beide Stränge von den *beads* und werden Bestandteil der *library*. Die *beads* werden mithilfe eines *magnetic particle collectors* konzentriert und von der gelösten *library* separiert.

Zum Abschluss werden noch Qualität und Quantität der *library* ermittelt. Mithilfe einer Messung mit einem *bioanalyzer* können Spannweite und Durchschnitt der Fragmentlängen in der *library* erfasst werden. Für eine funktionale Quantifizierung zur Ermittlung des optimalen Volumens der *library* für deren Einsatz in der folgenden Emulsions-PCR (EmPCR) (d. h. das Volumen mit dem größten Anteil an *reads*, die es nach der Sequenzierung durch den Qualitätsfilter schaffen), werden Verdünnungsreihen der *library* erstellt und ein Titrations-*assay* durchgeführt. Da dies jedoch sehr zeit- und kostenaufwendig ist, kann alternativ eine *real-time quantitative PCR* (RT qPCR) durchgeführt werden.

7.2.1.2 Emulsions-PCR

Die emulsionsbasierte klonale Amplifizierung der *library* ersetzt das herkömmliche bakterielle Klonieren der *template*-Sequenzen. Die Reaktion findet hierbei in sich in einer Ölphase befindlichen aquatischen Micellen statt. Diese Wasserbläschen sollten möglichst ein einzelnes DNA-*template* und die für die Amplifikation notwendigen Reagenzien enthalten. Das Erreichen dieses Zustands ist ein rein stochastischer Prozess, weshalb die oben beschriebene funktionale Quantifizierung der *library* unumgänglich ist.

Zunächst werden die DNA-Fragmente der *library* an sogenannten *DNA capture beads* immobilisiert. An deren Oberfläche sind Oligonukleotide kovalent fixiert, die komplementär zur PCR-Primer-Region von Adapter B sind¹⁶. Die Konzentrationen sind so gewählt, dass möglichst ein DNA-Fragment an ein *capture bead* hybridisiert. Die an *beads* gebundenen *templates* werden zusammen mit einem Amplifikationsmix, der alle notwendigen PCR-Reagenzien enthält, in Emulsions-Öl gegeben und durch starkes Schütteln in einem *tissue lyser* emulsifiziert.

Die eigentliche Amplifikation findet in jeder Micelle isoliert und unabhängig von allen anderen – mit einem DNA-Einzelstrang als *template* – statt. Die im Überschuss an das *bead* verankerten Oligonukleotide dienen als Primer, so dass die bei der Elongation entstehenden DNA-Stränge ebenfalls an das *bead* gebunden sind. Der Ansatz enthält auch in Lösung vorliegende Primer. Der an Adapter A hybridisierende Primer ist biotinyliert. Neu synthetisierte Stränge, die nicht direkt an das *bead* gebunden sind, können wiederum als *template* an festsitzenden Primern dienen. Am Ende der EmPCR trägt ein *DNA capture bead* zwischen 10 und 30 Millionen Kopien des *templates*.

Die Emulsion wird chemisch durch Zugabe von Isopropanol aufgebrochen, die *beads* mithilfe einer Filtereinheit aufgefangen, gewaschen und resuspendiert. Um die *beads* zu entfernen, an denen keine Amplifikation stattgefunden hat, wird ein *DNA library bead enrichment* durchgeführt. Erneut werden Streptavidin-beschichtete *magnetic beads* eingesetzt, welche die erfolgreich amplifizierte Moleküle über die Biotin-Markierung am A-Primer binden. Sie werden mithilfe des *magnetic particle collectors* pelletiert. *DNA capture beads*, an denen keine Amplifikation stattfand, werden durch Waschen entfernt.

Durch Zufügen einer Natriumhydroxid-haltigen Schmelzlösung werden anschließend die doppelsträngigen PCR-Produkte denaturiert. Für die Sequenzierung werden die an den *DNA capture beads* fixierten Stränge verwendet. Zusammen mit einem *annealing*-Puffer wird der zu Adapter A komplementäre Sequenzier-Primer

¹⁶ Dies gilt für das Standardprotokoll für *shotgun*-Sequenzierungen. Je nach Applikation kann auch Adapter A als *target* für entsprechende *capture beads* zur Hybridisierung dienen. Bei *paired-end* oder *amplicon libraries* findet die Hybridisierung erst während der Amplifikation statt.

(Standard-*shotgun*-Protokoll) an die DNA hybridisiert. Mithilfe eines *Coulter counters* kann die Anzahl erfolgreich amplifizierter *templates* abgeschätzt werden. Etwa 10 – 15 % der ursprünglich eingesetzten *beads* sollten nach dem *enrichment* noch vorhanden sein.

7.2.1.3 Pyrosequenzierung

Die Sequenzierung wurde auf dem ROCHE Genome Sequencer FLX+ der 454 LIFE SCIENCES CORPORATION, später übernommen von ROCHE APPLIED SCIENCE, durchgeführt. Es handelt sich um eine Pyrosequenzierung, eine „*sequencing by synthesis*“-Methode (Ronaghi *et al.* 1996). Anders als bei der Sanger-Sequenzierung (Sanger *et al.* 1977) wird hier die Sequenz „live“ erfasst, während die DNA synthetisiert wird. Bei der DNA-Synthese wird bei der Verknüpfung zweier Nukleotide Pyrophosphat freigesetzt. Dieses wird über eine Enzymkaskade für die Erzeugung eines Lichtsignals verwendet.

Die Durchführung einer Sequenzierung ist in Margulies *et al.* (2005) eingehend beschrieben. Zunächst wird eine Pikotiterplatte mit den *template*-tragenden *DNA beads* beladen. Sie werden zusammen mit einem Puffer über die Platte gespült und dann mittels Zentrifugation in die *wells* gedrückt. Diese haben einen Durchmesser von 44 µm, so dass maximal ein *DNA bead* hineinpasst. Die Menge an eingesetzten *beads* ist basierend auf der Zählung nach der EmPCR so gewählt, dass viele *wells* unbesetzt bleiben, um eine Überstrahlung zu vermeiden. Mithilfe von *packing beads* werden die Reaktionskomponenten im *well* stabilisiert. Hinzu kommen in einer dritten Schicht die *enzyme beads*, an denen Sulfurylase und Luciferase, die Enzyme des chemilumineszierenden Systems fixiert sind. Der Puffer in den *wells* enthält die Polymerase zur DNA-Synthese, den entsprechenden Cofaktor und Apyrase, welche die Degenerierung von Nukleosidtriphosphaten zu Nukleosidmonophosphaten katalysiert.

Nach dem Beladen wird die Pikotiterplatte senkrecht in den Genome Sequencer installiert. Gegenüber der Platte befindet sich eine CCD-Kamera zur Detektierung der Lichtsignale. Während des *sequencing runs* werden in wiederholter Abfolge die vier

Nukleotidtypen für die DNA-Synthese einzeln über die Pikotiterplatte gespült. Dabei folgen nach jedem dNTP noch mehrere Chemikalien-Flutungen:

dNTP \Rightarrow Adenosinphosphosulfat (APS) + Luciferin \Rightarrow Apyrase \Rightarrow APS + Luciferin

Diese Folge wird auch als „*flow*“ bezeichnet. Ein Zyklus bei der Sequenzierung entspricht einer viermaligen Durchführung dieser Abfolge, also vier *flows*, mit jedem dNTP einmal. Insgesamt finden mittlerweile 100 Zyklen beim GS FLX System bzw. 200 mit dem GS FLX Titanium statt. Die Synthese des Komplementärstrangs zum *template* aus der *library* und damit die Sequenzierung beginnen am 3'-Ende des Sequenzier-Primers. Bei jeder Zugabe von dNTPs tritt eines der folgenden Szenarios ein:

- (1) Die in die *wells* gespülten dNTPs tragen nicht die nächst einzubauende, zum *template* komplementäre Base.
- (2) Die in die *wells* gespülten dNTPs tragen die nächst einzubauende, zum *template* komplementäre Base.
- (3) Im *template* folgt ein Homodimer, -trimer oder -polymer (ein sogenannter „*stretch*“). Die in die *wells* gespülten dNTPs tragen die nächsten einzubauenden, zum *template* komplementären Basen.

In Fall (1) findet keine Synthese statt. Die dNTPs und weiteren in die *wells* gespülten Chemikalien werden unverbraucht gewaschen bzw. vorher noch degradiert. Dies übernimmt eine ATP-Diphosphohydrolase, kurz Apyrase. Sie katalysiert die Abspaltung des γ -Phosphats von ATP, und die des β -Phosphats von ADP, so dass schließlich AMP entsteht. Auch andere Nukleosidtri- und -diphosphate dienen als Substrate. Hierdurch wird gewährleistet, dass bei der Zugabe des nächsten dNTPs keine verwendbaren Überreste aus der vorigen Abfolge mehr im Ansatz zur Verfügung stehen.

In den Fällen (2) und (3) wird die entsprechende Anzahl an Nukleotiden von der Polymerase in den Strang eingebaut. Dabei werden die beiden äußeren Phosphatgruppen (Pyrophosphat) des anzuhängenden Nukleotids freigesetzt. Mit dem Pyrophosphat bildet die ATP-Sulfurylase aus dem ins *well* gespülte Adenosinphosphosulfat (APS) Adenosintri-phosphat (ATP). Dieses wiederum dient der

Luciferase als Cofaktor für die oxidative Carboxylierung von Luciferin zu Oxiluciferin. Bei dieser Reaktion befindet sich das oxidierte Luciferin kurzzeitig in einem elektronisch angeregten Zustand. Beim Übergang in den Grundzustand wird ein Lichtquant emittiert. Die Millionen von gleichzeitigen Lichtblitzen in einem *well* – da sich an einem *DNA capture bead* Millionen von Kopien desselben Moleküls befinden – werden von der Kamera erfasst, ebenso wie die Koordinaten des entsprechenden *well*s in der Sequenzierplatte. Da die Reihenfolge der dNTP-Zugabe fest definiert ist, ist durch die Zeitpunkte der Photonenemissionen in einem *well* automatisch die Basenabfolge der DNA bekannt. Die Intensität des Lichts ist dabei annähernd proportional zur Menge an freigesetztem Pyrophosphat und damit zur Anzahl eingebauter Nukleotide, so dass auch in Fall (3) die Länge von Homopolymeren direkt erfasst wird. Schließlich werden nicht verbrauchte dNTPs und ATP, welches unter Umständen von der Sulfurylase produziert, aber anschließend nicht verwertet wurde, von der Apyrase zu Monophosphaten degradiert und gewaschen. Die Reaktionen im Verlauf dieser Enzymkaskade sind in Abbildung 7.1 skizziert.

Wie oben beschrieben ist die Anzahl der Zyklen und damit die *read length* begrenzt. Zum einen wirken sich mit zunehmender Zyklenzahl nicht-synchronisierte Extensionen verstärkt auf die Sequenzierung aus. Zu diesem Phänomen kommt es, wenn bei einigen Molekülen an einem *bead* während eines *flows* zufällig kein Nukleotid eingebaut wird. Insbesondere bei Homopolymeren ist die Gefahr solcher *frameshifts* groß. Die Apyrase baut freie dNTPs bereits ab, während die Polymerase noch den Strang synthetisiert. Im Laufe des *sequencing runs* akkumulieren die nicht-synchronen Stränge. Die maximale *read length* wird außerdem durch die „Müllanhäufung“ von anorganischen Phosphaten (P_i), Nukleosidmonophosphaten (NMPs), Sulfaten etc., beeinflusst, die inhibierend auf die Enzyme wirken kann (Ronaghi *et al.* 1996).

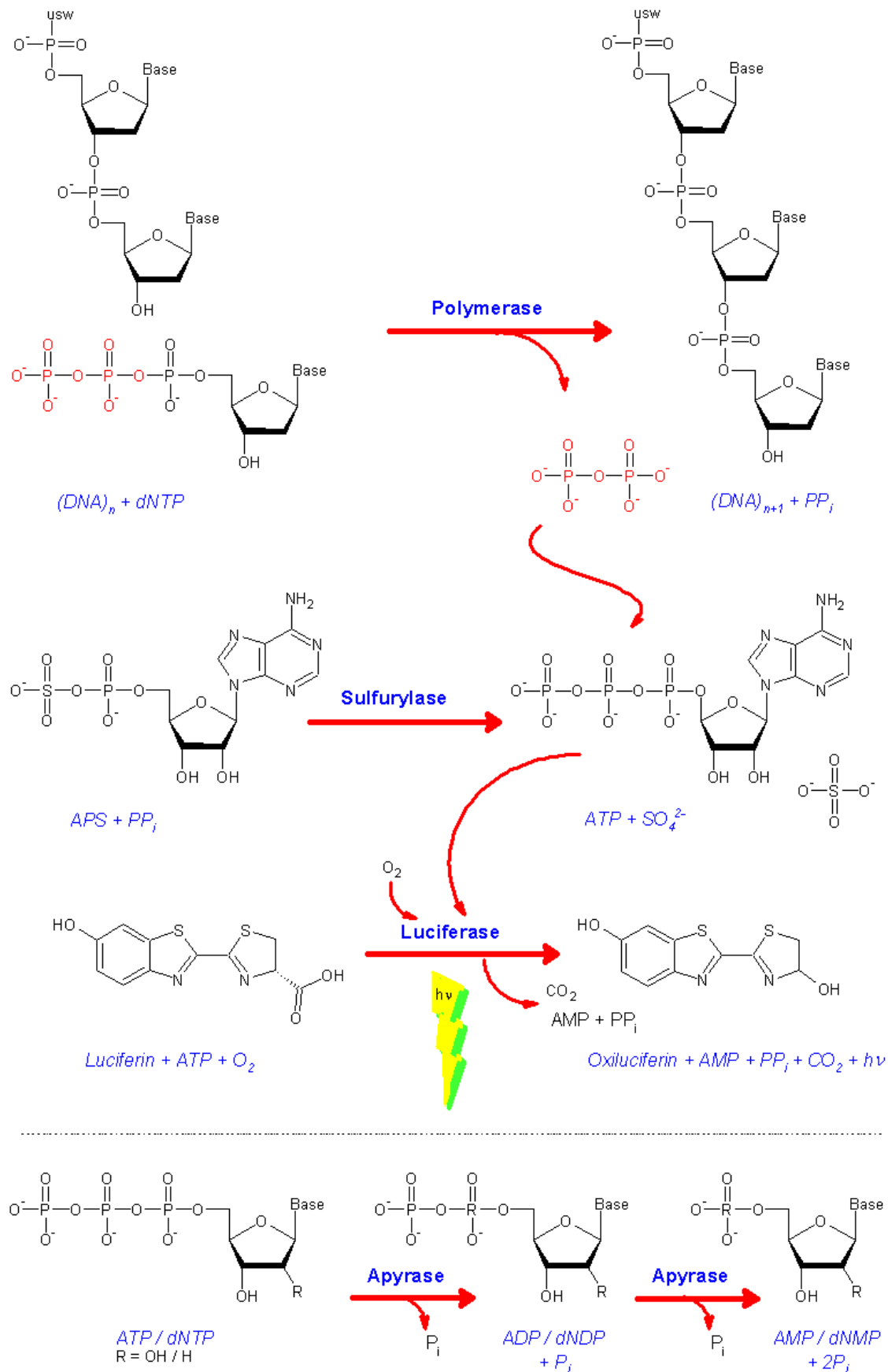


Abbildung 7.1: Enzymkaskade der photochemischen Reaktion bei der Pyrosequenzierung

7.2.2 Parallelsequenzierung durch *Tagging* von Proben

Mit dem Genome Sequencer FLX System werden derzeit bei einem Sequenzierlauf bis zu einer Millionen DNA-Sequenzen bzw. *reads* produziert. Die Sequenzierplatte kann in maximal 16 von einander separate Felder mit jeweils bis zu 50 000 *reads* unterteilt werden. Bezogen auf die vorliegende Arbeit würde rein rechnerisch bei der Sequenzierung der PCR-Produkte eines Individuums jeder der 37 mitochondrialen und 23 nukleären *loci* bei einer $\frac{1}{16}$ Platte über 800-fach abgedeckt, bei einer vollen Platte sogar über 16 000-fach. Eine solche Abdeckung ist überreichlich und weder finanziell noch arbeitstechnisch tragbar.

Um mehrere Proben parallel zu sequenzieren, können diese zuvor mit sogenannten Barcodes bzw. *tags* versehen werden. Dabei handelt es sich um Oligonukleotide, die an die DNA-Fragmente ligiert werden. Jede Probe erhält einen individuellen *tag*, der sich in seiner Basenabfolge von allen anderen *tags* unterscheidet. Anschließend können die Proben gemischt und dieser Probenpool dann sequenziert werden. Vor der eigentlichen bioinformatischen Auswertung werden die DNA-Sequenzen anhand ihrer *tag*-Sequenzen sortiert und den entsprechenden Proben zugeordnet. Bei einer parallelen Sequenzierung von 50 mitochondrialen und 50 nukleären Multiplex-PCR-Produkten auf einer $\frac{1}{4}$ Platte (ca. 104 – 160 000 *reads*¹⁷) des Genome Sequencer FLX würde die theoretische Abdeckung je *locus* 35 – 53 *reads* betragen.

Eine andere Möglichkeit für die parallele Sequenzierung mehrerer Individuen bietet die Verwendung sogenannter Roche *454 multiplex identifiers (MIDs)*. Hierbei werden bei der *library*-Erstellung Adapter verwendet, in die Erkennungssequenzen integriert sind. 2008 gab es sechs verschiedene *MID*-Adapter; mittlerweile können bis zu 132 Proben mit individuellen Adaptern ausgestattet werden¹⁸.

Alternativ könnten bereits bei der vorgeschalteten Multiplex-PCR probenspezifische Primer verwendet werden (Binladen *et al.* 2007). Der eigentliche Primer ist dabei am 5'-Ende durch eine Erkennungssequenz erweitert. Da für jeden einzelnen *locus* in der PCR ein probenspezifischer Primer benötigt wird, ist diese

¹⁷ Angaben aus den Sequenzierungsangeboten von GATC Biotech entnommen.

¹⁸ <http://454.com/products/gs-flx-system/index.asp> [Stand: 01.10.2014]

Methode bei einer Multiplex-PCR und einer größeren Anzahl an Proben jedoch kosten- und arbeitstechnisch impraktikabel. Bei der Untersuchung degradierter Proben käme zusätzlich noch das Problem der vielfach erhöhten Anzahl benötigter Negativkontrollen hinzu. Zudem ist bei der PCR die Hybridisierung von mit *tags* versehenen Primern möglicherweise weniger effizient.

7.2.2.1 Das Tagging nach Meyer et al. 2008

Für diese Arbeit wurde die Verwendung von Barcodes gewählt, mit denen die Proben im Anschluss an die Multiplex-PCRs gekennzeichnet wurden. Das *tagging* wurde nach einer modifizierten Variante des Protokolls von Meyer *et al.* (2008) durchgeführt. Es ist allgemein für doppelsträngige Nukleinsäure konzipiert, also sowohl für reine DNA-Extrakte für eine *shotgun*-Sequenzierung als auch für PCR-Produkte. Hier sollen das Prinzip der Methode (mit Fokus auf Amplikons) und die biochemischen Hintergründe erläutert werden. Das Protokoll für die praktische Durchführung wird im Methoden-Teil dieser Arbeit beschrieben (siehe 8.2.5.1 und 8.2.5.2).

Die von Meyer *et al.* (2007, 2008) entwickelten *tags* bestehen aus 20 – 24 Basen langen palindromischen DNA-Sequenzen. Von 5' nach 3' beginnt der *tag* mit der eigentlichen Barcode-Sequenz, anhand derer die Identifizierung der Proben bei der Datenauswertung erfolgt. Sie besteht aus sechs bis acht Basen und enthält keine Homopolymere, da diese bei der 454-Sequenzierung zu Fehlern führen können. Dem Barcode folgt eine Erkennungssequenz für das Restriktionsenzym Srf I (GCCCGGGC) und dann zum Abschluss die revers komplementäre Sequenz zum Barcode. Aufgrund dieses Aufbaus wird nur eine Oligonukleotidsequenz pro Probe benötigt. Zwei Moleküle des gleichen *tags* hybridisieren miteinander und bilden somit einen Doppelstrang für die Ligation. Außerdem sind die *tags* am 5'-Ende dephosphoryliert. Hierdurch wird die Bildung von *tag*-Concatemeren bei der Ligation verhindert.



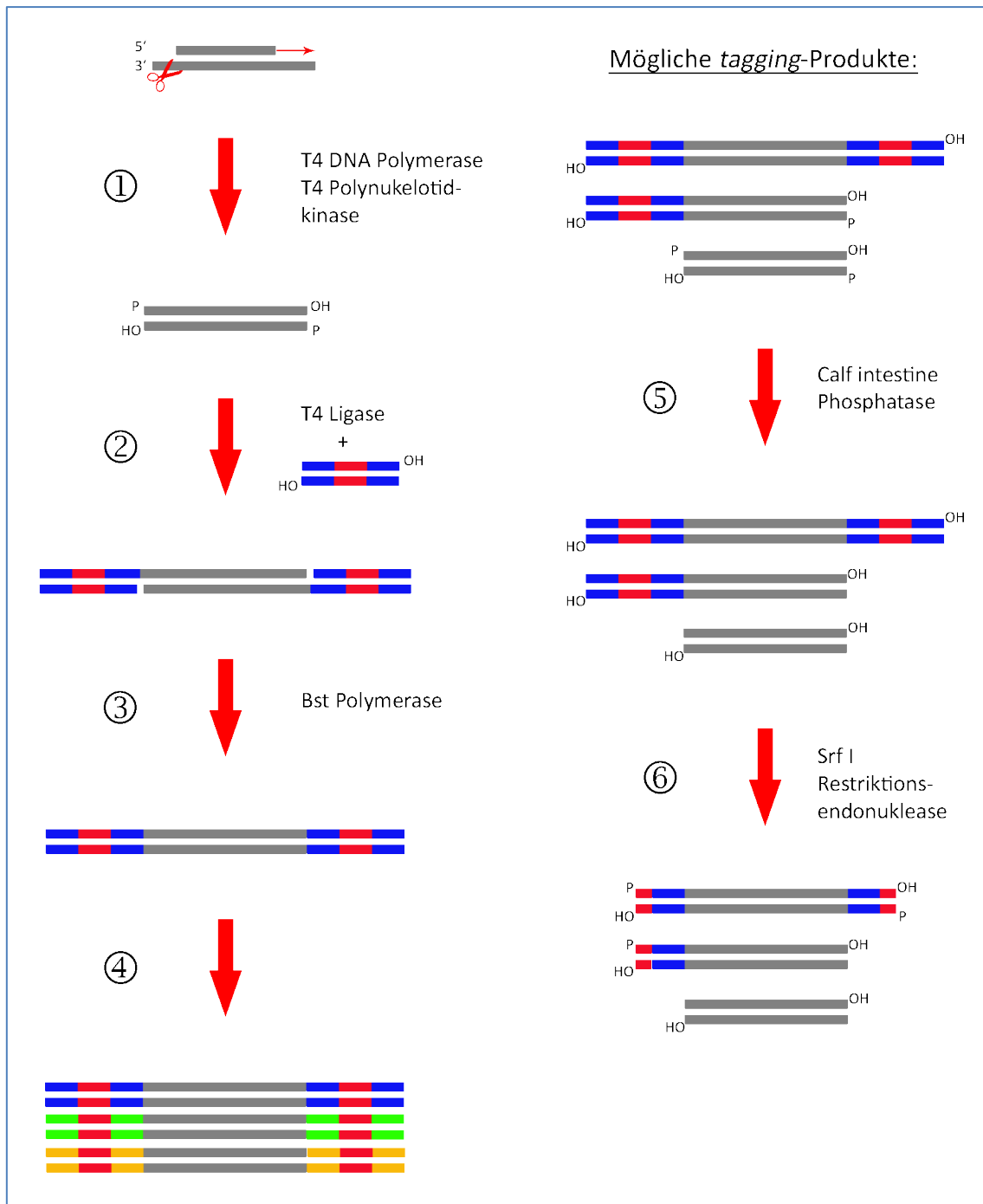


Abbildung 7.2: Das Prinzip des Taggings nach Meyer et al.

Grauer Doppelbalken = DNA-Fragment; blauer (grüner, gelber) Balken = Erkennungssequenz des tags, roter Balken = Restriktionsstelle im tag. (1) *Blunt-end repair*, (2) *Ligation*, (3) *Fill-in (nick repair)*, (4) *Pooling* mehrerer Proben mit unterschiedlichen tags, (5) *Desphosphorylierung*, (6) *Restriction cut*.

Das *tagging* besteht aus einer Reihe von enzymatischen Reaktionen, die an die Vorgehensweise bei der *library*-Erstellung angelehnt ist (siehe Abbildung 7.2). Im ersten Schritt, dem *blunt-end repair*, werden die Enden von den zu sequenzierenden doppelsträngigen DNA-Fragmenten geglättet. Zum Beispiel produziert die in der PCR eingesetzte AmpliTaq Gold Polymerase Adenin-Überhänge am 3'-Ende. Oder bei der Verwendung von dephosphorylierten Primern in der PCR tragen die Produkte 5'-Hydroxyl-Enden. Die T4 DNA-Polymerase weist neben der 5' > 3'-Polymeraseaktivität eine hohe 3' > 5'-Exonukleaseaktivität auf und entfernt Einzelstrangüberhänge bzw. füllt sie auf. Die T4 Polynukleotidkinase repariert mithilfe von ATP dephosphorylierte 5'-Enden und kann gegebenenfalls 3'-Phosphatgruppen entfernen.

Im nächsten Schritt werden die hybridisierten *tags* durch die T4 Ligase über eine Phosphodiesterbindung an die Amplikons gebunden. Da die *tags* keine Phosphatgruppe am 5'-Ende tragen, kann nur das 3'-Ende eine kovalente Bindung mit der DNA eingehen. *Blunt-end*-Ligationen laufen wesentlich ineffizienter und langsamer ab als welche mit *sticky ends* (komplementäre Überhänge). Daher wird Polyethylenglycol¹⁹ (PEG 4000) als Reaktionszusatz eingesetzt. Es stimuliert die Ligation und fördert so den Ligationserfolg.

Anders als bei der *454 library*-Erstellung (siehe Kapitel 7.2.1.1) werden die Ligationenprodukte nicht an *beads* fixiert, sondern die *fill-in*-Reaktion findet frei in der Lösung statt. Die aus *Geobacillus stearothermophilus* gewonnene Bst DNA-Polymerase (*large fragment*) ist eine *strand displacement*-Polymerase. Neben ihrer 5' > 3'-Polymerase-Aktivität hat sie noch Helicase-ähnliche Eigenschaften. Die OH-Gruppe des DNA-Inserts bei der Stelle des *nicks* dient dem Enzym als Primer, von wo aus es die Wasserstoffbrücken des doppelsträngigen Meyer-*tags* aufbricht und den Strang neu synthetisiert. Der ursprüngliche Strang des *tags* auf der Seite des *nicks* wird dabei verdrängt. Nach dem *fill-in* sind die Meyer-*tags* kovalent mit den Amplikons

¹⁹ Polyethylenglycole (PEG) sind wasserlösliche Polymere, die zahlreiche Anwendungen findet, z. B. als Wirkstoffträger in pharmazeutischen Produkten oder als Konservierungsmittel. Die Eigenschaften und Anwendungsmöglichkeiten hängen von der relativen Molekülmasse ab, die als Zahlenwert hinter der Bezeichnung PEG angegeben wird. Pfeiffer und Zimmerman (1983) zeigten, dass bei der Ligation über *blunt-ends* die Anwesenheit von PEG 6000 die Ligrationsrate über tausendfach erhöht, wahrscheinlich indem es die Sedimentierbarkeit der DNA beeinflusst. Beim *tagging* wird mit PEG 4000 eine kürzere Variante des Polymers verwendet.

verbunden und die Proben können nun quantifiziert und äquimolar vermischelt werden (*sample pooling*).

Theoretisch sollten nun alle DNA-Fragmente an beiden Enden mit einem Meyer-*tag* versehen sein. Da eine Ligation jedoch nie hundertprozentig erfolgreich ist, tragen einige Amplikons nur einen oder gar keinen *tag*. Um nicht bei der 454-Sequenzierung unzählige DNA-Stränge zu sequenzieren, die anschließend aufgrund des fehlenden *tags* nicht einem Individuum zugeordnet werden können, schließen Meyer *et al.* (2008) nach dem *pooling* der Proben noch zwei Schritte zum Ausschluss dieser Sequenzen bei der Sequenzierung an.

Zunächst findet eine Dephosphorylierung des Probenpools durch eine *calf intestinal phosphatase* (CIP) statt. Sie gehört zur Gruppe der alkalischen Phosphatasen und hydrolysiert die freien Phosphatgruppen an den 5'-Enden der DNA. Anschließend wird der *restriction cut* durchgeführt. Die aus einem *Streptomyces*-Stamm gewonnene Typ II Restriktionsendonuklease Srf I schneidet die Meyer-*tags* exakt in der Mitte, so dass die geglätteten Enden erhalten bleiben. An der Schnittstelle sind die 5'-Phosphatgruppen erhalten. D. h. beidseitig *tags*-tragende Ligationsprodukte sind an beiden Strängen wieder 5'-phosphoryliert. Amplikons ohne bzw. mit nur einem *tag* tragen aufgrund der vorigen Dephosphorylierung keine bzw. nur an einem Strang die äußere Phosphatgruppe.

Für die weitere Vorbereitung des Probenpools auf die 454-Sequenzierung (7.2.1) bedeutet das Endergebnis des *tagging* nun folgendes: Bei der *library*-Erstellung (Kapitel 7.2.1.1) wird direkt mit der Adapter-Ligation begonnen. Das *end-polishing* ist zum einen nicht nötig, da die Proben bereits glatte Enden haben, zum anderen würden mit diesem Schritt die DNA-Fragmente ohne *tags* wieder phosphoryliert und damit zugänglich für die *library*-Erstellung gemacht. Die Roche 454 Adapter können aufgrund ihrer Struktur ausschließlich an DNA-Fragmente mit freien Phosphatgruppen ligieren, d. h. sie binden nur an Produkte mit mindestens einem (geschnittenen) Meyer-*tag*. An Sequenzen ohne *tags* findet also gar keine Ligation statt. Einseitig *tag*-tragende Amplikons können entweder einen A- oder einen B-Adapter binden. Im ersten Fall gehen sie direkt bei der Aufreinigung mit Streptavidin-*beads* verloren. Im

zweiten Fall gehen sie zwar in die *library* mit ein, jedoch findet bei der Emulsions-PCR (Kapitel 7.2.1.2) keine exponentielle Amplifizierung statt. Der biotinylierte A-Primer kann nicht binden und das Produkt wird bei der Aufreinigung der PCR entfernt. In die 454-Sequenzierung gehen also ausschließlich beidseitig mit *Meyer-tags* versehene Produkte ein, die schließlich anhand der Barcode-Sequenz bioinformatisch separiert werden können.

7.3 Populationsgenetische Datenanalyse

Das Feld der populationsgenetischen Methoden lässt sich in drei interkorrelierende Kategorien einteilen: die (kladistische) Phylogeographie, *summary patterns* und Modellierungen mit computerbasierten Simulationen. Sie alle tragen zur Aufschlüsselung der demographischen Geschichte von Populationen bei.

7.3.1 Phylogeographie

Die Phylogeographie ist ein Teilgebiet der Biogeographie, das zwischen der Phylogenetik und Populationsgenetik angesiedelt ist. Informationen über phylogenetische Verwandtschaftsbeziehungen und geographische Herkunft werden kombiniert, um mithilfe induktiver statistischer Methoden die Verbreitungsgeschichte von Arten und Populationen zu rekonstruieren (Avice *et al.* 1987). Beispielsweise werden das abgeschätzte Alter und das Gebiet der stärksten Verbreitung einer Klade (z. B. definiert als Haplogruppe) mit einer alten Population in Beziehung gesetzt. Ihre heutige Gesamtverbreitung wäre dann die Folge einer demographischen Entwicklung dieser Population (Underhill & Kivisild 2007, Thomas *et al.* 2013).

Ein weitverbreiteter Ansatz in der Phylogeographie ist die *nested clade phylogeographic analysis* (NCPA; Templeton *et al.* 1995, Templeton 1998). Zunächst wird z. B. im Fall der mtDNA ein parsimonisches Haplotypen-Netzwerk erstellt. Dieses

wird ausgehend von den Spitzen schrittweise in hierarchisch geschachtelte Kladen unterteilt. Jede Klade wird auf geographische Differenzierung getestet. Mithilfe eines *inference keys* wird für signifikant differenzierte Kladen eine qualitative Erklärung ermittelt, z. B. kontinuierliche Ausbreitung, Isolation durch Distanz etc. Panchal & Beaumont (2007) und Nielsen & Beaumont (2009) zeigten in Simulationsstudien die Schwächen der *NCPA* auf, die vor allem in der Subjektivität der Interpretation liegen bzw. schon bereits in der Erstellung des Netzwerkes.

7.3.2 *Summary Patterns*

Phylogeographie, wie sie ursprünglich definiert ist, beruht auf der Betrachtung der Phylogenie, um auf geschichtliche Ereignisse zu schließen. Eine andere Herangehensweise ist die Berechnung und Interpretation von Parametern, die die genetische Variation in und zwischen Populationen beschreiben bzw. zusammenfassen, sogenannte „*summary statistics*“ oder „*summary patterns*“. Hierbei handelt es sich um einen klassischen populationsgenetischen Ansatz. Die einfachsten Maße beschreiben die Diversität einer Population als die Anzahl von Allelen, die Anzahl oder der Anteil variabler Positionen in DNA-Sequenzen (*segregating sites*), die durchschnittliche Anzahl an Unterschieden zwischen jeweils zwei Sequenzen etc. Solche Maße hängen von Faktoren wie Sequenzlängen und/oder Stichprobengröße ab. Andere *summary statistics*, z. B. die Nukleotiddiversität, berücksichtigen diese Größen. Die Verteilung der paarweisen Unterschiede (*mismatch distribution*) kann verwendet werden, um Rückschlüsse auf Populationswachstum zu ziehen. Durch Neutralitätstests lassen sich demographische oder Selektionsereignisse nachweisen. Genetische Distanzen zwischen Populationen oder Subpopulationen geben Hinweise auf Populationsstrukturen und Beziehungen zwischen Gruppierungen. Sie können mithilfe von dimensionsreduzierenden Methoden wie einer Hauptkomponentenanalyse oder einer multidimensionalen Skalierung visuell dargestellt werden. Da der Schwerpunkt der datenanalytischen Auswertung dieser Arbeit auf *summary patterns* liegt, sollen an dieser Stelle die den verwendeten Statistiken zugrundeliegenden mathematischen Formeln aufgeführt werden. Sie wurden dem Anwenderhandbuch des Programms

Arlequin 3.5.1.3²⁰ und den darin enthaltenen Referenzen entnommen (Excoffier & Lischer 2010).

Haplotypendiversität (\hat{H})

Die Haplotypendiversität (\hat{H}) gibt die Wahrscheinlichkeit an, dass sich zwei zufällig aus der Probe ausgewählte Haplotypen unterscheiden.

$$\hat{H} = \frac{n}{n-1} \left(1 - \sum_{i=1}^k p_i^2\right)$$

Dabei ist n die Anzahl der Sequenzen, k die Anzahl der Haplotypen und p_i die relative Häufigkeit des i -ten Haplotypen. Die Formel entspricht Nei's Gendiversität (*gene diversity*; Nei 1987: 180). Der Quotient dient als Korrektur für den *sampling bias* bei kleinen Stichprobengrößen.

Mean Number of Pairwise Differences ($\hat{\pi}$)

Die durchschnittliche Anzahl der Unterschiede bei allen paarweisen Vergleichen von Haplotypen bzw. *mean number of pairwise differences* ($\hat{\pi}$) (Nei & Li 1979, Tajima 1983) berücksichtigt im Gegensatz zur Haplotypendiversität (\hat{H}) die Distanzen zwischen den Haplotypen.

$$\hat{\pi} = \frac{\sum_{j < i=1}^n d_{ij}}{\binom{n}{k}} = \frac{n}{n-1} \sum_{i,j=1}^k p_i p_j \hat{d}_{ij}$$

Diese Distanz wird durch \hat{d}_{ij} ausgedrückt, eine Abschätzung der Anzahl von Mutationen, die seit der Divergenz der Haplotypen i und j stattgefunden haben. Analog zur Formel der Haplotypendiversität ist k die Anzahl der Haplotypen, p_i und p_j die relativen Häufigkeiten der entsprechenden Haplotypen.

²⁰ <http://cmpg.unibe.ch/software/arlequin35/> [Stand: 01.10.2014]

Nukleotiddiversität ($\hat{\pi}_n$)

Die Nukleotiddiversität ($\hat{\pi}_n$) gibt die Wahrscheinlichkeit an, dass sich zwei zufällig ausgewählte homologe Nukleotide unterscheiden (Tajima 1983, Nei 1987: 257). Dieses Maß ist eine Art Normierung von $\hat{\pi}$ durch Einbeziehung der Sequenzlänge (L)²¹.

$$\hat{\pi}_n = \frac{1}{L} \sum_{i=1, j < i}^k p_i p_j \hat{d}_{ij}$$

Tajima's D

Die Neutralitätstests Tajima's D und Fu's F_S wurden auf die Sequenzen der Hypervariablen Region I der mtDNA angewandt, um Hinweise auf Populationswachstum zu erhalten. Der Populations-Mutations-Parameter θ (θ) beschreibt die erwartete genetische Diversität in einem neutralen Populationsmodell, bei dem sich Mutation und Gendrift im Gleichgewicht befinden, mit $\theta = 4N_e\mu$ (diploide *loci*) bzw. $\theta = 2N_e\mu$ (mtDNA), wobei N_e die zur Gendrift invers proportionale effektive Populationsgröße und μ die Mutationsrate pro *site* und Generation bezeichnet (Jobling *et al.* 2004: 155). Tajima's D (Tajima 1989) vergleicht unter Annahme eines *infinite sites*-Modells (Kimura 1968) zwei Abschätzungen von θ . $\hat{\theta}_\pi = \hat{\pi}$ basiert auf der bereits vorgestellten *mean number of pairwise differences*. $\hat{\theta}_S$ wird abgeschätzt als die Anzahl an polymorphen Positionen (S , *number of segregating sites*), dividiert durch die Gesamt-Astlänge der Genealogie der n Proben (Watterson 1975, Hamilton 2009: 248 f).

$$\hat{\theta}_S = S / \sum_{i=0}^{n-1} \frac{1}{i}$$

Entsprechend der Nullhypothese einer neutralen Evolution sind für $\hat{\theta}_\pi$ und $\hat{\theta}_S$ ungefähr gleiche Abschätzungen zu erwarten. Da $\hat{\theta}_\pi$ von Allelfrequenzen abhängt, $\hat{\theta}_S$ jedoch nicht, führen Prozesse, die die Frequenzen beeinflussen, zu Diskrepanzen

²¹ Es scheint, als enthielte die im Arlequin-Handbuch angegebene Formel einen Fehler. Die Werte, die Arlequin für $\hat{\pi}_n$ berechnet, sind tatsächlich gleich $\hat{\pi}/L$. Aber entsprechend der angegebenen Formeln müsste $\hat{\pi}_n = \hat{\pi}/2L$ sein.

zwischen den beiden Werten. Tajima's D gibt die Differenz der beiden Abschätzungen von $\hat{\theta}$ relativ zu deren Standardabweichung an.

$$\text{Tajima's } D = \frac{\hat{\theta}_{\pi} - \hat{\theta}_S}{\sqrt{\text{Var}(\hat{\theta}_{\pi} - \hat{\theta}_S)}}$$

$D = 0$ für $\hat{\theta}_{\pi} = \hat{\theta}_S$, d. h., N_e ist konstant und die Allele neutral,

$D < 0$ für $\hat{\theta}_{\pi} < \hat{\theta}_S$, d. h., es gibt einen Überschuss an seltenen Allelen und somit einen Hinweis auf ein Populationswachstum oder positive Selektion,

$D > 0$ für $\hat{\theta}_{\pi} > \hat{\theta}_S$, d. h., es gibt einen Überschuss an mittleren Allelfrequenzen und somit einen Hinweis auf einen rezenten *bottleneck* oder ausgleichende Selektion.

Fu's F_S

Fu's F_S (Fu 1997) basiert ebenfalls auf dem *infinite site*-Modell ohne Rekombination. Der Neutralitätstest macht sich zunutze, dass die Abschätzung von θ basierend auf der *mean number of pairwise differences* ($\hat{\theta}_{\pi}$) bei einem Überschuss an kürzlichen / seltenen Mutationen wahrscheinlich kleiner ist, als die Abschätzung basierend auf der Anzahl von Allelen ($\hat{\theta}_k$). Angenommen, die untersuchte Probe weist k_{obs} Allele auf, mit $\hat{\pi} = \hat{\theta}_{\pi}$. Fu's F_S schätzt die Wahrscheinlichkeit (S') ab, eine zufällige Stichprobe mit mindestens genauso vielen Allelen (K) zu beobachten, unter der Bedingung, dass sie die gleiche Anzahl an durchschnittlichen paarweisen Unterschieden ($\hat{\pi}$) aufweist wie die untersuchte Probe, und alle Allele neutral sind.

$$S' = \Pr(K \geq k_{obs} | \theta = \hat{\theta}_{\pi})$$

Details zur Berechnung dieser Wahrscheinlichkeit können in Fu (1997) oder dem Arlequin-Handbuch nachgelesen werden. Da S' in vielen Fällen Werte extrem nahe 0 annimmt, wird Fu's F_S als Logitfunktion dargestellt:

$$F_S = \ln\left(\frac{S'}{1 - S'}\right)$$

Folglich ist $F_S \geq 0$ für $S' \geq 0,5$, und $F_S < 0$ für $S' < 0,5$. Ist F_S signifikant negativ, also die Wahrscheinlichkeit S' extrem nahe Null, so deutet dies auf Selektion oder Populationsexpansion hin. Während standardmäßig der Signifikanzwert (p -Wert) zur Ablehnung der Nullhypothese auf 5 % festgesetzt wird, so sollte bei Fu's F_S der Schwellenwert auf 2 % gesetzt werden (Fu 1997, Excoffier & Lischer 2010).

F_{ST} und AMOVA

Als Maß für die genetische Distanz zwischen zwei Populationen, d. h. ihr Verwandtschafts- bzw. Isolationsgrad, dient der zur Familie der F -Statistiken gehörende F_{ST} -Wert (*fixation index* bzw. Fixationsindex) (Wright 1951). Es gibt verschiedene Möglichkeiten den F_{ST} -Wert bzw. analoge Parameter zu ermitteln oder abzuschätzen, wobei im Prinzip die Reduktion der genetischen Diversität einer theoretisch panmiktischen Gesamtpopulation aufgrund ihrer Unterteilung in die beiden Subpopulationen berechnet wird:

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

H_S und H_T seien dabei die Wahrscheinlichkeiten, dass zwei zufällig aus der gleichen Subpopulation (S), bzw. aus der Gesamtpopulation (T) gezogene Allele unterschiedlich sind. Der Fixationsindex kann demnach Werte zwischen 0 (keine Distanz, Panmixie) und 1 (vollständige Isolation der Populationen, kein Genfluss) annehmen (Weir & Cockerham 1984). Für einen Überblick über F_{ST} und analoge Statistiken siehe Holsinger & Weir (2009).

In Arlequin ist die Berechnung von F_{ST} -Werten Teil der Analyse der molekularen Varianz (AMOVA; Excoffier *et al.* 1992). Eine AMOVA untersucht den Grad der Differenzierung einer Population auf verschiedenen Ebenen, Subpopulationen (S), Gruppen von Subpopulationen (C) (z. B. geographische Regionen) und die

Metapopulation (T). F_{SC} , F_{ST} und F_{CT} geben dabei die die Aufteilung der Variation zwischen den beiden jeweiligen Hierarchien basierend auf Kovarianzkomponenten an.

Zunächst wird ein *minimum spanning*-Netzwerk aller Individuen erstellt und daraufhin die Haplotypen als Vektoren von Mutationsereignissen dargestellt. Es wird ein hierarchisches Modell mit Populationen in *a priori* definierten Gruppen (basierend auf nicht-genetischen Kriterien) angenommen, bei dem der i -te Haplotyp in der j -ten Population der k -ten Gruppe durch eine lineare Gleichung beschrieben werden kann: $x_{ijk} = x + a_k + b_{jk} + c_{ijk}$. Der Vektor x ist dabei der erwartete Vektor von x_{ijk} (Mittelwert der Gesamtpopulation); a , b , und c sind die zufälligen, voneinander unabhängigen, additiven Einflussfaktoren bzw. Effekte von Gruppe, Population und Haplotypen innerhalb einer Population in einer Gruppe. Damit assoziiert sind die Kovarianzkomponenten σ_a^2 , σ_b^2 und σ_c^2 . Die Summe der Komponenten ergibt die Gesamtvarianz der Population: $\sigma_T^2 = \sigma_a^2 + \sigma_b^2 + \sigma_c^2$. Für jede Hierarchieebene wird die Summe quadratischer Abweichungen von x (*sum of squared deviations, SSD*) und die jeweilige mittlere quadratische Abweichung (Varianz, *mean squared deviation, MSD*) mithilfe der entsprechenden Freiheitsgrade berechnet. Die jeweilige *MSD* kann auch durch Kombination der Kovarianzen beschrieben werden, so dass diese ermittelt werden können. Dabei werden unterschiedliche Populations- und Gruppengrößen berücksichtigt. Die Aufteilung der Gesamtvarianz zwischen den verschiedenen Ebenen der Populationshierarchie lassen sich wie folgt darstellen:

$$\Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_T^2} = \frac{\sigma_T^2 - \sigma_c^2}{\sigma_T^2} = F_{ST}$$

$$\Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2} = \frac{(\sigma_b^2 + \sigma_c^2) - \sigma_b^2}{(\sigma_b^2 + \sigma_c^2)} = F_{SC}$$

$$\Phi_{CT} = \frac{\sigma_a^2}{\sigma_T^2} = \frac{\sigma_T^2 - (\sigma_b^2 + \sigma_c^2)}{\sigma_T^2} = F_{CT}$$

Φ (*Phi*) ist ein molekulares Analogon zu Wrights F -Statistiken. Links sind die Formeln aus Excoffier *et al.* (1992) bzw. dem Arlequin-Handbuch angegeben. Auf der rechten Seite ist zur Veranschaulichung die Umwandlung dieser Formeln in das Format

für die allgemeine F_{ST} - (bzw. F_{SC} - und F_{CT} -) Berechnung, wie sie weiter oben ($F_{ST} = (H_T - H_S)/H_T$) aufgeführt ist, dargestellt.

Zur Berechnung der paarweisen genetischen Distanzen zwischen verschiedenen Populationen ermittelt Arlequin den Φ_{ST} -Wert mit einer AMOVA, als Spezialfall mit zwei Subpopulationen. Eigentlich nimmt die F_{ST} -Statistik, wie bereits angemerkt, Werte zwischen 0 und 1 ein. Aus algorithmischen Gründen können die in Arlequin berechneten Φ_{ST} -Werte auch negativ sein, was im Prinzip dem Wert 0 (keine genetische Distanz) entspricht.

Nicht-metrische multidimensionale Skalierung

Mithilfe eines Plots einer multidimensionalen Skalierung (MDS) können die Kombinationen von paarweisen genetischen Distanzen zwischen den Populationen oder Gruppen graphisch dargestellt werden. Es gibt mehrere Methoden und Algorithmen; für diese Arbeit wurde die nicht-metrische multidimensionale Skalierung von Kruskal (1964) verwendet. Die Reduktion der maximal $(n - 1)$ Dimensionen, wobei n der Anzahl der Populationen entspricht, auf zwei oder drei Dimensionen geschieht durch ein iteratives Verfahren. Die Populationen werden zunächst zufällig in einem zwei- oder dreidimensionalen Koordinatensystem angeordnet und ihre räumlichen Distanzen ermittelt. Es folgt eine wiederholte Verschiebung der Punkte mithilfe eines Algorithmus, bis die räumlichen Distanzen die F_{ST} -Werte (oder allgemein Disparitäten) zwischen den Populationen so gut wie möglich widerspiegeln. Das heißt, wenn die F_{ST} -Werte und die räumlichen Distanzen jeweils in einer monoton ansteigenden Rangfolge sortiert sind, sollten die beiden Reihenfolgen im optimalen Fall gleich sein. Der Stressfaktor gibt an, wie stark die Abweichungen der Konfiguration von den tatsächlichen Daten sind. Im optimalen Fall einer perfekten monotonen Beziehung wäre der Stresswert gleich Null. Kruskal (1964) gibt eine Richtlinie für die Bewertung eines MDS-Plots an:

<i>Stress</i>	<i>Goodness of fit</i> (Güte der Anpassung)
20 %	<i>poor</i> (schlecht)
10 %	<i>fair</i> (ausreichend)
5 %	<i>good</i> (gut)
2,5 %	<i>excellent</i> (exzellent)
0 %	„ <i>perfect</i> “ (perfekte Monotoniebeziehung)

Haplotype Sharing

Ein Maß für die genetische Nähe oder Ähnlichkeit von zwei Populationen ist der Anteil gemeinsamer Haplotypen, also welche, die in beiden Populationen vorkommen (*shared haplotypes*). Eine mögliche Formel hierfür ist die folgende, die die Wahrscheinlichkeit angibt, dass zwei aus zwei Populationen zufällig ausgewählte Haplotypen identisch sind:

$$ht_S = \sum_i X_i Y_i,$$

wobei X_i und Y_i die Häufigkeiten des gemeinsamen Haplotypen i in den beiden Populationen X und Y sind.

Exact Hardy Weinberg Test

Die bisher beschriebenen *summary statistics* wurden für die populationsgenetische Auswertung der mtDNA verwendet. Die autosomalen *loci* wurden dem *Exact Hardy Weinberg Test* in R²² unterzogen (Graffelman 2013, R Core Team 2013). Die Hardy-Weinberg-Gleichung bzw. das -Gleichgewicht (HWG) beschreibt den Zusammenhang zwischen den Häufigkeiten von Allelen und Genotypen (Hardy 1908, Weinberg 1908). In einer idealen Population bleiben diese konstant. Seien A und B zwei Allele in einem diploiden Organismus und seien p_A und p_B die entsprechenden Allelfrequenzen in der Population. Ist die Population für diesen *locus* im HWG-Gleichgewicht, so lässt sich die Genotypenverteilung ermitteln: $p_{AA} = p_A^2$, $p_{AB} = 2p_A p_B$ und $p_{BB} = p_B^2$. Durch Vergleich der beobachteten Allel- und Genotypenfrequenzen mit denen, die zu erwarten sind,

²² <http://www.r-project.org/> [Stand: 01.10.2014]

wenn die Population sich im HWG befindet, können evolutionäre Prozesse (z. B. Populationsstrukturen, Selektion etc.) untersucht werden.

Am einfachsten lassen sich Abweichungen vom HWG durch Tests wie einem Chi-Quadrat- (χ^2 -) Test überprüfen. Diese sind allerdings sehr fehleranfällig, z. B. bei kleinen Populationsgrößen, oder einer niedrigen Frequenz des selteneren Allels (Guo & Thompson 1992, Wigginton *et al.* 2005). Alternativ kommen exakte HWG-Tests in Frage, die allerdings sehr rechenintensiv sein können.

Der für die Überprüfung der in dieser Arbeit generierten Genotypenfrequenzen verwendete exakte HWG-Test beruht auf der Methode von Wigginton *et al.* (2005). Sie kombiniert etablierte Tests und Rechenmethoden, die sich durch Robustheit auch bei kleinen Stichproben und extrem niedrigen Allelfrequenzen auszeichnen. Zunächst werden die Häufigkeiten der Genotypen und Allele im zu untersuchenden Datensatz bestimmt. n_A, n_B = absolute Häufigkeiten der Allele A (Allel mit der niedrigeren Frequenz) und B; n_{AA}, n_{BB} = Anzahl der Homozygoten für Allel A bzw. B; n_{AB} = Anzahl der Heterozygoten; N = Anzahl der Individuen in der Stichprobe. Es wird die Wahrscheinlichkeit berechnet, diese tatsächliche Häufigkeitsverteilung unter HWE-Bedingungen zu beobachten, insbesondere die Anzahl an Heterozygoten:

$$P(N_{AB} = n_{AB} | N, n_A) = \frac{2^{n_{AB}N!}}{n_{AA}! n_{AB}! n_{BB}!} \cdot \frac{n_A! n_B!}{(2N)!}$$

Dann werden alle möglichen Kombinationen von Genotypenverteilungen mit gleicher Individuenzahl (N) und gleicher Frequenz von Allel A (n_A) ermittelt. Ausgehend von dem Wert für n_{AB}^* , für den das HWG ($p_{AB} = 2p_A(1-p_A)$) erfüllt wäre,

$$n_{AB}^* = \frac{n_A(2N-n_A)}{2N},$$

wird n_{AB}^* bei gleichbleibendem N und n_A schrittweise um jeweils 2 erhöht bzw. verringert und infolgedessen n_{AA}^* und n_{BB}^* um 1 verringert bzw. erhöht. Für jede²³ Kombination von Genotypenhäufigkeiten wird die Wahrscheinlichkeit berechnet, dass sie im HWG zustande kommen könnte. Zur Reduzierung des Rechenaufwandes wird

²³ Aufgrund des Rechenaufwandes gibt es auch Verfahren, die lediglich per Permutation zufällige Stichproben an Kombinationen von Genotypenfrequenzen berechnen.

nicht die obige Formel sondern ein *Markov-Chain-Monte-Carlo*-Verfahren angewandt (Guo & Thompson 1992, Wigginton *et al.* 2005), bei dem die Wahrscheinlichkeit jeder Kombination basierend auf der vorigen berechnet wird:

$$P(N_{AB}^* = n_{AB}^* + 2 | N, n_A) = P(N_{AB}^* = n_{AB}^* | N, n_A) \frac{4n_{AA}^* n_{BB}^*}{(n_{AB}^* + 2)(n_{AB}^* + 1)}$$

und

$$P(N_{AB}^* = n_{AB}^* - 2 | N, n_A) = P(N_{AB}^* = n_{AB}^* | N, n_A) \frac{n_{AB}^* (n_{AB}^* - 1)}{4(n_{AA}^* + 1)(n_{BB}^* + 1)}$$

Es werden die Kombinationen ermittelt, die mindestens genauso stark vom HWG abweichen wie die beobachtete Genotypenverteilung, d. h. alle Wahrscheinlichkeiten, die gleich oder kleiner sind als $P(N_{AB} = n_{AB} | N, n_A)$. Diese Wahrscheinlichkeiten werden summiert und ergeben den *P*-Wert des Tests:

$$P_{HWE} = \sum_{n_{AB}^*} I[P(N_{AB} = n_{AB} | N, n_A) \geq P(N_{AB}^* = n_{AB}^* | N, n_A)] \cdot P(N_{AB}^* = n_{AB}^* | N, n_A)$$

Bei $I[x]$ handelt es sich um eine Indikatorfunktion. Sie nimmt den Wert 1 an, wenn die Ungleichung in der Klammer erfüllt ist, 0 wenn nicht.

Das hier beschriebene Verfahren stellt den zweiseitigen HWG-Test dar, d. h. die Überprüfung sowohl eines Überschusses an Homozygoten als auch Heterozygoten. Bei einem einseitigen Test auf Homozygotenüberschuss, wie er in Wilde *et al.* (2014) durchgeführt wurde, wird bei der Ermittlung aller möglicher Kombinationen an Genotypenfrequenzen lediglich die vom HWG linksseitige Suche $P(N_{AB}^* = n_{AB}^* - 2 | N, n_A)$ gerechnet.

7.3.3 Explizite Modellierungen

Summary statistics geben gute Anhaltspunkte zur Populationsgeschichte, aber ihre Interpretation ist nicht immer eindeutig, da verschiedene demographische Prozesse und zufällige Entwicklungen, aber auch Selektion die gleichen genetischen Muster hervorbringen können. Immer mehr an Bedeutung gewinnen computergestützte Modellierungen, sowohl zur Beantwortung spezieller phylogeographischer Fragestellungen als auch allgemein zur Untersuchung der demographischen Geschichte von Populationen. Sie umgehen das Problem der Interpretation, indem verschiedene explizite populationsgeschichtliche Szenarien statistisch verglichen werden. Dabei unterscheidet man sogenannte *forward*- und *backward-in-time*-Simulationen. Ihr Einsatz hängt von der Fragestellung ab. Für detaillierte Informationen zu Simulationsmethoden und Anwendungsgebiete sei der Übersichtsartikel von Hoban *et al.* (2012) empfohlen.

Die *backward-in-time*-Methode basiert auf der Koaleszenztheorie. Sie betrachtet eine genetische Genealogie ausgehend von der gegenwärtigen Generation als eine Reihe von Koaleszenzereignissen bis zum letzten gemeinsamen Vorfahren (*MRCA*, *most recent common ancestor*) (Kingman 1982, Hudson 1990). Die Wahrscheinlichkeit einer Koaleszenz hängt von der Anzahl an Linien und Individuen in der Population oder Subpopulation ab, welche wiederum von verschiedenen populationsgenetischen Faktoren beeinflusst werden. Ein Unterschied zu *forward*-Simulationen, die von einem bestimmten Zeitpunkt aus *lineage branching*-Ereignisse generieren, liegt darin, dass bei dem koaleszenzbasierten Ansatz die Genealogie einer Stichprobe von Individuen bzw. Genen betrachtet wird und nicht die gesamte Population (Rosenberg & Nordborg 2002, Hoban *et al.* 2012).

Bei Simulationsanalysen werden zunächst Modelle zur genetischen und demographischen Populationsgeschichte entwickelt und die dazugehörigen Parameter definiert. In diese Modelle können geographische, ökologische, klimatische, archäologische und genetische Informationen einfließen. Zu den Parametern gehören je nach Fragestellung z. B. Divergenzzeiten, effektive Populationsgrößen, Migrationsraten, Selektionskoeffizienten etc. Für sie müssen sinnvolle Spannweiten

angegeben werden. Da nicht alle evolutionären Faktoren parametrisiert werden können bzw. sollen, müssen auch bei den Modellierungen einige Vorannahmen getroffen werden, beispielsweise eine konstante Populationsgröße oder -wachstumsrate. Außerdem wird dem Simulationsprogramm vorgegeben, für wie viele Gene Genealogien simuliert werden sollen (entsprechend der Zahl der Individuen im beobachteten Datensatz). Es gibt zurzeit mehrere Dutzend Programme, die für unterschiedliche Datensätze und Fragestellungen geeignet sind (Hoban *et al.* 2012).

Ein solches Programm ist SIMCOAL (Excoffier *et al.* 2000). Zunächst generiert es nach den beschriebenen Vorgaben eine Genealogie. Wenn der *MRCA* erreicht ist, wird diesem Gen eine zufällige Nukleotidsequenz (im Falle eines DNA-Sequenz-Datensatzes) zugewiesen. Dann wird der Baum ausgehend von der Wurzel bis zur Spitze zufällig mit Mutationen belegt, abhängig von der vorgegebenen Mutationsrate. Auf diese Weise erhält man die vorgegebene Anzahl an „modernen“ Pseudosequenzen. Je nachdem wie groß der Parameterraum, also wie viele Kombinationen von Werten angenommen werden können, liegt die Anzahl der Simulationen und generierten Sequenzen schnell im Bereich von Millionen.

Sowohl für die simulierten als auch für den beobachteten Datensatz werden *summary statistics* berechnet. Es müssen solche Statistiken gewählt werden, die stark von den Modellparametern beeinflusst werden. Insgesamt ist darauf zu achten, dass die Modelle, Parameter und *summary statistics* aufeinander abgestimmt und an die Fragestellung angepasst sind.

Für jedes Modell erhält man eine Verteilung der simulierten Werte der Statistiken, und es wird überprüft, wo die Werte des beobachteten Datensatzes in diese Verteilung fallen. Dies wird durch *p*-Werte angegeben. Auf diese Weise kann überprüft werden, wie gut simulierte und beobachtete Daten übereinstimmen, und ob ein Modell abgelehnt oder angenommen wird. Mittlerweile werden *approximate Bayesian computation*- (ABC) Algorithmen angewendet, um a) den Parameterraum bei den Simulationen besser zu sondieren und b) Modelle miteinander zu vergleichen (Beaumont 2010, Hoban *et al.* 2012). Programme wie Serial SimCoal (Anderson *et al.* 2005) und fastsimcoal (Excoffier & Foll 2011) erlauben das Inkorporieren von Daten

aus verschiedenen Zeitschichten. So können Datensätze simuliert werden, die mit aDNA-Daten verglichen werden können. Mit *fastsimcoal* wurde beispielsweise getestet, ob mitochondriale Kontinuität zwischen den bronzezeitlichen Steppenbewohnern und den modernen Bevölkerungen des nordpontischen Raumes besteht (Wilde *et al.* 2014). Im gleichen Artikel wurde mithilfe von *forward-in-time*-Simulationen überprüft, ob die Unterschiede in den Allelfrequenzen für drei mit Pigmentierung assoziierte SNPs zwischen dem aDNA-Datensatz und modernen Proben durch Gendrift erklärbar sind, oder ob Selektion eine Rolle gespielt haben muss (siehe auch Kapitel 10.4.1).

8 Material & Methoden

8.1 Probenmaterial

Insgesamt lagen 180 menschliche und eine Tierprobe als Kontaminationskontrolle für die molekulargenetische Bearbeitung vor. Sie stammen von 39 archäologischen Fundplätzen aus dem west- und nordpontischen Raum (Tabelle 8.1). Kolonia Człczyce, Kolonia Depultczyce Nowe und Zauschwitz gehören nicht dem eigentlichen Untersuchungsgebiet an, sondern lieferten Proben der Kugelamphorenkultur, zu der einige nordpontische Steppengruppen nachweislich Kontakte hatten. Abbildung 8.1 zeigt die geographische Lage der 28 Fundorte mit erfolgreicher DNA-Extraktion. Von den 26 „erfolgreichen“ Fundplätzen um das Schwarzmeer liegt lediglich Tetçani in der Waldsteppe. Kirovograd und Moljučov Bugor am Dnjepr, sowie die drei Fundorte der Poltavka-Kultur – Kalinovka, Nikolaevka und Peščanij – befinden sich am Übergang zwischen der Waldsteppe und dem reinen Steppengürtel. In diesem verteilen sich die restlichen 20 Fundorte.

Das Skelettmaterial wurde von verschiedenen Archäologen aus der Ukraine, Bulgarien, Russland und Moldawien zur Verfügung gestellt. Die Akquirierung der Proben wurde von E. K. koordiniert. Es handelt sich vor allem um Knochenfragmente, weniger um Zähne, da diese für parallel durchgeführte Isotopenanalysen verwendet wurden (Gerling im Druck). Das Material stammt vorwiegend aus Kurganen, lediglich einige wenige äneolithische Individuen waren in Flachgräbern bestattet. Ziel war es, Proben aus dem Ausbreitungsgebiet der Jamnaja-Kultur im nordpontischen Steppenraum und entlang der westlichen Schwarzmeerküste zu sammeln, die aus der Zeit vor, während und nach der Ausbreitung dieser Kultur stammen. Dabei wurden möglichst viele verschiedene Kurgane beprobt, um einem etwaigen *sampling bias* durch Verwandtschaftsbeziehungen der Bestatteten entgegenzuwirken. Die Informationen zum archäologischen Kontext der einzelnen Individuen sowie deren kulturelle Einordnung wurden von E. K. recherchiert und zur Verfügung gestellt. Sie sind der Tabelle 8.1 zu entnehmen.



Abbildung 8.1: Geographische Lage der Fundplätze

Die Fundplätze sind entsprechend ihrer Laborcodes gekennzeichnet. Fundorte, aus denen lediglich nukleäre Daten gewonnen wurden, sind kursiv beschriftet; die 11, die keine Ergebnisse lieferten, sind nicht dargestellt. Die vollständigen Namen und Geokoordinaten aller Fundplätze sind in Tabelle 8.1 angegeben.

Tabelle 8.1: Probenliste

Die Proben stammen aus den Gebieten des heutigen Westruslands (R), der Ukraine (U), Bulgarien (B), Moldawien (M), Polen (P) und Deutschland (D). N = Anzahl der Individuen bei Mehrfachbestattungen. KGK = Katakombengrabkultur. Fettgedruckte Labornummern: Proben, deren HVR I eindeutig bestimmt werden konnte. In grauer Schrift: 17 Proben, die aufgrund ihres schlechten Erhaltungszustandes nicht bearbeitet wurden und drei, die sich als wesentlich jünger als der Untersuchungszeitraum herausstellten.

Fundort (Land)	Breiten-grad	Längen-grad	Labor-Nr.	Kurgan	Grab (N)	Zeitschnitt / Kultur
Bădragii Veche (M)	48,0396	27,0947	BAD 1	27	2	Frühe Bronzezeit - Jamnaja
Bezeda (M)	48,1608	27,0011	BED 1	2	10	Frühe Bronzezeit - Jamnaja
			BED 2	2	11	Mittlere Bronzezeit - entwickelte KGK
Benkovski (B)	42,4090	25,9018	BEN 1	1	2	Frühe Bronzezeit
			BEN 2	2	1	Frühe Bronzezeit
			BEN 3	2	2	Frühe Bronzezeit
			BEN 4	2	5	=> Wesentlich jünger
			BEN 5	2	6	Frühe Bronzezeit
Berězovo (U)	49,8739	34,9344	BEZ 1	1	1	Frühe/mittlere Bronzezeit - frühe KGK
Bojanovo (B)	42,2636	26,6133	BOY 1	Lozjanskata	4	Mittlere Bronzezeit
			BOY 2	Lozjanskata	5	Frühe Bronzezeit - Jamnaja
			BOY 3	Lozjanskata	6	Frühe Bronzezeit - Jamnaja
			BOY 4	Lozjanskata	14 (2)	Frühe Bronzezeit - Jamnaja
			BOY 5	Lozjanskata	14 (2)	Frühe Bronzezeit - Jamnaja
Kolonia Czuczycze (P)	51,2083	23,4341	CZU 1	Steinkistengrab (Steinkiste)	st. 6 südl. Grabteil (2)	Frühe Bronzezeit - Kugelamphorenkultur

Fundort (Land)	Breiten-grad	Längen-grad	Labor-Nr.	Kurgan	Grab (N)	Zeitabschnitt / Kultur
Kolonia Czulczyce (P)	51,2083	23,4341	CZU 2	Flachgrab (Steinkiste)	st. 6 südl. Grabteil (2)	Frühe Bronzezeit - Kugelamphorenkultur
Kolonia Depultczyce Nowe (P)	51,0692	23,3956	DEP 1	Flachgrab (Steinkiste)	st. 12	Frühe Bronzezeit - Kugelamphorenkultur
Durankulak (B)	43,6933	28,5267	DUR 1		15	Spätes Äneolithikum - Protojamnaja
			DUR 2		448	Äneolithikum - Cemavodä I
			DUR 3		1061	Spätes Äneolithikum - Protojamnaja
			DUR 4		24	keine Angaben
Goljam kairak (B)	43,2132	27,5170	GOL 1	1	22	Frühe Bronzezeit - Jamnaja
			GOL 2	1	23	Frühe Bronzezeit - Jamnaja
			GOL 3	1	25	Frühe Bronzezeit - Jamnaja
Kalinovka I (R)	53,8338	51,2890	KAL 1	1	4	Frühe/mittlere Bronzezeit - Poltavka
			KAL 2	1	6	Frühe/mittlere Bronzezeit - Poltavka
Krasnorečensk (U)	49,2106	38,2072	KNO 1	1	2	Mittlere Bronzezeit - entwickelte KGK
			KNO 2	1	6	Mittlere Bronzezeit - entwickelte KGK
			KNO 3	1	7 (3)	Mittlere Bronzezeit - entwickelte KGK
			KNO 4	1	13	Mittlere Bronzezeit - entwickelte KGK
			KNO 5	2	2	Frühe/mittlere? Bronzezeit - KGK
			KNO 6	1	5 (3)	keine Angaben
Košary (U)	46,6667	31,1551	KOS 1	Flachgrab	11	Spätes Äneolithikum - Usatovo
			KOS 2	Flachgrab	12	Spätes Äneolithikum - Usatovo
			KOS 3	Flachgrab	13	Spätes Äneolithikum - Usatovo
Krasnaja Zarja (U)	48,3824	38,6842	KRZ 1	6	2	Frühe/mittlere? Bronzezeit - KGK
			KRZ 2	6	9	Frühe Bronzezeit - Jamnaja
			KRZ 3	6	3	Frühe/mittlere? Bronzezeit - KGK
Lisičansk (U)	48,9130	38,4358	LIS 1	1	2 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			LIS 2	2	1	Frühe/mittlere Bronzezeit - frühe KGK
			LIS 3	3	2	Mittlere Bronzezeit - entwickelte KGK
			LIS 4	3	5	keine Angaben
Majaki (U)	46,4135	30,2758	MAJ 1	Flachgrab	K8	=> Späte Eisenzeit
			MAJ 2	?	G5	Spätes Äneolithikum oder frühe Bronzezeit
			MAJ 3	1	7	Frühe Bronzezeit - Jamnaja
			MAJ 4	1	13	Frühe Bronzezeit - Jamnaja
			MAJ 5	1	15	Frühe Bronzezeit - Jamnaja
			MAJ 6	1	18	=> Späte Eisenzeit
			MAJ 7	7	2	Spätes Äneolithikum - Usatovo
			MAJ 8	8	2	Spätes Äneolithikum - Usatovo
			MAJ 9	8	4	Spätes Äneolithikum - Usatovo
			MAJ 10	9	2	Spätes Äneolithikum - Usatovo
			MAJ 11	10	2 (2)	Spätes Äneolithikum - Usatovo
Mandzikiny I (R)	45,5747	44,6160	MAZ 1	14	13	Spätes Äneolithikum - Majkop
Mednikarovo-Iskrica (B)	42,1708	26,0194	MED 1	6	2	Frühe Bronzezeit - Jamnaja
			MED 2	6	3	keine Angaben
			MED 3	6	5	Frühe Bronzezeit - Jamnaja
Moljučov Bugor (U)	49,1269	32,4611	MOB 1	Flachgrab	3	Spätes Äneolithikum
			MOB 2	Flachgrab	4 (2)	Spätes Äneolithikum
			MOB 3	Flachgrab	4 (2)	Spätes Äneolithikum
Nevskoe (U)	49,1701	37,9765	NEV 1	2	10 (2)	Mittlere Bronzezeit - entwickelte KGK
			NEV 2	3	3	Mittlere Bronzezeit - entwickelte KGK
			NEV 3	5	9	Mittlere Bronzezeit - entwickelte KGK
Nikolaevka III (R)	53,0915	50,3146	NIK 1	1	2	Frühe/mittlere Bronzezeit - Poltavka
			NIK 2	1	3	Frühe/mittlere Bronzezeit - Poltavka
			NIK 3	3	1 (2)	Frühe/mittlere Bronzezeit - Poltavka
			NIK 4	3	1 (2)	Frühe/mittlere Bronzezeit - Poltavka
			NIK 5	3	3	Frühe/mittlere Bronzezeit - Poltavka
			NIK 6	3	4	Frühe/mittlere Bronzezeit - Poltavka
			NIK 7	5	1	Frühe/mittlere Bronzezeit - Poltavka
Novozvanovka 2 (U)	48,5827	38,3564	NOZ 1	1	2	Mittlere Bronzezeit - entwickelte KGK
			NOZ 2	1	3	Mittlere Bronzezeit - entwickelte KGK

Fundort (Land)	Breiten-grad	Längen-grad	Labor-Nr.	Kurgan	Grab (N)	Zeitschnitt / Kultur
Olennij (R)	45,6003	38,5304	OLE 1	1	9	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 2	1	21 (2)	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 3	2	25 (2)	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 4	2	26	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 5	2	35 (2)	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 6	3	3	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 7	3	14 (2)	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 8	1	21 (2)	Frühe Bronzezeit - Jamnaja (Novotitarovka)
			OLE 9	3	14 (2)	Frühe Bronzezeit - Jamnaja (Novotitarovka)
Ovčarci (B)	42,2009	26,0685	OVI 1	The Big Tumulus	9	Frühe Bronzezeit - Jamnaja
			OVI 2	The Big Tumulus	10	Frühe Bronzezeit - Jamnaja
			OVI 3	2	1	Frühe Bronzezeit - Jamnaja
			OVI 4	Tumulus in the vineyard	3	Frühe Bronzezeit - Jamnaja
			OVI 5 TK	The Big Tumulus	10	Frühe Bronzezeit - Jamnaja (Tierkontrolle)
			OVI 6	The Big Tumulus	15	keine Angaben
			OVI 7	The Big Tumulus	15	keine Angaben
			OVI 8	2	2	keine Angaben
			OVI 9	2	2	keine Angaben
Pesčanj V (R)	46,5560	43,6769	PEJ 1	1	3	Frühe Bronzezeit - Jamnaja
			PEJ 2	1	4	Mittlere Bronzezeit - entwickelte KGK
			PEJ 3	2	3 (3)	Frühe/mittlere Bronzezeit - frühe KGK
			PEJ 4	2	3 (3)	Frühe/mittlere Bronzezeit - frühe KGK
			PEJ 5	2	3 (3)	Frühe/mittlere Bronzezeit - frühe KGK
			PEJ 6	3	2	Frühe/mittlere Bronzezeit - frühe KGK
			PEJ 7	2	1	keine Angaben
Pesčanka II (U)	48,5996	35,3120	PES 1	1	4	Mittlere Bronzezeit - entwickelte KGK
			PES 2	1	6	Äneolithikum
			PES 3	1	7	Äneolithikum
			PES 4	1	10	Frühe Bronzezeit - Jamnaja
			PES 5	1	11	Frühe Bronzezeit - Jamnaja
			PES 6	1	14	Äneolithikum
			PES 7	1	15	Frühe Bronzezeit - Jamnaja
Podlesnij (R)	53,1383	50,6664	POD 1	1	3	Frühe/mittlere Bronzezeit - Poltavka
			POD 2	1	5	Frühe/mittlere Bronzezeit - Poltavka
			POD 3	1	6	Frühe/mittlere Bronzezeit - Poltavka
			POD 4	1	7	Frühe/mittlere Bronzezeit - Poltavka
Popovo (B)	42,2036	26,7284	POP 1	Goljamata Mogila	29	Frühe Bronzezeit - Jamnaja
			POP 2	Goljamata Mogila	33	keine Angaben
			POP 3	Goljamata Mogila	36	Frühe Bronzezeit - Jamnaja
			POP 4	Goljamata Mogila	37	Frühe Bronzezeit - Jamnaja?
			POP 5	Goljamata Mogila	38	Frühe Bronzezeit - Jamnaja
Rilci (B)	43,5923	27,7786	RIL 1	260	1	Frühe Bronzezeit - Jamnaja
			RIL 2	264	4	Frühe Bronzezeit - Jamnaja
			RIL 3	264	5	Frühe Bronzezeit - Jamnaja
Šachta Stepnaja (U)	48,3830	36,3716	SAC 1	2	4	Mittlere Bronzezeit - entwickelte KGK
			SAC 2	3	3 (3)	Frühe/mittlere Bronzezeit - frühe KGK
			SAC 3	4	5	Äneolithikum
			SAC 4	4	10	Mittlere Bronzezeit - entwickelte KGK
			SAC 5	4	11	Frühe Bronzezeit - Jamnaja
Smjadovo (B)	43,0526	26,9951	SMY 1	Flachgrab	6	Frühes Äneolithikum
			SMY 2	Flachgrab	8	Frühes Äneolithikum
			SMY 3	Flachgrab	10	Frühes Äneolithikum
			SMY 4	Flachgrab	11	Frühes Äneolithikum
			SMY 5	Flachgrab	13	Frühes Äneolithikum
			SMY 6	Flachgrab	15	Frühes Äneolithikum
			SMY 7	Flachgrab	16	Frühes Äneolithikum
			SMY 8	Flachgrab	18	Frühes Äneolithikum
			SMY 9	Flachgrab	22	Frühes Äneolithikum

Fundort (Land)	Breiten-grad	Längen-grad	Labor-Nr.	Kurgan	Grab (N)	Zeitabschnitt / Kultur
Smjadovo (B)	43,0526	26,9951	SMY 10	Flachgrab	24 (5)	Frühes Äneolithikum
			SMY 11	Flachgrab	24 (5)	Frühes Äneolithikum
			SMY 12	Flachgrab	24 (5)	Frühes Äneolithikum
			SMY 13	Flachgrab	26	keine Angaben
Starye Beljary (U)	46,6238	31,0311	STB 1	1	14 (Steinkiste)	Frühe Bronzezeit - Kemi-Oba Kultur
Kirovograd (U)	48,5215	32,2450	SUG 1	Sugokleja	5	Frühe Bronzezeit - Jamnaja
			SUG 2	Sugokleja	8	Frühe Bronzezeit - Jamnaja
			SUG 3	Sugokleja	10 (2)	Frühe Bronzezeit - Jamnaja
			SUG 4	Sugokleja	10 (2)	Frühe Bronzezeit - Jamnaja
			SUG 5	Sugokleja	13	Mittlere Bronzezeit - entwickelte KGK
			SUG 6	Sugokleja	16	Frühe Bronzezeit - Jamnaja
			SUG 7	Sugokleja	20	Frühe Bronzezeit - Jamnaja
			SUG 8	Sugokleja	24	Frühe Bronzezeit - Jamnaja
Temrta III (R)	46,5508	43,6649	TEM 1	1	2	Frühe/mittlere Bronzezeit - frühe KGK
			TEM 2	1	4	Frühe/mittlere Bronzezeit - frühe KGK
			TEM 3	1	6	Mittlere Bronzezeit - entwickelte KGK
Temrta V	46,5508	43,6649	TEM 4	1	2 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			TEM 5	1	2 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			TEM 6	1	3 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			TEM 7	1	3 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			TEM 8	1	4	Mittlere Bronzezeit - entwickelte KGK
Tetčani (M)	48,1817	26,9756	TET 1	3	22	Mittlere Bronzezeit - entwickelte KGK
			TET 2	5	5	Frühe Bronzezeit - Jamnaja
			TET 3	5	11	Frühe?/mittlere Bronzezeit - KGK
Vapnjarka (U)	46,9588	30,0314	VAP 1	4	3	Frühe?/mittlere Bronzezeit - KGK
			VAP 2	4	4	Spätes Äneolithikum
			VAP 3	4	16	Frühe Bronzezeit - Jamnaja
			VAP 4	4	14	Frühe Bronzezeit - Jamnaja
Venec (B)	42,6333	26,8000	VEN 1	Tončova	3	keine Angaben
			VEN 2	Tončova	4	keine Angaben
Vinogradnoe (U)	47,2038	35,5592	VIN 1	3	15	Frühes Äneolithikum
			VIN 2	3	25	Frühe Bronzezeit - Jamnaja
			VIN 3	3	36 (2)	Mittlere Bronzezeit - entwickelte KGK
			VIN 4	24	8 (2)	Frühe Bronzezeit - Jamnaja
			VIN 5	24	8 (2)	Frühe Bronzezeit - Jamnaja
			VIN 6	24	15 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			VIN 7	24	15 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			VIN 8	24	17 (2)	Mittlere Bronzezeit - entwickelte KGK
			VIN 9	24	19 (2)	Frühe Bronzezeit - Jamnaja
			VIN 10	24	22	Frühe Bronzezeit - Jamnaja
			VIN 11	24	30	Spätes Äneolithikum – Untere Mikhailovka
			VIN 12	24	31	Frühe Bronzezeit - Jamnaja
			VIN 13	24	32	Mittlere Bronzezeit - entwickelte KGK
			VIN 14	24	34 (2)	Frühe/mittlere Bronzezeit - frühe KGK
			VIN 15	24	34 (2)	Frühe/mittlere Bronzezeit - frühe KGK
Sabrano (B)	42,4380	25,9401	ZAG 1	Zagorci Objekt 12	1 (4)	keine Angaben
			ZAG 2	Zagorci Objekt 12	1 (4)	keine Angaben
			ZAG 3	Zagorci Objekt 12	1 (4)	keine Angaben
			ZAG 4	Zagorci Objekt 12	1 (4)	Frühe Bronzezeit - Jamnaja?
Zauschwitz (D)	51,1807	12,2625	ZAU 1	Flachgrab	P3 (2)	Frühe Bronzezeit - Kugelamphorenkultur
			ZAU 2	Flachgrab	P5 (2)	Frühe Bronzezeit - Kugelamphorenkultur

8.2 Methoden

8.2.1 Maßnahmen zur Kontaminationsvermeidung

Aufgrund der geringen Mengen an genetischem Ausgangsmaterial und einem damit einhergehenden hohen Kontaminationsrisiko durch moderne DNA wurden alle Probenbearbeitungsschritte bis zur Amplifikation der aDNA unter Reinstraumbedingungen durchgeführt. Das Palaeogenetiklabor im Institut für Anthropologie in Mainz bietet für jeden Schritt speziell dafür vorgesehene Räumlichkeiten. Außerdem gibt es eine Reihe von Kriterien zur Kontaminationsvermeidung, die stetig ergänzt und optimiert werden, und die strikt eingehalten werden müssen. Es dürfen daher nur entsprechend geschulte Mitarbeiter das Spurenlabor betreten. Die wichtigsten Punkte seien hier aufgeführt:

Die Prä- und Post-PCR-Bereiche befinden sich in separaten Gebäuden²⁴. Es herrscht ein sogenanntes „Einbahnstraßensystem“. Nach einem Aufenthalt in einer molekulargenetischen Einrichtung oder den zugehörigen Büroräumen darf der Prä-PCR-Bereich (das aDNA-Spurenlabor) nicht mehr betreten werden. Um den DNA-Eintrag von außen zu minimieren, ist ein Aufenthalt im Spurenlabor nur nach einer Dusche und Haarwäsche gestattet. Auch die Kleidung muss frisch gewaschen sein. Bereits vor Eintritt in das Labor werden Taschen, Jacke, Schuhe und sonstige Gegenstände abgelegt und in Spinden aufbewahrt. Direkt beim Betreten der ersten Schleuse, wird an der Tür ein paar frische Socken über die bereits getragenen angezogen. In dem Raum entledigt man sich der Oberbekleidung und zieht ein gewaschenes Shirt und kurze Hosen (z. B. Boxershorts) an. Ebenfalls angezogen werden zwei Paar Einweghandschuhe, ein chirurgischer Mundschutz, ein sauberes Kopftuch und eine Papierhaube. In Schleuse 2 werden ein Tyvek®-Reinstraumanzug mit Kapuze, Überziehschuhe und ein drittes Paar Handschuhe angelegt. Prinzipiell wird das oberste Handschuhpaar bei jedem Betreten eines Raumes, sowie zwischen Arbeitsschritten oder bei sonstigem Bedarf gewechselt. Vor Arbeitsbeginn wird noch

²⁴ Kurze Zeit nachdem die Laborarbeiten zu dieser Arbeit abgeschlossen waren, wurde der Bau eines neuen Institut-Gebäudes mit Palaeogenetiklabor abgeschlossen. Seitdem gibt es diese Gebäudetrennung nicht mehr. Bis auf einige Anpassungen aufgrund der mittlerweile mehr auf NGS ausgerichtete Methoden, sind die hier beschriebenen Regeln weiterhin gültig.

ein Gesichtsschutzvisier angelegt. Anzüge werden ein bis drei Mal getragen, und jeweils nur für einen bestimmten Arbeitsschritt.

Gegenstände dürfen nur nach gründlicher Reinigung mit einer Seifenlösung und Natriumhypochlorit (2,8 %; Danklorix Hygiene-Reiniger) oder DNA-ExitusPlus™ IF ins Palaeogenetiklabor eingeführt werden. Sie werden für mehrere Stunden von allen Seiten in einem eigens dafür vorgesehenen Schrank mit UV-Licht (UV-C, 254 nm) bestrahlt. Dies gilt auch für Pipettenspitzen, Reaktionsgefäße und allgemeine Gebrauchsgegenstände die mehrfach benutzt werden, z.B. *racks* (Reaktionsgefäßständer/-halter), Glasgefäße und Kopfhalterungen (Helme) der Visiere. Prinzipiell werden alle Materialien und Oberflächen nach jeder Verwendung, d. h. nach jedem Arbeitsschritt und am Ende des Arbeitstages wie oben beschrieben gereinigt und bis auf wenige Ausnahmen über Nacht für mehrere Stunden UV-bestrahlt. Hierdurch soll das Risiko der Kreuzkontaminationen minimiert werden. Das Labor ist mit einer zeitschaltuhrgesteuerten UV-Lichtanlage versehen, so dass die wichtigsten Bereiche jede Nacht sechs Stunden den Strahlen ausgesetzt werden. Auch das Wasser, das für die Putzarbeiten, z. B. für die Seifenlösung oder zum Abspülen von Gegenständen nach der Inkubation mit Natriumhypochlorit oder DNA-Exitus, verwendet wird, wird zunächst durch Umkehrosiose aufbereitet und für mindestens 24 Stunden mit einem Tauchstrahler unter Umwälzung UV-bestrahlt.

Jedem Arbeitsschritt im Prä-PCR-Bereich ist ein spezieller Raum zugewiesen, und die Durchführung findet in zweckbestimmten Boxen mit UV-Installation statt. Hierdurch werden die Proben vor den Luftverwirbelungen, die durch Bewegungen des Bearbeiters entstehen, bis zu einem gewissen Maß geschützt. Prinzipiell muss darauf geachtet werden, Bewegungen im Spurenlabor so ruhig und minimal wie möglich zu halten. Alle vier bis sechs Wochen findet eine Grundreinigung des Labors statt, bei der alle Oberflächen, auch Innenräume von Schränken, und Gegenstände gereinigt werden. Der Fußboden wird wöchentlich bzw. bei Säge- und Mahlarbeiten am Ende des Arbeitstages geputzt.

Der Post-PCR-Bereich ist in zwei Labore unterteilt, die ebenfalls räumlich voneinander getrennt sind. Im eigentlichen Post-PCR-Labor werden Agarose-

Gelelektrophoresen, Aufreinigungen von PCR-Produkten, Klonierungen, Sanger-Sequenzierungen und alle Arbeiten mit rezenter DNA durchgeführt. Das zweite Labor war ursprünglich ausschließlich dem *tagging* von Proben gewidmet. Mittlerweile finden dort auch Versuche zu *DNA capture*-Methoden statt. Das *tagging*-Labor ist eine Art Semi-Reinraumlabor. Zwar handelte es sich bei den Proben, die mit *tags* versehen werden sollten, um bereits amplifizierte DNA, doch die hohe Moleküldichte im Post-PCR-Labor und dessen tägliche Nutzung von drei Arbeitsgruppen, stellte doch ein Risiko für das *tagging* dar. Insbesondere die Gefahr von Kreuzkontaminationen zwischen den Meyer-*tag*-Lösungen muss minimiert werden. Da die Proben im weiteren Ablauf bei der Emulsions-PCR noch einmal amplifiziert werden, sollte natürlich auch ein Übertrag zwischen den Proben möglichst vermieden werden. Auch für das *tagging*-Labor gibt es daher einige Regeln zur Kontaminationsvermeidung:

Das Labor darf, wie auch das Spurenlabor, nur geduscht, mit gewaschenen Haaren und frisch gewaschener Kleidung betreten werden. Es ist nicht erlaubt, vorher ein molekulargenetisches Labor zu besuchen, ein voriger Aufenthalt in den Büroräumen des Instituts ist jedoch gestattet. Aufgrund der vielen Inkubations- und Zentrifugierschritte ist das *tagging* mit vielen Luftverwirbelungen verbunden. Daher müssen insbesondere beim Hantieren geöffneter Gefäße unnötige Bewegungen minimiert werden. Handschuhe sollten zwischen den einzelnen Arbeitsschritten gewechselt werden. Am Ende des Arbeitstages werden alle Gebrauchsgegenstände und Oberflächen mit Seifenlösung (mit Osmosewasser) und Natriumhypochlorit bzw. DNA-Exitus gereinigt. Mit Beginn der *DNA capture*-Versuche wurde auch dieses Labor mit Boxen mit UV-Installation für die Durchführung der Arbeiten ausgestattet.

8.2.2 Probenvorbereitung

Vor ihrer Einführung in das palaeogenetische Labor wurde das äußere Erscheinungsbild der Proben digitalphotographisch dokumentiert. Im Labor wurde das Skelettmaterial zunächst von zwei Seiten für jeweils 45 – 60 min mit UV-Licht bestrahlt. Bei Knochenproben erfolgte die Probenentnahme mit einer Rotationssäge [KAVO DENTAL] mit Diamanttrennscheibe [DREMEL]. Je nach Ausgangsmaterial wurde in

einem Bereich von 4 bis 8 cm² die Oberfläche mechanisch entfernt und ein etwa 2 bis 4 g schweres Stück heraus gesägt. Dieses wurde nach Entfernen der *Substantia spongiosa* in kleine Würfel zersägt. Sofern die vorliegenden Knochenfragmente groß genug waren, wurden vorzugsweise unversehrte Bereiche (d. h. keine Risse) mit fester *Substantia compacta* und geringem mikrobiellem Befall für die Probenentnahme gewählt. Bei Zähnen oder sehr weichem, porösen Knochenmaterial wurde lediglich die Oberfläche mithilfe eines Sandstrahlgeräts [HARNISCH & RIETH] entfernt, die Knochen anschließend noch mit der Säge zerkleinert. Der Staubentwicklung bei dieser Arbeit wurde mit einer Absauganlage [HARNISCH & RIETH] entgegengewirkt.

Nach einer erneuten UV-Bestrahlung von 45 – 60 min pro Seite wurde das Probenmaterial mit einer Kugelschwingmühle [RETSCH] pulverisiert. Es wurden 2x 0,5 g Pulver abgewogen und zusammen mit einem dritten Aliquot mit dem restlichen Pulver bei -20 °C gelagert. Neben der üblichen Reinigung mit Seifenlösung und Natriumhypochlorit wurden die Zirkonoxid-Mahlbecher zusätzlich durch Mahlen von ausgeglühtem Seesand gesäubert. Als Negativkontrolle wurde in regelmäßigen Abständen Hydroxylapatitpulver gemahlen und später extrahiert.

8.2.3 DNA-Extraktion

Im ersten Schritt der DNA-Extraktion wurde das Probenmaterial lysiert. Dazu wurden ca. 0,5 g Knochen- bzw. Zahnpulver mit 2,5 ml 0,5 M EDTA (pH 8), 250 µl N-Laurylsarcosin (0,5 % finale Konzentration) und 30 µl Proteinase K (27 U) versetzt und über Nacht bei 37 °C unter konstanter Rotation inkubiert. In Fällen, in denen sich aus dem ersten Extrakt keine oder kaum DNA amplifizieren ließ, wurde bei der zweiten Extraktion ca. 1 g Probenpulver eingesetzt und die Volumina der Lysereagenzien verdoppelt.

Um die DNA zu isolieren, wurde 1 Volumen Phenol/Chloroform/Isoamylalkohol (im Verhältnis 25:24:1) zur Lysesuspension zugegeben, der Ansatz ca. 1 min geschüttelt und anschließend bei 3 300 xg für 10 min zentrifugiert. Die wässrige Phase wurde in ein neues Reaktionsgefäß überführt. Wies die Probe eine starke Verfärbung oder eine

ausgeprägte Interphase auf, wurde dieser Schritt wiederholt. Anschließend wurde die wässrige Lösung mit 1 Volumen Chloroform gut durchmischt und erneut bei 3 300 xg für 10 min zentrifugiert. Die wässrige Phase wurde abgenommen und auf eine vorher für 30 min UV-bestrahlte Amicon® Ultra-15 50K Filtereinheit übertragen.

In mehreren Schritten wurde die Probe durch Zugabe von insgesamt 8 – 10 ml UV-bestrahltem HPLC-H₂O mit anschließendem Abzentrifugieren bei 5 000 xg auf 200 – 250 µl gewaschen und aufkonzentriert. Das Retentat wurde abgenommen und in drei Aliquots bis zur weiteren Verwendung bei -20 °C gelagert.

8.2.4 DNA-Amplifikation

8.2.4.1 *Mitochondriale Primersysteme*

Die mitochondrialen Primersysteme (Tabelle 8.2) und die zugehörige Multiplex-PCR wurden von Martina Unterländer im Rahmen ihres zeitgleich stattfindenden Promotionsvorhabens entwickelt (2014). Der Ansatz ist so konzipiert, dass die Nukleotidpositionen (np) 16011 – 16413 (Nummerierung entsprechend der *revised Cambridge Reference Sequence* (rCRS); nach Andrews *et al.* 1999)²⁵ und damit die HVR I durch die Primersysteme M31 bis M37 vollständig abgedeckt wird. Dabei überlappen benachbarte Amplikons, so dass sich die Sequenzen bezüglich der Authentifizierung gegenseitig stützen. Mit den Primerpaaren M01 bis M30 werden 32 zusätzliche SNPs in der *coding region* investigiert. Die Zusammenstellung der *loci* enthält die wesentlichen Zweige des hierarchischen Systems bei der Definition der mitochondrialen Haplogruppen. Sie untermauern die Ergebnisse der HVR I und ermöglichen teilweise eine tiefere Aufschlüsselung der genetischen Verwandtschaften. Daher gibt es einige theoretisch redundante Primersysteme, aber auch hier dienen die Kombinationen der Mutationen der Authentifizierung. Außerdem können Ausfälle einzelner *loci* (*locus dropout*) durch das Vorhandensein anderer kompensiert werden. Aufgrund der Überlappungsbereiche einiger Primersysteme

²⁵ Im Folgenden bezieht sich die Nummerierung der Nukleotidpositionen der mitochondrialen DNA stets auf die der *revised Cambridge Reference Sequence* (rCRS) nach Andrews *et al.* 1999.

wurde sie in drei Primersets A, B und C aufgeteilt, die in getrennten Multiplex-PCR-Ansätzen (MPA, MPB und MPC) amplifiziert wurden.

Tabelle 8.2: Mitochondriale Primersysteme

In der Spalte „Locus“ sind die Positionen der untersuchten SNPs angegeben. Die Nummerierung richtet sich nach der rCRS. Die zweite Spalte enthält die Primernamen (U = upper; L = lower). Die Zahl gibt dabei die Position der ersten auf das 3'-Ende des Primers folgende Base an. Die Spalte „Code“ gibt die alternativen Primernamen für das Computer-Skript „sort“ an, mit dem die 454-Sequenzdaten sortiert wurden. Es ist jeweils die Länge der Amplikons mit und ohne Primer angegeben, sowie in welchem Multiplex-PCR-Ansatz (MP) sie amplifiziert wurden. Das Primersystem für SNP 10034 (grau) amplifizierte nicht.

Locus	Name	Code	Primersequenz 5' > 3'	Länge [bp]	MP
456	423U	M01U	AAT TTT ATC TTT TGG CGG TAT GCA CTT	110	A
	485L	M01L	GAT GGG CGG GGG TTG TAT TG	63	
663	654U	M02U	CTC ACA TCA CCC CAT AAA CAA ATA GG	94	B
	699L	M02L	AAC TCA CTG GAA CGG GGA TGC T	46	
3010	2992U	M03U	CAA CAA TAG GGT TTA CGA CCT CGA T	116	C
	3057L	M03L	CTC CGG TCT GAA CTC AGA TCA CGT A	66	
4216	4155U	M04U	TAC CCC CGA TTC CGC TAC GA	113	A
	4221L	M04L	ATG CTG GAG ATT GTA ATG GGT ATG GA	67	
4529	4499U	M05U	CTG GCC CAA CCC GTC ATC TA	103	B
	4554L	M05L	GCA TGT TTA TTT CTA GGC CTA CTC AGG	56	
4580	4549U	M06U	ACA GCG CTA AGC TCG CAC TGA T	113	C
	4617L	M06L	ATG GCA GCT TCT GTG GAA CGA G	69	
4833	4815U	M07U	GAA TAG CCC CCT TTC ACT TCT GAG TC	101	A
	4864L	M07L	TGA GAT GGG GGC TAG TTT TTG TCA T	50	
4917	4871U	M08U	GGC CTG CTT CTT CTC ACA TGA CA	120	B
	4940L	M08L	ACT GCC TGC TAT GAT GGA TAA GAT TGA	70	
5178	5163U	M09U	CCA GCA CCA CGA CCC TAC TAC TAT CT	71	C
	5179L	M09L	GGA TGG AAT TAA GGG TGT TAG TCA TGT T	17	
5843	5836U	M10U	AAA TCA CCT CGG AGC TGG TAA AAA G	88	A
	5875L	M10L	GGG GTG AGG TAA AAT GGC TGA GT	40	
6371 6392	6336U	M11U	CAC CCT GGA GCC TCC GTA GAC	116	B
	6403L	M11L	ATG GCA GGG GGT TTT ATA TTG ATA ATT	68	
6776	6764U	M12U	CAA TTG GCT TCC TAG GGT TTA TCG T	101	C
	6814L	M12L	GAT GAT TAT GGT AGC GGA GGT GAA A	51	

Locus	Name	Code	Primersequenz 5' > 3'	Länge [bp]	MP
7028	6975U	M13U	GGT GGC CTG ACT GGC ATT GTA	120	A
	7046L	M13L	TAT GAT GGC AAA TAC AGC TCC TAT TGA	72	
8271 -8281	8226U	M14U	CAT GCC CAT CGT CCT AGA ATT AA	112	B
	8287L	M14L	GCT AAG TTA GCT TTA CAG TGG GCT CTA	62	
8392	8385U	M15U	TAC AGT GAA ATG CCC CAA CTA AAT ACT A	89	C
	8417L	M15L	TTT AGT TGG GTG ATG AGG AAT AGT GTA A	33	
8994	8932U	M16U	ACT TCT TAC CAC AAG GCA CAC CTA CA	116	A
	8996L	M16L	AGT AAT GTT AGC GGT TAG GCG TAC G	65	
9090	9072U	M17U	GAA GCG CCA CCC TAG CAA TAT C	101	B
	9124L	M17L	TAA GGC GAC AGC GAT TTC TAG GAT AG	53	
10034	10000U	M18U	CAT CTA TTG ATG AGG GTC TTA CTC TTT TA	107	C
	10048L	M18L	AAA TTA AGG CGA AGT TTA TTA CTC TTT TT	49	
10115 10118	10105U	M19U	TTA ATA ATC AAC ACC CTC CTA GCC TTA C	109	A
	10166L	M19L	GGT CGA AGC CGC ACT CGT A	62	
10398 10400	10387U	M20U	TCT GGC CTA TGA GTG ACT ACA AAA AG	117	B
	10451L	M20L	AGG GGC ATT TGG TAA ATA TGA TTA TC	65	
10873	10865U	M21U	CAA CCA CCC ACA GCC TAA TTA TTA GC	83	C
	10895L	M21L	TGG GGA ACA GCT AAA TAG GTT GTT GT	31	
11719	11700U	M22U	AGC TTC ACC GGC GCA GTC A	89	A
	11743L	M22L	GTG CGT TCG TAG TTT GAG TTT GCT AG	44	
11947	11935U	M23U	ACC ACG TTC TCC TGA TCA AAT ATC AC	101	B
	11983L	M23L	CCC CAT TGT GTT GTG GTA AAT ATG TA	49	
12308	12303U	M24U	GAT AAC AGC TAT CCA TTG GTC TTA GGC	103	C
	12352L	M24L	GGA AGT CAG GGT TAG GGT GGT TAT AG	50	
12705	12692U	M25U	CAG ACC CAA ACA TTA ATC AGT TCT TCA	110	A
	12754L	M25L	GCC CTC TCA GCC GAT GAA CA	63	
13263	13231U	M26U	GCG CCC TTA CAC AAA ATG ACA TC	94	B
	13275L	M26L	GGT TGG TTG ATG CCG ATT GTA ACT AT	45	
13626	13620U	M27U	AAG CGC CTA TAG CAC TCG AAT AAT TCT	116	C
	13683L	M27L	CCA GGC GTT TAA TGG GGT TTA GTA G	64	
13708	13701U	M28U	ACC CCA CCC TAC TAA ACC CCA TTA A	88	A
	13740L	M28L	GAT GCG GGG GAA ATG TTG TTA GT	40	
14766	14717U	M29U	CAA CCA CGA CCA ATG ATA TGA AAA AC	120	B
	14784L	M29L	GGA GGT CGA TGA ATG AGT GGT TAA TT	68	

<i>Locus</i>	Name	Code	Primersequenz 5' > 3'	Länge [bp]	MP
14798	14783U	M30U	ATA CGC AAA ACT AAC CCC CTA ATA AAA	105	C
	14839L	M30L	GCC AAG GAG TGA GCC GAA GTT	57	
HVR I I	16011U	M31U	AGC ACC CAA AGC TAA GAT TCT AAT TT	130	A
	16088L	M31L	GTG GCT GGC AGT AAT GTA CGA AAT AC	78	
HVR I II	16071U	M32U	GGG TAC CAC CCA AGT ATT GAC TCA	134	B
	16153L	M32L	TGA TGT GGA TTG GGT TTT TAT GTA CTA	83	
HVR I III	16119U	M33U	GTA CAT TAC TGC CAG CCA CCA TG	138	C
	16207L	M33L	TGA TAG TTG AGG GTT GAT TGC TGT AC	89	
HVR I IV	16185U	M34U	TAC ATA AAA ACC CAA TCC ACA TCA AAA C	139	A
	16271L	M34L	GGT GGG TAG GTT TGT TGG TAT CCT	87	
HVR I V	16233U	M35U	AGT ACA GCA ATC AAC CCT CAA CTA TC	127	B
	16305L	M35L	TGT ACG GTA AAT GGC TTT ATG TAC TAT G	73	
HVR I VI	16274U	M36U	AAA GCC ACC CCT CAC CCA CTA G	116	C
	16345L	M36L	TGG GGA CGA GAA GGG ATT TGA C	72	
HVR I VII	16340U	M37U	ACA TAA AGC CAT TTA CCG TAC ATA GCA C	127	A
	16413L	M37L	CAC TCT TGT GCG GGA TAT TGA TTT C	74	

8.2.4.2 *Mitochondriale Multiplex-PCR*

Die mitochondriale Multiplex-PCR wurde mit dem QIAGEN Multiplex PCR Kit durchgeführt. Das grundlegende Protokoll wurde von M. U. entwickelt. Das Primerset für Multiplex-PCR-Ansatz A enthielt 13, die Sets für die Ansätze B und C jeweils 12 äquimolar konzentrierte Primerpaare (siehe Tabelle 8.2 und Tabelle 8.3).

Tabelle 8.3: Konzentrationen der mitochondrialen Primer

(Vgl. Tabelle 8.2)

Primermix	Anzahl Primer	Stammkonzentration je Primer	Finale Konzentration je Primer im Mix
Set A	26	10 pmol/μl	0,385 pmol/μl
Set B	24	10 pmol/μl	0,417 pmol/μl
Set C	24	10 pmol/μl	0,417 pmol/μl

Tabelle 8.4: Mitochondriale Multiplex-PCR – Protokolle 1 & 2

Verwendet wurde das QIAGEN Multiplex PCR Kit. Der Primer-Mix enthält alle in Tabelle 8.2 aufgeführten Primer des jeweiligen Ansatzes. Es sind jeweils die Volumina für einen 70 µl (jeweils obere Zeile = Protokoll 1) und einen 60 µl Ansatz (jeweils untere Zeile = Protokoll 2) angegeben. n. best. = nicht bestimmt. *Hierbei handelt es sich um die Konzentration der einzelnen Primer im PCR-Ansatz.

<i>Reaktionsansätze:</i>					
Reagenzien	Ansatz MPA	Ansatz MPB	Ansatz MPC	finale Konzentration	Stamm- Konzentration
Multiplex Master Mix	35,0 µl 30,0 µl	35,0 µl 30,0 µl	35,0 µl 30,0 µl	1x 1x	2x
Primer Mix	9,1 µl 6,5 µl	8,4 µl 6,0 µl	8,4 µl 6,0 µl	*0,05 pmol/µl *0,04 pmol/µl	10 pmol/µl
BSA	0,0-1,0 µl 0,0-1,0 µl	0,0-1,0 µl 0,0-1,0 µl	0,0-1,0 µl 0,0-1,0 µl	0,29 µg/µl 0,33 µg/µl	20 mg/ml
DNA-Extrakt	5,0-8,0 µl	5,0-8,0 µl	5,0-8,0 µl	n. best.	n. best.
H ₂ O dest.	ad 70,0 µl ad 60,0 µl	ad 70,0 µl ad 60,0 µl	ad 70,0 µl ad 60,0 µl		
<i>PCR-Parameter:</i>					
Initiale Denaturierung		95 °C	15 min		
Denaturierung		94 °C	40 s	39-40 Zyklen	
Hybridisierung		56 °C	40 s		
Elongation		72 °C	40 s		
Finale Elongation		-----	-----		

Bis auf die Variierung der Einsatzmenge an DNA-Extrakt wurde das M. U.s Protokoll zunächst unverändert übernommen. Ursprünglich wurde die PCR in Reaktionsansätzen von 70 µl durchgeführt, um sicher zu sein, ausreichend PCR-Produkt für die Sequenzierung zu haben. Im Zuge der ersten Anwendungen wurde die Volumina der Primer-Mixe und der Gesamtansätze (auf 60 µl) reduziert. Zur Stabilisierung der Reaktion wurde Bovines Serum Albumin (BSA) zugesetzt. Die beiden Protokolle sind in Tabelle 8.4 dargestellt. Während Protokoll 1 bei M. U.s Proben erfolgreich angewendet werden konnte, wurde es aufgrund der schlechteren Probenerhaltung und daraus resultierenden unzureichenden Amplifikationserfolgen nach den ersten Sequenzierläufen nochmals modifiziert. Die Abwandlungen betreffen vor allem die PCR-Parameter. Das in Tabelle 8.5 gezeigte Protokoll 3 wurde ab seiner Etablierung (PCR Nr. M11) ausschließlich verwendet. Alle PCR-Reaktionen (auch die im

folgenden beschriebenen nukleären Multiplex-PCRs und die Einzel-PCRs) fanden in Thermocyclern [EPPENDORF] statt.

Tabelle 8.5: Mitochondriale Multiplex-PCR – Protokoll 3

Verwendet wurde das QIAGEN Multiplex PCR Kit. Der Primer-Mix enthält alle in Tabelle 8.2 aufgeführten Primer des jeweiligen Ansatzes (MP A, B bzw. C). *Hierbei handelt es sich um die Konzentration der einzelnen Primer im PCR-Ansatz.

<i>Reaktionsansätze:</i>					
Reagenzien	Ansatz MPA	Ansatz MPB	Ansatz MPC	finale Konzentration	Stamm-Konzentration
Multiplex Master Mix	25,0 µl	25,0 µl	25,0 µl	1 x	2 x
Primer Mix	6,5 µl	6,0 µl	6,0 µl	*0,05 pmol/µl	10 pmol/µl
BSA	1,0 µl	1,0 µl	1,0 µl	0,40 µg/µl	20 mg/ml
DNA-Extrakt	6,0-8,0 µl	6,0-8,0 µl	6,0-8,0 µl	n. best.	n. best.
UV-HPLC-H ₂ O	ad 50,0 µl	ad 50,0 µl	ad 50,0 µl		
<i>PCR-Parameter:</i>					
Initiale Denaturierung		95 °C	15 min		
Denaturierung		94 °C	30 s	39-41 Zyklen	
Hybridisierung		56 °C	90 s		
Elongation		72 °C	90 s		
Finale Elongation		72 °C	10 min		

8.2.4.3 Nukleäre Primersysteme

Die Auswahl der nukleären Marker erfolgte im Zuge einer Literaturrecherche nach den in Kapitel 5.4 erläuterten Kriterien – zusammengefasst: Punktmutationen mit phänotypischen Auswirkungen, die unter dem Verdacht stehen, positiv selektiert zu sein. Als Referenzsequenzen dienten die Eintragungen in der NCBI SNP-Datenbank (NCBI - dbSNP). Da die SNPs in den Genen *AGT* und *CASP12* zum Zeitpunkt des Primerdesigns nicht in der Datenbank gelistet waren, wurden sie durch den Abgleich mehrerer Sequenzen (*alignment*) identifiziert. Bei ersterem stellte sich im Nachhinein

heraus, dass die falsche Stelle amplifiziert wurde. Daher wurde der *locus* aus der weiteren Betrachtung ausgeschlossen.

Die nukleären Primersysteme (Tabelle 8.6) wurden mit dem Programm PrimerSelect™ des Softwarepakets Lasergene® Versionen 8 und 9 [GATC BIOTECH AG und DNASTAR, INC.] entwickelt. Bei der Auswahl der Primer spielten mehrere Faktoren eine Rolle: Es wurden kürzere Amplikons als bei der mtDNA angestrebt. Das ΔG -Profil der Primer sollte möglichst ausgeglichen, zum 3'-Ende ansteigend sein. Die Spanne der Schmelztemperaturen und, davon abhängig, die der jeweils optimalen *annealing*-Temperaturen sollte möglichst klein sein. Die Tendenz zur Ausbildung von Sekundärstrukturen wie Dimer- oder *hairpin*-Strukturen sollte ebenfalls möglichst gering sein. Dies spielte vor allem im Hinblick auf den Einsatz der Primer in einer Multiplex-PCR eine große Rolle, da hier Kompromisse eingegangen werden mussten. Ein an sich optimales Primerpaar war nicht unbedingt für die Verwendung in der Multiplex-PCR geeignet. Ein SNP im Pigmentiergen *KITLG* wurde zwar erfolgreich in einer Einzel-PCR amplifiziert, doch aufgrund der AT-reichen Sequenz des *locus* erwies sich die in der Multiplex-PCR verwendete Hybridisierungstemperatur als zu hoch.

Tabelle 8.6: Nukleäre Primersysteme

Die Spalte „Locus“ gibt die Gene an, in denen die untersuchten SNPs liegen. Mit a, b und c werden mehrere im gleichen Gen liegende SNPs unterschieden. Es gibt *locus*-basierte Primernamen (U = *upper*, L = *lower*) und Codes für die bioinformatische Sortierung der 454-Sequenzen nach Primersystemen. Die Länge der Amplikons ist mit und ohne Primer angegeben. Bis auf das Paar für den *locus AMELX/Y* (Graefen 2010) wurden alle Systeme im Rahmen dieser Dissertation entwickelt. *Die jeweiligen SNPs liegen in Introns von *HERC2* bzw. *MCM6* und gleichzeitig in regulatorischen Elementen von *OCA2* bzw. *LCT*. S07 und S19 wurden zwar amplifiziert, gingen jedoch nicht in die Datenanalyse ein.

<i>Locus</i>	Name	Code	Primersequenz 5' > 3'	Länge [bp]
<i>ABCB1</i> a	ABBa_U	S01U	TGT CTG TGA ATT GCC TTG AAG TTT TT	92
	ABBa_L	S01L	GCC ACC GTC TGC CCA CTC T	47
<i>ABCB1</i> b	ABBb_U	S02U	GGA CAA GCA CTG AAA GAT AAG AAA GAA CT	74
	ABBb_L	S02L	ATT AAT CAA TCA TAT TTA GTT TGA CTC ACC T	14
<i>ABCB1</i> c	ABBc_U	S03U	AAC AGC CGG GTG GTG TCA CA	74
	ABBc_L	S03L	GAC TCG ATG AAG GCA TGT ATG TTG G	29

<i>Locus</i>	Name	Code	Primersequenz 5' > 3'	Länge [bp]
<i>ABCC11</i>	ABC_U	S04U	AGC CTA GAG TCC CCC AAA CCT CAC	84
	ABC_L	S04L	CAC TTC TGG GCA TCT GCT TCT GC	37
<i>ADH1B a</i>	ADHa_U	S05U	AAA TGA AAT TCA AGA ACA TCT CTG ACC	104
	ADHa_L	S05L	GTT TTA GGG CCC ATT CTC TTT TGT	53
<i>ADH1B b</i>	ADHb_U	S06U	TCT AGT TAC AAT CTT TTC TGA ATC TGA ACA G	110
	ADHb_L	S06L	GGG TCA CCA GGT TGC CAC TAA	58
<i>AGT</i>	AGT_U	S07U	CTG AGG GGT GGG GAT GGA G	84
	AGT_L	S07L	CTG TGG TCT GGC CAA GTG ATG TA	45
<i>CASP12</i>	CAS_U	S08U	AAT CGA AAT GGT TCT GAA CTT GAC CT	107
	CAS_L	S08L	GGA GCC ATT ACC TGA GCT GTG AGA TT	55
<i>CYP3A4</i>	CY4_U	S09U	GTA GGT GTG GCT TGT TGG GAT GAA	110
	CY4_L	S09L	AAG TGG AGC CAT TGG CAT AAA ATC TAT	59
<i>CYP3A5</i>	CY5_U	S10U	CCA CCC AGC TTA ACG AAT GCT CTA	108
	CY5_L	S10L	GAT GAA GGG TAA TGT GGT CCA AAC AG	58
<i>EDAR</i>	EDA_U	S11U	AAT GCT CAG CTC CAC GTA CAA CTC TG	97
	EDA_L	S11L	GCC CCC AAT CTC ATC CCT CTT C	49
<i>HERC2*</i> <i>OCA2</i>	HER_U	S12U	CCA AGA GGC GAG GCC AGT TTC A	76
	HER_L	S12L	AAG CCT CGG CCC CTG ATG ATG	33
<i>MCM6*</i> <i>LCT a</i>	LCTa_U	S13U	CTG CGC TGG CAA TAC AGA TAA GA	72
	LCTa_L	S13L	CAA ATG CAA CCT AAG GAG GAG AGT T	24
<i>MCM6*</i> <i>LCT b</i>	LCTb_U	S14U	AAG ATG TCC TTA AAA ACA GCA TTC TCA	67
	LCTb_L	S14L	CCA AAG TAC TGG GAC AAA GGT GTG	16
<i>NAT2</i>	NAT_U	S15U	CAC TGG CAT GGT TCA CCT TCT CC	72
	NAT_L	S15L	CCA GAC CCA GCA TCG ACA ATG TAA	25
<i>SLC12A3</i>	S12_U	S16U	CGT GGA CCC CAT TAA CGA CAT C	107
	S12_L	S16L	ATG GCC TCC TCA CCT TGG ACT C	63
<i>SLC24A5</i>	S24_U	S17U	GTT CAG CCC TTG GAT TGT CTC AGG	104
	S24_L	S17L	GGA GCA ATA TTT ACC TAG GAA AGC AGT	53
<i>SLC45A2</i>	S45_U2	S18U	AGA ATA AAG TGA GGA AAA CAC	72
	S45_L2	S18L	GAA AGA GGA GTC GAG GTT GGA	24
<i>TRPV6 a</i>	TV6a_U	S19U	AGC CAC AGG CAC TGC CTT CC	65
	TV6a_L	S19L	CAG CCC TGC TCT CAC CAA AGT AGA T	20
<i>TRPV6 b</i>	TV6b_U2	S20U	AGG CCT GAA AGC CCT GAG CAT	84
	TV6b_L2	S20L	ACC AGC CGG ATA TCG TCC TTA GG	40

<i>Locus</i>	Name	Code	Primersequenz 5' > 3'	Länge [bp]
<i>TRPV6 c</i>	TV6c_U	S21U	GGA AAA ACT AGA GCT GGG CTG TCC	88
	TV6c_L	S21L	GGC ACT GCT GCG GGA GGT A	45
<i>TYR</i>	TYR_U	S22U	TTT GTC TGG ATG CAT TAT TAT GTG TCA	85
	TYR_L	S22L	CTT CAT GGG CAA AAT CAA TGT CTC	34
<i>AMEL X/Y</i>	AME_U	S23U	CCC TGG GCT CTG TAA AGA ATA GTG	106 / 112
	AME_L	S23L	ATC AGA GCT TAA ACT GGG AAG CTG	58 / 64

Desweiteren musste beim Primerdesign der Evolution der Gene Rechnung getragen werden. Einige der Gene gehören zu Genfamilien, die sich in ihren Sequenzen stark ähneln können. Die Primer und Amplikons wurden mittels BLAST²⁶ (NCBI - BLAST , Altschul *et al.* 1990) mit den Einträgen in der NCBI-Nukleotiddatenbank abgeglichen (blastn) um eine spezifische Amplifikation zu gewährleisten. Gegebenenfalls wurden Primer so positioniert, dass z. B. das 3'-Ende auf einer variablen Stelle der sich ähnelnden Sequenzen lag, wie beispielsweise bei *ADH_U*, um nicht *ADH1A* oder *ADH1C* zu amplifizieren. Im Fall von *TRPV6* wurde erst im Verlauf des Projekts die Sequenz des homologen *TRPV5* in der NCBI-Datenbank archiviert. Daher musste im Nachhinein der *locus TRPV6 a* aufgrund von unspezifischer Amplifikation von der Untersuchung ausgeschlossen werden. Das Primerpaar für *Amelogenin* wurde von Angela Graefen (2010) entwickelt..

8.2.4.4 Nukleäre Multiplex-PCR

Für die Etablierung der nukleären Multiplex-PCR wurden mehrere Testreihen sowohl an rezenter als auch alter DNA durchgeführt. Es wurden das QIAGEN Multiplex PCR Kit mit und ohne Q-Faktor und verschiedene Ansätze mit AmpliTaq Gold® DNA Polymerase getestet. Zur Evaluierung wurden zunächst die Photos der Agarose-Gelelektrophorese herangezogen. Die augenscheinlich besten Produkte, d. h. ein

²⁶ BLAST, *Basic Local Alignment Search Tool*, ist ein Suchprogramm, das eine Sequenz mit bereits in Datenbanken eingetragenen Sequenzen vergleicht und durch ein *alignment* Regionen lokaler Ähnlichkeit findet. Es gibt verschiedene Versionen des Programms, z. B. für DNA- (blastn) oder Proteinsequenzen (blastp), oder es können bestimmte Datenbanken, z. B. ausgewählte Organismen, durchsucht werden (Altschul *et al.* 1990).

kräftiges, definiertes Bandenfeld auf dem Gel, wurden 454-sequenziert. Anhand der Gesamtanzahl von Sequenzen und ihrer Verteilung auf die einzelnen *loci* wurde das am besten geeignete Protokoll ermittelt. Dieses wurde während der Datenproduktion weiter optimiert. Ein parametrisiertes Protokoll, in dem die Spannweiten der getesteten Konzentrationen der Reagenzien im Verlauf der Optimierungsphase angegeben sind, sowie das finale Protokoll sind in Tabelle 8.7 dargestellt.

Die 46 Primer (Tabelle 8.2) wurden äquimolar in den Primer-Mix eingesetzt, da bei den Vorversuchen unterschiedliche Konzentrationen der einzelnen Primerpaare nicht zu einer Verbesserung der Ausbeute führten.

Tabelle 8.7: Nukleäre Multiplex-PCR-Protokolle

Es wurde die *hotstart*-Polymerase AmpliTaq Gold® [APPLIED BIOSYSTEMS INC.] eingesetzt.

Protokoll 1: zunächst wurde parallel zur Probenbearbeitung das Protokoll weiter optimiert.

Protokoll 2: das finale Standardprotokoll (ab PCR S32).

n. best. = nicht bestimmt.

*Hierbei handelt es sich um die Konzentration der einzelnen Primer im PCR-Ansatz.

**PCR S38 – S40: 8,4 µl, da die Primersysteme AGT und TV6a nicht mehr verwendet wurden.

†ab PCR S35: Verkürzung der initialen Denaturierungsphase, um die Polymerase nach und nach freizusetzen, und so das Verhältnis zwischen *template*- und Enzym-Menge zu verbessern.

<i>Reaktionsansätze:</i>					
Reagenzien	Protokoll 1	finale Konzentration	Protokoll 2	finale Konzentration	Stamm-Konzentration
Puffer	15,0 µl	1,5 x	12,0 µl	1,5 x	10 x
MgCl ₂	25,0-30,0 µl	6,25-7,50 mM	20,0 µl	6,25 mM	25 mM
dNTPs	4,0 µl	0,4 mM	3,2 µl	0,4 mM	10 mM
Taq	2,0-3,0 µl	10-15 U	1,6 µl	8 U	5 U/µl
BSA	3,0-4,0 µl	0,6-0,8 µg/µl	2,4 µl	0,6 µg/µl	20 mg/ml
Primer Mix	11,5-15,0 µl	*25-33 fmol/µl	**9,2 µl	*25 fmol/µl	10 pmol/µl
DNA-Extrakt	14,0-16,0 µl	n. best.	10,0-18,0 µl	n. best.	n. best.
UV-HPLC-H ₂ O	ad 100,0 µl		ad 80,0 µl		
<i>PCR-Parameter:</i>					
Initiale Denaturierung		94 °C	6 / 3 [†] min		
Denaturierung		94 °C	40 s	Protokoll 1: 45-50 Zyklen Protokoll 2: 45 Zyklen	
Hybridisierung		58 °C	40 s		
Elongation		72 °C	40 s		

8.2.4.5 Einzel-PCR

Tabelle 8.8 zeigt die PCR-Protokolle zur Amplifikation einzelner *loci*. Auch hier kam die AmpliTaq Gold® DNA Polymerase zum Einsatz. Einzel-PCRs wurden zum Testen von Chemikalien auf DNA-Kontaminationen (z. B. Primerlösungen, UV-bestrahltes HPLC-Wasser für Extraktion und PCR) und zur Feststellung von Spezifität und Effizienz von Primersystemen verwendet. Außerdem wurde damit die Amplifikation einzelner *loci*, die in der Multiplex-PCR ausfielen, wiederholt. Diese Maßnahme kam insbesondere für die Primersysteme der HVR I zum Tragen. Alternativ zu den Primern M31 bis M37 (Tabelle 8.2) wurden zur Amplifikation auch die vier überlappenden Primersysteme aus Tabelle 8.9 eingesetzt.

Tabelle 8.8: Einzel-PCR-Protokolle

Es wurde die *hotstart*-Polymerase AmpliTaq Gold® [APPLIED BIOSYSTEMS INC.] eingesetzt. Die Hybridisierungstemperatur war primerabhängig (56 °C für die mitochondrialen, 58 °C für die nukleären Primer). U = *upper/forward primer*, L = *lower/reverse primer*. n. best. = nicht bestimmt.

Protokoll aDNA: Nachamplifikationen fehlender *loci*

Protokoll rezent: Primer- und Chemikaliestests mit Speichelextrakt als Positivkontrolle.

Reaktionsansätze:					
Reagenzien	Protokoll aDNA	finale Konzentration	Protokoll rezent	finale Konzentration	Stamm-Konzentration
Puffer	5,0 µl	1 x	2,5 µl	1 x	10 x
MgCl ₂	5,0 µl	2,5 mM	2,5 µl	2,5 mM	25 mM
dNTPs	1,0 µl	0,2 mM	0,5 µl	0,2 mM	10 mM
Taq	0,5 µl	2,5 U	0,1 µl	0,5 U	5 U/µl
BSA	1,0 µl	0,4 µg/µl			20 mg/ml
Primer U	0,5-1,0 µl	0,1-0,2 pmol/µl	0,5 µl	0,2 pmol/µl	10 pmol/µl
Primer L	0,5-1,0 µl	0,1-0,2 pmol/µl	0,5 µl	0,2 pmol/µl	10 pmol/µl
DNA-Extrakt	3,0-5,0 µl	n. best.	0,5-1,0 µl	n. best.	n. best.
UV-HPLC-H ₂ O	ad 50,0 µl		ad 25,0 µl		
PCR-Parameter:					
Initiale Denaturierung		94 °C	6 min		
Denaturierung		94 °C	40 s	aDNA: 40 Zyklen rezent: 30 Zyklen	
Hybridisierung		56-58 °C	40 s		
Elongation		72 °C	40 s		

Tabelle 8.9: Alternative Primersysteme zur Amplifikation der HVR I

Die Spalte „Locus“ gibt die vier überlappenden Fragmente der HVR I an (a = alternativ; zur Abgrenzung von den Multiplex-PCR-Primern). Die Primer sind nach dem Strang, an den sie binden (L = *light strand*; H = *heavy strand*) und der Position ihres 3'-Endes in Bezug auf die rCRS benannt. Es ist jeweils die Länge der Amplikons mit und ohne Primer angegeben. Die Primer wurden aus den aufgelisteten Studien entnommen.

Locus	Name	Primersequenz 5' > 3'	Länge [bp]	Quelle
HVR I Ia	L15996	CTC CAC CAT TAG CAC CCA AAG C	187	Endicott <i>et al.</i> 2003
	H16142	ATG TAC TAC AGG TGG TCA AG	145	Stone & Stoneking 1998
HVR I IIa	L16117	TAC ATT ACT GCC AGC CAC CAT	162	Haak <i>et al.</i> 2005
	H16233	GCT TTG GAG TTG CAG TTG ATG TGT	115	Haak <i>et al.</i> 2005
HVR I IIIa	L16209	CCC CAT GCT TAC AAG CAA GT	179	Handt <i>et al.</i> 1996
	H16348	ATG GGG ACG AGA AGG GAT TTG	138	Haak <i>et al.</i> 2005
HVR I IVa	L16287	CAC TAG GAT ACC AAC AAA CC	162	Handt <i>et al.</i> 1996
	H16410	GCG GGA TAT TGA TTT CAC GG	122	Handt <i>et al.</i> 1996

8.2.4.6 Agarosegelelektrophorese

Zur Überprüfung des Amplifikationserfolgs wurden die PCR-Produkte auf eine mit Ethidiumbromid versetzte 5 – 6 mm dicke Agarose-Gelmatrix aufgetragen. Die Agarosekonzentration betrug 2 % (w/v) in 1x Tris-Borat-EDTA-Puffer (TBE) für Einzel- und mitochondriale Multiplex-PCR-Produkte, und 3 – 3,5 % für die kürzeren nukleären PCR-Produkte. Der TBE-Puffer diente auch als Laufpuffer. Jeweils 10 µl DNA wurden mit ca. 1 µl Bromphenolblau oder *loading dye* gemischt und in die Geltaschen pipettiert. Außerdem wurde eine 50 bp DNA-Leiter in einer separaten Spur mitgeführt. Die Laufzeit betrug für die mtDNA ca. 20 – 25 min bei 130 – 145 V, für nukleäre DNA (ncDNA) ca. 65 – 75 min bei 110 – 120 V. Die digitale Dokumentation erfolgte mit einer CCD-Kamera unter UV-Belichtung bei 254 nm [INTAS SCIENCE IMAGING INSTRUMENTS].

Da die Amplikonlängen innerhalb einer Multiplex-PCR zum Teil identisch waren oder sich nur um wenige Basenpaare unterschieden, wurde der Amplifikationserfolg basierend auf der Intensität und Abdeckung des Bandenfeldes grob abgeschätzt. Eine Aussage über die Amplifikation individueller *loci* konnte dabei nicht getroffen werden.

8.2.4.7 *Aufreinigung der PCR-Produkte*

Vorbereitend auf das *tagging* nach Meyer *et al.* (2008) wurden die PCR-Produkte mit dem MinElute® PCR Purification Kit nach dem „*PCR Purification Spin Protocol*“ aufgereinigt. Abweichend zu diesem Protokoll wurde der Elutionspuffer (EB) mit 0,05 % Tween 20 (in Tris-EDTA-Puffer) versetzt und auf 60 °C erwärmt. Die Inkubation der auf der Silicamembran der MinElute®-Säule gebundenen DNA mit dem Elutionspuffer fand ebenfalls bei 60 °C statt, und die Dauer wurde auf 5 – 10 min erhöht, um die Ausbeute zu optimieren.

Es wurden jeweils die Produkte einer mitochondrialen – im Falle der Multiplex-PCR von drei, im Falle der HVR I-Einzel-PCRs von sieben Reaktionsansätzen – und einer nukleären Multiplex-PCR zusammen über eine Filtereinheit aufgereinigt. Gegebenenfalls wurde das Gemisch von Probe und Bindepuffer (PB) in zwei Schritten auf die Säule gegeben und zentrifugiert. Das Volumen des Eluats war abhängig vom nächsten Bearbeitungsschritt und betrug mindestens 20 µl. In einigen Fällen wurden vor der Aufreinigung von jedem Reaktionsansatz 15 µl entnommen, die stattdessen enzymatisch aufgereinigt wurden. Sie dienten als Reserve für etwaige Nachsequenzierungen.

PCR-Produkte, die für die klassische Sequenzierung nach Sanger *et al.* (1977) vorgesehen waren, wurden durch die Enzyme Exonuklease I (EXO I) – zum Primerabbau – und *shrimp alkaline phosphatase* (SAP) bzw. *FastAP thermosensitive alkaline phosphatase* (FastAP) – zur Entfernung von freien Phosphatgruppen – aufgereinigt. Dem PCR-Produkt wurden 2 µl des folgenden Enzymgemischs zugefügt: 0,5x EXO I-Puffer, 2 U EXO I, 3,3x SAP- bzw. FastAP-Puffer und 0,33 U SAP bzw. FastAP. Nach einer 45-minütigen (SAP) bzw. 30-minütigen (FastAP) Inkubation wurden die Enzyme bei 85 °C für 15 min inaktiviert. Alternativ wurde eine Säulenaufreinigung mit dem MSB® Spin PCRapace Kit nach Herstellerangaben, aber mit einer Inkubation des Elutionspuffers von 10 anstatt 1 min, durchgeführt.

8.2.5 „Tagging“ der Proben

8.2.5.1 Barcodes

Es kamen 60 verschiedene *tags* mit 8 Basen langen Erkennungssequenzen, die sich in mindestens 3 Nukleotiden unterscheiden, zum Einsatz (siehe Tabelle 8.10).

Vor der Verwendung wurden die Oligonukleotid-Aliquots (500 µM) einmalig durch Erhitzen auf 95 °C und kontrollierten (ca. 0,1 °C / s) Abkühlen auf Raumtemperatur hybridisiert. Sie wurden bei -20 °C gelagert und konnten ohne erneute Hybridisierung eingesetzt werden.

Tabelle 8.10: Barcode-Sequenzen

Es wurden 60 *tags* aus Meyer *et al.* (2008) ausgewählt, die sich an mindestens drei Positionen der acht Basen langen Identifikationssequenz unterscheiden, um Fehlzunordnungen aufgrund von Sequenzierfehlern zu vermeiden.

<i>Tag</i>	Sequenz 5' > 3'	<i>Tag</i>	Sequenz 5' > 3'
001	ACA CAC ACG CCC GGG CGT GTG TGT	031	AGC ATC ACG CCC GGG CGT GAT GCT
002	ACA CGA GCG CCC GGG CGC TCG TGT	032	AGC GAG TCG CCC GGG CGA CTC GCT
003	ACA GAT CGG CCC GGG CCG ATC TGT	033	AGC GTG AGG CCC GGG CCT CAC GCT
004	ACA GCA TCG CCC GGG CGA TGC TGT	034	AGC TAT ACG CCC GGG CGT ATA GCT
005	ACA GTC AGG CCC GGG CCT GAC TGT	035	AGC TGA TGG CCC GGG CCA TCA GCT
006	ACA TCG AGG CCC GGG CCT CGA TGT	036	AGT ACA TCG CCC GGG CGA TGT ACT
007	ACG ACA GCG CCC GGG CGC TGT CGT	037	AGT AGT ACG CCC GGG CGT ACT ACT
008	ACG AGA CGG CCC GGG CCG TCT CGT	038	AGT ATC TGG CCC GGG CCA GAT ACT
009	ACG AGC ACG CCC GGG CGT GCT CGT	039	AGT CAG ACG CCC GGG CGT CTG ACT
010	ACG CAG AGG CCC GGG CCT CTG CGT	040	AGT CTA GCG CCC GGG CGC TAG ACT
011	ACG CTA TCG CCC GGG CGA TAG CGT	041	AGT GAC AGG CCC GGG CCT GTC ACT
012	ACG TAC TCG CCC GGG CGA GTA CGT	042	AGT GCT GCG CCC GGG CGC AGC ACT
013	ACG TCA TGG CCC GGG CCA TGA CGT	043	ATA CAG TCG CCC GGG CGA CTG TAT
014	ACG TCT ACG CCC GGG CGT AGA CGT	044	ATA CGT ACG CCC GGG CGT ACG TAT
015	ACT ACG ACG CCC GGG CGT CGT AGT	045	ATA CTG CGG CCC GGG CCG CAG TAT
016	ACT ATC GCG CCC GGG CGC GAT AGT	046	ATA GAG AGG CCC GGG CCT CTC TAT
017	ACT CAC TGG CCC GGG CCA GTG AGT	047	ATA TAC GCG CCC GGG CGC GTA TAT
018	ACT CGT AGG CCC GGG CCT ACG AGT	048	ATA TCA CGG CCC GGG CCG TGA TAT

Tag	Sequenz 5' > 3'	Tag	Sequenz 5' > 3'
019	ACT GAT ACG CCC GGG CGT ATC AGT	049	ATC ACT ACG CCC GGG CGT AGT GAT
020	ACT GTA CGG CCC GGG CCG TAC AGT	050	ATC AGC AGG CCC GGG CCT GCT GAT
021	ACT GTG TCG CCC GGG CGA CAC AGT	051	ATC ATA TCG CCC GGG CGA TAT GAT
022	AGA CAT AGG CCC GGG CCT ATG TCT	052	ATC GAC ACG CCC GGG CGT GTC GAT
023	AGA CGA CGG CCC GGG CCG TCG TCT	053	ATC GCA GCG CCC GGG CGC TGC GAT
024	AGA GCG ACG CCC GGG CGT CGC TCT	054	ATC GCG CGG CCC GGG CCG CGC GAT
025	AGA GTA TGG CCC GGG CCA TAC TCT	055	ATC GTC TGG CCC GGG CCA GAC GAT
026	AGA GTC GCG CCC GGG CGC GAC TCT	056	ATC TAT CGG CCC GGG CCG ATA GAT
027	AGA TAG CGG CCC GGG CCG CTA TCT	057	ATC TGC TCG CCC GGG CGA GCA GAT
028	AGA TGC ACG CCC GGG CGT GCA TCT	058	ATG ACG TCG CCC GGG CGA CGT CAT
029	AGC ACA CGG CCC GGG CCG TGT GCT	059	ATG ACT CGG CCC GGG CCG AGT CAT
030	AGC AGA GCG CCC GGG CGC TCT GCT	060	ATG ATG AGG CCC GGG CCT CAT CAT

8.2.5.2 Das Tagging-Protokoll

Für den ersten Schritt beim *tagging*, dem *blunt-end repair* wurden 20 µl aufgereinigtes PCR-Produkt benötigt. Dabei waren mitochondriale und nukleäre Amplikons jeweils eines Individuums bereits gemischt. Die Aufreinigungsschritte zwischen den einzelnen enzymatischen Reaktionen erfolgten mit dem MinElute® PCR Purification Kit wie unter 8.2.4.7 beschrieben. Das Volumen des Elutionspuffers entsprach jeweils der in der nachfolgenden Reaktion benötigten Menge an Produkt plus 2 µl, um Verluste bei der Filtration auszugleichen.

Das Prinzip der *tagging*-Methode und die biochemischen Aspekte wurden bereits ausführlich in Kapitel 7.2.2 erläutert. Tabelle 8.11 (A – E) zeigt die Reaktionsansätze der einzelnen Schritte. Die Inkubationen fanden in Thermocyclern [BIOMETRA, EPPENDORF] statt. Stichprobenartig wurden mit dem Qubit® Fluorometer [LIFE TECHNOLOGIES™ CORP.] und dem dsDNA HS Assay Kit die DNA-Konzentrationen von Proben vor dem *tagging* und vor dem *pooling* gemessen. Eine Anpassung der DNA-Mengen, wie von Meyer *et al.* (2008) empfohlen, wurde jedoch nur bei einem ersten 454-Testlauf im Department of Biology an der University of Copenhagen (T. G.

Tabelle 8.11: Reaktionsansätze beim Tagging

Konz. = Konzentration; Temp. = Temperatur; Pol = Polymerase; PNK = Polynukleotidkinase; CIP = *calf intestinal phosphatase*; ü. N. = über Nacht

Reagenz	Volumen	finale Konz.	Stammkonz.		Temp.	Zeit
<i>(A) Blunt-End Repair</i>						
T4 DNA Pol Buffer	8,0 µl	1 x	5 x		12 °C	15 min
BSA	0,5 µl	0,25 µg/µl	20 mg/ml		25 °C	15 min
dNTPs	0,4 µl	0,1 mM	10 mM			
ATP	4,0 µl	1,0 mM	10 mM			
T4 PNK	2,0 µl	20 U	10 U/µl			
T4 DNA Pol.	0,8 µl	4 U	5 U/µl			
PCR-Produkt	20,0 µl					
HPLC-H ₂ O	ad 40 µl					

<i>(B) Ligation</i>						
T4 Ligase Buffer	4,0 µl	1 x	10 x		22 °C	ü. N.
PEG 4000	4,0 µl	5 %	50 %			
T4 DNA Ligase	1,0-1,1 µl	5,0-5,5 U	5 U/µl			
Tag (hybridisiert)	1,6-1,8 µl	20,0-22,5 µM	500 µM			
Produkt aus (A)	29,0 µl					
HPLC-H ₂ O	ad 40 µl					

<i>(C) Adapter Fill-In</i>						
Thermo Buffer	3,0 µl	1 x	10 x		37 °C	30 min
dNTPs	0,75 µl	250 µM	10 mM			
Bst Polymerase	1,0 µl	8 U	8 U/µl			
Produkt aus (B)	23,0 µl					
HPLC-H ₂ O	ad 30 µl					

<i>(D) Dephosphorylation</i>						
NEBuffer	3,0 µl	1 x	10 x		37 °C	30 min
CIP	1,5 µl	15 U	10 U/µl			
Produkt aus (C)	20,0 µl					
HPLC-H ₂ O	ad 30 µl					

<i>(E) Restriction Cut</i>						
Universal Buffer	3,0 µl	1 x	10 x		37 °C	60 min
Srf I	1,0 µl	10 U	10 U/µl			
Produkt aus (D)	25,0 µl					
HPLC-H ₂ O	ad 30 µl					

und Kollegen) vorgenommen. Da die DNA-Mengen relativ gering und die Konzentrationsmessungen äußerst ungenau waren, wurde daraufhin so viel Produkt wie möglich in die Reaktionen eingesetzt.

Nach dem *adapter fill-in* (C) wurden jeweils 4 – 5 Proben über eine gemeinsame MinElute® Säule aufgereinigt, so dass sich die Anzahl der Reaktionsansätze entsprechend auf 8 – 12 Probenpools reduzierte. Bei den ersten Durchführungen des *taggings* wurde das Protokoll bis zum *restriction cut* (E) umgesetzt. Nach der Dephosphorylierung (D) wurde die Probenanzahl im Zuge der Aufreinigung nochmals reduziert, so dass der letzte Schritt (E) nur noch mit einem oder zwei Probenpools durchgeführt wurde. Nach einer finalen Aufreinigung befanden sich alle *tagging*-Produkte, gelöst in 100 µl Elutionspuffer, in einem Reaktionsgefäß, welches dann für die *library*-Erstellung und die 454-Sequenzierung zur GATC BIOTECH AG [Konstanz] geschickt wurde. Ein run (T8) wurde von der LGC GENOMICS GMBH [Berlin] sequenziert.

Die Dephosphorylierung (D) und der *restriction cut* (E) dienen dazu, dass nur beidseitig mit *tags* versehene DNA-Fragmente 5'-Phosphatreste tragen, und die Roche 454-Adapter somit nur an diese Moleküle ligieren können. Voraussetzung ist jedoch, dass der erste Schritt der *library*-Erstellung, das *end polishing*, ausgelassen wird. Da dieser Schritt bei der GATC BIOTECH AG allerdings standardmäßig durchgeführt wird, erwiesen sich die Schritte (D) und (E) des *taggings* als unzweckmäßig. Auf sie wurde im Verlauf der Arbeit verzichtet (ab Sequenzierung T4). Stattdessen wurden die Proben bei der Aufreinigung nach dem *adapter fill-in* (C) in ca. 15 µl Elutionspuffer gelöst. Die Eluate wurden in ein einziges Reaktionsgefäß überführt (Gesamtvolumen ca. 100 – 150 µl), und dieses dann an das Sequenzierunternehmen geschickt.

8.2.6 DNA-Sequenzierung

Das Sequenzierunternehmen GATC BIOTECH AG in Konstanz führte zunächst eine Qualitätskontrolle des Probenpools durch. Anschließend wurde eine Amplikon-*library* erstellt, diese mittels Emulsions-PCR amplifiziert und schließlich auf dem ROCHE 454 Genome Sequencer FLX und FLX+ sequenziert (siehe hierzu Kapitel 7.2.1).

8.2.6.1 Direktsequenzierung

Einzel-PCRs mit anschließender Direktsequenzierung dienten dem Schließen von Lücken im Datensatz aufgrund von *loci*-Ausfällen bei der Multiplex-PCR und der 454-Sequenzierung. Tabelle 8.12 zeigt den Reaktionsansatz und den Programmablauf der Direktsequenzierungen, die in Thermocyclern [EPPENDORF] stattfanden.

Tabelle 8.12: Cycle Sequencing-Protokoll

Es wurde das BigDye® Terminator v3.1 Cycle Sequencing Kit mit zusätzlichem BigDye® Terminator 5x Sequencing Buffer [APPLIED BIOSYSTEMS INC.] verwendet. Als Sequenzier-Primer diente einer der beiden PCR-Primer. Die Hybridisierungstemperatur betrug 56 °C für mitochondriale, 58 °C für nukleäre Primer. Die Menge an eingesetztem PCR-Produkt richtete sich nach dessen Bandenintensität auf dem Agarose-Gelbild. n. best. = nicht bestimmt.

Reaktionsansatz:				CycleSeq-Parameter:		
Reagenzien	Volumen	Finale Konz.	Stamm-Konz.			
Puffer	1,5 µl	0,75 x	5 x	Denaturierung	92 °C	30 s
BigDye-Mix	1,0 µl	0,25 x	2,5 x	Hybridisierung	56-58 °C	15 s
Primer	1,0 µl	1,0 µM	10 µM	Elongation	60 °C	2 min 30 s
PCR Produkt	1,0-4,0 µl	n. best.	n. best.			
HPLC-H ₂ O	ad 10,0 µl			Zyklen		25

8.2.6.2 Aufreinigung der Sequenzierprodukte

Die Aufreinigung der Sequenzierprodukte erfolgte mit Sephadex™ G-50 Fine nach dem Prinzip einer Säulenausschlusschromatographie bzw. Gelfiltration. Zunächst wurde mithilfe eines *column loaders* je 45 mm³ des Dextrans in die *wells* einer *MultiScreen*-Platte gegeben und mit 300 µl HPLC-H₂O für mindestens 2,5 Stunden bei Raumtemperatur gequillt. Das überschüssige Wasser wurde für 3 min bei 2 750 min⁻¹ abzentrifugiert, wobei eine Mikrotiterplatte als Auffanggefäß diente. Zum Waschen der Säulen wurden erneut jeweils 150 µl HPLC-H₂O aufgetragen und der Zentrifugationsschritt wiederholt. In eine *PCR Detection Plate* (Sequenzierplatte) wurden 10 µl Formamid vorgelegt und diese unter der *MultiScreen*-Platte befestigt. Die Sequenzierprodukte (10 µl) wurden auf die Sephadex™-Säulen gegeben und für 5 min bei 2 750 min⁻¹ in die Sequenzierplatte zentrifugiert.

8.2.6.3 Kapillar-Gelelektrophorese

Die DNA-Sequenzen wurden durch eine Kapillar-Gelelektrophorese mit dem ABI PRISM® 3130 Genetic Analyzer [APPLIED BIOSYSTEMS INC.] unter Verwendung von Polymer POP-6™ als Trennmedium gewonnen. Die Injektionszeiten betragen zwischen 10 und 15 s, die Laufzeiten zwischen 15 und 25 min.

8.2.7 Kontrollen und Replikationen

Da es trotz Maßnahmen zur Kontaminationsvermeidung (siehe 8.2.1) zum Eintrag von Fremd-DNA in die Proben kommen kann, wurden bei jedem Bearbeitungsschritt kontinuierlich Leerkontrollen mitgeführt. Wie bereits erwähnt, wurde bei der Pulverisierung auch Hydroxylapatit gemahlen, und dann genauso wie das Pulver von Proben weiterbearbeitet. Bei den Extraktionsleerkontrollen befanden sich nur die Chemikalien für die Extraktion in den Reaktionsgefäßen. Neben den Hydroxylapatit- und den Extraktionskontrollen wurden PCR-Leerkontrollen, in denen anstatt DNA-Extrakt durch reines UV-bestrahltes HPLC-Wasser eingesetzt wurde, als Reaktionsansätze mitgeführt. Mit diesen Maßnahmen können Kontaminationen von Reagenzien und DNA-Einträge bei der Bearbeitung (Kreuz- und *carry-over*-Kontaminationen) erfasst werden.

DNA-Schäden und ins Knochenmaterial eingedrungene Fremd-DNA, z. B. von Mikroorganismen, die nicht entfernt werden konnte, kann Sequenziererergebnisse verfälschen. Daher wurden die Ergebnisse mehrfach reproduziert: Aus jeder Probe wurde zwei Mal unabhängig voneinander DNA extrahiert. Mit jedem Extrakt wurde mindestens eine mtDNA-Multiplex-PCR durchgeführt. Durch eine weitere Multiplex-PCR und/oder durch Einzel-PCRs wurde eine mindestens dreifache Abdeckung der HVR I-Region erzielt. Bei Sequenzunstimmigkeiten in der HVR I wurden zusätzliche PCRs durchgeführt. Für die Sequenzen der *coding region* wurde auf eine gezielte Ergebnisreproduktion verzichtet. Die nukleären *loci* wurden ebenfalls aus beiden Extrakten in insgesamt mindestens vier unabhängigen Multiplex-PCRs amplifiziert. Für die in Wilde *et al.* (2014) behandelten SNPs sowie die C-15 910T-Mutation in *LCT* (S13)

wurden zusätzliche PCRs direktsequenziert. Die Regeln zur Bestimmung der Haplo- bzw. Genotypen im Zuge der Sequenzanalyse sind in Kapitel 8.2.8.1 erklärt.

8.2.8 Datenanalyse

8.2.8.1 DNA-Sequenzanalyse und Reproduktion der Ergebnisse

Die Rohdatenpakete von GATC BIOTECH enthielten drei Dateien: die DNA-Sequenzen des 454 runs in einem *fasta*-Format (*.fna); die *base quality scores* (*.qual), d. h. für jedes sequenzierte Nukleotid ist die Verlässlichkeit des *base calls* durch einen Qualitätswert angegeben; und eine der *trace file* bei der Sanger-Sequenzierung analoge Datei im *standard flowgram format* (*.sff) zur Angabe von Signalintensitäten für jeden Nukleotid-*flow*.

Zunächst wurden die DNA-Sequenzen mit einem von Benjamin Rieger (Bioinformatiker am Institut für Molekulargenetik, Mainz) entwickelten *Perl*²⁷-Skript, „sort“, in den Versionen 1 bis 3, nach Meyer-*tags* und Primersequenzen sortiert und beschnitten. Das Programm benötigt die Datei mit den 454-Sequenzen (*.fna) und Text-Dateien (*.txt) mit den Barcode- und Primersequenzen. Aufgrund der hohen Fehlerraten der 454-Sequenzierung bei Homopolymeren, kollabiert das Programm beim sogenannten „*reduced scan*“ solche Abschnitte zu Einzelbasen in allen drei Dateien, bevor es nach Übereinstimmungen sucht. Es wurden jeweils die ersten 25 – 30 Basen der Sequenzen nach dem *tag* durchsucht, dieser dann abgeschnitten und erneut die ersten 25 – 30 Basen nach Primersequenzen gesucht. Diese konnten optional ebenfalls abgeschnitten werden. Für jeden *tag* wurde ein Ordner angelegt, in dem sich für jeden Primer eine separate *fasta*-Datei mit den entsprechenden Amplikons befindet.

Die primäre Auswertung der sortierten Sequenzen erfolgte mit dem Programm Seqman ProTM des Softwarepakets DNASTAR Lasergene®, Versionen 8 bis 10 [GATC BIOTECH AG und DNASTAR, Inc.]. Für jedes Primersystem wurde ein eigenes *contig*

²⁷ Eine kostenfreie, plattformunabhängige Programmiersprache; <http://www.perl.org/>

erstellt. Die Sequenzen wurden überprüft. Fehlerhaft einsortierte Sequenzen wurden mittels BLAST (NCBI - BLAST , Altschul *et al.* 1990) identifiziert und aus den *contigs* entfernt. Unvollständig beschnittene Sequenzen und artifizielle *gaps* (Lücken) die bei Homopolymeren entstehen und beim *assembling* aus algorithmischen Gründen in den einzelnen Sequenzen an unterschiedlichen Positionen eingefügt werden, wurden gegebenenfalls per Hand korrigiert. Die zu untersuchenden SNPs und etwaige weitere Mutationen wurden identifiziert und die Anzahl der Sequenzen (*reads*) pro Allel ermittelt.

Die *consensus*-Bildung der HVR I-Sequenzen fand in zwei Schritten statt. Zunächst wurde für jeden 454-Lauf eine *consensus*-Sequenz für jedes Amplikon, das durch mindestens drei *reads* repräsentiert war, erstellt. Dabei wurde für unstimmmige Positionen eine $\frac{2}{3}$ -Mehrheitsregel angewandt. Wurde diese nicht erfüllt, so wurde die Ambivalenz entsprechend des IUPAC-IUB-Codes (Nomenclature Committee of the International Union of Biochemistry 1985) bezeichnet. Produkte mit weniger als drei *reads* wurden nicht berücksichtigt, sondern als *dropout* betrachtet. Auch bei Sanger-Sequenzierungen kam für das *base calling* die $\frac{2}{3}$ -Regel bei Doppel-*peaks* zum Tragen. Für die HVR I wurden, wie in Kapitel 8.2.7 erwähnt, für jedes Fragment mindestens drei solcher Sequenzen aus unabhängigen PCRs ermittelt. Mit diesen überlappenden Abschnitten wurde für jede Probe eine endgültige *consensus*-Sequenz der HVR I von np 16 013 bis 16 409 erstellt, die im Nachhinein für Vergleiche mit anderen Datensätzen auf np 16 059 – 16 390 beschnitten wurde. Bei Diskrepanzen wurden weitere PCR-Produkte hinzugezogen. Ein Nukleotid an einer variablen Position galt als bestätigt, wenn es in einem Verhältnis von 3:1 oder höher auftrat. Konnte dies nicht erreicht werden, so wurde die Probe von der Analyse ausgeschlossen. Für die SNPs der *coding region* war bereits eine PCR ausreichend, da sie lediglich als Untermauerung der Authentizität der HVR I-Daten dienten.

Bei den 454-Sequenzierungen der autosomalen *loci* wurden aus technischen Gründen keine *consensus*-Sequenzen gebildet, sondern lediglich die Position des zu untersuchenden SNPs analysiert. Ein Allel galt als authentisches Produkt der entsprechenden PCR, wenn es durch mindestens 10 *reads* repräsentiert wurde. Allele mit einer geringeren Abdeckung wurden als Allel- oder *locus dropouts* behandelt.

Es wurde zwischen einer „sicheren“ und einer „vorbehaltlichen“ Authentifizierung unterschieden. Für eine sichere Authentifizierung von Homozygotie waren mindestens vier widerspruchsfreie Sequenzierungen aus beiden Extrakten notwendig. Ein Individuum galt demnach als sicher bzw. vorbehaltlich homozygot, wenn alle vier (teilweise fünf) bzw. drei PCRs das gleiche Allel (A) hervorbrachten. Falls eine einzelne PCR ein heterozygotes Ergebnis (AB) lieferte, mussten mindestens fünf PCRs, davon zwei vom gleichen Extrakt wie die abweichende PCR, homozygot für das gleiche Allel sein, d. h. ein Verhältnis von 11:1 zwischen authentischem (A) und „falschem“ Allel (B) bestehen. Außerdem musste die 454-Sequenzierung des heterozygoten PCR-Produkts mindestens doppelt so viele *reads* von Allel A wie Allel B aufweisen. Vorbehaltlich anerkannt wurde ein Genotyp bei drei widerspruchsfreien oder einem Allelverhältnis von 9:1, wobei die weiteren Regeln bzgl. Extrakten und *read*-Zahlen bestehen blieben, vorbehaltlich anerkannt. Ein Individuum galt als sicher / vorbehaltlich heterozygot für einen *locus*, wenn beide Allele in jeweils mindestens drei / zwei PCR-Produkten aus beiden Extrakten auftreten und mindestens ein Ergebnis heterozygot war. Tabelle 8.13 zeigt das Reproduktionsschema.

Tabelle 8.13: Mindestanforderung an die Authentifizierung von Genotypen

Jedes Buchstabenpaar steht für ein PCR-Produkt, wobei A und B die beiden Allele repräsentieren. *Bei dieser PCR musste die Anzahl der *reads* von Allel A mindestens doppelt so hoch sein wie von Allel B. Traf keiner der fünf bzw. vier Fälle der Mindestanforderungen für eine sichere bzw. vorbehaltliche Authentifizierung zu, so galt der *locus* als nicht genotypisiert.

Sichere Authentifizierung:			Vorbehaltliche Authentifizierung:		
Extrakt 1	Extrakt 2	Extrakt 1 oder 2	Extrakt 1	Extrakt 2	Extrakt 1 oder 2
AA	AA	AA, AA	AA	AA	AA
AB*, AA, AA	AA	AA, AA	AB*, AA, AA	AA	AA
↪ sicherer Genotyp: AA			↪ vorbehaltlicher Genotyp: AA		
AB	AB	AB	AB	AB	
AB	AB	AA, BB	AB	AA, BB	
AB	AA, BB	AA, BB			
↪ sicherer Genotyp: AB			↪ vorbehaltlicher Genotyp: AB		

8.2.8.2 Biostatistische Analysen

Die biostatistische Auswertung der mitochondrialen Sequenzdaten erfolgte mit dem Programm Arlequin 3.5.1.3 (Excoffier & Lischer 2010). Der Gesamtdatensatz wurde in zeitliche, kulturelle und räumliche Gruppen aufgeteilt. Diese sind im Ergebnisteil (Kapitel 9.3) aufgelistet. Die einzelnen Populationen wurden durch Bestimmung der Haplotypdiversität (\hat{H} , Nei 1987: 180), *mean number of pairwise differences* ($\hat{\pi}$, Tajima 1983), Nukleotiddiversität ($\hat{\pi}_n$, Tajima 1983), Tajima's D (Tajima 1989) und Fu's F_S (Fu 1997) charakterisiert. Zur Feststellung von Ähnlichkeiten bzw. der genetischen Nähe zwischen Populationen wurde der Anteil der paarweisen gemeinsamen Haplotypen (*shared haplotypes*) in Excel 2007 [MICROSOFT] ermittelt. Als Maß für die genetischen Distanzen wurden mit Arlequin paarweise F_{ST} -Werte (Wright 1951, Excoffier *et al.* 1992) basierend auf paarweisen Unterschieden berechnet. Die zugehörigen p -Werte wurden jeweils mit 10 000 Permutationen ermittelt. Das Signifikanzniveau wurde für die F_{ST} -Statistik und Tajima's D auf 5 % festgesetzt, bei Fu's F_S auf 2 %.

Für einen auf den F_{ST} -Werten zwischen allen prähistorischen Populationen Westeurasiens basierenden MDS-Plot wurde die Funktion „isoMDS“ aus dem Paket „MASS“ (Venables & Ripley 2002) der „R“-Sammlung verwendet und in der „R“-Umgebung Version 3.0.2 ausgeführt (R Core Team 2013). Die in dieser Arbeit generierten Genotypenfrequenzen wurden mit dem *HWExact*-Test aus dem ebenfalls zur „R“-Sammlung gehörenden Paket „HardyWeinberg“ (Graffelman 2013) bezüglich ihrer Hardy-Weinberg-Gleichgewichtszustandes überprüft. Die zugrundeliegenden mathematischen Formeln und Algorithmen der hier aufgeführten Statistiken und Tests sind in Kapitel 7.3 erläutert.

8.2.8.3 Vergleichsdatensätze

Um die genetischen Beziehungen der nordpontischen Steppengruppen des Äneolithikums und der Bronzezeit in den Kontext der mittel- und osteuropäischen Prähistorie einzuordnen, wurden eine Reihe von prähistorischen und modernen HVR I-Datensätzen in die biostatistischen Untersuchungen mit einbezogen (Tabelle 8.14). Darunter befinden sich zwei prä-neolithische (HGE, HGC) und zwei neolithische Jäger/Sammler-Populationen (NHG bzw. NHG2 und PWC). Da bei den Proben aus dem Fundort Ostorf Unklarheit besteht, ob es sich um Jäger/Sammler oder um Angehörige der Trichterbecherkultur (TBK bzw. FBC in Tabelle 8.14) handelt (Bramanti *et al.* 2009, Bollongino *et al.* 2013, Brandt *et al.* 2013), wurde die Population „neolithische Jäger/Sammler/Fischer“ doppelt betrachtet, als NHG (mit Ostorf) und als NHG2 (ohne Ostorf). Die Ostorf-Proben sind auch im Datensatz der FBC enthalten. Ein Großteil der neolithischen Sequenzen stammt aus dem Artikel von Brandt *et al.* (2013), der sich der Prähistorie Mitteleuropas vom Mesolithikum bis zur frühen Bronzezeit am Beispiel der Mittelelbe-Saale-Region widmet. Das Untersuchungsgebiet „nordpontischer Steppenraum“, wird lediglich durch sechs Individuen der Dnjepr-Donetz-Kultur (DDC) repräsentiert. Aus dem Raum östlich der nordpontischen Steppe wurden bisher nur etwa zeitgleiche und jüngere HVR I-Sequenzen veröffentlicht. Die Proben aus Westsibirien decken ungefähr den Zeitraum von 4 000 bis 900 v. Chr. ab. Genauere Informationen zu den Kulturen und Gruppen können den zugehörigen, in Tabelle 8.14 aufgeführten Publikationen entnommen werden.

Die Allelfrequenzen der untersuchten nukleären Marker in der heutigen Ukraine stammen aus der Magisterarbeit von Nina Hollfelder (2013), die eine Stichprobe von insgesamt 60 Individuen bearbeitete.

Tabelle 8.14: HVR I-Vergleichsdatensätze

Da ein Großteil der Sequenzen bereits in der umfangreichen Studie von Brandt *et al.* (2013) zusammengetragen wurden, wurden die dort verwendeten Abkürzungen (basierend auf den englischen Kulturnamen) übernommen und die aus anderen Populationen angepasst. Bei den zeitlichen Angaben handelt es sich um die groben Zeitrahmen aus denen die verwendeten Proben stammen, entnommen aus den jeweiligen Publikationen. N = Anzahl der Individuen.

Code	(N)	Kultur, Herkunft der Proben, ungefähres Alter der Proben (Publikationen)
HGE	(14)	prä-neolithische Jäger/Sammler , Westrussland, ca. 36 700 – 5 800 v. Chr. (Bramanti <i>et al.</i> 2009, Krause <i>et al.</i> 2010, Der Sarkissian <i>et al.</i> 2013)
HGC	(16)	prä-neolithische Jäger/Sammler , Mitteleuropa, ca. 31 250 – 6 250 v. Chr. (Bramanti <i>et al.</i> 2009, Fu <i>et al.</i> 2013)
NHG	(20)	neolithische Jäger/Sammler/Fischer , Mitteleuropa (inklusive Ostorf), ca. 4 500 – 2 100 v. Chr. (Bramanti <i>et al.</i> 2009, Bollongino <i>et al.</i> 2013)
NHG2	(13)	neolithische Jäger/Sammler/Fischer , Mitteleuropa, ca. 4 500 – 2 100 v. Chr. (exklusive Ostorf => siehe FBC) (Bramanti <i>et al.</i> 2009, Bollongino <i>et al.</i> 2013)
PWC	(19)	Grübchenkeramische Kultur (neolithische Jäger/Sammler), Südschweden, ca. 3 300 – 2 150 v. Chr. (Malmström <i>et al.</i> 2009)
LBK	(102)	Linearbandkeramische Kultur , Mitteleuropa, ca. 5 500 – 4 775 v. Chr. (Haak <i>et al.</i> 2005, Haak 2006, Haak <i>et al.</i> 2010, Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
RSC	(11)	Rössener Kultur , Mittelbe-Saale, ca. 4 625 – 4 250 v. Chr. (Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
SCG	(33)	Schöninger Gruppe , Mittelbe-Saale-Region, ca. 4 100 – 3 950 v. Chr. (Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
LNF	(11)	Spätneolithische Bauern , Blätterhöhle (Sauerland), ca. 4 000 – 2 900 v. Chr. (Bollongino <i>et al.</i> 2013)
BAC	(19)	Baalberger Kultur , Mittelbe-Saale-Region, ca. 3 950 – 3 400 v. Chr. (Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
SMC	(29)	Salzmünder Kultur , Mittelbe-Saale-Region, ca. 3 400 – 3 025 v. Chr. (Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
BEC	(17)	Bernburger Kultur , Mittelbe-Saale-Region, ca. 3 100 – 2 650 v. Chr. (Brandt <i>et al.</i> 2013)
FBC	(10)	Trichterbecherkultur (TBK), Ostorf, Südschweden, ca. 3 400 – 2 900 v. Chr. (Bramanti <i>et al.</i> 2009, Malmström <i>et al.</i> 2009)
CWC	(44)	Schnurkeramische Kultur , Mittelbe-Saale-Region, ca. 2 800 – 2 200 v. Chr. (Haak <i>et al.</i> 2008, Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
BBC	(35)	Glockenbecherkultur , Mittelbe-Saale-Region, ca. 2 500 – 2 050 v. Chr. (Lee <i>et al.</i> 2012, Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)

Code	(N)	Kultur, Herkunft der Proben, ungefähres Alter der Proben (Publikationen)
UNC	(94)	Aunjetitzer Kultur , Mittelbe-Saale-Region, ca. 2 200 – 1 550 v. Chr. (Brandt <i>et al.</i> 2013, Brotherton <i>et al.</i> 2013)
DDC	(6)	Dnjepr-Donetz-Kultur , nordpontische Steppe, ca. 5 600 – 4 800 v. Chr. (Nikitin <i>et al.</i> 2012)
UTC	(17)	Ust-Tartas-Kultur , Westsibirisches Tiefland, ca. 4 000 – 2 500 v. Chr. (Molodin <i>et al.</i> 2012)
ODI	(10)	Odinovo-Kultur , Westsibirisches Tiefland, ca. 3 000 – 2 500 v. Chr. (Molodin <i>et al.</i> 2012)
KRO	(26)	Krotovo-Kultur , Westsibirisches Tiefland, ca. 2 250 – 1 750 v. Chr. (Molodin <i>et al.</i> 2012)
ANF	(20)	Andronovo-Kultur (Fedorovo-Stufe), Westsibirisches Tiefland, ca. 2 000 – 1 500 v. Chr. (Molodin <i>et al.</i> 2012)
AND	(9)	Andronovo-Kultur , Westsibirisches Tiefland, ca. 1 800 – 1 400 v. Chr. (Keyser <i>et al.</i> 2009)
BAK	(8)	Bronzezeit Kasachstan , Kasachstan, ca. 1 400 – 900 v. Chr. (Lalueza-Fox <i>et al.</i> 2004)
mBul	(30)	Osteuropäer , Bulgarien, heute (Calafell <i>et al.</i> 1996)
mUkr	(18)	Osteuropäer , Ukraine, heute (Malyarchuk & Derenko 2001)
mRus	(198)	Osteuropäer , Westrussland, heute (Malyarchuk <i>et al.</i> 2002)
mGer	(262)	Mitteleuropäer , Deutschland, heute (Bramanti <i>et al.</i> 2009)
mLit	(97)	Nord-/Mitteleuropäer , Litauen, heute (Bramanti <i>et al.</i> 2009)
mPol	(75)	Mitteleuropäer , Polen, heute (Bramanti <i>et al.</i> 2009)

9 Ergebnisse

Insgesamt stand Skelettmaterial von 180 menschlichen und einer Tierprobe (OVI 5) zur Kontaminationskontrolle aus 39 Fundorten für die molekulargenetische Untersuchung zur Verfügung. Nach der äußerlichen Begutachtung und aufgrund von Desintegration während der Probenvorbereitung wurden 17 Proben nicht weiter bearbeitet. Bei drei Proben, BEN 4, MAJ 1 und MAJ 6 stellte sich im Verlauf der Untersuchungen heraus, dass sie nicht in den zeitlichen Kontext dieser Studie gehören. MAJ 1 und MAJ 6 sind wahrscheinlich in die Späteisenzeit einzuordnen. Sie wurden von den weiteren Analysen ausgeschlossen. In die folgenden Evaluationen der analytischen Methoden und der Probenerhaltung gingen 160 humane Proben von 38 Fundorten ein (siehe Abbildung 8.1 und Tabelle 8.1).

Die Auswertung des Datenmaterials fand auf drei Ebenen statt: die reine Bewertung der Methoden durch Erfassung der Amplifikations- und Sequenzierungserfolge, ohne Berücksichtigung der Reproduzierbarkeit; der Reproduktionserfolg und die Erhaltung der Proben; und die inhaltliche Analyse.

9.1 Methodische Ergebnisse

9.1.1 Amplifikationserfolg der mitochondrialen Multiplex-PCR

Bei der Evaluation des Amplifikationserfolges der mitochondrialen Multiplex-PCR wird zwischen dem theoretischen Amplifikationserfolg (basierend auf **allen** PCR-Reaktionen), dem praktischen Amplifikationserfolg (basierend auf **allen sequenzierten** PCR-Produkten), im Folgenden mit (°) gekennzeichnet, und dem Sequenziererfolg, der auf der Anzahl der *reads* pro *locus* beruht, unterschieden. Im Prinzip ist der Amplifikationserfolg^s eine Kombination aus Amplifikations- und Sequenziererfolg, wobei die Anzahl der *reads* unberücksichtigt bleibt.

Von 1104 (368 Proben à 3 MP-Ansätze) Reaktionsansätzen wurden 699 (233 x 3) 454-sequenziert. Bei den ersten Testläufen wurde beobachtet, dass bei geringen Mengen an DNA die 454-Sequenzierung oft ergebnislos blieb. Daher wurden nicht alle positiven PCR-Produkte sequenziert, sondern zunächst die Erfolgchancen anhand der Bandenintensitäten auf dem Agarosegel abgewogen. Einer der 454 runs (T4) mit 66 (22 x 3) PCR-Produkten erbrachte keine Ergebnisse, und 30 (10 x 3) PCR-Produkte wurden für einen Illumina-Testlauf verwendet. Diese 96 Ansätze sind in der Gruppe der 505 nicht sequenzierten Proben enthalten.

Der Amplifikationserfolg der mitochondrialen Multiplex-PCR wurde wie folgt ermittelt: Bei den 454-sequenzierten Produkten wurde jeder *locus* als binäre Variable dargestellt: 1 = mindestens 3 *reads*, d. h. erfolgreich amplifiziert; 0 = 0 bis 2 *reads*, d. h. nicht amplifiziert. Die Anzahl der Amplikons, die mindestens drei *reads* aufwiesen, wurde durch die Anzahl der Primersysteme in der PCR-Reaktion (13 für MPA, 12 für MPB und C) dividiert. Die tatsächliche Anzahl von *reads* blieb also unberücksichtigt. Nicht sequenzierte PCRs wurden anhand ihrer Agarosegelbilder bewertet. Je nach Bandenintensität und -vollständigkeit erhielt eine PCR-Reaktion zwischen 0 und 3 Punkten, wobei 0 für keine und 3 für eine starke Bandenmenge steht. Der theoretische bzw. zu erwartende Amplifikationserfolg dieser nicht sequenzierten Produkte ergab sich durch eine Division durch 3, der maximal erreichbaren Punktzahl. Damit lag der allgemeine Amplifikationserfolg gemittelt über alle (sequenzierten und nichtsequenzierten) PCRs und unabhängig von der inhaltlichen Analyse bei 57,11 %.

Die Beurteilung der einzelnen Primersysteme beruhte auf den 233 sequenzierten PCR-Produkten je Multiplex-PCR-Ansatz. Die Amplifikationserfolge^S sind in Abbildung 9.1 aufgetragen. Der Gesamterfolg der sequenzierten PCRs lag bei 68,62 %. Dabei sind Multiplex-PCR-Ansätze A und B mit durchschnittlichen Erfolgen^S von 70,19 % und 72,75 % ähnlich effizient. Ansatz C hat mit 62,80 % den niedrigsten Erfolg^S (siehe Tabelle 9.1). Dies lag vor allem an dem Primerpaar M18, das komplett ausfiel. Würde man diesen *locus* bei der Erfolgsanalyse auslassen, läge die durchschnittliche Amplifikationsrate^S in Ansatz C bei 68,47 %. Sieben der insgesamt 37 Primersysteme wiesen Raten^S von über 90 % auf. Von den sieben Primerpaaren der HVR I (M31 – M37) lagen alle außer M36 deutlich unter dem Durchschnitt. M32 schnitt von allen

loci, mit Ausnahme von M18, mit 35,19 % am schlechtesten ab. Die Korrelationskoeffizienten²⁸ zwischen dem Amplifikationserfolg⁵ der einzelnen Primersysteme und der Fragmentlänge bzw. dem GC-Gehalt der Amplikons betragen -0,521 und 0,408.

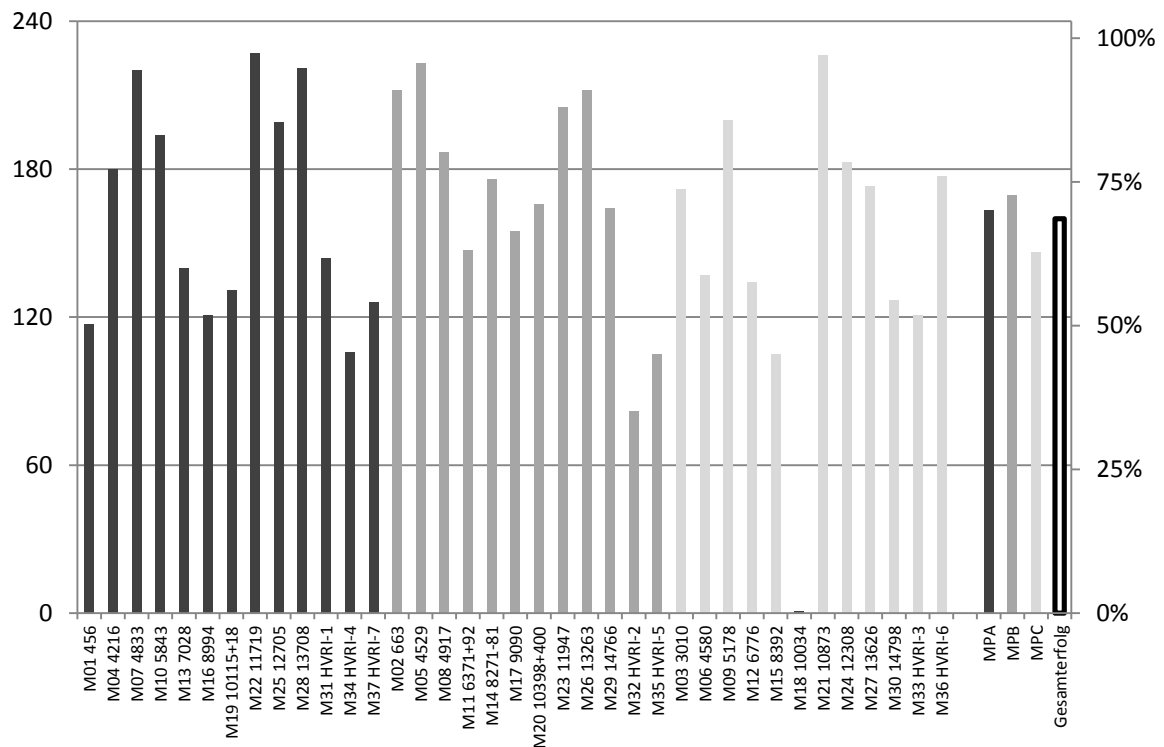


Abbildung 9.1: Amplifikationserfolg⁵ der mitochondrialen Multiplex-PCR

Die Werte beruhen rein auf den 233 Tripel-Ansätzen, die sequenziert wurden. Die linke Ordinatenachse gibt die Anzahl von PCR-Produkten an, aus denen der entsprechende *locus* erfolgreich sequenziert wurde. Auf der rechten Achse ist der entsprechende prozentuale Erfolg⁵ skaliert. Die unterschiedlichen Grautöne repräsentieren die drei Reaktionsansätze der Multiplex-PCR. Die vier Balken rechts geben den durchschnittlichen Erfolg⁵ der einzelnen Ansätze und den Gesamterfolg⁵ der mitochondrialen Multiplex-PCR aller drei Ansätze an.

²⁸ Korrelationskoeffizient $r = \text{Korrel}(A, X) = \frac{\sum(a - \bar{a})(x - \bar{x})}{\sqrt{\sum(a - \bar{a})^2 \sum(x - \bar{x})^2}}$ (aus Microsoft Excel 2007), mit a = Amplifikationserfolg eines Primersystems und x = Länge des Amplikons bzw. x = GC-Gehalt des Amplikons, \bar{a} bzw. \bar{x} = Durchschnittswerte. r kann Werte zwischen -1 und +1 annehmen, die totaler negativer bzw. positiver Korrelation entsprechen. $r = 0$ bedeutet, dass kein Zusammenhang zwischen zwei Merkmalen (A, X) besteht.

Tabelle 9.1: Amplifikationserfolge^S der einzelnen Multiplex-PCR-Ansätze

Berücksichtigt wurden nur sequenzierte PCR-Ansätze. Ein *locus* gilt als amplifiziert, wenn er mindestens 3 (mtDNA) bzw. 10 (ncDNA) *reads* aufweist.

Multiplex-PCR Ansatz	Anzahl Primersysteme (N)	Anzahl amplifizierter <i>loci</i> /PCR ($\bar{X}_n \pm \sigma_n$)	Amplifikations-Erfolgsrate ^S $E(X) \pm \sigma_x$ ($X = n/N$)
mtDNA - MPA	13	9,1 ± 2,5	70,19 % ± 19,06 %
mtDNA - MPB	12	8,7 ± 2,2	72,75 % ± 18,60 %
mtDNA - MPC	12	7,5 ± 3,0	62,80 % ± 24,88 %
mtDNA gesamt	37	25,4 ± 7,7	68,62 % ± 20,83 %
ncDNA	23 (21)	13,9 ± 3,0	60,54 % ± 13,14 %

Für die Ermittlung des Amplifikationserfolgs^S war lediglich das Vorhandensein von Sequenzen (≥ 3 *reads/locus*) maßgebend. Als Sequenziererfolg sei die durchschnittliche Anzahl von *reads* pro *locus* bezeichnet. Die Gesamtzahlen von *reads* der einzelnen Multiplex-PCRs schwankten zwischen 16 und 19117. Die Streuweite zwischen den einzelnen *loci* innerhalb einer PCR betrug in der Hälfte aller Sequenzierungen über 150, in einigen über 1000 *reads* (Maximum: 3540 *reads*). Die durchschnittlichen Sequenziererfolge sind in Tabelle 9.2 bzw. nach *loci* aufgetrennt in Abbildung 9.2 dargestellt. Aufgrund der großen Spannweite der *read*-Zahlen wurde zur Normierung der Streuung der Variationskoeffizient als relative Standardabweichung berechnet. Er betrug für alle *loci* (ausgenommen M18) zwischen 1,3 und 3,5.

Tabelle 9.2: Sequenziererfolge der einzelnen Multiplex-PCR-Ansätze

\bar{X} = durchschnittliche Anzahl an *reads* je PCR-Ansatz; X = *coverage* eines *locus* in einer PCR. $E(X)$ = Erfolgsrate (*reads/locus/PCR*); σ = Standardabweichung; $\text{VarK}(X)$ = Variationskoeffizient; min/max = Mindest-/maximale *coverage*; $Q_{0,25/0,5/0,75}$ = Anzahl von *reads* mit denen 25 / 50 / 75 % der *loci* höchstens abgedeckt sind (z. B. in MPA haben 50 % der *loci* eine *coverage* von ≤ 12 *reads*).

Multiplex Ansatz	<i>reads</i> gesamt	\bar{X} <i>reads/PCR</i>	$E(X)$	σ_x	$\text{VarK}(X)$	Streuweite $\text{min}(X)$ - $\text{max}(X)$	$Q_{0,25} - Q_{0,5} - Q_{0,75}$ (X)
mt - MPA	178 453	765,9	58,9	147,8	2,5	0-3540	1 - 12 - 54
mt - MPB	200 410	860,1	71,7	152,7	2,1	0-1910	2 - 16 - 69
mt - MPC	161 880	694,8	57,9	139,5	2,4	0-1964	0 - 9 - 49
mt gesamt	540 743	2 320,8	62,7	146,9	2,3	0-3540	1 - 12 - 57
ncDNA	823 200	2 487,0	109,4	250,8	2,3	0-4116	0 - 23 - 105

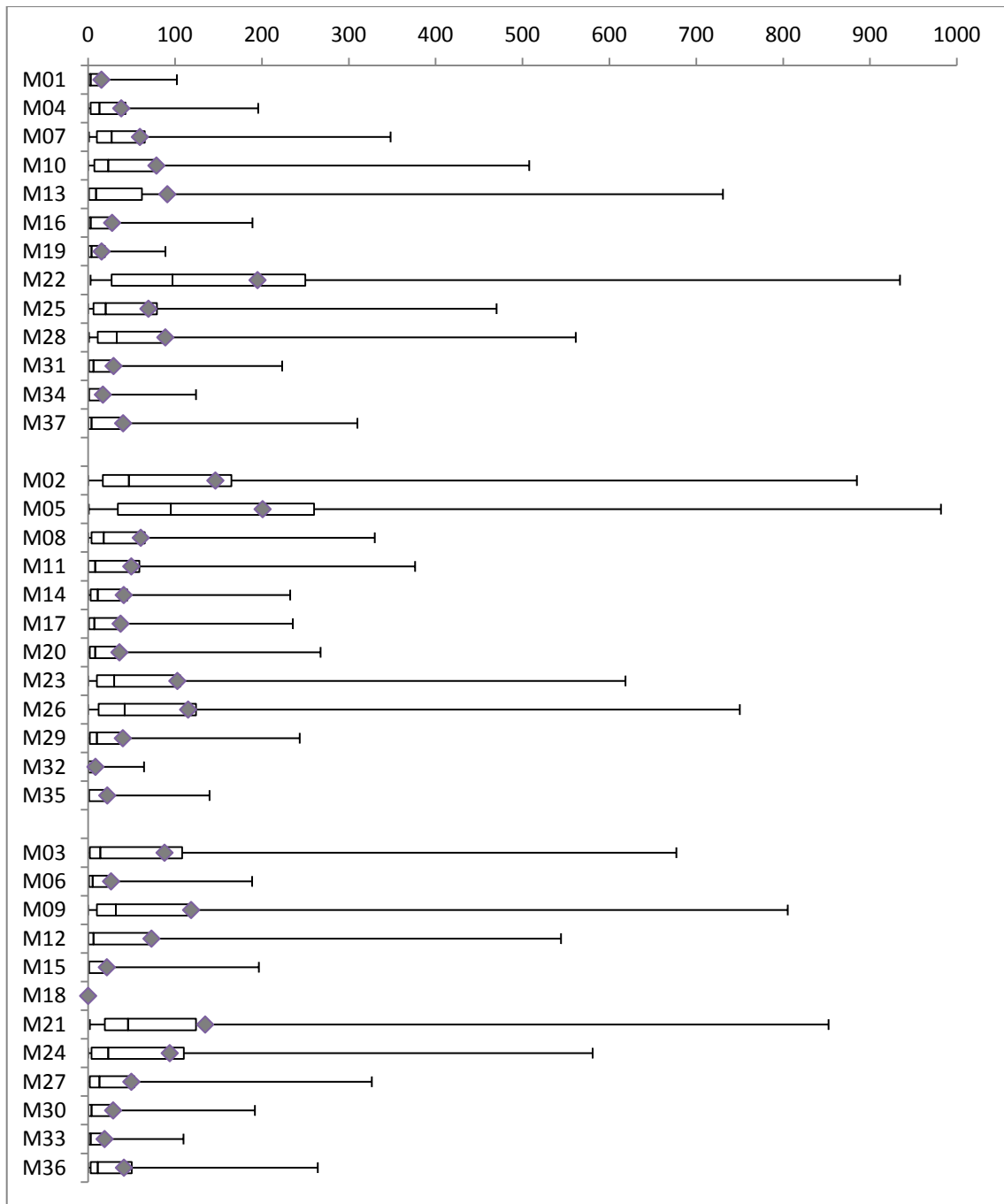


Abbildung 9.2: Boxplot der reads pro locus in der mitochondrialen Multiplex-PCR

Die *loci* sind entsprechend ihrer Zugehörigkeit in die Reaktionsansätze A, B und C sortiert. Das Rechteck, unterteilt durch den Median, stellt den Bereich vom unteren bis oberen Quartil der jeweiligen *coverage* dar, und enthält damit die mittleren 50 % der Daten. Die Antennen geben das 2,5 %- und das 97,5 %-Quantil an. Der Mittelwert ist durch eine graue Raute repräsentiert.

9.1.2 Amplifikationserfolg der nukleären Multiplex-PCR

Bei der Evaluation des Amplifikationserfolges der nukleären Multiplex-PCR wird zwischen dem theoretischen Amplifikationserfolg (basierend auf **allen** PCR-Reaktionen), dem praktischen Amplifikationserfolg (basierend auf **allen sequenzierten** PCR-Produkten), im Folgenden mit (^S) gekennzeichnet, und dem Sequenziererfolg, der auf der Anzahl der *reads* pro *locus* beruht, unterschieden. Im Prinzip ist der Amplifikationserfolg^S eine Kombination aus Amplifikations- und Sequenziererfolg, wobei die Anzahl der *reads* unberücksichtigt bleibt. Die Abschätzung der Amplifikationsrate basierend auf unterschiedlichen Grundlagen trägt dem Umstand Rechnung, dass nicht alle positiven PCR-Produkte sequenziert wurden.

331 der 704 nukleären Multiplex-PCR Reaktionen wurden mit der 454-Technologie sequenziert. Zu den 373 nicht sequenzierten Produkten zählen Proben, die keine oder nur sehr schwache Banden auf dem Agarosegel aufwiesen und 22 Proben eines Sequenzierlaufs (T4), der keine Ergebnisse lieferte.

Zur Evaluierung des nukleären Amplifikationserfolges wurde bei den 454-sequenzierten Produkten jeder *locus* in einer PCR als binäre Variable dargestellt (1 = mindestens 10 *reads*, d. h. erfolgreich amplifiziert; 0 = 0 bis 9 *reads*, d. h. Amplifikation nicht erfolgreich). Die Zahl der Amplikons, die eine Abdeckung von mindestens 10 *reads* hatten, wurde bestimmt und durch die Anzahl der Primersysteme im Reaktionsansatz dividiert (23 bzw. 21 *loci*). Die tatsächliche Anzahl an *reads* oberhalb dieses Schwellenwertes blieb unberücksichtigt. Nicht sequenzierte PCR-Produkte wurden anhand ihrer Agarosegelbilder bewertet. Sie erhielten zwischen 0 (keine Banden) und 5 Punkte (starke Banden). Geteilt durch die maximal erreichbare Punktzahl 5 ergab dies den gelbasierten theoretischen Amplifikationserfolg. Unter Berücksichtigung aller sequenzierter und nicht-sequenzierter PCRs lag der mittlere Amplifikationserfolg, unabhängig von der inhaltlichen Analyse, bei 40,83 %.

Betrachtet man rein die 331 sequenzierten PCR-Produkte, liegt der Amplifikationserfolg^S bei 60,43 %. Die Verteilung auf die einzelnen *loci* zeigt Abbildung 9.3. Die *loci* *ABCC11* und *NAT2* wurden in 256 PCRs und damit am häufigsten erfolgreich amplifiziert und sequenziert. Bei *SLC45A2* handelte es sich um einen

negativen Ausreißer mit einer Erfolgsrate von 16,29 %. Bei fast allen *loci* lagen die Raten zwischen 50 und 70 %. Die Korrelation²⁹ zwischen Amplifikationserfolg^S und Fragmentlänge ($r = -0,280$) bzw. GC-Gehalt ($r = 0,282$) ist nur schwach.

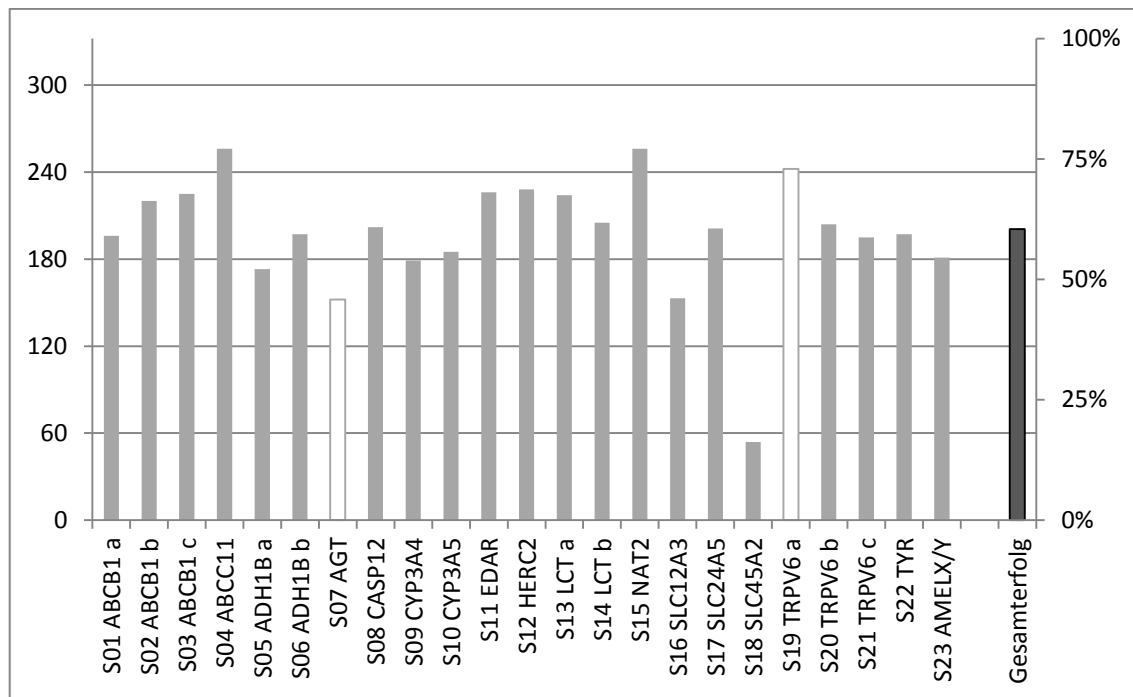


Abbildung 9.3: Amplifikationserfolg^S der nukleären Multiplex-PCR

Die Werte beruhen rein auf den 331 Ansätzen, die sequenziert wurden. Die linke Ordinatenachse gibt die Anzahl von PCR-Produkten an, aus denen der entsprechende *locus* erfolgreich sequenziert wurde. Auf der rechten Achse ist der entsprechende prozentuale Erfolg^S skaliert. Die weißen Balken geben die beiden *loci* an, die nicht in die populationsgenetische Untersuchung einfließen.

Bezüglich des Sequenziererfolgs war, wie auch bei der mtDNA, die Streuung der *read*-Zahlen sehr groß (0 bis 4116). Durchschnittlich brachte die Sequenzierung eines Multiplex-PCR-Ansatzes 109 *reads* pro *locus* hervor (Tabelle 9.2). Abbildung 9.4 zeigt die Schwankungen sowohl zwischen den einzelnen Fragmenten als auch zwischen verschiedenen PCRs.

²⁹ Korrelationskoeffizient $r = \text{Korrel}(A, X) = \frac{\sum(a - \bar{a})(x - \bar{x})}{\sqrt{\sum(a - \bar{a})^2 \sum(x - \bar{x})^2}}$ (aus Microsoft Excel 2007), mit a = Amplifikationserfolg eines Primersystems und x = Länge des Amplikons bzw. x = GC-Gehalt des Amplikons, \bar{a} bzw. \bar{x} = Durchschnittswerte.

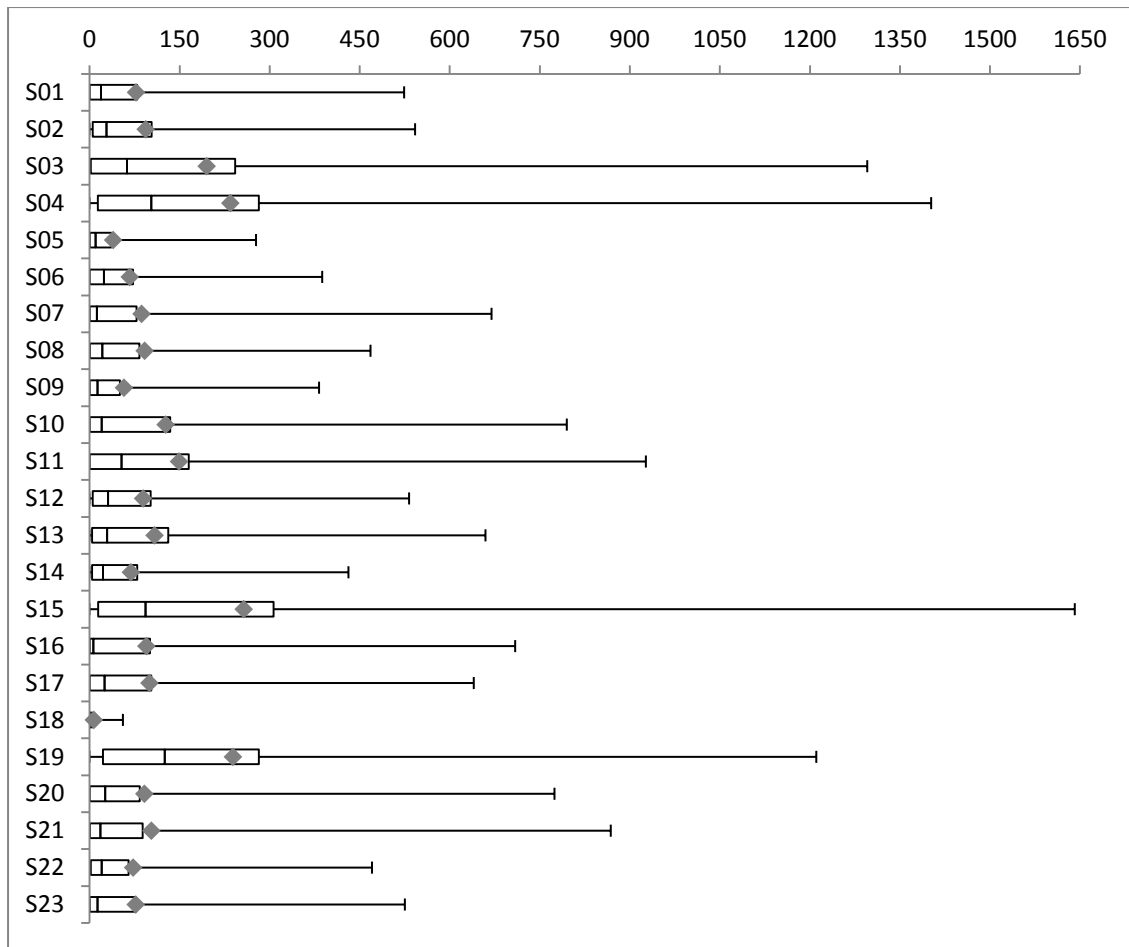


Abbildung 9.4: Boxplot der reads pro locus in der nukleären Multiplex-PCR

Das Rechteck, unterteilt durch den Median, stellt den Bereich vom unteren bis oberen Quartil der *coverage* dar, und enthält damit die mittleren 50 % der Daten. Die Antennen geben das 2,5 % und das 97,5 % Quantil an. Der Mittelwert ist durch eine graue Raute repräsentiert.

9.1.3 Amplifikationserfolg der Einzel-PCR

Aufgrund der hohen Ausfallrate der HVR I-Fragmente in der Multiplex-PCR wurden für 110 Proben zusätzliche Sequenzen durch Einzel-PCRs gewonnen. 50 Proben wurden bereits zuvor aufgrund mangelnder Erfolge aussortiert. Verwendet wurden die sieben HVR I-Primersysteme, die auch in der Multiplex-PCR zum Einsatz kamen, sowie in geringem Umfang vier Paare, die die HVR I in vier etwas längeren Fragmenten amplifizieren. Letztere wurden ausschließlich mit der Methode nach Sanger sequenziert, erstere zum Teil auch durch 454-Sequenzierung. Dabei wurden die sieben

PCR-Produkte zusammengefügt (*pooling*) und einem *tag* zugeordnet. Als erfolgreich amplifiziert galten Produkte, die mindestens drei *reads* aufwiesen oder bei der Sanger Sequenzierung eine Sequenz ergaben.

Die durchschnittliche Amplifikationsrate in den 1486 mitochondrialen Einzel-PCR-Reaktionen lag bei 65,01 %. Von 403 *next-generation*-Sequenzierungen waren 331 erfolgreich (82,13 %), bei den Sanger-Sequenzierungen 570 von 641 (88,92 %). 65 der 442 nicht sequenzierten PCR-Produkte zeigten eine schwache Bande auf dem Agarosegel (14,71 %).

Von den nukleären Fragmenten wurden S12 (*HERC2*), S18 (*SLC45A2*) und S22 (*TYR*) für eine sichere Authentifizierung für Wilde *et al.* (2014) in Einzel-PCRs amplifiziert und direktsequenziert. Auch für S13 (*LCT*) wurden auf diese Weise die Reproduktion einiger Individuen vervollständigt. *HERC2* wurde aus 51 von 98 PCR-Reaktionen, *LCT* aus 16 von 31, *SLC45A2* aus 90 von 141, und *TYR* aus 58 von 126 Reaktionen erfolgreich sequenziert. Die eigentliche Amplifikationsrate (nach Gelbildern) war höher, doch die sehr kurzen Fragmente führten zu Verlusten bei der Aufreinigung der PCR-Produkte und Problemen bei der Kapillar-Gelelektrophorese. Abbildung 9.5 zeigt die Amplifikationserfolge der einzelnen HVR I- und der nukleären Primersysteme.

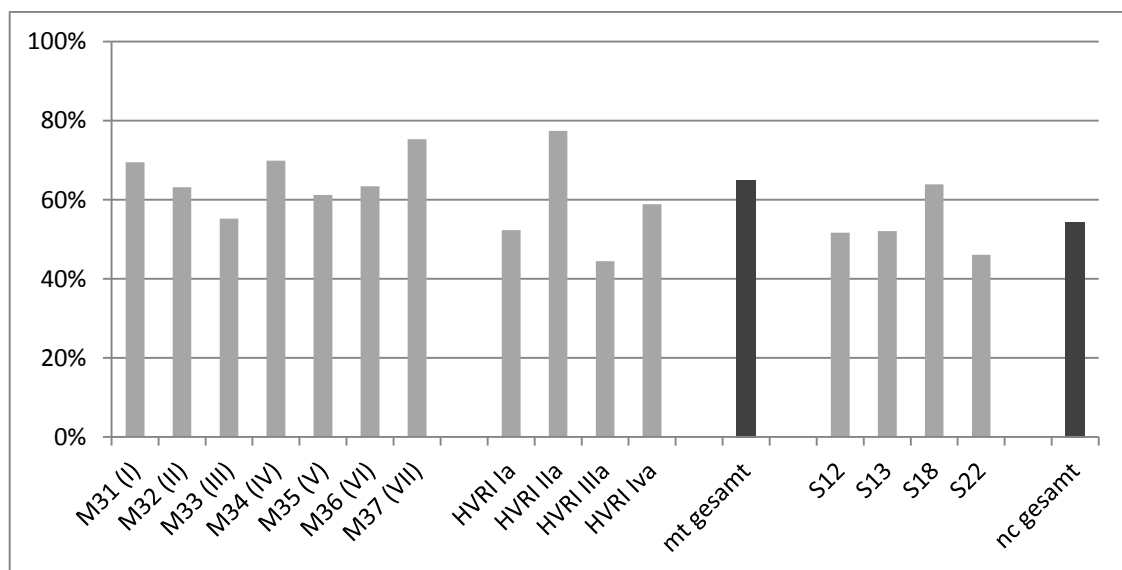


Abbildung 9.5: Amplifikationserfolg der Einzel-PCRs

9.1.4 Kontaminationen in den Negativkontrollen

Im Rahmen dieser Dissertation wurden 644 Negativkontrollen parallel zu den Proben amplifiziert. Die 23 Hydroxylapatit- und 48 Extraktionskontrollen wurden sowohl mit der mitochondrialen als auch der nukleären Multiplex-PCR auf das Vorhandensein von DNA überprüft. 110 der 139 Reaktionsansätze, die auf dem Agarose-Gel nicht absolut eindeutig leer erschienen, wurden sequenziert. Die Kontaminationsraten für die Multiplex-PCRs wurden wie folgt berechnet: Basierend auf der Anzahl an Fragmenten in den Leerkontrollen, die bei der Sequenzierung *reads* auswiesen, wurde eine Hochrechnung auf die erwartete Anzahl solcher Fragmente in allen Kontrollen, die auf dem Agarosegel nicht eindeutig leer erschienen, durchgeführt. Die Kontaminationsrate wurde als der Anteil dieser Fragmente an der Gesamtanzahl amplifizierter Kontroll-*loci* beschrieben. Für mitochondriale DNA lag sie bei 4,56 %, nukleär bei 2,23 % (Tabelle 9.3). Berücksichtigt man bei den mitochondrialen Kontrollen lediglich *loci*, die mindestens 3 *reads* und bei den nukleären Kontrollen nur solche, die mindestens 10 *reads* aufwiesen, dann reduzieren sich die Kontaminationsraten auf 2,40 % bzw. 0,35 %³⁰.

Bei Negativkontrollen, aus denen erfolgreich mitochondriale DNA amplifiziert werden konnte, wurden die auftretenden Mutationen mit denen der Proben verglichen, die am gleichen Tag gemahlen, extrahiert und/oder in der gleichen PCR amplifiziert wurden. In mehreren Kontrollen trat in Fragment M03 der Haplotyp 3009T 3019A 3027C auf. Diese Austausche wurden in keiner Probe beobachtet. In den anderen Fällen, in denen es Übereinstimmungen einzelner Mutationen gab, konnte die Authentizität der Proben durch Diskrepanzen an weiteren Positionen und durch die Konsistenz bei der Reproduktion belegt werden. Aus der Tierprobe (OVI 5 TK) wurde mit den humanspezifischen Primersystemen keine DNA amplifiziert.

³⁰ Die minimalen Abweichung zu den Angaben in Wilde et al. (2014) ergeben sich durch drei 454-Sequenzierungs-Produkten, die erst nach der Publikation dieses Artikel ausgewertet wurden.

Tabelle 9.3: Negativkontrollen und Kontaminationen

n_p = Gesamtanzahl an Kontroll-PCR-Reaktionen; n_L = Anzahl an *loci* (Primersystemen) in den Kontrollen; k_p = Kontroll-PCR-Produkte unter Kontaminationsverdacht, d. h. Agarose-Gel nicht definitiv „leer“; k_{ps} = k_p , die sequenziert wurden; k_{LS} = Anzahl der *loci* in allen k_{ps} , die mind. einen 454 *read* oder eine Sanger-Sequenz lieferten; k_L = erwarteter Gesamtzahl kontaminierter *loci* in allen k_p basierend auf den k_{LS} in den k_{ps} . HAP = Hydroxylapatitpulver (Negativkontrolle beim Mahlen der Proben), ExLK = Extraktionsleerkontrolle, PCR LK = PCR-Leerkontrolle. *Kursiv in Klammer die Kontaminationsrate, falls nur *loci* mit mindestens 3 (mtDNA) bzw. 10 (ncDNA) reads gezählt werden.

Kontrollen	PCR-Ansätze n_p	<i>Loci</i> n_L	Kontaminierte(?) PCRs k_p (k_{ps})	Kontaminierte <i>loci</i> k_L (k_{LS})	Kontaminations- rate k_L / n_L (k_L / n_L)*
Mitochondriale Multiplex-PCR (454)					
HAP	102	1258	33 (27)	89 (73)	7,07 % (3,49 %)
ExLK	150	1850	10 (10)	16 (16)	0,86 % (0,76 %)
PCR LK	96	1184	41 (33)	91 (73)	7,67 % (3,78 %)
Gesamt	348	4292	84 (70)	194 (162)	4,56 % (2,40 %)
Nukleäre Multiplex-PCR (454)					
HAP	26	596	8 (6)	13 (10)	2,24 % (0,67 %)
ExLK	52	1196	19 (15)	27 (21)	2,22 % (0,32 %)
PCR LK	67	1529	21 (16)	34 (26)	2,23 % (0,26 %)
Gesamt	145	3321	48 (37)	74 (57)	2,23 % (0,35 %)
Einzel-PCR (Sanger)					
PCR LK (mt)	115	115	7 (4)	7 (2)	6,09 %
PCR LK (nc)	36	36	0 (0)	0 (0)	0,00 %

In den nukleären Negativkontrollen traten fast ausschließlich vereinzelte Sequenzen bei der 454-Sequenzierung auf. Neun *loci* brachten dabei mehr als neun *reads* hervor. Zweimal betraf dies das Fragment S19 (*TRPV6* a), welches von der Auswertung ausgeschlossen wurde. Eine Extraktionskontrolle wies für S22 (*TYR*) das abgeleitete A-Allel auf. Alle Individuen, die in dieser Extraktion bearbeitet wurden, tragen das C-Allel. Dies gilt auch für die entsprechende PCR, mit der Ausnahme von RIL 3, welches konsistent in dieser und anderen PCRs heterozygot für diesen SNP ist. In einer PCR-Leerkontrolle wurde für diesen *locus* das C-Allel amplifiziert. Alle Proben dieser PCR zeigten dieses Allel noch in mindestens vier weiteren unabhängigen Reaktionen. In einer anderen PCR-Leerkontrolle wurde das A-Allel des zweiten SNPs von *ADH1B* (S06) amplifiziert, während alle Individuen dieser PCR homozygot für das G-Allel sind. Das Primersystem S15 (*NAT2*) amplifizierte in einer Kontrolle das

tatsächliche *target*, in einer weiteren eine Region des *NAT1*-Gens. Jeweils einmal von Kontaminationen betroffen sind die *loci* S01 (*ABCB1 a*) und S09 (*CYP3A4*). Die parallel bearbeiteten Proben konnten entweder durch Reproduktion oder durch *locus dropout* oder durch Vorhandensein des alternativen Allels für S01 authentifiziert werden. Bei *CYP3A4* wies eine PCR-Kontrolle das G-Allel auf, welches im Gesamtdatensatz fixiert ist.

Im Post-PCR-Bereich wurden dem ersten Anschein nach ebenfalls Kontaminationsphänomene beobachtet. Bei den 454 *runs* T5 bis T15 wurde bei der Sortierung der Rohdaten mithilfe des „sort3“-Skripts nach allen 60 Meyer-*tags* gesucht, auch denen, die in der entsprechenden Sequenzierung nicht zum Einsatz kamen. Mitochondrial wurden 31 von 167 (18,56 %) nicht verwendeten *tags reads* zugeordnet. Nukleär waren es 48 von 200 (24 %). Durchschnittlich zeigten 0,54 *loci* pro nicht verwendetem *tag* Sequenzen, bzw. ein solcher *tag* wies im Schnitt 1,89 *reads* auf. Wie Martina Unterländer (2014) feststellte, handelte es sich dabei jedoch nicht um Kreuzkontaminationen von *tags*. Vielmehr führt die Kollabierung der Homopolymere bei der Sortierung (siehe 8.2.8.1) zu kurzen DNA-Stücken mit identischen Sequenzen zwischen einigen *tags* und Primern.

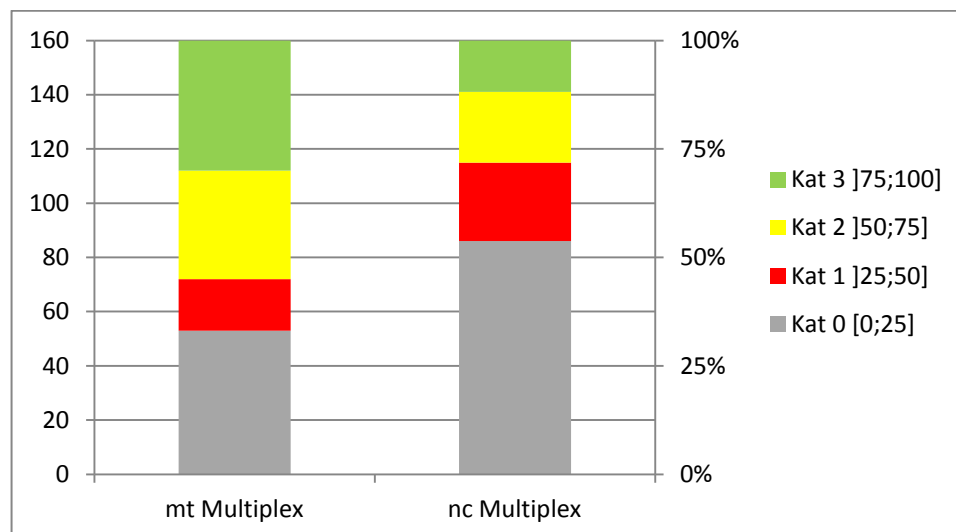
9.2 Ergebnisse der Probenerhaltung

Der Einschätzung des Erhaltungszustandes der Proben liegen lediglich die Amplifikationserfolge der Multiplex-PCRs zugrunde, da Einzel-PCRs vorwiegend an Proben durchgeführt wurden, für die bereits amplifizierbare DNA nachgewiesen wurden. Je Probe wurde der Gesamterfolg als Mittelwert der Amplifikationsraten aller mitochondrialen bzw. nukleären Multiplex-PCRs ermittelt (siehe 9.1.1 & 9.1.2). Es erfolgte eine Einteilung der Proben in vier Kategorien entsprechend ihres Amplifikationserfolges (siehe Tabelle 9.4).

Tabelle 9.4: Kategorisierung der Proben entsprechend ihres Erhaltungszustands

Einteilung	Amplifikationserfolg	DNA-Erhaltung	Anzahl Proben mitochondrial	Anzahl Proben nukleär
Kategorie 0	0 - 25 %	keine / kaum	53	86
Kategorie 1	25 - 50 %	schlecht	19	29
Kategorie 2	50 - 75 %	mäßig	40	26
Kategorie 3	75 - 100 %	gut bis sehr gut	48	19

Aus etwa einem Drittel der 160 Proben konnte keine oder nur sporadisch mitochondriale DNA amplifiziert werden (Kategorie 0). Nukleär war dies bei über der Hälfte der Proben der Fall. 55 % der Proben wiesen einen mitochondrialen Amplifikationserfolg von über 50 % auf (Kategorien 2 und 3). Bei nukleären PCRs reduzierte sich der Anzahl auf 28,13 % (siehe Abbildung 9.6). Die Korrelation zwischen mitochondrialer und nukleärer Erhaltung ist erwartungsgemäß hoch, mit einem Korrelationskoeffizienten³¹ von $r = 0,72$. Bei diesen Betrachtungen wurde noch nicht die Reproduzierbarkeit der Ergebnisse berücksichtigt.

**Abbildung 9.6: Vergleich der mitochondrialen und nukleären Erhaltung**

Die linke Ordinate gibt den absoluten, die rechte den relativen Anteil an Proben je Kategorie an. Die Intervallangaben in der Legende geben den Amplifikationserfolg in % an, der jeder Kategorie zugeordnet ist. Bei den Kategorien 1-3 sind die Intervalle je linksseitig offen, das heißt z. B. für Kategorie 1: $25 \% < \text{Erfolg} \leq 50 \%$.

³¹ $r = \text{Correl}(M, N) = \frac{\sum(m - \bar{m})(n - \bar{n})}{\sqrt{\sum(m - \bar{m})^2 \sum(n - \bar{n})^2}}$ (aus Microsoft Excel 2007), mit m und n = mitochondrialen und nukleären Probenerhaltungswerte. \bar{m} bzw. \bar{n} = Durchschnittswerte.

Von 66 der 160 bearbeiteten Proben konnten die Ergebnisse der Hypervariablen Region I reproduziert und eindeutige Haplotypen bestimmt werden (siehe Kapitel 9.3.1). Dies entspricht einem Erfolg von 41,25 %. Einzelne Proben, für die die Authentifizierungskriterien (siehe 8.2.7 und 8.2.8.1) nicht strikt eingehalten wurden, die aber dennoch als reproduziert angesehen werden, werden im folgenden Kapitel (9.3.1) besprochen. Von den 53 Proben der Kategorie 0 konnten keine reproduzierbaren Daten gewonnen werden. In Kategorie 1 war MAJ 4 das einzige der 19 Individuen, dessen Haplotyp bestimmt werden konnte. In den Kategorien 2 und 3 war die Ergebnisreproduktion in 23 der 40 und 42 der 48 Proben erfolgreich. Für die *coding region* wurde auf eine gezielte Reproduktion verzichtet. Hier wurde lediglich kontrolliert, ob die erfolgreich sequenzierten *loci* die Ergebnisse der HVR I stützen, was in den 66 Proben der Fall war.

Bei der nukleären DNA lassen sich drei Reproduktionsstufen unterscheiden: sicher, vorbehaltlich und nicht reproduziert. Die entsprechenden Kriterien sind den Abschnitten 8.2.7 und 8.2.8.1 zu entnehmen. Insgesamt wurden 3 360 individuelle SNPs untersucht (je 21 *loci* in 160 Individuen). 771 (22,95 %) Genotypen konnten reproduziert werden, davon 526 (15,65 %) sicher. Die restlichen 245 Genotypen sind vorbehaltlich reproduziert.

Setzt man den Reproduktions- mit dem Amplifikationserfolg in Beziehung, so ergibt sich folgendes Bild: 61 der insgesamt 160 untersuchten Individuen enthielten ausreichend amplifizierbare DNA, um zumindest für einen Teil der analysierten *loci* reproduzierbar genotypisiert zu werden. Von den 86 Proben der Kategorie 0 konnten keine reproduzierbaren Daten gewonnen werden. 17 der 29 Individuen aus Kategorie 1, 25 der 26 Proben aus Kategorie 2 und alle 19 aus Kategorie 3 lieferten reproduzierte *loci* (siehe Tabelle 9.5).

Tabelle 9.5: Nukleärer Reproduktionserfolg in Abhängigkeit vom Erhaltungszustand

N = Anzahl untersuchter Proben (siehe auch Abbildung 9.6); N_{rep} = Anzahl der Proben, die (rep+ = sicher) reproduzierte *loci* aufwiesen; \bar{L}_{rep} = durchschnittliche Anzahl (rep+ = sicher) reproduzierter *loci* in einer Probe aus N. Die Spannweite von $L_{rep/rep+}$ unter den reproduzierten Proben $N_{rep/rep+}$ ist in jeweils in Klammern angegeben.

Amplifikationserfolg	N	N_{rep}	\bar{L}_{rep}	N_{rep+}	\bar{L}_{rep+}
Gesamt	160	61	4,8 (0-21)	54	3,3 (0-20)
Kategorie 0	86	0	0,0	0	0,0
Kategorie 1	29	17	2,9 (1-12)	10	0,8 (0-5)
Kategorie 2	26	25	13,1 (8-19)	25	8,0 (3-14)
Kategorie 3	19	19	18,2 (11-21)	19	15,4 (8-20)

Die Feststellung geographischer und temporaler Unterschiede in der Probenerhaltung basierte auf der mitochondrialen Kategorisierung. Die regionalen Unterschiede sind Tabelle 9.7 zu entnehmen. Hierfür wurden Gruppen von Fundorten basierend auf ihrer geographischen und klimatischen Lage gebildet.

Tabelle 9.6: Regionale Unterschiede in der Probenerhaltung

Vegetationszone	Geographische Region	Fundplätze	Kategorie
Steppe-Trockensteppe	Ost-pontisch	MAZ (1), PEJ (6), TEM (8)	2,60
Gemäßigte Laubwälder	Mitteleuropa	CZU (2), DEP (1), ZAU (2)	2,40
Steppe-Waldsteppe	Nord-kaspisch	KAL (2), NIK (7), POD (4)	1,85
Steppe-Waldsteppe	Nord-pontisch	BEZ (1), MOB (3), SUG (8)	1,58
Trockensteppe	Nord-Pontisch	VIN (13)	1,46
Waldsteppe	Nordwest-pontisch	BAD (1), BED (2), TET (3)	1,33
Steppe	Nord-pontisch	KRZ (2) KNO (5), LIS (3), NEV (3), NOZ (2), PES (7), SAC (5)	1,33
Steppe	Nordost-pontisch	OLE (7)	1,14
Steppe	West-pontisch	BEN (4), BOY (5), DUR (3), GOL (3), MED (3), OVI (4), POP (5), RIL (3), SMY (12); ZAG (4)	1,02
Trockensteppe	Nordwest-pontisch	KOS (3), MAJ (9), STB (1), VAP (3)	0,75

Den besten Erhaltungszustand wiesen die Proben aus der Region des nördlichen Kaukasus auf, wobei es sich hierbei fast ausschließlich um Individuen aus der jüngsten Epoche, der Katakombengrabkultur, handelt. Deutlich schlechter erhalten waren die Proben an der westlichen und nordwestlichen Schwarzmeerküste, von wo allerdings auch die meisten äneolithischen Proben stammen.

Bei der Kategorisierung des durchschnittlichen Erhaltungszustands der drei untersuchten Zeitschichten (Tabelle 9.7) erscheinen die äneolithischen Proben etwas besser erhalten als die Jamnaja. Bei den letzteren war jedoch der Anteil an erfolgreich reproduzierten Proben höher. Die aus der gleichen Zeitscheibe aber anderen geographischen Regionen stammenden Proben der Kugelamphorenkultur lagen durchschnittlich zwischen Kategorie 2 und 3, und wiesen somit den besten Erhaltungszustand auf.

Tabelle 9.7: Temporale Unterschiede in der Probenerhaltung

In Klammern ist die Anzahl der Individuen angegeben, die analysiert wurden (1. Spalte), und die die erfolgreich reproduziert wurden (4. Spalte). Die DNA-Erhaltung ist entsprechend der Kategorien angegeben. Bei diesen handelt es sich um den Durchschnitt der entsprechenden Zeitscheibe.

Alter	Kategorie	DNA-Erhaltung	Reproduzierte Proben
Äneolithikum (32)	1,44	schlecht bis mäßig	31,25 % (10)
Frühe Bronzezeit (63)	1,29	schlecht bis mäßig	41,27 % (26)
Jamnaja-Kultur (58)	1,19	schlecht	37,93 % (22)
Kugelamphorenkultur (5)	2,40	gut	80,00 % (4)
Frühe - Mittlere Bronzezeit (58)	1,78	mäßig	51,72 % (30)
Katakombengrabkultur (45)	1,84	mäßig	52,27 % (24)
Poltavka-Kultur (13)	1,92	mäßig	46,15 % (6)
unbestimmt (7)	1,14	schlecht	0,00 % (0)

9.3 Ergebnisse der populationsgenetischen Analysen

9.3.1 Haplotypisierung der Proben

Die Haplotypisierung der Proben erfolgte nach dem in den Abschnitten 8.2.7 und 8.2.8.1 beschriebenen Authentifizierungsschema. Tabelle 9.8 zeigt die Haplotypen der 66 reproduzierten Individuen und ihre Zuordnung in die entsprechenden Haplogruppen. Das für die HVR I geltende Kriterium von mindestens drei PCRs wurde dabei nicht immer für alle Positionen erfüllt, insbesondere für die ersten zwei bis vier (np 16011 – 16014) und die letzten vier bis neun Nukleotide (np 16405 – 16413) der HVR I. In die Datenanalyse ging der Bereich von np 16013 bis 16409 ein.

Die 3'-Enden der Primer M34L und M36U überschneiden sich an den zwei Nukleotiden 16272 – 16273. Im Falle eines *locus dropout* von Primersystem M35 sind nur diese beiden Positionen betroffen. Da diese im gesamten Datensatz monomorph sind, wurde teilweise auf eine wiederholte Amplifizierung von Fragment M35 verzichtet. Für 26 Individuen fehlt die dritte Abdeckung der beiden Positionen, in 14 Fällen ist auch Position 16274 betroffen.

Abgesehen von diesen Bereichen gibt es in 6 der 66 Individuen kurze Abschnitte, in denen die *consensus*-Sequenz auf zwei anstatt drei PCR-Produkten basiert. Bei MOB 3 betrifft dies die Nukleotide 16154 – 16158 (5 nt), bei VIN 3 den Bereich von 16177 – 16184 (8 nt), bei TET 1 die Positionen 16182 – 16184 (3 nt) und bei OLE 1 16374 – 16377 (4 nt). Alle untersuchten Individuen entsprechen in diesen Bereichen der rCRS. Bei SMY 4 ist die Abdeckung ab Position 16399 doppelt, ab 16405 nur noch einfach. Bei NIK 4 fehlt die dritte Sequenz von Fragment M36. Hierdurch ist der Bereich von 16306 – 16332 (27 nt) nur zweifach und die folgenden sieben Nukleotide nur einfach repräsentiert. Mit Ausnahme von OLE 1 und VIN 3 existieren für alle diese Proben PCR-Produkte mit weniger als drei *454 reads*, die somit „verworfen“ wurden, allerdings die Ergebnisse stützen.

Von der Probe PES 7 konnte nur ein Extrakt gewonnen werden. Da alle sonstigen Kriterien erfüllt wurden, wurde ihr Haplotyp als authentisch betrachtet. Bei der Probe MAJ 4 ist die HVR I zwar hundertprozentig mit drei oder mehr Sequenzen abgedeckt,

doch an Position 16311 konnte das erforderliche Verhältnis von mindestens 3:1 zur Anerkennung der Transition nicht erreicht werden. Das der rCRS entsprechende T-Allel wurde in zwei Ansätzen, für die ein gemeinsamer Mastermix verwendet wurde, amplifiziert. Das C-Allel hingegen wurde in fünf unabhängigen PCRs amplifiziert. Aufgrund des schlechten Erhaltungszustandes, der sich auch in dem hohen Degradierungsgrad widerspiegelte, blieben Versuche, ein weiteres PCR-Produkt zu gewinnen, erfolglos.

Von den 94 Individuen, die nicht reproduziert werden konnten, wurden sechs aufgrund von Kontaminationen in den Extrakten verworfen. BOY 4 stimmt in der Hypervariablen Region I mit der rCRS überein. In der *coding region* gibt es mehrfach Unstimmigkeiten an Kladen-spezifischen Positionen, wobei keine Eingrenzung auf zwei oder drei Linien möglich ist. KAL 2 entspricht der rCRS. Als Nebenlinie tritt ein Haplotyp der Gruppe K auf. Zwar setzt sich dieser Typ nie durch, aber in einem Großteil der PCRs erscheinen die Positionen 16224, 16258 und 16311 „heterozygot“. Die *coding region* wäre der Haplogruppe H (rCRS) zuzuordnen. KNO 3 zeigt in beiden Extrakten Linien der Haplogruppen T und H. Bei SAC 3 weisen die Mutationen 16129A und 16256T auf die Haplogruppe H3b1 hin. Dies wird in der *coding region* jedoch nicht bestätigt, da mehrfach entscheidende Positionen ambig sind, ohne dass eine weitere Linie identifiziert werden kann. In der HVR I ist außerdem die Position 16051 sowohl durch Adenosin als auch Guanin repräsentiert. TEM 8 ist definitiv der Haplogruppe U5a zuzuordnen. Zusätzlich zu den definierenden Mutationen 16256T und 16270T trägt TEM 8 ein C an Position 16189. Position 16192 zeigte in mehreren PCRs entweder T oder Y. Auch Position 16234 zeigte sich mehrfach „heterozygot“ oder entsprach der rCRS. VIN 13 gehört wahrscheinlich auch der Haplogruppe U5 an. Allerdings konnte keine eindeutige *consensus*-Sequenz erstellt werden, da sowohl innerhalb als auch zwischen den PCRs Diskrepanzen an den entsprechenden Positionen bestanden.

Tabelle 9.8: Mitochondriale Haplotypen der 66 reproduzierten Individuen

Die Mutationen sind entsprechend der Nummerierung der rCRS (Andrews *et al.* 1999) benannt, bei der HVR I minus 16000. Sofern im Text nicht anders angegeben, basieren HVR I-Haplotypen auf mindestens drei unabhängigen PCRs, *coding region*-SNPs auf einer oder mehr PCRs mit jeweils mindestens drei *reads*. Fehlende Amplikons sind als \Mxx\ gekennzeichnet. HG = Haplogruppe, ermittelt mithilfe von HaploGrep (basierend auf *PhyloTree build 16*) (van Oven & Kayser 2009, Kloss-Brandstätter *et al.* 2011).

Probe	HVR I [16011-16413]	Coding Region [M01-M30]	HG
<i>Frühes Äneolithikum</i>			
SMY 3	rCRS	rCRS	H2a2
SMY 4	304C	456Y	H5
SMY 9	261T	rCRS	H7a1
SMY 11	126C 153A 269G 294T 296T	4216C 4917G 7028T 11719A 14766T	T2e
VIN 1	rCRS	rCRS	H2a2
<i>Spätes Äneolithikum</i>			
DUR 1	114A 192T 256T 270T 294T	7028T 11719A 12308G 14766T 14793G	U5a2a
MAJ 8	126C 294T 296T 304C	4216C 4917G 7028T 11719A 14766T	T2b
MAJ 9	223T 292T	7028T 8251A 11719A 11947G 12705Y 14766T	W
MOB 1	192T 256T 270T 399G	7028T 11719A 12308G 14766T 14793G	U5a1
MOB 3	192T 256T 270T 399G	\M06\ 7028T 8251R 11719A 12308G 14766T \M30\	U5a1
<i>Frühe Bronzezeit – Jamnaja-Kultur</i>			
BEN 3	188T	4937C 8935T 13269G	H
MAJ 3	256T 270T 399G	\M01\ \M03\ \M06\ \M12\ 7028T \M14- -M20\ 11719A \M24\ 14766T \M30\	U5a1
MAJ 4	192T 311C	4194T 7028T 11719A 12308G 13637G 14766T	U5b2a1a
MAJ 5	136C 189C 223T 278T	6371T 7028T 11719A 12705T 14766T	X(2 or 3)
OLE 1	126C 292T 294T	4216C 4917G 7028T 11719A 14766T	T
OLE 7	048A 069T 126C 193T	4216C 7028T 10398G 11719A 13708A 14766T 14769G	J2b1d
OVI 2	224C 290T 311C	7028T \M15\ 10398G 11719A 12308G 14766T 14798C	K1
PEJ 1	192T 256T 270T 399G	7028T 11719A 12308G 14766T 14793G	U5a1
PES 4	304C	456T 7028Y	H5

Probe	HVR I [16011-16413]	Coding Region [M01-M30]	HG
PES 7	042A 209C	481T	H
POP 1	126C 294T 296T 324C	4216C 4917G 7028T 9117C 11719A 14766T	T2a1b1a
POP 3	051G 129C 362C	7028T 10876G 11719A 11732C 12308G 13734C 14766T	U2e1a
POP 4	256T 270T 399G	7028T 11719A 12308G 14766T 14793G	U5a1
RIL 3	192T 224C 311C	7028T 10398G 11719A 12308G 14766T 14798C	K1
SUG 2	129A 172C 223T 311C 391A	4529T 7028T 8251A 8260C 10398G 11719A 12705T 14766T	I
SUG 6	362C	rCRS	H6
SUG 7	rCRS	rCRS	H2a2
SUG 8	rCRS	3010R \M15\	H2a2
TET 2	134T 356C	7028T 11719A 12308G 14766T	U4a1
VIN 2	126C 163G 186T 189C 294T	\M01\ 4216C \M06\ 4917G 7028T \M15\ \M19\ 11719A 14766T	T1a
VIN 5	126C 163G 186T 189C 294T	4216C 4917G 7028T 11719A 14766T	T1a
VIN 12	126C 163G 294T	4216C 4917G 7028T 11719A 14766T	T1
<i>Frühe/mittlere Bronzezeit – Frühe Katakombengrabkultur</i>			
LIS 1	256T 270T 399G	7028T 11719A 12308G 14766T 14793G	U5a1
LIS 2	356C	6366G 7028T 11719A 12308G 13708A 14766T	U4b1a2
PEJ 3	362C	rCRS	H6
PEJ 4	362C	rCRS	H6
PEJ 5	356C	7028T 11719A 12308G 14766T	U4
SAC 2	069T 126C 193T	4216C 7028T 10398G 11719A 13708A 14766T	J2b
TEM 1	356C	7028T 11719A 12308R 14766T	U4
TEM 2	rCRS	rCRS \M15\	H2a2
TEM 4	256T 270T 399G	7028T 11719A 12308G 14766T 14769A \M30\	U5a1
TEM 5	311C	4917G 7028T 8388C 11719A 14766T	R1
TEM 6	311C	4917G 7028T 8388C 11719A \M29\	R1
TEM 7	356C	7028T 11719A 12308G 14766T	U4
<i>Mittlere Bronzezeit – Entwickelte Katakombengrabkultur</i>			
KNO 4	356C	7028T 11719A 12308G 14766T	U4
LIS 3	354T	rCRS	H2a1

Probe	HVR I [16011-16413]	Coding Region [M01-M30]	HG
NEV 1	192T 256T 270T 399G	7028T 11719A 12308G 14766T 14793G	U5a1
NEV 3	362C	rCRS	H6
NOZ 1	356C	7028T 11719A 12308G 14766T	U4
NOZ 2	356C	7028T 11719A 12308G 14766T	U4
PEJ 2	234T	rCRS	H24a1
SUG 5	362C	rCRS	H6
TEM 3	069T 126C 145A 172C 222T 261T	462T 3010A 4216C 7028T 8269A 10398G 11719A 13708A 14766T	J1b1a1
TET 1	223T 391A	4529T 7028T 8251A 10398G 11719A 12705T 14766T	N1a1b
VIN 3	256T 270T 399G	7028T 11719A 12308G 14766T 14793G	U5a1
VIN 8	069T 126C 145A 172C 222T 261T	462T 3010A 4216C 7028T 8269A 10398G 11719A 13708A 14766T	J1b1a1
<i>Frühe/mittlere Bronzezeit – Poltavka-Kultur</i>			
KAL 1	147G 172C 223T 248T 344T 355T	669C 7028T 10398G 11719A 12705T 14766T	N1a1a
NIK 1	126C 163G 186T 189C 294T	4216C 4917G 7028T 11719A 12346T 14766T	T1a
NIK 4	192T 256T 270T	4596A 7028T 11719A 12308G 14766T	U5
NIK 7	rCRS	13680T	H13a1a
POD 1	192T 223T 292T 325C	7028T 8251A \M15\ 8994A 11719A 11947G 12705T 14766T	W6
POD 2	126C 292T 294T	4216C 4917G 7028T 11719A 14766T	T
<i>Frühe Bronzezeit – Kugelamphorenkultur</i>			
CZU 1	080G 189C 356C	3010A 8251A	H1b2
CZU 2	093C 224C 311C	7028T 10398G 11719A 12308G 14766T \M30\	K1a1
ZAU 1	rCRS	rCRS	H2a2
ZAU 2	069T 126C	462Y 3010A 4216C 7028T 10398G 11719A 13708A 14766T	J1

9.3.2 Genetische Charakterisierung der prähistorischen Populationen

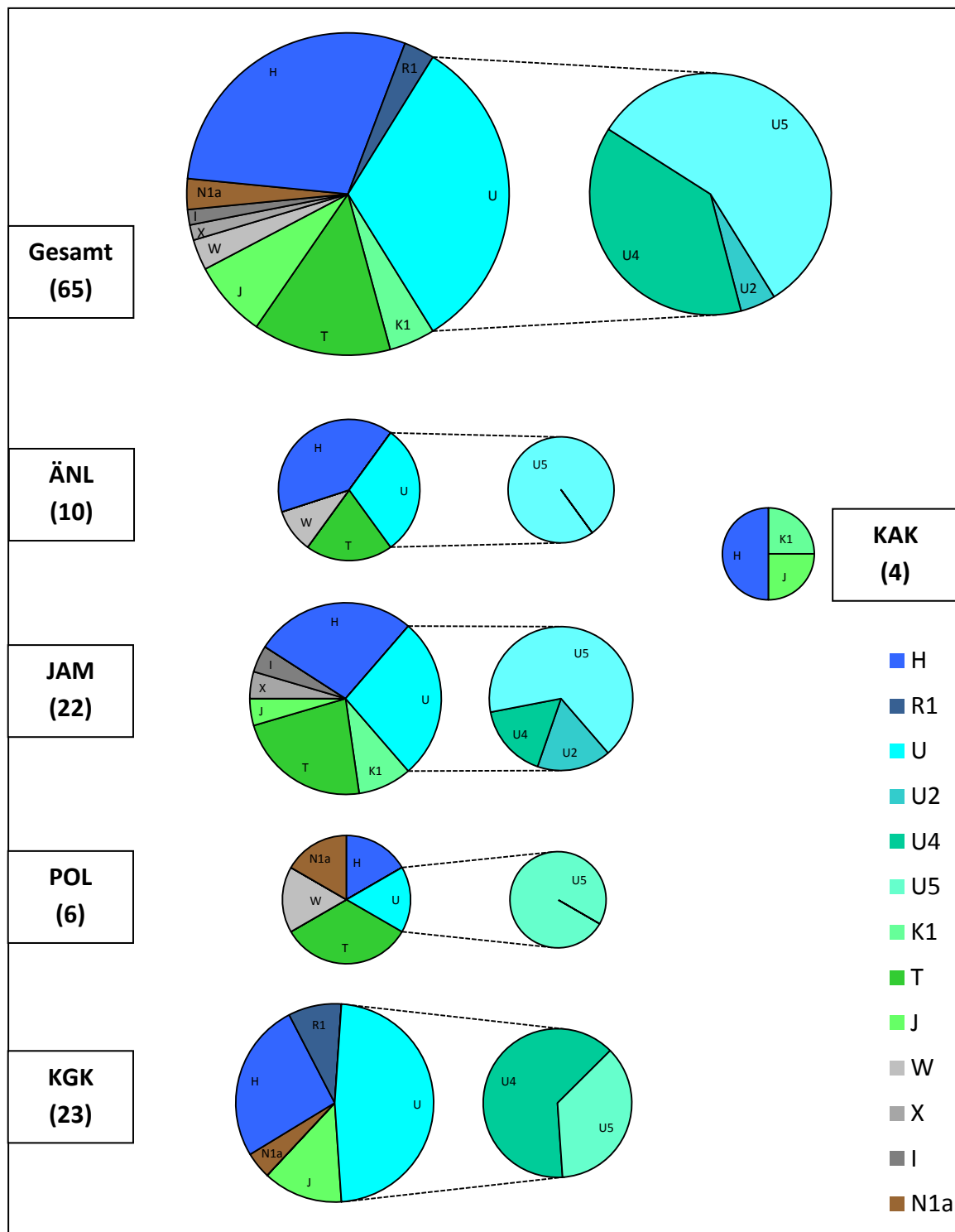


Abbildung 9.7: Haplogruppenverteilung in den prähistorischen Kulturen

Gesamt = alle typisierten Proben, ÄNL = Äneolithikum, JAM = Jamnaja-Kultur, POL = Poltavskakultur, KGK = Katakombengrabbkultur, KAK = Kugelamphorenkultur. Zahl in Klammer = Anzahl der Individuen. Die Probe PEJ 4 ist nicht enthalten (betrifft „Gesamt“ und „KGK“), siehe Text nächste Seite.

Unter den 66 analysierten Proben wurden mehrfach aus einem Fundplatz stammende identische Haplotypen beobachtet. In einem dieser Fälle, den Proben PEJ 3 und PEJ 4, handelt es sich um ein adultes Individuum und ein Kind, die in einem gemeinsamen Grab bestattet wurden. Die morphologische Geschlechtsbestimmung ergab bei PEJ 4 einen Mann, die genetische deutet aber auf eine Frau hin. Daher ist es wahrscheinlich, dass es sich um Mutter und Kind handelt. Aufgrund der direkten Abstammung wurde PEJ 4³² bei den populationsgenetischen Auswertungen nicht berücksichtigt. Ansonsten stammen die Proben mit identischen Haplotypen aus verschiedenen Gräbern.

Die restlichen 65 analysierten Proben weisen insgesamt 36 verschiedene Haplotypen auf, von denen 65 % in die Haplogruppen H (und R1³³) und U fallen. In Abbildung 9.7 sind die Untergruppen der stark vertretenen Haplogruppe U gesondert herausgestellt. Die Zuordnung der Individuen TEM 5 und TEM 6 zur Haplogruppe R1 erfolgte basierend auf der *coding region*. Des Weiteren treten die Haplogruppen J, T, K1, W, N1a, I und X auf. Mit Ausnahme der Haplogruppe R1 konnten bei allen in Abbildung 9.7 aufgezeigten Haplogruppen, die durch mehr als ein Individuum vertreten sind, die Sequenzen in verschiedene Sub-Haplogruppen unterteilt werden (siehe auch Tabelle 9.8).

Für die populationsgenetischen Analysen wurden die archäologisch definierten Kulturen als Populationen betrachtet³⁴, wobei zwischen äneolithischen Gruppen (ÄNL), Jamnaja- (JAM), Katakombengrab- (KGK) und Poltavka-Kultur (POL) unterschieden wurde. Da die Region südlich der Donau zur der Zeit, aus der die Proben aus Smjadovo (SMY) stammen, noch nicht in das Steppenäneolithikum integriert war (näheres dazu in Kapitel 10.3.1), wurde dieser Fundort in einigen Analysen aus dem ÄNL-Datensatz ausgeklammert (SÄN).

Für einen ersten Anhaltspunkt bei der Überprüfung der archäologischen Ursprungshypothese zu den nordpontischen Viehhirten wurden die Proben nach

³² In Wilde *et al.* 2014 gingen beide Individuen in die Analysen ein.

³³ Die Einordnung der Individuen in Haplogruppe R1 erfolgte auf Basis der *coding region*. Blicke diese unberücksichtigt, könnte die Mutation 16311C in der HVR I zu einer Zuordnung zu Haplogruppe HV oder H führen.

³⁴ Zur Vereinfachung werden daher im Folgenden die Kulturnamen stellvertretend für die Populationen, d. h. die Träger der entsprechenden Kulturen verwendet.

ihrem Haplotyp in zwei Kategorien aufgeteilt und ihre Anteile in den Steppenpopulationen festgestellt (Abbildung 9.8). Linien der Haplogruppe U (U2, U4, U5) werden mit Jäger/Sammler-Populationen assoziiert, während bei den anderen Typen vermutet wird, dass sie im Verlauf des Neolithikums nach Europa eingetragen wurden (Haak *et al.* 2005, Bramanti *et al.* 2009, Bollongino *et al.* 2013, Brandt *et al.* 2013). Im Folgenden werden an einigen Stellen die Begriffe „mesolithische“ oder „Jäger/Sammler-Linien“ und „neolithische Linien“ verwendet, um diese Unterscheidung vereinfacht, wenn auch nicht ganz korrekt, zu benennen.

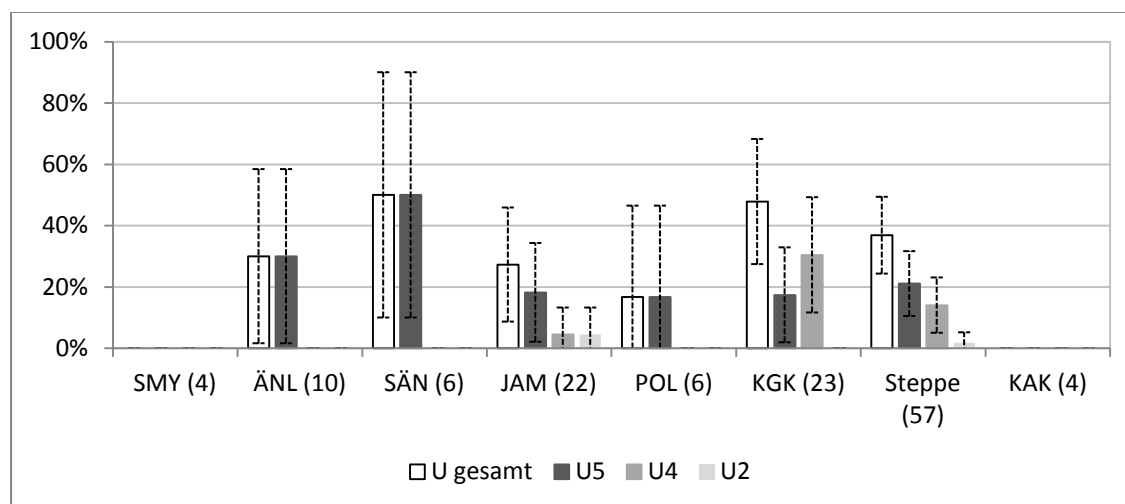


Abbildung 9.8: Die U-Linien (assoziiert mit Jägern/Sammlern) in den Steppenkulturen

Aufgetragen ist der relative Anteil der Haplogruppe U (weiß) und der Subhaplogruppen U5 (dunkelgrau), U4 (mittelgrau) und U2 (hellgrau) in den Steppenpopulationen. SMY = Fundplatz Smjadovo, ÄNL = Äneolithikum, SÄN = Steppenäneolithikum (= ÄNL ohne SMY), JAM = Jamnaja-Kultur, KGK = Katakombengrabbkultur, POL = Poltavka-Kultur, Steppe = Gesamtdatensatz ohne SMY und KAK, KAK = Kugelamphorenkultur. Die Antennen zeigen das jeweilige 95 % Konfidenzintervall an.

Die zur näheren genetischen Charakterisierung der Kulturen herangezogenen Größen sind in Tabelle 9.9 aufgeführt. Unter den drei Steppenkulturen ÄNL, JAM und KGK weist der Jamnaja-Komplex sowohl auf Basis der Haplotypen als auch auf Nukleotidebene die größte Diversität auf. Während der wesentlich kleinere Datensatz der äneolithischen Gruppen annähernd hohe Werte aufweist, weichen die der Katakombengrabbkultur deutlich nach unten ab. Die Haplotypendiversität der

Katakombengrabkultur (0,8393 – 0,9289) ist signifikant geringer als die der Jamnaja-Kultur (0,9600 – 0,9924). Vor allem die Werte der frühen Katakombengrabkultur sind mit $\hat{H} = 0,8636$ und $\hat{\pi}_n = 6,61 \times 10^{-3}$ auffallend niedrig. Mit Ausnahme der fünf Proben aus der Zeit des späten Äneolithikums sind Tajima's D und Fu's F_S für alle Gruppen negativ. Statistische Signifikanz wird nur im Rahmen des gesamten Datensatzes und für die Jamnaja-Kultur ($D = -1,6254$; $F_S = -11,8073$) erreicht.

Tabelle 9.9: Genetische Diversität und Demographie der prähistorischen Kulturen

Die Werte basieren auf den HVR I-Sequenzen, Nukleotidpositionen 16013 – 16409. N = Anzahl der Individuen, Ht = Anzahl der Haplotypen, Ht_p = Anzahl der ausschließlich in der einen Population vorkommenden („privaten“) Haplotypen; \hat{H} = Haplotypendiversität; $\hat{\pi}_n$ = Nukleotiddiversität; σ = Standardabweichung. Die Werte sind auf vier Nachkommastellen gerundet. Angaben zur Berechnung der Statistiken sind in Kapitel 7.3 und Excoffier & Lischer (2010) nachzulesen. Die p -Werte basieren auf 10 000 Permutationen. Für Tajima's D wurde gemäß Konvention das Signifikanzniveau auf 5 %, für Fu's F_S auf 2 % festgesetzt. Signifikante Werte sind fettgedruckt.

Gruppe	N	Ht Ht_p	\hat{H} ($\pm \sigma$)	$\hat{\pi}_n$ ($\pm \sigma$)	Tajima's D (p -Wert)	Fu's F_S (p -Wert)
Steppenkulturen gesamt (ohne Kugelamphoren)	61	33	0,9623 $\pm 0,0107$	0,0115 $\pm 0,0063$	-1,7613 (0,0149)	-22,3212 (0,0000)
Äneolithikum	10	8 5	0,9556 $\pm 0,0594$	0,0113 $\pm 0,0069$	-0,4256 (0,3571)	-2,2899 (0,0762)
früh	5	4 2	0,9000 $\pm 0,1610$	0,0071 $\pm 0,0052$	-1,1617 (0,0771)	-0,4448 (0,2837)
mittel	0					
spät/final	5	4 3	0,9000 $\pm 0,1610$	0,0136 $\pm 0,0092$	0,1636 (0,5948)	0,6123 (0,5323)
Steppenäneolithikum (ohne Smjadovo)	6	5 3	0,9333 $\pm 0,1270$	0,0123 $\pm 0,0080$	0,0617 (0,5480)	-0,3973 (0,3188)
frühe Bronzezeit – Jamnaja-Kultur	22	19 12	0,9870 $\pm 0,0175$	0,0128 $\pm 0,0072$	-1,6232 (0,0368)	-12,9234 (0,0000)
Frühe/mittlere Bronzezeit – Poltavka-Kultur	6	6 2	1,0000 $\pm 0,0962$	0,0168 $\pm 0,0107$	-0,64866 (0,3221)	-1,6233 (0,1012)
frühe/mittlere Bronzezeit - Katakombengrabkultur	23	11 7	0,8854 $\pm 0,0483$	0,0086 $\pm 0,0051$	-1,0976 (0,1363)	-2,6770 (0,0871)
früh	11	6 2	0,8545 $\pm 0,0852$	0,0064 $\pm 0,0042$	-0,7212 (0,2646)	-0,9401 (0,2418)
entwickelt	12	8 4	0,9242 $\pm 0,0575$	0,0108 $\pm 0,0065$	-0,8357 (0,2047)	-1,4880 (0,1776)
Kugelamphorenkultur	4	4 3	1,0000 $\pm 0,1768$	0,0101 $\pm 0,0076$	-0,8241 (0,1352)	-0,8247 (0,1571)

9.3.3 Genetische Distanzen im nordpontischen Steppenraum

Um Antworten auf die Fragen zu den steppeninternen Entwicklungen vom Äneolithikum bis zur mittleren Bronzezeit zu finden, wurden F_{ST} -Werte basierend auf paarweisen Unterschieden berechnet. Im Vergleich der Kulturen untereinander ist zwischen dem Äneolithikum und der Jamnaja-Phase keine genetische Distanz festzustellen. Zwischen Äneolithikum und Katakombengrabkultur liegt der F_{ST} -Wert bei 0,05438, erreicht jedoch nur marginale Signifikanz ($p < 0,1$). Die direkt aufeinanderfolgenden bzw. sich zeitlich überschneidenden Jamnaja- und Katakombengrabkultur haben eine geringe, nicht signifikante Distanz von 0,02287. Signifikante F_{ST} -Werte ($p < 0,05$) um 0,1 wurden nur zwischen der Katakombengrab- und der durch lediglich sechs Individuen vertretene Poltavka-Kultur beobachtet. Die Werte zwischen allen Steppenkulturen sind in Tabelle 9.10 als Matrix aufgeführt.

Tabelle 9.10: Paarweise F_{ST} -Werte zwischen den Kulturen der Steppe

Die F_{ST} -Werte basieren auf paarweisen Unterschieden in der HVR I, np 16013 – 16409. In Klammer ist die Individuenanzahl der entsprechenden Kultur angegeben. Fettgedruckte Werte sind signifikant ($p < 0,5$). Die Kugelamphorenkultur ist nur der Vollständigkeit halber aufgeführt. Da sie nur durch vier Individuen vertreten ist, sind die F_{ST} -Werte nicht verwendbar.

	Kultur	ÄNL (10)	JAM (22)	KGK (23)	fKGK (11)	eKGK (12)	POL (6)
Äneolithikum	ÄNL (10)	0					
Jamnaja-Kultur	JAM (22)	-0,00794	0				
Katakombengrabkultur	KGK (23)	0,05438	0,02287	0			
früh	fKGK (11)	0,05953	0,00544	-0,05073	0		
entwickelt	eKGK (12)	0,02696	0,01167	-0,05254	-0,02453	0	
Poltavka-Kultur	POL (6)	-0,03268	-0,02341	0,09135	0,10078	0,0368	0
Kugelamphorenkultur	KAK (4)	-0,03448	-0,05774	-0,01545	-0,02026	-0,02907	-0,00741

Für eine differenziertere Betrachtung der Vorgänge in der Steppe wurden die Fundplätze in regionale Gruppen unterteilt, um geographische Strukturen innerhalb der Kulturen oder räumlich begrenzte Entwicklungen aufzudecken. Dafür wurden elf lokale Ausgangsgruppen bzw. Fundplatz-Einheiten gebildet, die je nach Fragestellung zu verschiedenen Gruppen zusammengefasst wurden. Die Fundplatz-Einheiten sind in Tabelle 9.11 aufgelistet und auf der Karte in Abbildung 9.9 zu sehen.

Tabelle 9.11: Fundplatz-Einheiten als Grundlage für die Bildung von Gruppen

Die Nummerierung der Einheiten entspricht der in Abbildung 9.9. Die Namen der Fundplätze finden sich in Tabelle 8.1. *Der Fundort Smjadovo in Einheit 3 gehört zeitlich in das frühe Äneolithikum entsprechend der nordpontischen Zeitrechnung. Doch aufgrund der Lage und Datierung der Proben wäre er eigentlich dem südosteuropäischen Neolithikum zuzusprechen.

Nr.	Gebiet (Fluss)	Kultur	N	Fundorte
1	Mittleuropa	Kugelamphorenkultur	4	CZU, ZAU
2	Donau	Äneolithikum (früh)*	4	SMY*
		Äneolithikum (spät)	1	DUR
		Jamnaja-Kultur	6	BEN, OVI, POP, RIL
3	Pruth	Jamnaja-Kultur	1	TET
		Katakombengrabbkultur (entwickelt)	1	TET
4	Dnjestr	Äneolithikum (spät)	2	MAJ
		Jamnaja-Kultur	3	MAJ
5	Dnjepr (rechts)	Äneolithikum (spät)	2	MOB
		Jamnaja-Kultur	4	SUG
		Katakombengrabbkultur (entwickelt)	1	SUG
6	Dnjepr (links)	Jamnaja-Kultur	2	PES
		Katakombengrabbkultur (früh)	1	SAC
7	Dnjepr (links)	Äneolithikum (früh)	1	VIN
		Jamnaja-Kultur	3	VIN
		Katakombengrabbkultur (entwickelt)	2	VIN
8	Donez	Katakombengrabbkultur (früh)	2	LIS
		Katakombengrabbkultur (entwickelt)	6	KNO, LIS, NEV, NOZ
9	Kuban	Jamnaja-Kultur	2	OLE
10	Manych	Jamnaja-Kultur	1	PEJ
		Katakombengrabbkultur (früh)	8	PEJ, TEM
		Katakombengrabbkultur (entwickelt)	2	PEJ, TEM
11	Wolga	Poltavka-Kultur	6	KAL, NIK, POD



Abbildung 9.9: Fundplatz-Einheiten als Grundlage für die Bildung von Gruppen

(Abbildung auf voriger Seite) Jedes Symbol repräsentiert ein Individuum bzw. eine DNA-Sequenz. Die Farben geben die Kulturzugehörigkeit an: grün = Äneolithikum, rot = Jamnaja-Kultur, blau = Katakombengrabkultur, orange = Poltavka-Kultur, gelb = Kugelamphorenkultur. Die Form der Symbole unterscheidet zwischen Linien der Haplogruppe U, die typisch für Jäger/Sammler-Populationen ist (Kreis) und in neolithischen Kulturen beobachtete Linien (Stern), wobei die hochfrequente Haplogruppe H gesondert herausgestellt ist (Dreieck). Der schwarze Pfeil indiziert, dass der Fundplatz Zauschwitz (ZAU) weiter westlich des Kartenausschnitts (in Sachsen) liegt. Die Fundplatz-Einheiten bzw. Ausgangsgruppen sind durch grau hinterlegte Ovale gekennzeichnet und nummeriert (siehe hierzu Tabelle 9.11).

Im Neolithikum und Äneolithikum lassen sich in der nordpontischen Steppe archäologisch verschiedene Kultur fassen (siehe Kapitel 4.2). Falls es sich bei diesen um genetisch distinkte Subpopulationen gehandelt hat, könnten die entsprechenden Strukturen möglicherweise noch im Genom der bronzezeitlichen Bevölkerung nachzuweisen sein. Da die Jamnaja- und die Katakombengrabkultur unterschiedliche Verteilungen in der Steppe haben, wurden diese zunächst zu einer Population zusammengefasst (die äneolithischen Proben wurden nicht integriert), um einen groben Überblick über mögliche Strukturen zu bekommen. Es wurden geographische Gruppen gebildet, die den Verbreitungsgebieten der einzelnen Kulturen des Neolithikums (Tabelle 13.4) des frühen (Tabelle 13.5) oder des mittleren Äneolithikums (Tabelle 13.6) zuzuordnen sind. Da die Viehwirtschaft sich sukzessive ausbreitete, und Flüsse von Archäologen als zeitweilige kulturelle Grenzen identifiziert wurden, wurden die Fundplätze der Zwischenstromgebiete (Tabelle 13.7) bzw. Fundplätze entlang der einzelnen Flüsse (Tabelle 13.8) gruppiert. Außerdem wurden verschiedene West-Ost-Unterteilungen des Gebiets vorgenommen (Tabelle 13.9). Die angegebenen Tabellen im Anhang zeigen die Fundplatzgruppen und paarweisen F_{ST} -Werte bei den jeweiligen Vergleichen.

Die F_{ST} -Werte waren bis auf wenige Ausnahmen niedrig ($< 0,03$) und nicht-signifikant. Die Gruppe auf dem Gebiet der mitteläneolithischen Kvitanska-Kultur um den mittleren Dnjepr (Fundplatz-Einheiten, FPE 3-5-6-8) zeigte signifikante F_{ST} -Werte von 0,1306 bzw. 0,0887 zu den vormaligen Gebieten der Unteren Mikahilovka- (FPE 4-7) und der Khvalynsk-Kultur (FPE 11). Bei der Unterteilung des Untersuchungsraums nach Flüssen wurde eine große genetische Distanz von 0,0864 ($p < 0,05$) zwischen den

Fundplätzen entlang des Dnjeprs (FPE 5-6-7) und des Donez (FPE 8) festgestellt. Detailliertere Analysen fanden auf der kulturellen Ebene statt

Um die Übergänge zwischen den Kulturschichten genauer zu beleuchten, wurden die interkulturellen genetischen Distanzen zwischen verschiedenen geographischen Gruppen berechnet. Hier dargestellt ist lediglich eine Auswahl der F_{ST} -Werte, basierend auf den Fragestellungen. Eine vollständige Matrix mit den paarweisen Vergleichen zwischen allen geographischen und kulturellen Gruppen findet sich im Anhang (Tabelle 13.10).

Jamnaja-Gruppen westlich des Dnjepr weisen keine signifikanten F_{ST} -Werte zu den äneolithischen Populationen auf, die, bis auf eine Probe, ebenfalls alle aus diesem Gebiet stammen. Östlich des Dnjepr unterscheiden sich die Jamnaja-Population deutlich und signifikant von den äneolithischen Proben (F_{ST} -Werte zwischen 0,12 und 0,18; siehe Tabelle 9.12). Beim Vergleich zwischen der Jamnaja- und der Katakombengrabbkultur zeigt sich ein ähnliches Bild. Während die Kulturen insgesamt kaum voneinander distanziert sind, fallen erneut die Jamnaja-Proben aus der Region östlich des Dnjepr durch hohe F_{ST} -Werte auf.

Tabelle 9.12: F_{ST} -Werte zwischen dem Äneolithikum und Gruppen der Jamnaja-Kultur

Innerhalb des Untersuchungsraums wurden verschiedene regionale Gruppierungen untersucht. Sie orientieren sich grob an den Flüssen des Gebiets (W / O = Region am West- bzw. Ostufer). Die den Gruppen zugehörigen Fundplatz-Einheiten entsprechen der Nummerierung in Abbildung 9.9. N = Anzahl der Individuen, fettgedruckt = signifikante F_{ST} Werte ($p < 0,05$).

	Jamnaja		gesamtes Gebiet	Donau – Dnjepr (O)	Dnjepr (W) – Kuban	Dnjepr (O) – Kuban	Dnjepr (O) – Manych
Äneolithikum	Einheiten		2-7,9,10	2,3,4,5,6,7	5,6,7,9	6,7,9	6,7,9,10
		N	22	19	11	7	8
Gesamtgebiet	2,4,5,7	10	-0,00794	0,00641	0,05553	0,12452	0,04281
Steppen-ÄNL	2,4,5,7	6	0,02574	0,0498	0,12227	0,17856	0,07139

Tabelle 9.13 zeigt die genetischen Distanzen einer Auswahl von Gruppen der Katakombengrabkultur, die das vermeintliche Ursprungsgebiet, sowie die frühe und die späte Ausbreitung reflektieren, zu denen der Jamnaja-Kultur. Außerdem ist die Poltavka-Kultur in dieser Matrix enthalten, die zu allen Jamnaja-Gruppen geringe, nicht signifikante F_{ST} -Werte zeigt, aber hohe Distanzen insbesondere zur frühen Katakombengrabkultur aufweist.

Tabelle 9.13: F_{ST} -Werte zwischen Gruppen der Jamnaja-, der Katakombengrab- und der Poltavka-Kultur

Die den Gruppen zugehörigen Fundplatz-Einheiten entsprechen der Nummerierung in Abbildung 9.9. N = Anzahl der Individuen, fettgedruckt = signifikante F_{ST} -Werte ($p < 0,05$).

Katakombengrabkultur	Jamnaja		gesamtes Gebiet	Donau – Dnjepr	Pruth – Dnjepr (W)	Dnjepr (O) – Manych	Poltavka Wolga
	Einheiten		2-7,9,10	2,3,4	3,4,5	6,7,9,10	11
		N	22	10	8	8	6
Gesamtgebiet	3,5,6,7,8,10	23	0,02287	0,01026	-0,01057	0,1518	0,09135
frühe K.	6,8,10	11	0,00544	-0,01698	-0,01901	0,15772	0,10078
entwickelte K.	3,5,7,8,10	12	0,01167	0,00351	-0,02171	0,10959	0,0368
Donez	8	8	0,03131	-0,00265	0,00529	0,17834	0,08767
Kuban	10	10	-0,00231	-0,01806	-0,03716	0,13867	0,07536
POL: Wolga	11	6	-0,02341	-0,00083	0,00081	-0,03878	---

Zur Aufdeckung von Populationsstrukturen innerhalb der einzelnen Kulturschichten wurden die F_{ST} -Werte zwischen verschiedenen geographischen Regionen berechnet. Die Stichprobe des Äneolithikums war zu klein, um dort Gebiete sinnvoll miteinander zu vergleichen, daher sind die Analysen auf die bronzezeitlichen Proben beschränkt. Im Steppenraum wurden mehrere „kulturelle Grenzen“ im Verlauf des Neolithikums bis zur Bronzezeit postuliert (siehe Kapitel 4.2). Innerhalb der Jamnaja-Kultur wurden der Südliche Bug und der Dnjepr als „Grenze“ untersucht, indem die genetischen Distanzen der Gruppen westlich und östlich des jeweiligen Flusses zueinander berechnet wurden (Tabelle 9.14 und Tabelle 9.15). Außerdem wurden die Fundorte in Küstennähe mit denen im Landesinneren verglichen (Tabelle 9.16). Wie auch in den bereits beschriebenen interkulturellen Distanzberechnungen fällt die Region östlich des Dnjepr durch hohe Werte zu verschiedenen Gruppen auf. Eine kleinräumigere Analyse der Jamnaja-Kultur war aufgrund der Verteilung der Fundplätze nicht möglich.

Innerhalb der Katakombengrabkultur sind, zumindest mithilfe von F_{ST} -Werten, keine Populationsstrukturen zu erkennen. Alle durchgeführten Vergleiche ergaben negative Werte.

Tabelle 9.14: Jamnaja-Strukturen - westlich und östlich des Südlichen Bug (F_{ST} -Werte)

Die den Gruppen zugehörigen Fundplatz-Einheiten entsprechen der Nummerierung in Abbildung 9.9. N = Anzahl der Individuen, fettgedruckt = signifikante F_{ST} -Werte ($p < 0,05$).

		Osten	gesamtes Gebiet	Dnjepr (W) – Dnjepr (O)	Dnjepr (W) – Kuban	Dnjepr (O) – Manych	Dnjepr (O) – Kuban
Westen	Einheiten		5,6,7,9,10	5,6,7	5,6,7,9	6,7,9	6,7,9,10
		N	12	9	11	7	8
Gesamtgebiet	2,3,4	10	0,01774	0,03576	0,05138	0,14564	0,07945
Donau	2	6	0,00335	0,02497	0,03462	0,12634	0,06099

Tabelle 9.15: Jamnaja-Strukturen - westlich und östlich des Dnjepr (F_{ST} -Werte)

		Westen	gesamtes Gebiet	Dnjepr (W) – Dnjepr (O)	Dnjepr (W) – Pruth	Dnjepr – Donau	Donau
Osten	Einheiten		2,3,4,5	4,5	3,4,5	2,3,4	2
		N	14	7	8	10	6
Gesamtgebiet	6,7,9,10	8	0,10752	0,11614	0,11933	0,07945	0,06099
Dnjepr – Kuban	6,7,9	7	0,16841	0,18957	0,1886	0,14564	0,12634

Tabelle 9.16: Jamnaja-Strukturen - Küste versus Landesinneres (F_{ST} -Werte)

		Inland	gesamtes Gebiet	Pruth – Dnjepr (O)	Dnjepr (W) – Dnjepr (O)	Dnjepr (W) – Manych
Küsten-Nähe	Einheiten		3,5,6,10	3,5,6	5,6	5,6,10
		N	8	8	6	7
Gesamtgebiet	2,4,7,9	14	0,0279	0,04522	0,03386	0,01739
Pruth – Kuban	4,7,9	8	0,11933	0,13978	0,13062	0,11065
Pruth – Dnjepr	4,7	6	0,12713	0,15867	0,14945	0,11695
Donau–Dnjepr	2,4,7	12	0,01702	0,03975	0,02762	0,00512
Donau – Pruth	2,4	9	-0,03714	0,00235	-0,0117	-0,05309
Donau	2	6	-0,03304	-0,00097	-0,01477	-0,04719

9.3.4 Genetische Distanzen zwischen den Kulturen der nordpontischen Steppe und Mitteleuropas

Um herauszufinden, wie sich die nordpontischen Steppenkulturen in das europäische Neolithikum einfügen, wurden auf der HVR I (np 16059 – 16390) basierende F_{ST} -Werte zu früheren, zeitgleichen und jüngeren mitteleuropäischen Kulturen berechnet. Tabelle 9.17 zeigt außerdem die Distanzen zu bronzezeitlichen westsibirischen und modernen europäischen Populationen. Auffallend sind die großen Abstände zu der neolithischen Jäger/Sammler/Fischer-Gruppe (NHG bzw. NHG2, die sich durch die Zuordnung der Proben aus Ostorf unterscheiden), aber auch zu den Meso- und Palaeolithikern Mitteleuropas (HGC). Prinzipiell sind die F_{ST} -Werte der Katakombengrabkultur zu (fast) allen Vergleichspopulationen höher als die der äneolithischen, der Jamnaja- und der Poltavka-Kultur. Eine Ausnahme ist die Grübchenkeramik (PWC). Zwischen ihr und der Katakombengrabkultur ist die genetische Distanz sehr gering ($F_{ST} = 0,01798$, $p > 0,05$) während die anderen Steppenkulturen mit signifikanten Werten von 0,06 bis 0,15 deutlich von der Grübchenkeramischen Kultur distanziert sind. Ansonsten sind zwischen der Jamnaja-Kultur und den mitteleuropäischen Bauerngesellschaften, die vor allem durch Proben aus der Mittelelbe-Saale Region repräsentiert sind, keine bis geringe Distanzen festzustellen. Keiner der Werte ist signifikant.

Tabelle 9.17: F_{ST} -Werte zwischen den Populationen der Steppe und Westasiens

Die F_{ST} -Werte basieren auf paarweisen Unterschieden in der HVR I (16059-16390). Signifikante Werte ($p < 0,05$) sind fettgedruckt. Je dunkler der blaue Hintergrund, desto größer ist die genetische Distanz. N = Individuenzahl; ÄNL = Äneolithikum; JAM = Jamnaja-; POL = Poltavka-; fKGK/eKGK = frühe/entwickelte Katakombengrabbkultur; J/S/F = Jäger/Sammler/Fischer. Die drei Braun-/Sandtöne geben grob das zeitliche Verhältnis der entsprechenden Populationen zu den bronzezeitlichen Steppenkulturen (JAM, KGK) an; von dunkel nach hell: früher, (teilweise) überlappend, jünger. Für Informationen zu den Vergleichsdatensätzen siehe Referenzen in Tabelle 8.14. Die aus den englischen Kulturnamen abgeleiteten Abkürzungen wurden größtenteils aus den entsprechenden Veröffentlichungen übernommen. NHG2 = NHG ohne die Proben aus Ostorf, die auch der FBC (im Text auch TBK) zugeordnet sind.

Kultur			ÄNL	JAM	POL	KGK	fKGK	eKGK
		N	10	22	6	23	11	12
<i>Mittel- und nordeuropäische Kulturen</i>								
Prä-Neolithiker	HGC	16	0,04472	0,08493	0,10165	0,12369	0,13445	0,11063
Linienbandkeramik	LBK	102	0,05298	0,02807	0,03858	0,08035	0,06526	0,07009
Rössener Kultur	RSC	11	0,02492	-0,01224	-0,00036	0,05831	0,04851	0,03684
Schöninger Gruppe	SCG	33	0,06853	0,01625	0,05518	0,08153	0,05940	0,07249
Neolithische J/S/F	NHG	20	0,08320	0,10510	0,15340	0,15488	0,16615	0,14865
Neolithische J/S/F (2)	NHG2	13	0,20566	0,19688	0,27307	0,26187	0,33470	0,26360
Spätneolith. Bauern	LNF	11	0,05593	0,01546	0,08043	0,07794	0,06725	0,07710
Baalberger Kultur	BAC	19	0,02280	-0,01714	-0,02682	0,06796	0,05042	0,04677
Salzmünder Kultur	SMC	29	0,05452	0,02092	0,06496	0,05861	0,04793	0,04445
Bernburger Kultur	BEC	17	-0,01449	-0,00981	0,01712	0,05025	0,03103	0,03965
Trichterbecherkultur	FBC	10	-0,00385	-0,00478	0,01938	0,07749	0,08241	0,05158
Grübchenkeramik	PWC	19	0,09081	0,06301	0,15263	0,01798	-0,02303	0,04086
Schnurkeramik	CWC	44	0,01567	-0,02142	0,00645	0,03573	0,02073	0,02606
Glockenbecherkultur	BBC	35	0,00150	-0,01366	0,02193	0,01784	0,00152	0,01488
Aunjetitzer Kultur	UNC	34	0,01849	-0,00312	0,00703	0,03183	0,01416	0,02007
<i>Osteuropäische Kulturen</i>								
Prä-Neolithiker	HGE	14	0,06494	0,05142	0,03652	0,06640	0,04622	0,04780
Dnjepr-Donetz-Kultur	DDC	6	0,06006	0,04754	0,03750	0,05839	0,06392	0,03456
<i>West-/Zentralsibirische Kulturen</i>								
Ust-Tartas-Kultur	UTC	17	0,14116	0,12995	0,09426	0,17269	0,15664	0,13751
Odinovo-Kultur	ODI	10	0,25161	0,20705	0,17852	0,26573	0,31783	0,21451
Krotovo-Kultur	KRO	26	0,08904	0,09481	0,05270	0,14687	0,13435	0,11561
Andronovo/Fedorovo	ANF	20	0,11526	0,12992	0,07416	0,18147	0,16674	0,14201
Andronovo-Kultur	AND	9	0,01570	-0,03425	-0,02465	0,02843	0,00595	0,00585
Bronzezeit Kasachstan	BAK	8	-0,00217	-0,02002	-0,05767	0,13080	0,12318	0,08964
<i>Moderne Populationen</i>								
Osteuropäer:	mOEU	246	0,01982	-0,00618	0,03556	0,01966	0,00632	0,01194
Bulgaren	mBul	30	0,02927	-0,01283	0,03109	0,02566	0,00707	0,01694
Ukrainer	mUkr	18	0,00602	-0,02660	0,00445	0,01835	-0,00018	0,00876
Westrussen	mRus	198	0,01865	-0,00471	0,03551	0,01998	0,00795	0,01094
Mitteeuropäer:	mCEU	434	0,02768	-0,00405	0,04024	0,02185	0,00626	0,01683
Deutsche	mGer	262	0,03447	-0,00274	0,05059	0,02373	0,00787	0,01987
Litauer	mLit	97	0,02005	-0,00253	0,03359	0,02122	0,00348	0,01553
Polen	mPol	75	0,01755	-0,01001	0,01128	0,02418	0,01104	0,01152

Die F_{ST} -Werte zu den osteuropäischen Palaeo- und Mesolithikern (HGE) sind bis auf den des Äneolithikums geringer als zu den mitteleuropäischen. Die Distanzen zur direkt aus der Steppe, und damit aus dem Untersuchungsgebiet stammenden neolithischen Dnjepr-Donetz-Kultur hingegen sind größer als zu den mitteleuropäischen Neolithikern. Vor allem aufgrund der geringen Stichprobenprobengröße sind die Werte erwartungsgemäß jedoch nicht signifikant.

Von den ungefähr zeitgleichen westsibirischen Populationen sind die kupfer- und bronzezeitlichen Bewohner der nordpontischen Steppe, was die F_{ST} -Werte betrifft, weitestgehend distanziert. Lediglich zu den beiden Gruppen der Andronovo-Kultur sind die Ergebnisse der F_{ST} -Berechnungen gegensätzlich. Zu einer nicht genauer definierten bronzezeitlichen Kultur in Kasachstan weisen nur die Gruppen der Katakombengrabkultur hohe, signifikante F_{ST} -Werte auf.

Zu den modernen Populationen Mittel- und Osteuropas sind die genetischen Distanzen insgesamt gering. Zwar erreichen die die Katakombengrabkultur betreffenden Werte aufgrund der Stichprobengrößen Signifikanzniveau, dennoch sind die deutlich unter 0,03 liegenden F_{ST} -Werte als klein anzusehen. Einen höheren Wert ergibt nur der Vergleich zwischen dem Äneolithikum und der modernen Population aus Deutschland. Eine vollständige Matrix aller paarweisen genetischen Distanzen unter den Kulturen Westeurasiens findet sich im Anhang (Tabelle 13.11)

Mithilfe eines MDS-Plots wurde das durch paarweise F_{ST} -Werte repräsentierte Beziehungsgeflecht der 26 prähistorischen Populationen Westeurasiens zweidimensional veranschaulicht (Abbildung 9.10). Der Stresswert von 0,0475 zeigt, dass die geometrischen Distanzen die genetischen „gut“ widerspiegeln (siehe S. 100).

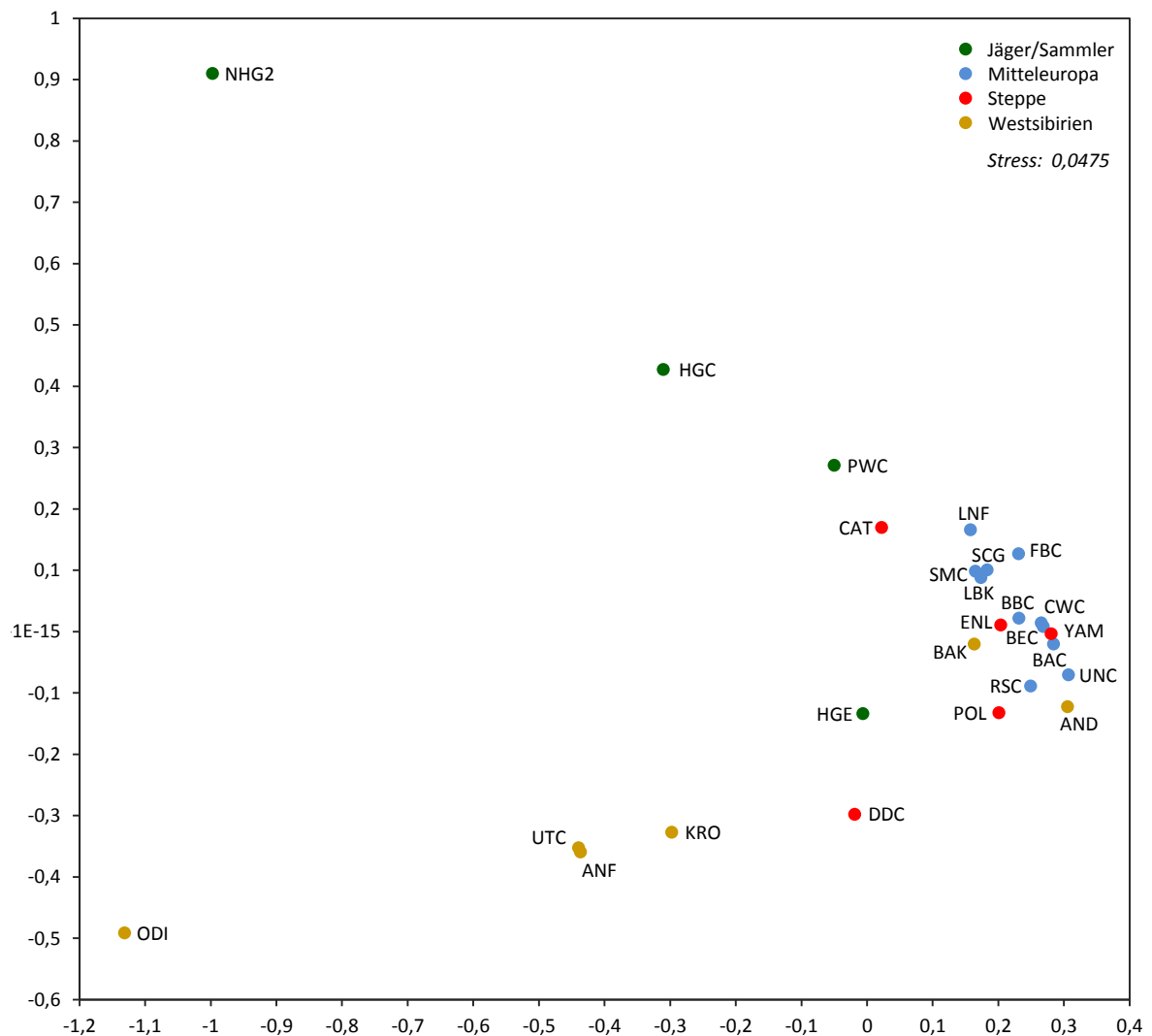


Abbildung 9.10: Multidimensionale Skalierung prähistorischer Populationen Westeurasiens

Der MDS Plot basiert auf paarweisen F_{ST} -Werten (siehe Tabelle 9.17 und Tabelle 13.11). Jäger/Sammler (grün): HGC – mitteleuropäische Prä-Neolithiker, HGE – osteuropäische Prä-Neolithiker, NHG2 – Neolithische Jäger/Sammler/Fischer, PWC – Grübchenkeramik; Bauernkulturen in Mitteleuropa (blau): LBK – Linearbandkeramik, RSC – Rössener Kultur, SCG – Schönninger Gruppe, LNF – Spätneolithische Bauern aus der Blätterhöhle, BAC – Baalberger Kultur, SMC – Salzmünder Kultur, BEC – Bernburger Kultur, FBC – Trichterbecherkultur, CWC – Schnurkeramik, BBC – Glockenbecherkultur, UNC – Aunjetitzer Kultur; Steppe (rot): DDC – Dnepr-Donetz-Kultur, ENL – äneolithische Gruppen, YAM – Jamnaja-Kultur, POL – Poltavka-Kultur, CAT – Katakombengrabkultur; Westsibirien (beige): UTC – Ust-Tartas Kultur, ODI – Odino- (Fedorovo-) Kultur, KRO – Krotovo-Kultur, ANF – Andronovo- (Fedorovo-) Kultur, AND – Andronovo-Kultur, BAK – Bronzezeit Kasachstan.

Ergänzend zu den F_{ST} -Werten wurden die Anteile der sogenannten „*shared haplotypes*“ zwischen Paaren von Populationen berechnet, um mögliche Verwandtschaftsbeziehungen offenzulegen. Die Wahrscheinlichkeit, dass zwei zufällig aus den beiden jeweils verglichenen Populationen gezogenen Haplotypen identisch sind ist in den Balkendiagrammen in Abbildung 9.11 illustriert. Aufgrund der Unterschiede in den Stichprobengrößen haben die Konfidenzintervalle eine große Spannweite, so dass kaum vergleichende Aussagen getroffen, sondern lediglich Tendenzen festgestellt werden können.

Das *haplotype sharing*-Profil der äneolithischen und der Jamnaja-Kultur sind sich sehr ähnlich, außer dass die Werte der ersteren tendenziell etwas höher sind. Beide Populationen teilen vergleichsweise „viele“ Haplotypen mit den Funden der spätneolithischen Bauern aus der Blätterhöhle aus Bollongino *et al.* (2013). Im Gegensatz zur Katakomben-grabkultur, deren Stichprobengröße vergleichbar ist, tragen innerhalb der Jamnaja-Kultur nur wenige Individuen identische Haplotypen. Dies zeigte sich bereits in den Diversitätsmaßen in Tabelle 9.9. Bei der Katakombengrabkultur fällt erneut die Grübchenkeramische Kultur durch einen hohen Anteil gemeinsamer Haplotypen von fast 8 % auf. Die Ergebnisse des *haplotype sharing* korrelieren mäßig mit den F_{ST} -Werten. Der Korrelationskoeffizient liegt zwischen -0,31 bei der Katakombengrab- und -0,63 bei der Jamnaja-Kultur.

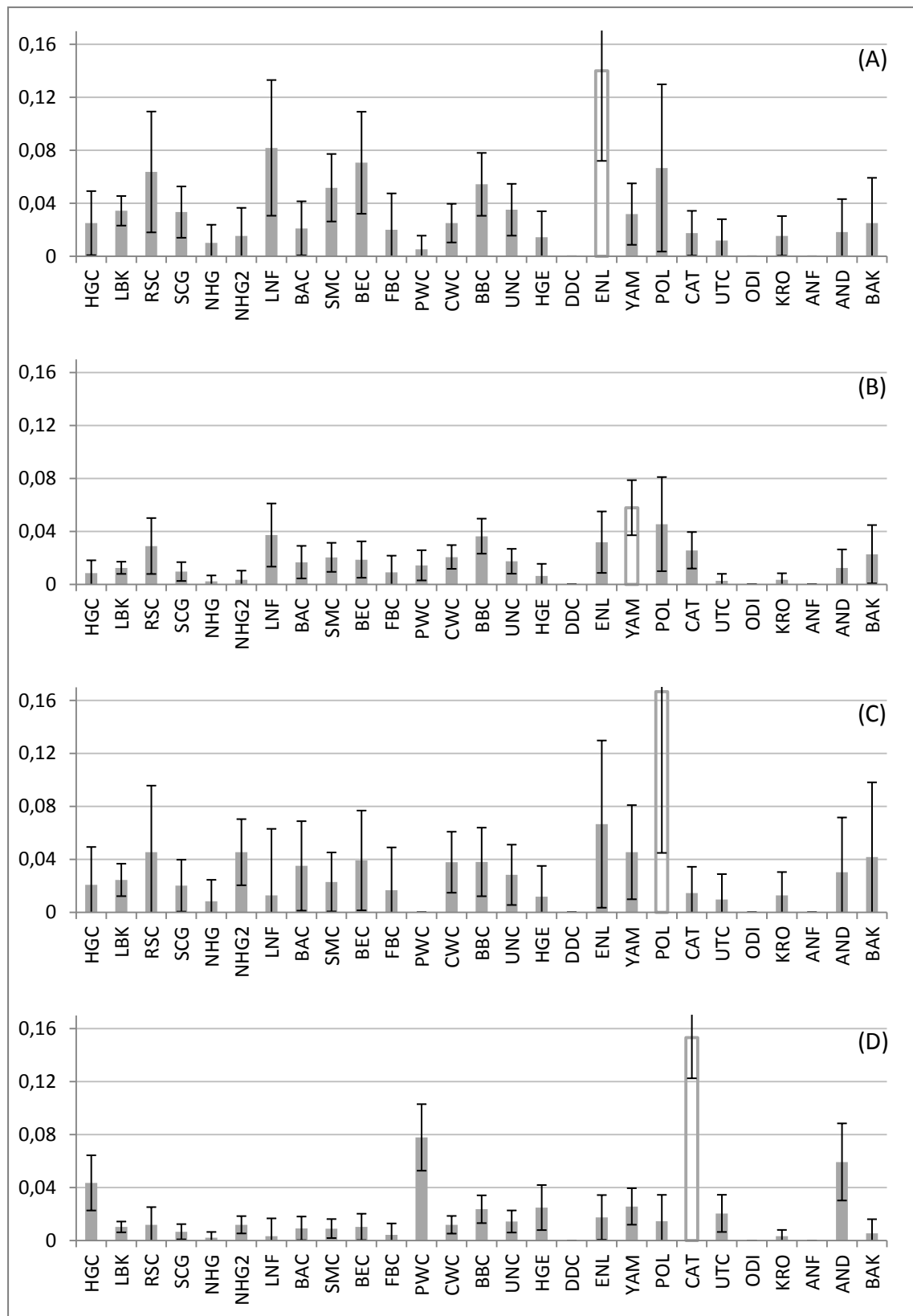


Abbildung 9.11: Haplotype Sharing zwischen den Steppopulationen und diversen Kulturen Westasiens

(A) Äneolithikum, (B) Jamnaja-Kultur, (C) Poltavka-Kultur, (D) Katakombengrabkultur.

Angegeben ist der relative Anteil identischer Haplotypen als Wahrscheinlichkeit, dass zwei zufällige ausgewählte Haplotypen identisch sind (basierend auf der HVR I, 16059 – 16390). Die Antennen zeigen das jeweilige 95 % Konfidenzintervall an.

9.4 Ergebnisse der Analyse der autosomalen Daten

9.4.1 Genotypen- und Allelfrequenzen

Bei der Datenanalyse wurden alle *loci* eines Individuums berücksichtigt, die die in den Abschnitten 8.2.7 und 8.2.8.1 erläuterten Authentizitätskriterien erfüllen, wobei zwischen vorbehaltlich und sicher reproduzierten Genotypen differenziert wurde. Die Genotypen aller Individuen und *loci* findet sich im Anhang (Tabelle 5.1).

Bei drei Individuen wich die Beurteilung jeweils eines *locus* vom festgesetzten Schema ab. OVI 3 wäre für *S04/ABCC11* eigentlich vorbehaltlich reproduziert, aber da eine der drei PCRs neben dem eigentlichen SNP noch vier weitere Mutationen in dem kurzen Fragment aufwies, wurde die Probe ausgeschlossen. POD 2 zeigte für *S12/HERC2* in sechs PCRs das ancestrale Allel (AA), in einer PCR beide Allele (AG). Allel A wurde durch 18, Allel G durch 11 *reads* repräsentiert, womit das 2:1-Kriterium für *reads* nicht erfüllt wurde. Da allerdings das ancestrale Allel 13 Mal beobachtet wurde, und somit das 11:1-Kriterium der Allelpräsenzen übertroffen wurde, gilt die Probe für diesen *locus* als sicher reproduziert. TEM 5 zeigte am *locus S23/Amelogenin* in vier PCRs jeweils XX und YY aus beiden Extrakten. Obwohl kein heterozygoten PCR-Produkt vorliegt, wurde das Individuum vorbehaltlich als männlich eingestuft. Da von PES 7 nur ein Extrakt vorlag, wurde für diese Probe das Authentifizierungsschema modifiziert. Für eine vorbehaltliche bzw. sichere Reproduktion musste jeweils ein zusätzliches widerspruchsfreies PCR-Produkt zu den bei zwei Extrakten geforderten vorliegen, d. h. zum Beispiel fünf anstatt vier PCRs, die ausschließlich Allel A zeigen für einen sicheren Genotyp AA.

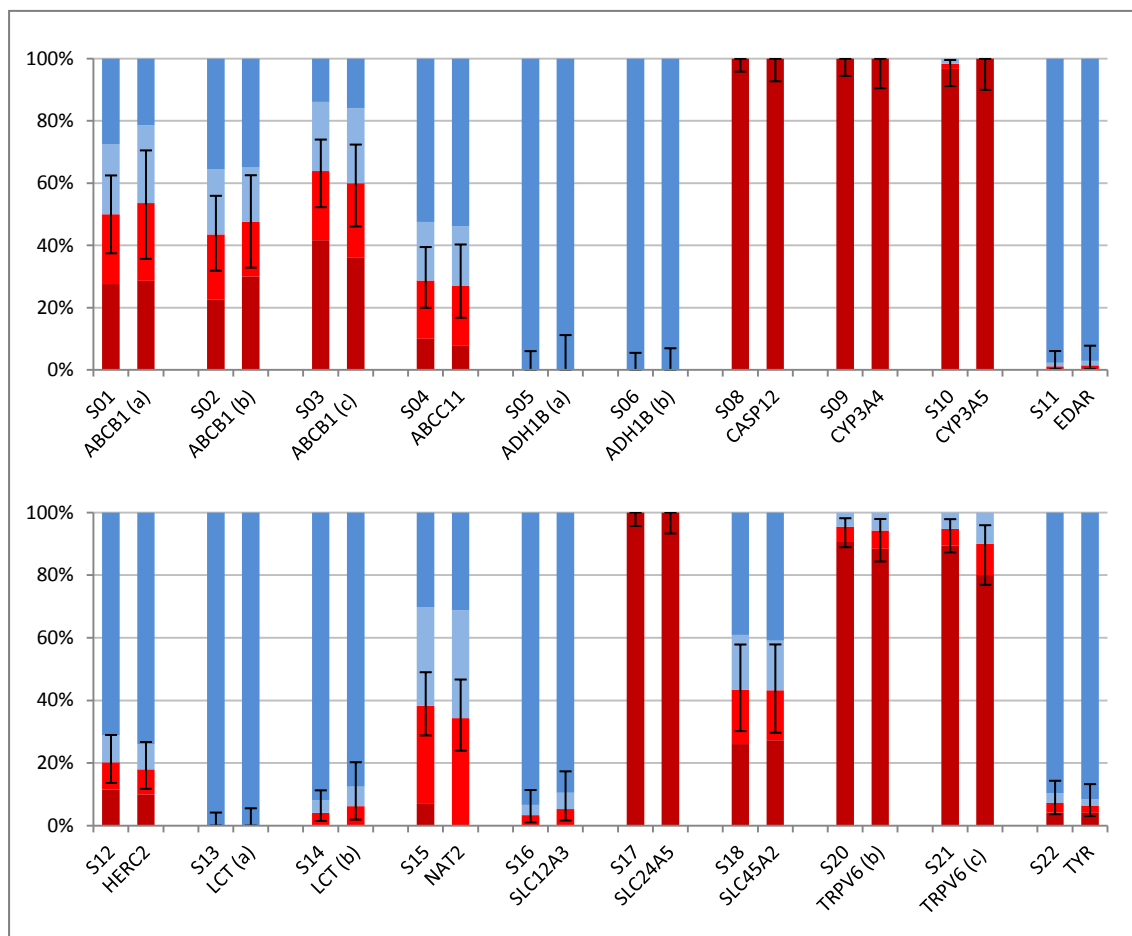


Abbildung 9.12: Vergleich der Allelfrequenzen nach Reproduktionsstatus

Linker Balken = „vorbehaltlich“ (≥ 3 PCRs), rechter Balken = „sicher“ (≥ 4 PCRs) reproduziert (für genauere Informationen siehe Kapitel 8.2.8.1).

rot = Frequenz des abgeleiteten Allels; blau = Frequenz des ancestralen Allels, dunkler Ton = Homozygote, heller Ton = Heterozygote. Antennen = 95% Konfidenzintervall der Allelfrequenz des abgeleiteten Allels (berechnet mit der `qbeta` Funktion in "R"; R Core Team 2013)

Abbildung 9.12 zeigt, dass sich die Genotypenfrequenzen der sicher reproduzierten Daten nur minimal von den nur dreifach reproduzierten unterscheiden. Sofern nicht anders angegeben beruhen die folgenden Beobachtungen und Berechnungen auf den mindestens dreifach reproduzierten Genotypen. Bei S02 in *ABCB1* handelt es sich um einen ternären SNP. Obwohl es abgeleitet ist, wurde das A-Allel als ancestrale gewertet, um den Sonderstatus der unter Selektionsverdacht stehenden T-Variante beizubehalten. In Tabelle 9.18 sind die Genotypen- und Allelfrequenzen im Gesamtdatensatz der im Rahmen dieser Studie analysierten Individuen aufgeführt.

Tabelle 9.18: Genotypen- und Allelfrequenzen der autosomalen loci

Die Frequenzen sind in Prozent angegeben. Grundlage ist der prähistorische Gesamtdatensatz. p = Häufigkeit des ancestralen Allels, q = Häufigkeit des abgeleiteten Allels. Sonderfall S23 (AMELX/Y): p = Häufigkeit des X-Chromosoms, q = Häufigkeit des Y-Chromosoms. In Klammer ist die Anzahl der Individuen (bei den Genotypen) bzw. die der Chromosomen (bei den Allelen) angegeben. HWE = Hardy-Weinberg-Equilibrium, basierend auf einem zweiseitigen Test (Graffelman 2013, R Core Team 2013). Fettgedruckt sind p -Werte $< 0,05$.

Locus (Gen)	Genotypenfrequenzen			Allelfrequenzen		HWE
	pp	pq	qq	p	q	p -Wert
S01 ABCB1 a	27,59 (8)	44,83 (13)	27,59 (8)	50,00 (29)	50,00 (29)	0,7099
S02 ABCB1 b	35,48 (11)	41,94 (13)	22,58 (7)	56,45 (35)	43,55 (27)	0,4692
S03 ABCB1 c	13,89 (5)	44,44 (16)	41,67 (15)	36,11 (26)	63,89 (46)	1,0000
S04 ABCC11	52,50 (21)	37,50 (15)	10,00 (4)	71,25 (57)	28,75 (23)	0,6985
S05 ADH1B a	100,00 (29)	0,00 (0)	0,00 (0)	100,00 (58)	0,00 (0)	1,0000
S06 ADH1B b	100,00 (32)	0,00 (0)	0,00 (0)	100,00 (64)	0,00 (0)	1,0000
S08 CASP12	0,00 (0)	0,00 (0)	100,00 (43)	0,00 (0)	100,00 (86)	1,0000
S09 CYP3A4	0,00 (0)	0,00 (0)	100,00 (32)	0,00 (0)	100,00 (64)	1,0000
S10 CYP3A5	0,00 (0)	3,33 (1)	96,67 (29)	1,67 (1)	98,33 (59)	1,0000
S11 EDAR	97,73 (43)	2,27 (1)	0,00 (0)	98,86 (87)	1,14 (1)	1,0000
S12 HERC2	71,15 (37)	17,31 (9)	11,54 (6)	79,81 (83)	20,19 (21)	0,0024
S13 LCT a	100,00 (43)	0,00 (0)	0,00 (0)	100,00 (86)	0,00 (0)	1,0000
S14 LCT b	91,89 (34)	8,11 (3)	0,00 (0)	95,95 (71)	4,05 (3)	1,0000
S15 NAT2	30,23 (13)	62,79 (27)	6,98 (3)	61,63 (53)	38,37 (33)	0,0534
S16 SLC12A3	93,33 (28)	6,67 (2)	0,00 (0)	96,67 (58)	3,33 (2)	1,0000
S17 SLC24A5	0,00 (0)	0,00 (0)	100,00 (41)	0,00 (0)	100,00 (82)	1,0000
S18 SLC45A2	39,13 (9)	34,78 (8)	26,09 (6)	56,52 (26)	43,48 (20)	0,2047
S20 TRPV6 b	0,00 (0)	9,09 (4)	90,91 (40)	4,55 (4)	95,45 (84)	1,0000
S21 TRPV6 c	0,00 (0)	10,53 (4)	89,47 (34)	5,26 (4)	94,74 (72)	1,0000
S22 TYR	89,58 (43)	6,25 (3)	4,17 (2)	92,71 (89)	7,29 (7)	0,0116
S23 AMELX/Y	42,31 (11)	57,69 (15)				

Die hier präsentierten Allel- und Genotypenfrequenzen sowie die p -Werte des Hardy-Weinberg-Tests weichen in geringem Maße zu den in Wilde *et al.* (2014) publizierten Daten ab. Dies liegt unter anderem daran, dass die Proben der Kugelamphorenkultur und vier weitere Genotypen nicht in die Publikation einfließen. Die Genotypen, die bereits in diesem Artikel veröffentlicht wurden, sind in Tabelle 13.12 des Anhangs gekennzeichnet.

Abbildung 9.13 auf den folgenden Seiten zeigt die Genotypen- und Allelfrequenzen der SNPs in separaten Panels, wobei zwischen den einzelnen Populationen (Kulturen) differenziert wird, und zum Vergleich der moderne Datensatz ukrainischer DNA-Proben aus Hollfelder (2013) aufgezeichnet ist.

Sechs *loci* sind im prähistorischen Datensatz fixiert: *ADH1B* (S05 und S06) und *LCT* (S13) zeigen lediglich die ancestrale Variante. Von *CASP12* (S08), *CYP3A4* (S09) und *SLC24A5* (S17) ist das jeweils abgeleitete Allel fixiert. Signifikante Unterschiede in den Allelfrequenzen zwischen alten und modernen Daten sind für *ABCC11* (S04), *HERC2* (S12), *LCT* (S13 und S14), *SLC45A2* (S18) und *TYR* (S22) zu beobachten. Mit Ausnahme von *ABCC11* ist die Häufigkeit des jeweils abgeleiteten Allels heute höher. Zwischen der Jamnaja- und der Katakombengrabkultur sind vereinzelt ebenfalls Veränderungen festzustellen, die allerdings nicht signifikant sind. Die drei *ABCB1*-SNPs (S01 – S03) zeigen eine leichte Abnahme des abgeleiteten Allels. Deutlicher ist diese Abnahme bei *ABCC11* (S04) und *SLC45A2* (S18), wobei vor allem die homozygot abgeleiteten Varianten verschwinden bzw. selten werden, während die Heterozygoten bei *ABCC11* zunehmen und bei *SLC45A2* ungefähr gleichbleiben. Die Häufigkeiten in den Stichproben des Äneolithikums, der Poltavka- und der Kugelamphorenkultur sind der Vollständigkeit halber und für einen Überblick in der Abbildung 9.13 mit aufgetragen. Aufgrund der geringen Chromosomenanzahlen sind hier keine interpretatorischen Aussagen zu machen.

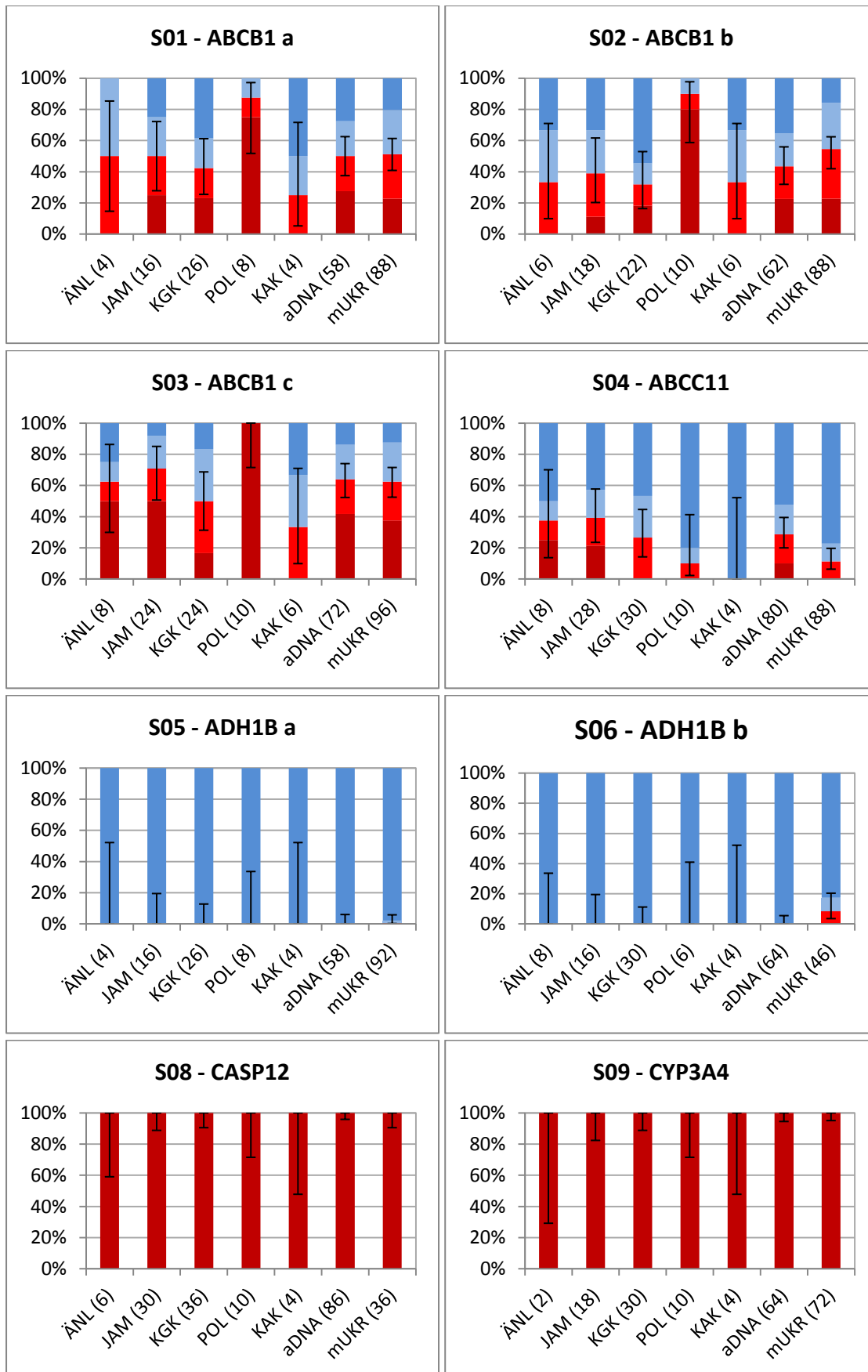


Abbildung 9.13: Genotypen- und Allelfrequenzen der kupfer- und bronzezeitlichen Kulturen sowie der modernen Population der Ukraine (S01 – S09; Fortsetzung auf der nächsten Seite)

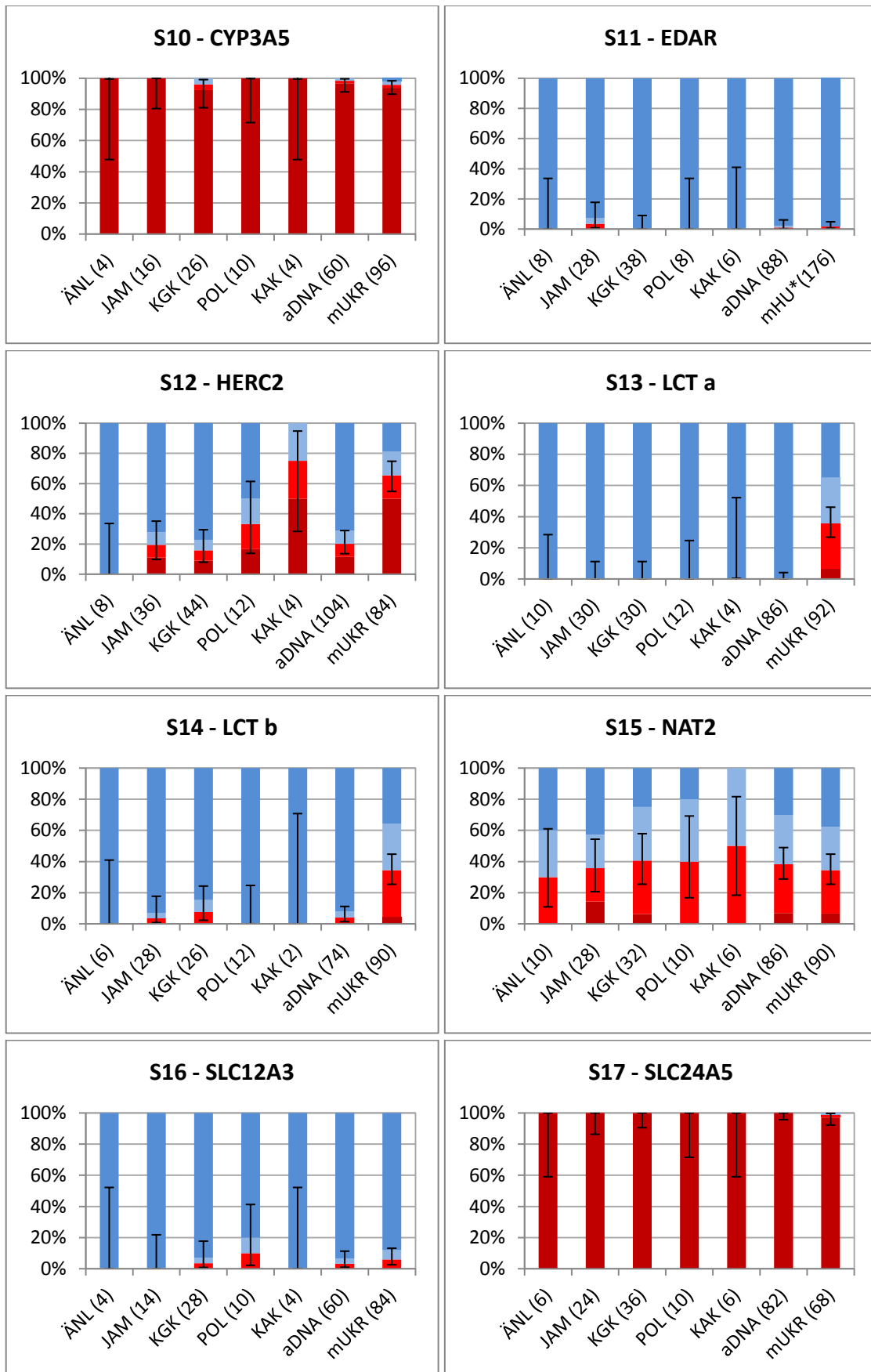


Abbildung 9.13: Genotypen- und Allelfrequenzen der kupfer- und bronzezeitlichen Kulturen sowie der modernen Population der Ukraine (S10 – S17; Fortsetzung auf der nächsten Seite)

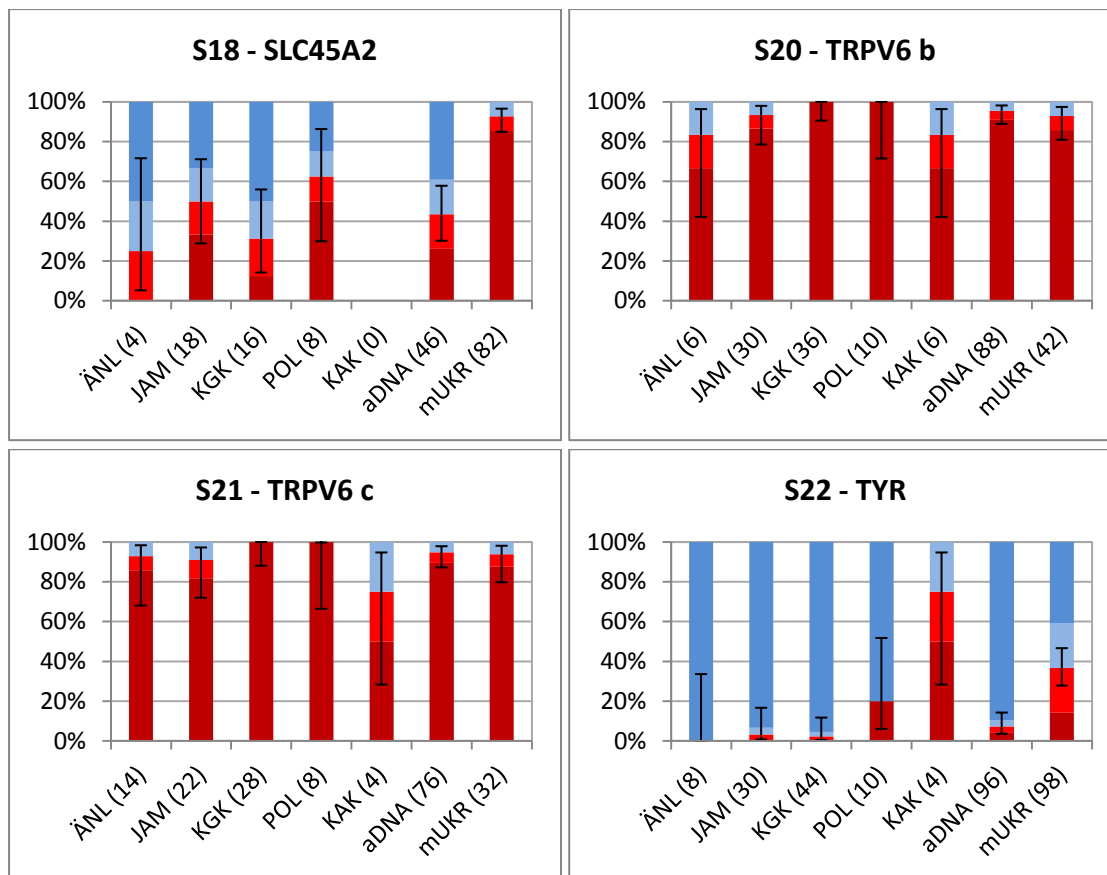


Abbildung 9.13: Genotypen- und Allelfrequenzen der kupfer- und bronzezeitlichen Kulturen sowie der modernen Population der Ukraine

rot = Frequenz des abgeleiteten Allels; blau = Frequenz des ancestralen Allels, dunkler Ton = Homozygote, heller Ton = Heterozygote. Antennen = 95 % Konfidenzintervall der Allelfrequenz des abgeleiteten Allels (berechnet mit der qbeta Funktion in "R"). ÄNL = Äneolithikum, JAM = Jamnaja-Kultur, KGK = Katakombengrabkultur, POL = Poltavka-Kultur, KAK = Kugelamphorenkultur, aDNA = Gesamtdatensatz der alten Proben, mUKR = moderne ukrainische Population (aus Hollfelder 2013), Zahl in Klammer = Anzahl der Chromosomen (2N).

In den Genen *ABCB1*, *ADH1B*, *LCT* und *TRPV6* wurden zwei oder drei SNPs untersucht. Innerhalb von Genen ist ein gewisses Maß an genetischer Kopplung zu erwarten. Daher wurden für *ABCB1* die beobachteten Allel-Kombinationen ausgezählt (Tabelle 9.19). Bei den anderen drei Genen wurde darauf verzichtet, da sie entweder für ein Allel fixiert, oder nahe der Fixierung sind.

Beobachteter Genotyp	N
TT – TT – TT	6
CC – GG – CC	1
CT – AT – CT	1
CT – GT – CT	6
CT – GT – TT	2
CC – GG – CT	3

Tabelle 9.19: Genetische Kopplung in ABCB1

Die Genotypen setzen sich von links nach rechts aus den Ergebnissen für S01 – S02 – S03 zusammen. Es wurden für diese Tabelle nur Proben berücksichtigt, von denen für alle drei SNPs Genotypen vorliegen. Die gametischen Phasen sind nicht bekannt. N = Anzahl der Individuen.

9.4.2 Genetische Distanzen

Basierend auf den Genotypenfrequenzen wurden die F_{ST} -Werte zwischen dem prähistorischen Gesamtdatensatz aus dieser Arbeit und der Stichprobe heutiger Ukrainer aus Hollfelder (2013) für die einzelnen SNPs berechnet. Sie sind in Tabelle 9.20 aufgeführt.

Tabelle 9.20: Genetische Distanzen der autosomalen loci

Angegeben sind die F_{ST} -Werte zwischen dem prähistorischen Gesamtdatensatz und den modernen Ukrainern (Hollfelder 2013). Sie basieren auf den Genotypen der einzelnen SNPs. Signifikante Werte sind fettgedruckt. Zu EDAR liegen keine modernen ukrainischen Daten vor.

<i>Locus</i>	F_{ST}	<i>Locus</i>	F_{ST}	<i>Locus</i>	F_{ST}
S01 ABCB1 a	-0,01425	S09 CYP3A4	0,00000	S16 SLC12A3	-0,00699
S02 ABCB1 b	0,00423	S10 CYP3A5	-0,00356	S17 SLC24A5	0,00278
S03 ABCB1 c	-0,01188	S11 EDAR		S18 SLC45A2	0,47018
S04 ABCC11	0,08040	S12 HERC2	0,34373	S20 TRPV6 b	-0,01117
S05 ADH1B a	-0,00524	S13 LCT a	0,33002	S21 TRPV6 c	-0,02176
S06 ADH1B b	0,08519	S14 LCT b	0,23672	S22 TYR	0,68731
S08 CASP12	0,00000	S15 NAT2	-0,00813		

10 Diskussion

Die Methodendiskussion umfasst die Evaluierung der einzelnen Schritte der Probenbearbeitung und Datenanalyse in Bezug auf Ihre Effizienz und Effektivität zur Beantwortung der populationsgenetischen Fragen.

Die inhaltliche Aufarbeitung der Ergebnisse dieser Arbeit unterteilt sich in drei sich überlappende Schwerpunkte. Zunächst wird die Steppe im Kontext des europäischen Neolithikums betrachtet, gefolgt von einer differenzierten Betrachtung der steppeninternen Entwicklungen. Im Anschluss an die populationsgenetischen Betrachtungen werden die Ergebnisse zu den autosomalen Markern und die zugehörigen Selektionshypothesen diskutiert.

10.1 Methodendiskussion

10.1.1 Die Kombination von Multiplex-PCR und 454-Sequenzierung

Das Standardverfahren zur Amplifikation und Sequenzierung von alter DNA war bisher die Polymerase-Kettenreaktion gefolgt von einer Sequenzierung nach Sanger *et al.* (1977), optional mit einer zwischengeschalteten Klonierung der PCR-Produkte. Die Anzahl von analysierbaren *loci* wird hier vor allem durch die Extraktmenge und den zeitlichen Aufwand begrenzt. Das ursprünglich als *shotgun*-Methode entwickelte *next-generation sequencing* erlaubt hingegen die Analyse großer Teile des Genoms. Zu Beginn dieses Projektes war das NGS auf alte DNA nicht effizient anzuwenden. Zum einen besteht das DNA-Extrakt einer prähistorischen Probe, falls es nicht gerade aus dem Permafrost stammt, zu 90 – 100 % aus mikrobieller DNA (Green *et al.* 2006, Briggs *et al.* 2009, Burbano *et al.* 2010, Green *et al.* 2010), wobei die Mengen an endogener DNA tendenziell zu gering für eine *shotgun*-Sequenzierung sind. Zum anderen ist das Risiko, Kontaminationen in die Probe einzutragen, extrem hoch, wenn mehrere enzymatische Reaktionen und Aufreinigungsschritte an reinem Extrakt

durchgeführt werden³⁵. Auch ist die Sequenzierung ganzer Populationen (anstatt einzelner Individuen) mit extrem hohen Kosten verbunden. In den letzten Jahren wurden Methoden entwickelt, *loci* von Interesse im Extrakt anzureichern, um eine spezifischere Sequenzierung zu erreichen. Damit wurde auch aDNA zugänglich für NGS.

Auf der methodischen Seite stellt die vorliegende Arbeit sozusagen die Vorstufe zu den aktuellen aDNA-NGS-Methoden dar. Als das *DNA capture* noch nicht auf aDNA angewandt werden konnte (vor allem aufgrund zu geringer Ausgangsmengen), wurde mit der kombinierten Multiplex-PCR-454-Sequenzierung (im Folgenden PCR-454) ein Mittelweg gefunden, eine Vielzahl von *loci* gleichzeitig zu amplifizieren und sequenzieren.

Für den Schritt der Amplifikation der mitochondrialen *loci* wurde das von M. U. entwickelte und zur Verfügung gestellte Protokoll mit dem QIAGEN Multiplex PCR Kit mit kleinen Modifikationen erfolgreich übernommen. Die Streuungen im Amplifikationserfolg^{5,36} zwischen den einzelnen Primersystemen korrelieren nur mittelmäßig mit den Längen oder dem GC-Gehalt der Amplikons. Auffällig ist jedoch, dass der komplett ausgefallene *locus* M18 einen GC-Gehalt von nur 28,97 % hat. Andererseits war M20 trotz ähnlich geringem GC-Gehalt überdurchschnittlich erfolgreich. Die Unterschiede im Amplifikationserfolg⁵ dürften neben Fragmentlänge und GC Gehalt wahrscheinlich auf mehrere andere Faktoren wie Zufall, Interaktionen zwischen den Primersystemen innerhalb der Multiplex-PCR-Ansätze und Erhaltungszustand der Proben zurückzuführen sein. Im Fall der HVR I, die bis auf M36 durchweg unterdurchschnittlich repräsentiert ist, könnte die direkte Nachbarschaft der Amplikons eine Rolle spielen.

Bei der nukleären Multiplex-PCR war der Reaktionsansatz mit der AmpliTaq Gold Polymerase effizienter als das Setup mit dem QIAGEN Multiplex Kit. Möglicherweise ist die AmpliTaq Gold sensitiver für geringe DNA-Mengen. Die Streuung zwischen den

³⁵ Mittlerweile wurde im Mainzer Palaeogenetik-Labor demonstriert, dass dieses Risiko mithilfe entsprechender Vorkehrungen minimiert werden kann.

³⁶ Anders als beim „Amplifikationserfolg“, in den alle PCRs eingingen, beruht der „Amplifikationserfolg“⁵ rein auf den sequenzierten PCRs (siehe Kapitel 9.1).

einzelnen Primersystemen war geringer als bei der mtDNA. Sie hing, insbesondere im Fall von S18 wahrscheinlich vor allem von Fehlbindungen aufgrund von Sequenzähnlichkeiten bzw. -komplementaritäten zwischen den verschiedenen Amplikons oder Primern ab. Stärker als bei der mtDNA-Multiplex-PCR wurden nach den nukleären PCRs Primerdimere im Gelbild und unter den *454 reads*, die nicht zugeordnet werden konnten, beobachtet.

Das *tagging*-Protokoll nach Meyer *et al.* (2008) beinhaltet eine Reihe von Aufreinigungsschritten, die mit deutlichen Verlusten von DNA einhergehen. Das Protokoll wurde modifiziert, um diese Schritte zu minimieren. Ob sich dies positiv oder negativ auf den Sequenziererfolg auswirkte, kann nicht beurteilt werden. Die Ergebnisse der einzelnen *454 runs* waren extrem variabel. Über die wahrscheinlich vielzähligen Gründe kann nur spekuliert werden, vor allem, da die Sequenzierung und deren vorbereitenden Schritte extern, in einem Labor, das auf moderne DNA ausgerichtet ist, und somit außerhalb der Kontrolle der Palaeogenetiker durchgeführt wurden. Aufgrund dieser externen Weiterbearbeitung der Proben mit erneutem *blunt-end repair* wurden auch Fragmente sequenziert, die keinen Meyer-*tag* trugen, oder „Sequenz-Chimäre“, d. h. zwei oder drei aneinander ligierte Amplikons. Eine direkte *454*-Adapter-Ligation nach dem *tagging*, ohne erneutes *blunt-end repair*, würde die Ausbeute erhöhen.

Auffallend waren auch die deutlichen Unterschiede in den Sequenziererfolgen innerhalb eines *runs* und innerhalb einer Probe. Bereits im DNA-Extrakt liegen, abhängig vom Erhaltungszustand und der Sequenzkomposition, homologe DNA Fragmente in unterschiedlicher Zahl vor. Diese Tatsache, sowie Zufall und die Kombination von Primersystemen führen in der Multiplex-PCR zu einer ungleichen Amplifikation der einzelnen Amplikons. Während man bei moderner DNA die Längen der Amplikons so wählen kann, dass die PCR-Produkte durch Gelelektrophorese unterschieden, und so die Effizienz kontrolliert werden kann, ist dies bei aDNA nicht möglich. Hier sollten die Amplikons möglichst 100 bis 200 Nukleotide nicht übersteigen. Selbst die Bestimmung der Gesamtkonzentration mittels Qubit® Fluorometer ist nicht aussagekräftig, da auch Primerdimere oder bakterielle DNA mitgemessen werden. Jedoch ist Äquimolarität der eingesetzten DNA-Mengen sowohl

beim *tagging*, als auch bei der 454-Sequenzierung die Grundlage für eine etwa gleichmäßige Abdeckung aller *loci*. Auf diese Weise potenzieren sich die Unterschiede vom Anfang der Bearbeitungskette bis zur Sequenzierung.

Überraschenderweise wurden bei 454-Sequenzierungen in einigen *reads* ambige Positionen beobachtet (Y, R, N), was durch die klonale Amplifikation und die Einzelzugabe der Nukleotide bei der Sequenzierung nicht zu erwarten wäre. Ein bekanntes Problem der 454-Sequenzierung hingegen ist die Identifizierung der Länge von Homopolymeren. Dieses Problem nicht, wie vom Hersteller angegeben, erst bei *stretches* von sechs bis acht Basen auf, sondern z. T. bereits bei zwei bis drei. Dadurch wurde die primäre Datenauswertung erschwert, insbesondere in Fällen bei denen der zu untersuchende SNP in oder direkt neben einem Homopolymer lag (z. B. beim Laktase SNP C-13 910T). Ein bedeutender Vorteil der 454-Sequenzierung wiederum betraf die nukleäre DNA. Die Sanger-Sequenzierung funktioniert schlecht bei sehr kurzen Fragmenten. Die „*peaks*“ sind oft nicht eindeutig, teilweise verschoben oder überlagern sich. NGS-Methoden sind (und waren es insbesondere in ihrer Anfangsphase) prädestiniert für die Sequenzierung kurzer Fragmente.

Insgesamt war die Etablierung der PCR-454-Methode als Brücke zwischen konventionellen und *next-generation sequencing*-Verfahren erfolgreich. Sie muss klar als Übergangslösung angesehen werden, die mittlerweile von effizienteren *target enrichment*-Methoden abgelöst wurde. Daher wurde der Fokus letztendlich von der Methodenoptimierung zur Datenproduktion verlagert. Zunächst wurden alle zu untersuchenden Proben mit dem PCR-454-Verfahren bearbeitet, anschließend mussten jedoch die in der aDNA-Forschung unausweichlichen *locus*-Ausfälle mit konventioneller PCR und Sanger-Sequenzierung kompensiert werden. Dies betraf vor allem die Fragmente der HVR I, auf der die Hauptlast der populationsgenetischen Analysen ruht. Bei fast alle Proben mussten mehrere Sanger-Sequenzierungen durchgeführt werden, um die HVR I-Haplotypen zu ermitteln oder zu verifizieren.

Die Beurteilung eines solchen Verfahrens wie PCR-454 kann nicht rein auf dem Amplifikations- und Sequenziererfolg beruhen, sondern auch die „Resistenz“ gegen Kontaminationen spielt eine Rolle. Hier scheint die PCR-454-Strategie gegenüber der

konventionellen Methode etwas im Vorteil zu sein. Im Prinzip ist das Risiko, Fremd-DNA in Proben einzutragen genauso hoch wie bei einer PCR mit Direktsequenzierung. Das *tagging* fand in einem zweckbestimmten Reinraum an den bereits amplifizierten Produkten statt, so dass eine Kontamination im Verlauf des *taggings* sich nicht durchsetzen wird. Ab dem zweiten Arbeitsschritt waren die Proben mit spezifischen *tags* versehen. Die größte Gefahr war eine Kreuzkontamination der Meyer-*tags*, weshalb diese im Spurenlabor angesetzt und aliquotiert wurden. Der Vorteil beim PCR-454 bezüglich Kontaminationen lag in der bereits erwähnten Streuung der Sequenziererfolge. Äquimolarität der gleichzeitig zu sequenzierenden Proben und *loci* ist eine wichtige Voraussetzung für eine möglichst gute Abdeckung bei der 454-Sequenzierung. Die Wahrscheinlichkeit ist hoch, dass eine durch wenige *templates* vertretene Probe unterdrückt bzw. verdrängt wird. Dies gilt dementsprechend ebenfalls für sporadische kontaminierende Sequenzen. Da außerdem kurze Sequenzen bevorzugt zum Sequenziererfolg führen, hat aDNA einen Vorteil gegenüber modernen Kontaminanten, die noch nicht oder wenig degradiert sind. In Fällen, in denen Sequenzen aus Leerkontrollen gewonnen wurden, lag die Anzahl der *reads* oft unter der festgesetzten Grenze von 3 (mtDNA) bzw. 10 (ncDNA) *reads* zur Anerkennung des Allels.

Das Phänomen der nicht verwendeten *tags*, die als Folge der Sequenzkollabierung bei der Sortierung der Rohdaten *reads* aufwiesen, führte nicht zu Fehltypisierungen. Zwar ist zu erwarten, dass sich auch verwendeten *tags* falsche *reads* zugeordnet wurden, aber basierend auf den Beobachtungen bei nicht verwendeten *tags* dürfte diese Fehlerrate verschwindend gering sein. Die durchschnittlich 1,88 Sequenzen pro unbenutzten *tag* (d. h. verteilt auf 21, 23 oder 37 *loci*) liegen weit unter dem Limit von 3 bzw. 10 *reads* pro *locus*. Mit einer Weiterentwicklung des Skripts könnte man dieses Phänomen wahrscheinlich verhindern.

Insgesamt konnte die standardmäßige Ergebnisreproduktion in den meisten Fällen Kontaminationen von authentischen Haplo- oder Genotypen diskriminieren (siehe Kapitel 9.1.4).

10.1.2 Die Auswahl der genetischen Marker

Die Hypervariable Region I ist Gegenstand fast jeder DNA-Studie an prähistorischem Material. Zu den laborseitigen Hintergründen zählen dabei hauptsächlich die hohe Kopienzahl, mit der sie in Zellen vorliegt, und die damit verbundene größere Chance auf Amplifizierbarkeit als nukleäre DNA. Auf der populationsgenetischen Ebene erlaubt die rein maternale Vererbung die relativ geradlinige Nachverfolgung demographischer Prozesse ohne Verkomplizierung durch biparentale Vererbung z. B. durch Rekombinationsereignisse. Auf der anderen Seite bedeutet dies, dass nur ein einseitiges, möglicherweise verzerrtes Bild der Geschichte einer Population gezeichnet wird. Bei den SNPs in der *coding region* handelte es sich um kladendeterminierende Mutationen. Daher dienten sie eher als unterstützender Authentizitätsnachweis, nicht als zusätzliche Information zur genetischen Konstitution eines Individuums. In einigen Fällen konnten sie das auf der HVR I basierende Ergebnis noch besser auflösen (z. B. die Einordnung von TEM 5 und TEM 6 in die Haplogruppe R1). In die populationsgenetische Analyse gingen die Sequenzen der *coding region* allerdings nicht ein, vor allem, da hier die entsprechenden Vergleiche in anderen aDNA-Datensätzen fehlten. Ihr Beitrag wäre minimal gewesen, da die kladendeterminierenden SNPs mit denen der HVR I gekoppelt sind. Lediglich einige zusätzliche Mutationen in den Amplikons einzelner Individuen erlauben noch genauere Charakterisierung.

Die nukleären Marker wurden unter den in Kapitel 5.4 beschriebenen Kriterien ausgewählt. Zwar sind in der Literatur viele Angaben zu genomischen Regionen mit Selektionssignalen zu finden, doch die Identifizierung betroffener Mutationen und ihrer phänotypischen Auswirkung ist nur in wenigen Fällen gelungen (z. B. *Laktase*, *SLC24A5*, *EPAS1*). Daher war die Auswahl an SNPs von vorneherein begrenzt. Populationsgenetisch konnten die Ergebnisse der Genotypisierungen Hinweise auf europäische oder asiatische Einflüsse geben. Außerdem wurden die in der Literatur für die SNPs aufgestellten Selektionshypothesen überprüft. Allerdings fehlen für eine effektive Auswertung neutrale Marker zum Vergleich. Die Untersuchung neutraler *loci* war für dieses Projekt geplant. Dabei sollte eine weitere Multiplex-PCR von einem

dritten Wissenschaftler entwickelt werden. Doch aus zeitlichen und organisatorischen Gründen kam dieses methodische Teilprojekt nicht zustande.

Für ein vollständiges Bild der Populationsgeschichte wäre noch zusätzlich die Betrachtung der paternalen Seite sinnvoll. Hierfür wurde von Silja Dillenberger eine Y-chromosomale Multiplex-PCR entwickelt (2013). Doch leider konnte diese aus zeitlichen Gründen nicht mehr integriert werden.

10.1.3 Datenanalyse

Die Vorbereitung der Sequenzierdaten auf die primäre Sequenzanalyse wurde mithilfe eines *Perl*-Skripts von Benjamin Rieger („sort“ 1 bis 3) durchgeführt. Damit wurden die *reads* nach Meyer-tags und Primersequenzen sortiert und diese dann abgeschnitten. Aufgrund des bekannten Problems der 454-Sequenzierung, die Länge von Homopolymeren zu erkennen, wurden diese sowohl in den Sequenzdaten als auch den Primern und Meyer-tags zu Einzelbasen kollabiert, bevor der Abgleich für die Sortierung und Beschneidung stattfand. Da lediglich vom 5'-Ende her nach diesen Oligonukleotiden gesucht wurde, war in den meisten Fällen der *tag* und der Primer am 3'-Ende noch vorhanden. Außerdem wurden dadurch *reads* verworfen, die nur am 3'-Ende einen Meyer-tag trugen. Obwohl das Skript noch nicht ausgereift war, erfüllte es seinen Zweck, und ein Großteil der Sequenzen konnte identifiziert werden.

Für die primäre Auswertung der sortierten Sequenzen zur Feststellung der Haplo- und Genotypen wurden verschiedene Vorgehensweisen mit dem Programm Seqman Pro™ [GATC BIOTECH AG und DNASTAR, Inc.] und mit dem Programm NextGENe® [SOFTGENETICS] getestet. Problematisch waren nicht abgeschnittene Meyer-tags, Homopolymere, Sequenzähnlichkeiten zwischen nukleären *loci*, Sequenzchimäre und Primerdimere. Die in Kapitel 8.2.8.1 beschriebene Vorgehensweise mit Seqman Pro™ war zwar zeit- und rechenaufwendig, lieferte aber die größte Sequenzausbeute und das genaueste *alignment* der *reads*.

Für die biostatistische Datenanalyse wurden, nach einer grundlegenden Charakterisierung des Gesamtdatensatzes auf der Haplogruppenebene und mit

Diversitätsstatistiken, die Proben in Gruppen eingeteilt. Diese basierten vorrangig auf archäologischen Grundlagen. Zunächst wurden kulturelle Gruppen gebildet. Die Zuordnung von Individuen zu Kulturen ist insbesondere zwischen den beiden Hauptgruppen Jamnaja- und Katakombengrabkultur durch die unterschiedlichen Grabtypen relativ eindeutig. Falls es nicht bei der Probenlagerung in den entsprechenden archäologischen Instituten zu Verwechslungen oder Vermischungen kam, kann man ziemlich sicher sein, dass die Proben korrekt klassifiziert wurden. Weiterhin wurden regionale Gruppen gebildet. Aufgrund der Verteilung der Funde im Untersuchungsgebiet konnte hier nur bedingt auf kulturhistorische Hintergrundinformationen zurückgegriffen werden. Die verschiedenen Gruppengruppenzusammenstellungen beruhten auf Vorgängen in den vorbronzezeitlichen Stufen wie z.B. der Ausbreitungsgebiete neo- oder äneolithischer Kulturen, und geographischen bzw. ökologischen Gegebenheiten im Steppenraum, z. B. eine Einordnung nach Einzugsgebieten der Flüsse oder Zwischenstromgebiete, oder nach Küstennähe. Problematisch erwiesen sich dabei die zum Teil sehr kleinen Gruppengrößen, die nur bedingt oder gar nicht für statistische Analysen geeignet sind. Bei kleinen Gruppen von beispielsweise fünf Individuen wurden mehrfach extrem hohe F_{ST} -Werte, die das Signifikanzniveau erreichten, beobachtet. Hier handelt es sich mit an Sicherheit grenzende Wahrscheinlichkeit um artifizielle Distanzen. Es wurde eine Mindestanzahl an Individuen von sieben als unteres Limit angesetzt. Ausnahme war die Poltavka-Kultur mit insgesamt nur sechs Individuen. Die vier Proben der Kugelamphorenkultur wurden lediglich im Gesamtdatensatz betrachtet, von den detaillierten Analysen jedoch ausgeschlossen.

Die F_{ST} -Werte wurden basierend auf paarweisen Unterschieden berechnet, die korrekte Bezeichnung wäre daher eigentlich Φ_{ST} . Im Gegensatz zu F_{ST} -Berechnungen auf der Grundlage von Haplotypenfrequenzen, bei denen lediglich festgestellt wird, ob zwei Haplotypen gleich oder ungleich sind, werden hier Sequenzähnlichkeiten und damit phylogenetische Verwandtschaften berücksichtigt. Brandt *et al.* (2013) verwendeten das Substitutionsmodell von Tamura & Nei (1993) für ihre F_{ST} -Berechnungen, wobei sie für die verschiedenen Vergleiche unterschiedliche durch das Programm jModeltest ermittelte γ -Werte verwendeten. Dieser Ansatz ist durchaus

geeignet für die Kontrollregion der mtDNA, da es die Heterogenität der Evolution berücksichtigt (Weiss & von Haeseler 1998, Excoffier & Yang 1999). Dennoch wurde (wie auch in Bramanti *et al.* (2009) und Bollongino *et al.* (2013)) die Methode der paarweisen Unterschiede gewählt, um eine direkte Vergleichbarkeit zwischen einzelnen F_{ST} -Werten zu erreichen und diese nicht mit populationspezifischen Parametern (γ -Werten) zu belegen. Ziel war es, ein Maß für die genetischen Distanzen zwischen paarweisen Populationen zu erhalten, und diese miteinander vergleichen zu können.

Für die Vergleiche des vorliegenden Datensatzes mit den prähistorischen Populationen Mitteleuropas und den modernen Bevölkerungsgruppen mussten die Sequenzen der HVR I auf den Bereich von np 16059 bis 16390 gekürzt werden. Dadurch blieben drei polymorphe Positionen (16042, 16048 und 16051) unberücksichtigt. Betroffen waren die Proben OLE 7, PES 7 und POP 3. In die steppeninternen Berechnungen gingen alle Polymorphismen zwischen np 16013 und 16409 ein. Im Vergleich zum verkürzten Datensatz waren dadurch die F_{ST} -Werte geringfügig kleiner, aber grundlegend änderte sich nichts.

10.1.4 Probenauswahl

Idealerweise würde man die Proben so wählen, dass Untersuchungsgebiet und -zeitraum gleichmäßig und durch eine statistisch aussagekräftige Stichprobe repräsentiert sind. In der vorliegenden Studie war dies nicht der Fall. Die äneolithischen Proben stammen fast ausschließlich aus dem west- und nordwestpontischen Raum. Die Jamnaja-Kultur ist westlich des Dnjeprs und in dessen Einzugsgebiet gut vertreten, doch weiter östlich so gut wie gar nicht. Die Stichprobe der Katakombengrabkultur ist mehr oder weniger auf das Gebiet östlich des Dnjepr beschränkt. Die Gründe für die Ungleichverteilung liegen hauptsächlich in der Verfügbarkeit der Proben und der unterschiedlichen Kooperation potentieller Probengeber. Eine weitere, aber eher untergeordnete Rolle spielte, dass bei der Probennahme parallel auch Material für Isotopenanalysen genommen wurde. Bei diesen lag der Fokus auf Fundplatz-internen Strukturen, so dass hierfür wenige

Kurgane intensiv zu beproben waren, im Gegensatz zur Populationsgenetik, für die flächendeckend Stichproben aus möglichst viele verschiedene Kurganen untersucht werden sollten, um direkte Verwandtschaften zu vermeiden. Als Strategie musste daher ein Kompromiss gefunden werden, um den Anforderungen beider Forschungszweige gerecht zu werden. Auf die Repräsentativität der einzelnen Zeitscheiben wirkte sich vor allem die unterschiedliche Probenerhaltung aus (siehe Abschnitt 10.1.5). Von der Tripolje-Kultur aus der nordwestpontischen Waldsteppe, die archäologischen Befunden zufolge intensive Kontakte mit den äneolithischen Steppenkulturen pflegte, gibt es leider kaum Skelettfunde. Nach der Archäologin E. K. (persönliche Kommunikation) ist anzuzweifeln, dass die Individuen, von denen bisher DNA-Daten publiziert wurden (Nikitin *et al.* 2010) tatsächlich dieser Kultur angehören.

Für einen Teil der Proben existieren Radiokarbondatierungen. Allerdings wurden diese von den jeweiligen Auftraggebern bis auf wenige Ausnahmen noch nicht publiziert und dürfen daher nicht in Veröffentlichungen verwendet werden. Auch scheinen einige Datierungen (insbesondere aus dem Labor Kiew) unzuverlässig zu sein. Die Einordnung der Proben in die verschiedenen Zeitscheiben wurde von E. K., einer Expertin für die Archäologie des Untersuchungsgebietes, hauptsächlich auf Basis des Fundgutes, unterstützt durch die vorliegenden absoluten Datierungen, vorgenommen.

10.1.5 Probenerhaltung

Die DNA-Erhaltung korrelierte negativ mit dem Alter der Proben, was dazu beitrug, dass nur zehn äneolithische Individuen für die biostatistischen Analysen zur Verfügung standen. Gleichzeitig wurden auch geographische Muster beobachtet. Mit Erhaltungswerten zwischen 0,75 und 1,14 (von maximal 3, siehe Tabelle 9.6) erschienen die Proben aus der Trockensteppe und Steppe in Küstennähe, insbesondere der Trockensteppe in der Region Odessa (Fundplatz-Einheit 4 in Abbildung 9.9), der bulgarischen Steppe (Einheit 2) und in Olennij (9) am schlechtesten erhalten. In der Trockensteppe Kalmückiens im Landesinneren (10) hingegen, von wo

allerdings auch vorwiegend die jüngeren Proben stammen, und im gemäßigten mitteleuropäischen Klima war die Erhaltung am besten (Erhaltungswert 2,6). Fast alle Skelettfunde wurden über mehrere Jahre bzw. Jahrzehnte in archäologischen Instituten bei wahrscheinlich schwankenden Temperaturen gelagert. Daher lässt sich nicht sagen, welcher Faktor (Alter, Klima, Boden, Lagerung) den stärksten Einfluss auf die DNA-Präservierung hatte.

Im Vergleich zu anderen aDNA-Studien ist die durchschnittliche Erhaltung (41,3 % reproduzierte HRV I-Haplotypen) in der nordpontischen Steppe gering. In Studien zum mitteleuropäischen Neolithikum lagen die Erfolgsraten beispielsweise bei 71 % (Haak *et al.* 2010), 82,8 % (Bollongino *et al.* 2013) oder 84,1 % (Brandt *et al.* 2013). Selbst von den wesentlich älteren, größtenteils mesolithischen Proben in Bramanti *et al.* (2009) konnten 55,5 % reproduziert werden. Dies bestätigt die Beobachtung der besseren DNA-Erhaltung in den Proben der Kugelamphorenkultur aus Polen und Deutschland (Erhaltungswert 2,4).

Auch die Bearbeitungsstrategie ist ein Faktor. So wirkte sich z. B. wiederholtes Auftauen und Einfrieren von DNA-Extrakten negativ auf den Amplifikations- und Reproduktionserfolg aus. Da zunächst die nukleäre DNA untersucht wurde, gibt es vereinzelt Fälle, in denen zwar Genotypen bestimmt werden konnten, aber die mtDNA Haplotypen nicht reproduzierbar waren (siehe nächster Abschnitt, 10.1.6).

10.1.6 Authentizität der Daten

In Einzel-PCRs gefolgt von Direktsequenzierungen sind Kontaminationen der Leerkontrollen relativ einfach zu identifizieren und quantifizieren. Bei der hier angewandten Methode mit Multiplex-PCR und 454-Sequenzierung können die positiven Negativkontrollen nicht einfach abgezählt werden, sondern es wurde ein „Erwartungswert“ der Kontaminationsrate berechnet. Dieser lag für die mtDNA bei 4,56 %, ein Wert, der z. B. mit Haak (2006) oder Bramanti *et al.* (2009) vergleichbar ist. Wurde das Limit von mindestens 3 *reads* für die Anerkennung einer Amplifizierung verwendet, sank der Wert auf 2,40 %. Unter der Annahme, dass die Kontamination

einer Kontrolle aus einer individuellen bzw. einzelnen Quelle stammt und sich die einzelnen Fragmente sich zu einem mitochondrialen Haplotypen zusammensetzen lassen, konnten alle Kontaminationen, die erfolgreich sequenziert wurden, von den Proben unterschieden werden. Es kann natürlich nicht ausgeschlossen werden, dass die einzelnen Fragmente aus einer Leerkontrolle verschiedene Ursprünge haben. Da vor allem in der *coding region* viele Amplikons mit rCRS-Sequenzen vorkommen, ist hier eine Unterscheidung nicht immer möglich. Wichtig war daher die wiederholte Replikation der Ergebnisse in unabhängigen PCRs.

Bei der nukleären DNA lag der Erwartungswert der Kontaminationsrate in den Leerkontrollen bei 2,23 %. Wurden nur *loci* mit mindestens 10 *reads* berücksichtigt, ist die Rate von 0,35 % fast zu vernachlässigen. Auch bei den nukleären Amplifikationen unterschied sich der kontaminierende Genotyp meist von denen der parallel bearbeiteten Proben. In den Fällen, wo dies nicht möglich war, gelang die Authentifizierung durch Reproduktion. Wie bereits erwähnt, dürfte die Fehlerrate bei der Sortierung der Rohdaten nach *tags* und Primer sich nicht auf die Typisierungen ausgewirkt haben (siehe 9.1.4 und Seite 190).

Bei alter DNA treten postmortale Schäden auf. Daher wurde ein striktes Reproduktionsschema zur Authentifizierung angewandt. Bei einer *shotgun*-Sequenzierung repräsentiert jeder *read* ein Ausgangsmolekül, so dass sporadische Schäden in einzelnen Molekülen direkt identifiziert werden können. Durch die vorgeschaltete Amplifikation ist dies nicht immer möglich. Bei schlecht erhaltenen Proben wurden in einigen Fällen, besonders bei den nukleären 454-Sequenzierprodukten, neben den eigentlichen SNPs scheinbar zusätzliche homozygote Polymorphismen beobachtet. Durch weitere PCRs konnten diese als Schäden (oder PCR-spezifische Kontaminanten) entweder sicher identifiziert werden, oder die entsprechenden Proben stellten sich als zu schlecht erhalten heraus, um überhaupt reproduzierbare Ergebnisse zu liefern.

In Einzelfällen konnten die Mindestanforderungen an die Reproduktion der Ergebnisse für kleinere Abschnitte von 1 bis 27 Nukleotiden nicht erfüllt werden. Aber die Identität der zwei (anstatt drei) Sequenzen an diesen Positionen, Konsistenz der

HVR I insgesamt, und zusätzliche unabhängige PCRs mit weniger als drei *reads* zeugen von der Authentizität dieser Sequenzen. Eine weitere Stützung der HVR I-Daten ist durch die Fragmente der *coding region* gegeben. Da diese allerdings nicht für die Datenanalyse verwendet wurden, wurde hier auf eine zwingende Reproduktion verzichtet.

Auffällig bei den 454-Sequenzierungen war, dass, wenn Amplikons nur durch vereinzelte *reads* repräsentiert sind, diese häufig von denen aus anderen PCRs mit stärkerer Abdeckung durch viele *reads* abweichen. Durch die festgelegte Mindestanzahl von drei (für mtDNA) bzw. zehn (ncDNA) *reads* pro Allel konnte ein daraus resultierender Fehler ausgeschlossen werden.

Bei der Sanger-Sequenzierung traten vereinzelt Positionen auf, an denen artifizielle Sequenzmodifikationen auftraten, z. B. in Fragment HVR I Ia. Diese konnten durch die Sequenzierung des Gegenstranges und dadurch, dass sie bei der 454-Sequenzierung nicht auftraten, identifiziert werden.

Für nukleäre aDNA-Daten ist die Authentifizierung prinzipiell etwas aufwendiger als für mitochondriale DNA. Dies liegt vor allem an der geringen Zahl an Ausgangsmolekülen, an der Diploidie und an der Unabhängigkeit der einzelnen *loci* zueinander (Lage auf unterschiedlichen Chromosomen, Rekombinationen). Daher sollten für die Authentifizierung eines nukleären Markers mindestens vier oder fünf unabhängige Amplifikationen zum gleichen Ergebnis führen. Ohne zusätzliche Einzel-PCRs (wie für die *loci* *HERC2*, *LCT*, *SLC45A2* und *TYR* durchgeführt) erfüllten nur gut erhaltene Proben und *loci* mit den effizientesten Primersystemen die strengen Vorgaben. Daher wurde die „vorbehaltliche“ Genotypisierung eingeführt. Die Kriterien wurden so gewählt, dass die Wahrscheinlichkeit, ein fehlerhaftes Ergebnis vorliegen zu haben, nur geringfügig höher als bei der „sicheren“ Authentifizierung ist. Wie Vergleiche zwischen vorbehaltlich und sicher reproduzierten Allel- und Genotypenfrequenzen zeigten, wurde damit ein Kompromiß zwischen Authentizität und Stichprobengröße gefunden.

10.2 Die Rolle der pontisch-kaspischen Steppe im europäischen Neolithikum

Nach dem derzeitigen Kenntnisstand liegt der Ursprung des Neolithikums im Nahen Osten, von wo aus es sich über den Balkan nach Mittel- und Osteuropa ausbreitete. Die ersten fassbaren frühneolithischen Kulturen in Südosteuropa gehören zum Starčevo-Körös-Criș-Kulturkomplex. Ihm folgte die Linearbandkeramik, mit deren Ausbreitung die Neolithisierung Europas assoziiert ist. Während sich eine Reihe von populations- und palaeogenetische Studien mit den Entwicklungen in Mitteleuropa befassen (Haak *et al.* 2005, Bollongino 2006, Bramanti *et al.* 2009, Haak *et al.* 2010, Brandt *et al.* 2013, u. a.), so wurde die Prähistorie des nordpontischen Steppenraums bisher fast ausschließlich archäologisch untersucht. Nachdem bis dato lediglich einige wenige DNA Sequenzen neolithischer Individuen veröffentlicht wurden (Nikitin *et al.* 2012), liefert die vorliegende Arbeit den ersten populationsgenetischen Datensatz prähistorischer Steppenulturen der lokalen Kupfer- und Bronzezeit, die synchron zum mitteleuropäischen Mittel- und Spätneolithikum ist.

Um diese Daten in das Gesamtbild der Entwicklungen in Europa einzubinden, wurden entsprechende Vergleichspopulationen in die Analysen einbezogen. Diese wurden vorwiegend aus den oben genannten Artikeln entnommen (für eine vollständige Auflistung siehe Abschnitt 8.2.8.3). Ein Großteil der Proben stammt aus Deutschland, insbesondere der Mittelsaale-Elbe-Region (Sachsen-Anhalt). Derzeit gibt es noch keine verlässlichen aDNA-Daten aus den dem Untersuchungsraum angrenzenden Regionen wie z. B. aus Südosteuropa. Daher können die interkulturellen Analysen nur Anhaltspunkte für mögliche Verbindungen zwischen der europäischen Prähistorie und der der Steppe bieten.

10.2.1 Der Ursprung der Viehhirten in der pontischen Steppe

Archäologischen Befunden zufolge waren wahrscheinlich Kontakte zwischen der Bug-Dnjestr- und der Criş-Kultur in der nordwestlichen Waldsteppe der Anstoß für die Neolithisierung der Steppe im Sinne der Verbreitung von Viehzucht (Anthony 2007: 138-146). Knochenfunde, die auf eine frühere Tierhaltung in östlicheren Gebieten der nordpontischen Steppe hindeuten, sind hinsichtlich der Authentizität ihrer Datierung umstritten (Anthony 2007: 158, 481 f). Nichtsdestotrotz sind auch Impulse aus dem transkaukasischen Neolithikum für die Entwicklungen in der Steppe denkbar (Wechler 1998). Übernahmen die indigenen Steppenkulturen die neue Subsistenzwirtschaft von benachbarten Bauern, oder wanderten Gruppen von außerhalb in die nordpontische Steppe ein und verdrängten die ansässigen Jäger und Sammler?

Während im Neolithikum und Äneolithikum in der Steppe Jagen, Sammeln und Fischen noch lange eine große Rolle in der Subsistenz spielten, nahm die Viehhaltung nach und nach zu, bis mit Beginn der Bronzezeit das Hirtentum voll etabliert war (Kaiser 2010). Damit handelt es sich bei dem vorliegenden Datensatz größtenteils um Populationen, die von Viehwirtschaft lebten. Um die Frage nach ihrem Ursprung bzw. ihrer Identität zu klären, wurden sie genetisch charakterisiert und mit möglichen Vorfahren verglichen.

Aus dem unmittelbaren Untersuchungsgebiet gibt es bisher keine prä-neolithischen Daten; lediglich in der nördlichen Peripherie der über 30 000 Jahre alte Mensch aus Kostenki am Don, der eine U2 Linie trägt (Krause *et al.* 2010). Im Nordwesten Russlands, 650 bis 950 km nördlich von Moskau weisen die Mesolithiker von zwei Fundplätzen neben sieben U noch drei C und ein H auf (Der Sarkissian *et al.* 2013). Die bisher haplotypisierten prä-neolithischen Jäger und Sammler Mitteleuropas tragen ausschließlich U-Linien (Bramanti *et al.* 2009, Bollongino *et al.* 2013, Fu *et al.* 2013).

Das Neolithikum des nordpontischen Raumes wird genetisch bisher durch sechs Individuen der Dnjepr-Donetz-Kultur repräsentiert (Nikitin *et al.* 2012). Deren Subsistenz beruhte zunächst auf Jagd und Fischfang, später soll sie zu einer agrikulturellen Lebensweise übergegangen sein (Mallory & Adams 1997: 166 f, Wechler 1998). Dabei treten zu gleichen Teilen die Haplogruppen U, C und H auf (je

zwei Individuen). Ein weiteres Individuum trägt ebenfalls eine C-Linie, allerdings ist dessen Alter unklar. In Ungarn wurde ebenfalls ein hoher Anteil an Haplogruppe C gefunden (Guba *et al.* 2011). Allerdings stellte sich im Nachhinein heraus, dass diese Proben fälschlicherweise als neolithisch datiert wurden und stattdessen in die sarmatische Zeit fallen (Bánffy *et al.* 2012)³⁷.

Linien der Haplogruppe U machen 34,4 % der im Rahmen dieser Arbeit untersuchten kupfer- und bronzzeitlichen Proben aus der Steppe aus³⁸ (27,3 % in der Jamnaja-Kultur). Haplogruppe C fehlt hingegen. Falls die wenigen Daten aus dem Neolithikum die tatsächliche Häufigkeit der Haplogruppe C von über 30 % repräsentieren sollten, dann wäre der Verlust dieser Linien innerhalb von 60 bis 70 Generationen kaum das Ergebnis einfacher Gendrift. Dies und die hohe Diversität an sonstigen Haplotypen und Gruppen lassen sich stattdessen durch eine Introgression von außen erklären.

Das vollständige Fehlen der Haplogruppe C oder anderer zentralasiatischer Linien (z. B. D oder F) im vorliegenden Datensatz lässt darauf schließen, dass der nordpontische Raum in der Zeit vom Neolithikum bis zur Bronzezeit genetisch vom westsibirischen Neolithikum isoliert war. Zumindest ist maternal kein Zufluss aus östlicher Richtung nachweisbar. Bestätigt wird dies durch die Unterschiede im Haplogruppenspektrum und den großen genetischen Distanzen zwischen der nordpontischen Steppe und der westsibirischen Ust-Tartas-Kultur (Barabawaldsteppe; 4. Jtsd. bis erste Hälfte des 3. Jtsd. v. Chr.) (Molodin *et al.* 2012). Den einzigen Hinweis auf die Möglichkeit asiatischer Anteile im Erbgut der Steppenbewohner liefert der autosomale Marker *ABCC11* (siehe 10.4.3.2), während die anderen zwischen Europa und Asien differenzierenden nukleären SNPs durchweg auf europäische Herkunft deuten. Allerdings fehlen bei den nukleären Analysen bis auf wenige Ausnahmen Daten anderer prähistorischer Populationen, wie z. B. der LBK.

³⁷ In einer noch nicht offiziell veröffentlichten Studie wurden aus dem mittleren Neolithikum in Ungarn lediglich „europäisch neolithische“ Linien gefunden (http://www.hungarianarchaeology.hu/wp-content/uploads/2013/08/eng_Banffy_13ny1.pdf) [Stand: 01.10.2014].

³⁸ Die auf den Gesamtdatensatz bezogenen Zahlen auf dieser und der nächsten Seite beruhen auf den 61 Proben der west- und nordpontischen Fundplätze. Die vier Individuen der Kugelamphorenkultur aus Deutschland und Polen sind nicht berücksichtigt.

Etwa ein Drittel der west- und nordpontischen Individuen tragen die mit Jägern/Sammlern assoziierten U-Linien. Bei den restlichen 65,6 % handelt es sich um Haplotypen, die bisher im europäischen Neolithikum beobachtet wurden (37,7 %) und Linien der Gruppe H (27,9 %), der auch heute in Europa mit Abstand häufigsten Haplogruppe. Brandt und Kollegen differenzierten in ihrer Publikation zwischen einem charakteristischen LBK-Profil aus N1a, T2, K, J, HV, V, W und X, und Linien, die ab dem mittleren Neolithikum in Europa nachgewiesen sind (Haak *et al.* 2005, Brandt *et al.* 2013). Die in diesem Artikel als mit der Linearbandkeramik assoziierten Linien würden demnach 26,2 % des vorliegenden Datensatz ausmachen, die restlichen (ohne H-Linien) 11,5 %. Allerdings ist eine solche Differenzierung aufgrund der hohen Diversität innerhalb der neolithischen Kulturen kritisch zu betrachten. Hier soll nur gezeigt werden, dass sich der Genpool der Steppenpopulationen größtenteils aus Linien der Haplogruppen zusammensetzt, die zum einen in den mesolithischen Jägern/Sammlern und zum anderen im Datensatz der linearbandkeramischen Individuen beobachtet wurden.

Die genetischen Ergebnisse stützen die archäologische These, dass die Viehhaltung mit der LBK oder deren Vorläufer, der Criş-Kultur, in den nordpontischen Raum gelang. Dass im Genpool der bronzezeitlichen Viehhirten sowohl die typischen Jäger- und Sammler-Linien, als auch die ab dem mitteleuropäischem Frühneolithikum nachgewiesenen Linien über den gesamten Untersuchungsraum verteilt vorkommen, spricht für ein Szenario, in dem europäische Bauern in die nordpontische Steppe einwanderten (demische Diffusion) und es vermutlich im Verlauf des Äneolithikums zu einer Vermischung der beiden Kulturträger kam. Eine reine Übernahme der neuen Wirtschaftsstrategie durch die nordpontischen Steppenbewohner von den westlichen Nachbarn (kulturelle Diffusion) ist als eher unwahrscheinlich anzusehen.

Unter der Voraussetzung, dass vor dem Einwandern der Bauern der mitochondriale Genpool hauptsächlich aus U-Linien (eventuell mit Anteilen von C und H), bestand, so muss es eine Einwanderungswelle von Frauen gegeben haben, um die Frequenz anderer Linien (nicht U und nicht H) von 0 % auf über 35 % zu steigern (die untere Grenze durch die Standardabweichung liegt bei etwa 25 %). Die hohe genetische Diversität der Jamnaja-Kultur, bei der man aufgrund der hohen Zahl an Kurganen von

einer großen Bevölkerung ausgeht, spricht nicht für die Zuwanderung einiger weniger Frauen mit vielen Nachkommen. Um hier wirklich Klarheit zu bekommen, müsste allerdings mehr über die Zusammensetzung des mesolithischen Genpools in der nordpontischen Region bekannt sein. Interessant wäre hierbei z. B. der Anteil der Haplogruppe C, die in den hier untersuchten Proben nicht mehr auftritt, aber wieder in Ungarn in der Sarmatenzeit (Guba *et al.* 2011, Bánffy *et al.* 2012); oder auch der Anteil der Haplogruppe H, die bisher in mitteleuropäischen Jägern/Sammlern nicht beobachtet wurde, die aber möglicherweise im osteuropäischen oder westsibirischen Mesolithikum / Epipaläolithikum vertreten war.

Die genetischen Distanzen der äneolithischen Proben zur LBK und zu den präneolithischen Jägern/Sammlern sind ähnlich und liegen bei 0,045 bis 0,065. Dies ändert sich in der Bronzezeit. Die genetische Distanz zwischen der Jamnaja-Kultur und der LBK ist mit einem F_{ST} -Wert von 0,02807 ($p < 0,05$) relativ niedrig. Dies ist ein Indiz für eine zunehmende Vermischung zwischen Bauern und den ursprünglichen Steppenbewohnern. Der geringe Anteil an identischen Haplotypen (*shared haplotypes* von 1,25 % (0,8 % – 1,7 %) zwischen Jamnaja und LBK könnte auf den zeitlichen Abstand der beiden Kulturen zurückzuführen sein. Die Katakombengrabkultur wiederum zeigt hohe, signifikante F_{ST} -Werte zwischen 0,5 und 0,8 zu den zeitlich vor ihr existierenden Bauerngesellschaften. Ihre Beziehungen mit den europäischen Kulturen werden in einem gesonderten Abschnitt weiter unten behandelt (siehe 10.2.5).

Ob es direkt nach Ankunft der Südosteuropäer zu *admixture* (Vermischung) innerhalb der Steppe kam oder die beiden Bevölkerungen zunächst noch parallel, genetisch isoliert voneinander lebten, wie es in Mitteleuropa der Fall war (Bramanti *et al.* 2009, Bollongino *et al.* 2013), kann nicht geklärt werden. Aus dem reinen Steppenäneolithikum (ohne den Fundort Smjadovo) liegen lediglich sechs Sequenzen von vier Fundorten vor. Interessanterweise stammen die beiden U5a1-Typen aus Moljučov Bugor, einer Siedlung, die in der spätäneolithischen Phase noch einen Wildtieranteil von etwa 40 % aufweist, wohingegen die beiden spätäneolithischen Individuen der Usatovo-Kultur aus Majaki mit T2b und W LBK-typische Linien tragen. Im Tierknochenspektrum dieses Fundortes sind weniger als 5 % Wildtiere (Kaiser

2010). Dies könnte ein Fingerzeig in Richtung getrennter Lebensweise zwischen Jägern/Sammlern und Bauern sein, allerdings sind vier Proben keine Basis für eine definitive Aussage bezüglich gesellschaftlicher Strukturen. DNA-Analysen, in Kombination mit isopenchemischen Untersuchungen, an weiteren neolithischen und äneolithischen Proben könnten zur Frage des Zusammenlebens Aufschluss geben, vorausgesetzt, deren Lebensgrundlage kann bestimmt und differenziert werden, wie es z. B. in Bollongino *et al.* (2013) demonstriert wurde.

Eine alternative Erklärung für die genetische Zusammensetzung der untersuchten Population wäre, dass erst eine der aus der LBK hervorgehenden Kulturen die typischen Linien in die Steppe brachte. Doch müsste dies mit einer noch größeren Bevölkerungseinwanderung und einer schnelleren Veränderung des Haplogruppenspektrums einhergegangen sein. Die mobile LBK selbst als Ursprung der Viehhüter passt sowohl in das genetische als auch das archäologische Bild.

10.2.2 Die Nachfolge der Linearbandkeramischen Kultur

Nach dem ersten Kontakt zwischen den Jägern und Sammlern der Steppe und den europäischen Bauern, deutet alles auf weiter andauernde Beziehungen zwischen den beiden Großräumen Mitteleuropa und Steppe hin. Auf dem Gebiet der Linearbandkeramik folgten die Rössener Kultur im Westen, die Stichbandkeramik im Osten und die Lengyel Kultur im südöstlichen Areal. In den östlichen Karpaten entstand der Cucuteni-Tripolje-Komplex, für den direkte Kontakte mit Kulturen der Steppe archäologisch nachgewiesen sind (Dergačev 1998, Rassamakin 1999, Müller 2009b). Von der Cucuteni-Tripolje Kultur sind bis auf ihre späte Phase Hinweise auf Bestattungsaktivitäten nahezu absent. Zwar veröffentlichten Nikitin *et al.* (2010) palaeogenetische Analysen von sieben Individuen, doch es ist strittig, ob es sich dabei tatsächlich um Angehörige der Tripolje-Kultur handelt (E. K., persönliche Kommunikation). Da der Artikel praktisch keine Angaben zur Probenbearbeitung macht und lediglich Haplogruppen (2x HV/V, 2x H, je 1x Prä-HV, T4, J), ohne Angabe der Polymorphismen präsentiert, wurde er bei den vorliegenden Untersuchungen nicht berücksichtigt. Von den der LBK direkt nachfolgenden Kulturen des fünften

Jahrtausends v. Chr. liegen lediglich Sequenzen aus dem westlichen Gebiet, der Rössener Kultur, sowie der anschließenden, geographisch eng auf die Mittelbe-Saale-Region begrenzten Schöninger Gruppe aus der Publikation von Brandt *et al.* (2013) vor.

Nach einem einmaligen Einzug von frühneolithischen Bauern in die Steppe ohne weitere Kontakte würde man mit der Zeit eine zunehmende Distanz zwischen mitteleuropäischen und Steppenkulturen aufgrund von genetischer Drift erwarten. Im Gegenteil deuten die durchweg geringen F_{ST} -Werte zwischen der Jamnaja-Kultur und den Mittelbe-Saale-Gruppen des frühen und mittleren Neolithikums Europas auf ständigen oder mehrfachen intensiven Beziehungen hin.

Die äneolithischen Proben haben höhere, allerdings nicht signifikante F_{ST} -Werte zur LBK, der Rössener Kultur und der Schöneberger Gruppe als die Individuen der Jamnaja-Kultur. Jedoch tragen sieben der zehn äneolithischen Individuen Haplotypen, die auch in diesen Kulturen vorkommen. Die anderen drei Individuen tragen mit U5, T2 und H7 zwar verbreitete Linien, allerdings mit „privaten“ Mutationen, die in den Stichproben der drei frühneolithischen Kulturen nicht vorkommen, was die hohen F_{ST} -Werte erklärt.

10.2.3 Die Jamnaja und die Trichterbecherkultur

Ab dem vierten vorchristlichen Jahrtausend verbreitete sich die Trichterbecherkultur (TBK) im nördlichen Mitteleuropa und Südkandinavien. Sie liegt damit zeitlich parallel zum Äneolithikum der Steppe, die jüngeren TBK-Gruppen synchron zur frühen Bronzezeit. Zur Zeit der frühen TBK sind auch die ersten Grabhügel für einzelne Individuen in Mitteldeutschland, wie auch der pontischen Steppe nachgewiesen (Müller 2009b). Erstere sind allerdings im Rahmen der sich entwickelnden westlichen Megalithtradition zu sehen. Sowohl in Mitteleuropa als auch in der Steppe sind Kupferimporte aus den balkanisch-karpatischen Abbaugebieten belegt (Häusler 1998, Chernykh 2008). Im Rahmen großräumiger

Handelskontakte in Europa, insbesondere mit der aufblühenden Metallurgie, könnten auch die Steppenpopulationen über Fernbeziehungen integriert gewesen sein.

Die negativen F_{ST} -Werte zwischen den nordpontischen Populationen Äneolithikum und Jamnaja und drei der vier Stichproben der Trichterbecherkultur sind ein deutliches Indiz für Kontakte, die sich anscheinend nicht nur auf den Austausch von Kupferwaren beschränkten. Die einzigen positiven Werte betreffen die sehr kleinräumige Salzmünder Gruppe (Brandt *et al.* 2013). Aufgrund der allgemein sehr ähnlichen genetischen Populationszusammensetzung lässt sich nicht sagen, in welche Richtung Genfluss stattgefunden hat.

Es ist auch nicht definitiv auszuschließen, dass es keine direkten Kontakte gab. Die Trichterbecherkultur basiert auf einem mesolithischen Substrat unter dem Einfluss und genetischem Beitrag von Bauern (Häusler 2003: 87, Brandt *et al.* 2013). Möglicherweise resultieren die genetischen Affinitäten auf mehr oder weniger unabhängigen, aber parallel ablaufenden Entwicklungen im neolithischen Mitteleuropa und der (ä)neolithischen Steppe, wobei beide ihren Anfang mit der Ausbreitung der LBK hatten. Hohe Populationsdichten, interkulturelle Beziehungen und die genetische Kontinuität könnten für beide Räume zu hohen effektiven Populationsgrößen (trotz kleinerer Subkulturen) geführt haben, so dass es in der Zeit von der LBK bis zur TBK nicht zum genetischen „Auseinanderdriften“ kam.

Der Plot der multidimensionalen Skalierung (Abbildung 9.10) veranschaulicht, wie eng alle europäischen Bauerngesellschaften, sowie die äneolithischen und Jamnaja-Proben der Steppe clustern. Eventuell wäre auf der paternalen Ebene eine genetische Unterscheidung zwischen den drei Komplexen (LBK, TBK, Steppe) möglich, und somit Unterschiede in den Beziehungen und Richtungen von Genfluss zu erkennen.

10.2.4 Die Genese der Becherkulturen

Ab etwa 2 900 und 2 800 v. Chr. prägten im Osten die Schnurkeramische und im Westen die Glockenbecherkultur das späte Neolithikum in Europa. Sie werden in der Literatur auch unter dem Sammelbegriff „Becherkulturen“ – nicht zu verwechseln mit der Trichterbecherkultur – zusammengefasst. Das Verbreitungsgebiet der Schnurkeramik überlagerte das der TBK und reichte im Osten bis nach Westrussland nördlich dessen der Jamnaja-Kultur. Über ihren Ursprung wird diskutiert. Neben der Trichterbecherkultur und ihrer östlichen Nachfolge, der Kugelamphorenkultur, wurde auch die Jamnaja-Kultur als genetischer oder kultureller Initiator der Kultur vermutet. Früher ging man sogar davon aus, dass es sich bei der Schnurkeramik um eine aus dem Osten eingewanderte Population handele. Von diesem Bild ist man mittlerweile wieder abgerückt (Mallory & Adams 1997: 127 f, Häusler 2003: u. a. 20, 69-74, Müller 2009a). Genetisch ist, wie auch schon bei den Trichterbecherkulturen keine Distanz zwischen den beiden Becherkulturen und den Jamnaja festzustellen. In Bezug auf *haplotype sharing* sind sich die Steppenvölker und die weiter westlich verbreitete Glockenbecherkultur (z. B. Jamnaja: 3,6 % ± 1,3 %) sogar etwas näher als erstere mit den Schnurkeramikern (2,1 % ± 0,9 %). Dies beruht wahrscheinlich auf der Stichprobe. In der archäologischen Literatur wird die Glockenbecherkultur nicht mit den Steppenkulturen in Verbindung gebracht.

In den beiden Becherkulturen taucht die in der Jamnaja-Kultur häufige Haplogruppe T1 auf, die zuvor in nur einem Individuum der frühesten Trichterbechergruppe beobachtet wurde. Aus dieser Zeit stammen auch die ersten Funde mit Haplogruppe I; jeweils eine in der Glocken- und der Schnurkeramischen Kultur. Haplogruppe I wurde außerdem in einem Individuum im bronzezeitlichen Kasachstan (1400 – 1000 v. Chr.) beobachtet. Haplogruppe U2 war zum Zeitpunkt der Veröffentlichung von Brandt *et al.* (2013) lediglich in präneolithischen Jägern und Sammlern aus Russland (Krause *et al.* 2010, Der Sarkissian *et al.* 2013) und in der westsibirischen Bronzezeit (Keyser *et al.* 2009) gefunden, in den Mittelsaale-Elbe-Kulturen ab dem späten Neolithikum. Die Autoren vermuteten einen genetischen Eintrag der Kurgankulturen der Steppe nach Mitteleuropa. Mittlerweile wurde die

Haplogruppe U2 auch in einem mitteleuropäischen Mesolithiker identifiziert (Bollongino *et al.* 2013). Die beiden Haplogruppen U2 und I treten jeweils einmal im Datensatz der Steppenproben, beide in der Jamnaja-Kultur, auf. Gesichert ist damit allerdings nicht, dass es sich hier tatsächlich um einen genetischen Eintrag aus der Steppe in die Schnurkeramik handelt. Dazu müssten mehr Individuen mit diesen Linien in der Steppe und aus vorbronzezeitlichen Proben aus dem Osten gefunden werden. Insbesondere der rasante Anstieg der I-Typen in der ab 2300 v. Chr. erfassten Aunjetitzer-Kultur ist bemerkenswert (12,8%). Eine multidimensionale Skalierung basierend auf *haplotype sharing* zeigt die enge genetische Verwandtschaft der Jamnaja-Kultur mit der Schnurkeramik, aber auch mit der Aunjetitzer-Kultur und älteren neolithischen Gruppen. Auch die Bronzezeit Kasachstans fällt in dieses Cluster.

10.2.5 Der Ursprung der Katakombengrabbkultur

Aus den bisherigen Ausführungen wurde die auf die Jamnaja-Kultur folgende Katakombengrabbkultur etwas ausgeklammert. Sie bedarf einer besonderen Betrachtung. Die Katakombengrabbkultur entwickelte sich wahrscheinlich auf der Grundlage der Jamnaja-Kultur (siehe Kapitel 10.3.3), allerdings gibt es einige Indizien, die für einen oder mehrere externe Einflüsse auf ihre Genese sprechen.

Während die Jamnaja-Kultur und in einem geringeren Maße auch die Proben des Äneolithikums große Affinitäten zu den mitteleuropäischen neolithischen Stichproben aufweisen, sind die genetischen Distanzen zwischen diesen und der Katakombengrabbkultur prinzipiell größer. Die F_{ST} -Werte liegen zwischen 0,03183 (Aunjetitzer-Kultur) und 0,08153 (Schöninger Gruppe) und sind alle signifikant. Lediglich die Distanz zur Glockenbecherkultur ist kleiner (0,01784, $p < 0,1$).

Unter den insgesamt hohen F_{ST} -Werten aller Steppenkulturen zu den prä-neolithischen und neolithischen Jäger/Sammler/(Fischer)-Stichproben gibt es ebenfalls eine Ausnahme: die geringe Distanz ($F_{ST} = 0,01798$, $p > 0,1$) zwischen der Katakombengrabbkultur und der Grübchenkeramik. Die Haplogruppenspektren dieser

beiden Kulturen weisen extrem hohe Anteile an U4-Linien auf (30,4 % und 42,1 %). Auch das *haplotype sharing* von $7,8 \% \pm 2,5$ deutet auf eine Verbindung hin.

In der Jamnaja-Kultur trägt nur die Probe TET 2 aus der nordwestpontischen Waldsteppe einen U4-Typen. Die rapide Zunahme der U4-Linien mit dem Auftauchen der Katakombengrabkultur spricht für einen Eintrag dieser Linien in die Steppe. Hierfür kommt aus dem Pool der Vergleichsdatensätze die Grübchenkeramische Kultur aufgrund der oben aufgeführten Beobachtungen einem möglichen Kandidaten für eine genetische Ursprungspopulation am nächsten. Basierend auf Analysen dieser Kultur bestätigten Malmström und Kollegen (2009) Vermutungen, dass es im östlichen Baltikum während des Neolithikums genetische Refugien für einige Jäger/Sammler-Populationen gab.

Der Anstoß zur Entstehung der Katakombengrabkultur könnte aus diesen Refugien gekommen sein. Vier der elf mesolithischen Proben aus der Peripherie des Baltikums tragen Linien der Haplogruppe U4 (Der Sarkissian *et al.* 2013). In den westlichen Kulturen hingegen dominieren die U5-Linien innerhalb der U-Gruppen, daher ist eher nicht davon auszugehen, dass der Anstieg der Haplogruppe U4 in der Katakombengrabkultur aus dieser Richtung initiiert wurde. Möglicherweise kam es bereits während der Jamnaja-Periode zu einem Eintrag der U4-Linien, und sie wurden in dieser Phase noch nicht erfasst. Entweder könnte die Häufigkeit noch zu niedrig gewesen sein, oder die Haplogruppe kam über die Wolga-Don-Region in die Steppe, von wo es keine repräsentative Stichprobe der Jamnaja-Kultur gibt. Eine etwas differenziertere Betrachtung der U4-Linien auf der Ebene der DNA-Sequenz zeigt, dass in den verschiedenen Populationen unterschiedliche Haplotypen dominieren. In der Katakombengrabkultur tritt ausschließlich die kladendeterminierende Mutation 16356C in der HVR I auf. In der Grübchenkeramik aus Malmström *et al.* (2009) kommen drei verschiedene Motive vor. Drei der acht Individuen entsprechen dabei der U4-Wurzel. Vier Individuen tragen dazu die Mutation 16093C, die auch drei der vier russischen Jäger/Sammler aus Der Sarkissian *et al.* (2013) aufweisen. In acht weiteren, noch unveröffentlichten, ebenfalls aus Nordwestrussland stammenden Mesolithikern (Anna Schulz vom Institut für Anthropologie in Mainz, persönliche Kommunikation) dominiert wiederum ein anderes Motiv, das in keiner der anderen drei

Populationen beobachtet wurde. Insgesamt sind in der HVR I nur 2 der 12 russischen Jäger/Sammler identisch mit den Individuen der Katakombengrabkultur. Das Bild in der Steppe könnte auf Gründereffekte zurückzuführen sein, falls es tatsächlich zu einer Einwanderung ausgehend vom baltischen Raum kam.

Möglicherweise spielte dabei die Fatjanovo-Kultur eine Rolle. Es handelt sich dabei vermutlich um eine östliche Variante der Schnurkeramik, vom Peipussee an der heutigen Grenze zwischen Estland und Russland bis zur mittleren Wolga. Für sie gibt es Hinweise auf Verbindungen zum Baltikum, zur Grübchenkeramik, zur Katakombengrab- und zur Poltavka-Kultur (Mallory & Adams 1997: 196 f, Anthony 2007: 377-383).

10.2.6 Beitrag der Steppenulturen zum modernen europäischen Genpool

Diverse palaeogenetische Studien zeigten, dass die heutige europäische Bevölkerung nicht rein auf die Vermischung von mesolithischen Jägern/Sammlern und frühen Neolithikern (LBK) zurückgeht. Es wird vermutet, dass demographische Entwicklungen im Verlauf des mittleren und späten Neolithikums den modernen mitteleuropäischen Genpool entscheidend mitprägten (Bramanti *et al.* 2009, Haak *et al.* 2010). Von Archäologen und Genetikern vermutet, aber nicht eindeutig bewiesen ist, dass dabei Einflüsse aus Osteuropa, insbesondere der Steppe eine wesentliche Rolle spielten. Diese Hypothese wird von den Ergebnissen von Brandt *et al.* (2013) gestützt. Allerdings gab es bis dato praktisch keine Sequenzdaten aus dem in Frage kommenden Zeitraum und Gebiet. Vor einigen Wochen zeigten Lazarides *et al.* (2014) anhand prähistorischer Genome, dass der genetische Influx während des Neolithikums auf einen jungpalaeolithischen, nordeurasischen Ursprung zurückzuführen ist.

Die prähistorischen Proben aus der Steppe wurden sowohl mit ost- als auch mitteleuropäischen modernen Populationen verglichen. Der osteuropäische Datensatz ist zusammengesetzt aus 246 Individuen aus Bulgarien, der Ukraine und Westrussland (Calafell *et al.* 1996, Malyarchuk & Derenko 2001, Malyarchuk *et al.* 2002) und umfasst

somit grob das Untersuchungsgebiet. Im Großen und Ganzen sind die Haplogruppenverteilungen in den alten und heutigen Proben ähnlich. Der einzig signifikante Unterschied betrifft die Haplogruppe U4, die in der Katakombengrabbkultur, wie bereits beschrieben, mit 30,4 % (11,6 % – 49,2 %) wesentlich häufiger ist als in der modernen Population (0,1 % – 3,4 %). Ansonsten gibt es kleinere, die Gruppen U5a, T1 und H betreffende Verschiebungen, die allesamt nicht signifikant sind, und sich durch Gendrift erklären lassen dürften. Einzelne Individuen im modernen Datensatz tragen asiatische Linien, die aber in der Summe nur 2,2 % ausmachen. Die Beobachtungen auf der Ebene der Haplogruppen werden auf der Nukleotidebene durch F_{ST} -Werte zwischen -0,006 und 0,020 bestätigt (Ausnahme: Poltavka Kultur mit 0,03556, $p > 0,5$). Die parsimonische Erklärung ist eine genetische Kontinuität zwischen den kupfer- und bronzezeitlichen Kulturen und den heutigen Osteuropäern aus diesem Gebiet. Zu diesem Ergebnis kam auch der auf Koaleszenzsimulationen beruhende Kontinuitätstest in Wilde *et al.* (2014), dessen Nullhypothese der Kontinuität nicht abgelehnt werden konnte.

Die genetischen Distanzen zum mittel- und nordeuropäischen Genpool mit Proben aus Deutschland, Polen und Litauen (Bramanti *et al.* 2009) sind vergleichbar mit denen zu Osteuropa (-0,004 – 0,028; Ausnahme Poltavka), so dass auch hier höchstwahrscheinlich Kontinuität gegeben ist. Innerhalb Mitteleuropas wurde von Brandt und Kollegen (2013) Kontinuität ab dem späten Neolithikum / der frühen Bronzezeit festgestellt (ungefähr 3 000 v. Chr. in der Mittelbe-Saale-Region). Nach den oben beschriebenen Affinitäten zwischen den mittel- und osteuropäischen Populationen könnten die Steppenpopulationen durchaus die oder eine der Komponenten sein, die der Mischung Mesolithiker und LBK noch fehlte (Bramanti *et al.* 2009), um den modernen Genpool zu erklären.

10.2.7 Östliche Kontakte der Steppenpopulationen

Es wurde bereits angedeutet, dass es praktisch keine Hinweise auf genetische Einflüsse aus Westsibirien auf die untersuchten Steppenpopulationen gab. Die einzigen älteren und zeitgleichen Populationen, von denen Sequenzdaten veröffentlicht wurden, stammen aus der Baraba-Waldsteppe im Südosten des westsibirischen Tieflands, den Ust-Tartas-, Odino- und Krotovo-Kulturen (Molodin *et al.* 2012). Aufgrund sehr hoher F_{ST} -Werte, und im Fall der Odino-Kultur keinen gemeinsamen Haplotypen mit den hier untersuchten Proben, sind diese Populationen im MDS-Plot deutlich separiert. Am nächsten kommt ihnen die (allerdings nur durch sechs Individuen repräsentierte) Dnjepr-Donetz-Kultur. Dies indiziert, wie bereits erwähnt, dass der nordpontische Raum bis zum Neolithikum durchaus von zentralasiatischen Elementen geprägt gewesen sein könnte, was mit den Erkenntnissen aus der Archäologie übereinstimmt (Parzinger 2006: 159, Nikitin *et al.* 2012).

In der Mittelbronzezeit folgte zunächst die Kultur mit Mehrwulstkeramik, dann die Srubna-Kultur im nordpontischen Raum auf die Katakombengrabkultur. Zwischen ihr und der östlich des Kaspischen Meeres angrenzenden Andronovo-Kultur werden Verbindungen gesehen, so dass sie von manchen Archäologen zu einem Komplex zusammengeführt werden (Mallory & Adams 1997: 541 f, Chernykh 2008). Von der Andronovo-Kultur liegen Sequenzdaten aus zwei Publikationen vor: die von asiatischen Linien (Haplogruppen A und C) geprägte Gruppe aus Molodin *et al.* (2012) und die überwiegend europäisch erscheinende Population aus Keyser *et al.* (2009). Dies spiegelt sich jeweils in sehr hohen ($0,11 - 0,18$, $p < 0,05$; Poltavka: $0,07416$, $p < 0,1$) und geringen, teils negativen F_{ST} -Werten zu den nordpontischen Kulturen wider. Laut den weit gefassten Angaben zum absolut-chronologischen Alter ist die zweite, durch europäische Linien ausgezeichnete Gruppe etwas jünger. Unter Annahme von Kontinuität zwischen den früheren Steppenkulturen Jamnaja- und Katakombengrabkultur und der Srubna-Kultur, scheint es auch auf der genetischen Ebene Verbindungen zur bestimmten Phasen oder Gruppen der Andronovo-Kultur gegeben zu haben.

Der Vergleichsdatensatz von Lalueza-Fox *et al.* (2004) aus dem bronzezeitlichen Kasachstan (BAK) ist noch etwas jünger (1 400 – 900 v. Chr.). Zur Jamnaja-Kultur ist hier keine genetische Distanz festzustellen, zur Katakombengrabkultur jedoch ist sie mit einem F_{ST} -Wert von 0,1308 signifikant hoch. Interessanterweise fallen sowohl die kasachischen als auch die Andronovo-Proben aus Keyser *et al.* (2009) in das Cluster mit den mitteleuropäischen Bauern und der Jamnaja-Kultur. Zur Repin-, einer Vorläuferkultur der Jamnaja gibt es Migrationshypothesen bezüglich einer Ostwärtsbewegung durch Kasachstan mit Genese der Afanasievo-Kultur im westlichen Altai-Gebirge (Anthony 2007: 307). Möglicherweise sind die Proben aus Kasachstan Nachkommen von Repin- oder Jamnaja-Leuten, die in diese Region kamen. Allerdings ist dies ohne tieferes Eindringen in die Materie der zentral- und osteurasischen Prähistorie sehr spekulativ.

10.3 Das Äneolithikum und die Bronzezeit in der osteuropäischen Steppe

Nachdem die großräumigen Kulturkomplexe der Steppe im Kontext des Neolithikums Europas betrachtet wurden, widmet sich dieser Abschnitt den Strukturen und demographischen Entwicklungen innerhalb dieser Kulturen des west- und nordpontischen Raumes vom fünften bis zum dritten vorchristlichen Jahrtausend.

10.3.1 Das Äneolithikum

Aus dem Äneolithikum konnten lediglich zehn Proben haplotypisiert werden, fünf aus der frühen (ca. 4 550 – 4 100 v. Chr.) und fünf aus der späten Phase (ca. 3 500 – 3 000 v. Chr.)³⁹. Vier der Individuen aus der Zeitschicht des frühen Äneolithikums stammen aus dem Fundort Smjadovo im heutigen Bulgarien. Die Proben datieren auf

³⁹ Chronologie nach Rassamakin (1999: 129)

etwa 4500–4500 v. Chr. Zu dieser Zeit gehörte Smjadovo dem Kodžadermen-Gumelnița-Karanovo VI-Komplex an und ist somit eher im Rahmen des südosteuropäischen und weniger des Steppenäneolithikums zu sehen. Erst in der darauffolgenden Cernavodă-Kulturphase sind Einflüsse aus der nordpontischen Steppe fassbar (Parzinger 1998b). Aus der gleichen Region stammt die spätäneolithische Probe aus Durankulak. Sie gehört archäologisch zu den Vorläufern der Jamnaja-Kultur und ist anders als die Proben aus Smjadovo dem Steppenäneolithikum anzurechnen. Im Folgenden umfasst der Begriff Äneolithikum alle 10 Individuen aus dieser Zeit, der Begriff Steppenäneolithikum wird verwendet, wenn die Proben aus Smjadovo ausgenommen sind.

Aufgrund der geringen Probenanzahl können zu Strukturen oder Entwicklungen im Verlauf des Äneolithikums keine gesicherten Aussagen gemacht werden. Es wurden zwar F_{ST} -Werte z. B. zwischen den Proben des frühen und des späten Äneolithikums, oder zwischen Smjadovo und dem Steppenäneolithikum berechnet, doch Gruppengrößen von vier bis fünf Individuen sind keine Basis für wissenschaftlich fundierte Erkenntnisse. Daher werden sie hier auch nicht weiter diskutiert.

In Abschnitt 10.2.1 zum Ursprung der Viehhaltung wurde bereits die Präsenz von „Jäger/Sammler“- und neolithischen Linien im Äneolithikum und das Zusammenleben der Menschen diskutiert. Erwähnenswert für die Charakterisierung der äneolithischen Stichprobe ist, dass von vier Individuen, die erfolgreich für den SNP in *ABCC11* genotypisiert wurden, zwei das „typisch asiatische“ Allel, das allerdings auch in modernen Europäern auftritt, tragen (VIN 1 homo- und MOB 1 heterozygot). Doch die Anzahl der Individuen ist viel zu niedrig, um Rückschlüsse zu ziehen. Nachdem in der neolithischen Dnjepr-Donetz-Kultur ein Drittel der Linien der Haplogruppe C zugeordnet sind, ist es nicht unerwartet, dass in irgendeiner Form noch „Überreste“ östlicher oder sibirischer Elemente in den nachfolgenden Kulturen vorhanden sind.

In der gesamtbronzezeitlichen Population wurde bei der Suche nach Spuren älterer Strukturen gesucht. Dabei fielen die Fundplätze, die im Verbreitungsgebiet der Kvitanska-Kultur des mittleren Äneolithikums liegen, durch hohe F_{ST} -Werte zur denen im Gebiet der Unteren Mikhailvoka- und der Khvalynsk-Kultur auf. Diese spiegeln die

Unterschiede zwischen der Schwarzmeerküste und dem Landesinneren in der Jamnaja-Kultur und die Unterschiede zwischen der Katakombengrabkultur und der Poltavka-Kultur wider. Aufgrund der archäologischen Befunde zu den weitreichenden Handelsbeziehungen im Äneolithikum und den verschiedenen Migrationen am Ende dieses Zeitabschnitts sind die Beobachtungen ist vermutlich nicht als Signatur einer genetisch isolierten, mitteläneolithischen Kvitanska-Kultur zu interpretieren. Die hier angedeuteten Strukturen in der Bronzezeit werden in den folgenden Abschnitten näher betrachtet.

10.3.2 Die Jamnaja-Kultur

Nachdem sich die Viehhaltung durch die Impulse und wahrscheinlich Einwanderungen aus dem Westen im Verlauf des Äneolithikums nach und nach durchgesetzt hatte, kam es um 3 000 v. Chr. zu einem relativ abrupten Wandel in der Tradition des Bestattungswesens, der den gesamten Steppenraum betraf. Als Träger dieser Neuerung – Grubengräber in Kurganen – sehen die Archäologen die Jamnaja-Kultur. Unsicher ist, wie es zu dieser rapiden Ausbreitung in dem weitgestreckten Steppengürtel kam, wo sie begann und die Identität der Bevölkerung. Als mögliche Ausgangspunkte werden das Wolga-Ural-Gebiet als einziger Ursprung (Masson 1998) mit anschließender Ausbreitung nach Westen oder zusätzlich noch das Gebiet um den Don (Anthony 2007: 304) oder den Dnjepr (Mallory & Adams 1997: 653) diskutiert. Genauso früh datierte Jamnaja-Funde im nordwestlichen Schwarzmeerraum und bisher ausschließlich jüngere Funde im Zentrum ziehen diese Hypothesen in Zweifel (Chernykh 2008). Dazu stellt sich die Frage, wie mobil die Gruppen waren. Basiert die Verbreitung des Grabbrauchs auf Migrationen von Bevölkerungsgruppen oder entwickelte er sich auf der jeweiligen lokalen Grundlage?

Mit dem vorliegenden Datensatz können die meisten dieser Fragen nicht beantwortet werden. Für eine mögliche Identifizierung des geographischen Ursprungs fehlen Daten insbesondere aus der Wolga-Don-Region. Es gibt allerdings einige Indizien auf die Vorgänge zur Zeit des Übergangs von der Kupfer- zur Bronzezeit. Aus genetischer Sicht, basierend auf diversen DNA-Studien (Haak 2006, Bramanti *et al.*

2009), würde man die äneolithischen Funde der vorliegenden Studie als Mischpopulation aus Jägern, Sammlern und Bauern charakterisieren, wobei ursprüngliche und eingewanderte Bewohner der Steppe möglicherweise noch voneinander getrennt lebten (siehe S. 203). Spätestens in der Bronzezeit allerdings finden sich sowohl mesolithische als auch neolithische Linien in gemeinsamen Fundplätzen mit gemeinsamer Lebensgrundlage, so dass man zu diesem Zeitpunkt wohl von einem Zusammenleben bzw. einer Vermischung ausgehen kann. Weitere Hinweise dazu könnte die Bestimmung der Y-chromosomalen Haplotypen geben.

Inwieweit sich die Haplogruppenzusammensetzung zwischen Äneolithikum und der darauffolgenden Jamnaja-Kultur ändert, lässt sich aufgrund der geringen Stichprobengröße der älteren Phase nicht sagen. Die U-Linien scheinen in der Häufigkeit abzunehmen, wenn man lediglich das Steppenäneolithikum (ohne Smjadovo) betrachtet (siehe Abbildung 9.8). Es ist nicht klar, ob die in der Jamnaja-Periode „neu“ auftauchenden Linien (z. B. T1, X, I) tatsächlich in dieser Phase erstmalig auftreten oder lediglich in der vorigen Zeitscheibe nicht erfasst wurden. Gut möglich ist ein Szenario andauernder Kontakte mit dem europäischen Neolithikum seit dem ersten Zusammentreffen der Steppenbevölkerungen mit der LBK bzw. deren Vorläufern. Weiteres zu diesen Beziehungen findet sich in den Abschnitten 10.2.2 bis 10.2.4.

Zu Beginn der Bronzezeit kam es zu einer relativ sprunghaften Veränderung auf der kulturellen Ebene, doch eine genetische Distanz zwischen den beiden Zeitscheiben ist nicht festzustellen ($F_{ST} = -0,00661$). Auch zwischen dem reinen Steppenäneolithikum und der Jamnaja-Kultur ist der F_{ST} -Wert von 0,03025 nicht signifikant. Zu den Haplotypen der Hälfte der zehn äneolithischen Individuen (2 von 4 aus Smjadovo, 3 von 6 aus der Steppe) gibt es Pendants in der Jamnaja-Kultur. Wie bereits gezeigt, scheinen, falls es äußere Einflüsse auf die Entwicklungen in der Steppe gab, diese aus westlicher Richtung gekommen zu sein.

Insgesamt deuten die fehlenden bzw. geringen genetischen Distanzen zwischen den jeweiligen Stichproben des Äneolithikums und der frühen Bronzezeit, zusammen mit den archäologischen Erkenntnissen (Rassamakin 1999: 125-127) auf eine

Kontinuität der Bevölkerung hin, die um etwa 3 000 v. Chr. eine neue Bestattungssitte etablierte. Die schnelle Verbreitung dieser Sitte steht im Einklang mit den signifikant negativen Werten von Tajima's D und Fu's F_s , die sich durch eine Expansion der Jamnaja-Population erklären lassen. Dafür spricht auch die sehr hohe genetische Diversität ($\hat{H} = 0,9870 \pm 0,0175$).

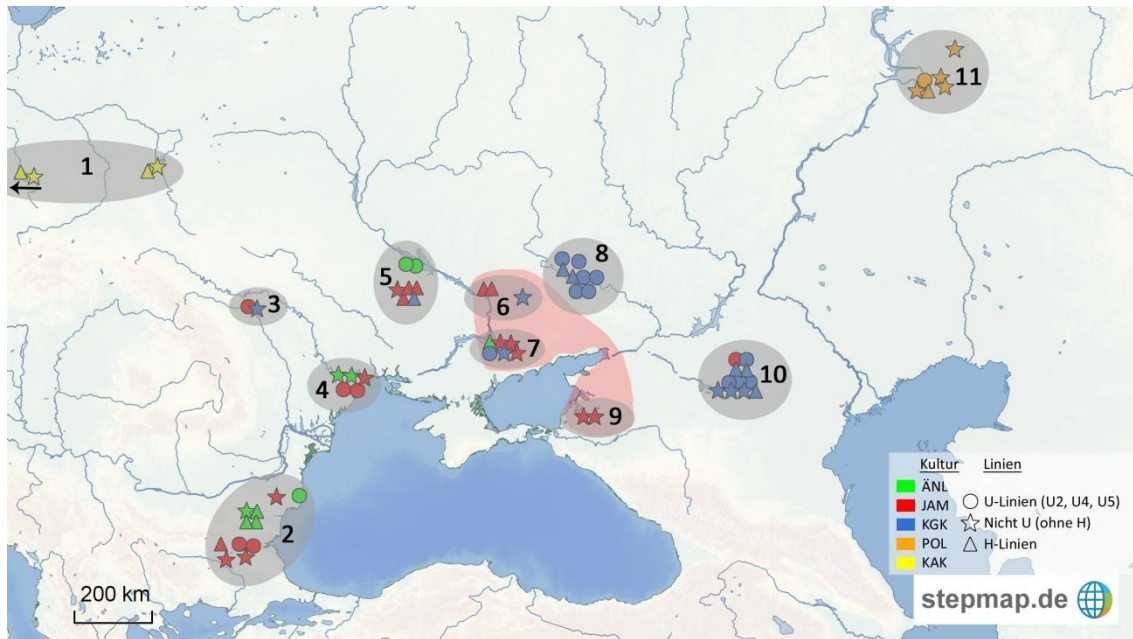


Abbildung 10.1: Fundplatz-Einheiten der geographischen Gruppen

Jedes Symbol repräsentiert ein Individuum bzw. eine DNA-Sequenz. Die Farben geben die Kulturzugehörigkeit an: grün = Äneolithikum, rot = Jamnaja-, blau = Katakombengrab-, orange = Poltavka-, gelb = Kugelamphorenkultur. Die Form der Symbole unterscheidet zwischen Linien der Haplogruppe U, die typisch für Jäger/Sammler-Populationen ist (Kreis) und in neolithischen Kulturen beobachtete Linien (Stern), wobei die hochfrequente Haplogruppe H gesondert herausgestellt ist (Dreieck). Der schwarze Pfeil indiziert, dass der Fundplatz Zauschwitz (ZAU) weiter westlich des Kartenausschnitts (in Sachsen) liegt. Die Fundplatz-Einheiten bzw. Ausgangsgruppen sind durch grau hinterlegte Ovale gekennzeichnet und nummeriert (vergleiche mit Tabelle 9.11 und Abbildung 9.9). Bloss-rot unterlegt ist die Region, aus der die Jamnaja-Gruppen mit hohen F_{ST} -Werten zu anderen Gruppen stammen.

Um den Verlauf der Ausbreitung näher zu charakterisieren, wurde der Untersuchungsraum auf regionale Unterschiede bzw. Muster überprüft. Es wurden verschiedene Gruppierungen basierend auf der geographischen, klimatischen oder ökologischen Lage der Fundorte, sowie auch auf den archäologisch festgelegten

Grenzen der Vorläuferkulturen aus dem Äneolithikum untersucht. Allerdings stößt man hier relativ schnell an die Grenzen des Datensatzes. Optimale Aufteilung(en) der Proben nach diesen Kriterien resultierten in der Regel in zu geringen Gruppengrößen, um robuste Ergebnisse zu erhalten und fundierte Aussagen zu treffen. Betrachtet werden hier lediglich Gruppen, die den Kompromiss zwischen einer sinnvollen spatialen und temporalen Einordnung und einer Mindestanforderung an die Stichprobengröße von mindestens sechs Individuen erfüllen.

Bei der differenzierten Analyse der Jamnaja-Zeitscheibe zeigen sich auf der Ebene der Haplogruppen regionale Unterschiede. In den entlang des Dnjepr gelegenen Fundplätzen (Einheiten 5, 6, 7 in Abbildung 10.1) treten im Gegensatz zu den anderen Jamnaja Fundorten keine U-Linien, sondern lediglich Typen der Haplogruppen H, T1 und I auf, von denen die beiden letzteren in Europa bisher erst ab dem mittleren Neolithikum nachgewiesen sind. Bei den F_{ST} -Werten tritt diese Gruppe im Vergleich mit den anderen Gegenden nicht besonders hervor. Dagegen fiel die Region nördlich bzw. um das Asowsche Meer (Einheiten 6, 7, 9), westlich begrenzt durch den Dnjepr, durch signifikant hohe genetische Distanzen zu den westlicheren Gebieten auf (je nach Gruppenkonstellation zwischen 0,138 bis 0,198). Es sei hier allerdings angemerkt, dass die Gruppe 6-7-9 lediglich aus sieben Individuen besteht. Der F_{ST} -Wert zwischen der Gruppe 4-7-9 entlang der Nordschwarzmeerküste und der Inlandsgruppe 3-5-6-10 beträgt 0,12522 ($p < 0,05$), exakt der gleiche Wert wie beim West-Ost-Vergleich zwischen den Fundplatz-Gruppen 3-4-5 und 6-7-9-10. Zu den Hintergründen dieses Musters kommt, neben einem etwaigen *sampling bias*, eine Reihe von Erklärungen in Frage:

Der Dnjepr bildete im Äneolithikum eine Art kulturelle Grenze. Zur Zeit ihrer maximalen Ausbreitung drangen die Ackerbauern der Tripolje-Kultur bis in die Dnjepr-Region vor. Östlich des Flusses lebten die viehzüchtenden Kulturen der Steppe (Dergačev 1998). Auch wenn es archäologisch nachgewiesene Kontakte und Austausche zwischen den beiden Kulturformen gab, könnte der Fluss dennoch eine genetische Grenze gewesen sein, die spätestens in der Bronzezeit zwar durchbrochen wurde, aber im mitochondrialen Erbgut Spuren hinterließ.

Gegen Ende des Äneolithikums kam es laut archäologischen Befunden zu zwei größeren Migrationen. Im nordwestlichen Schwarzmeerraum begann in der Pruth-Dnjestr-Waldsteppe eine Migration ausgehend von der Kasperovka-Gruppe (Životilovka-Volčansk-Migration): nach Süden in das Steppengebiet der Usatovo-Kultur (Fundort Majaki (4)) und nach Osten und Südosten in das Gebiet zwischen Bug und Dnjepr. Spuren dieser Migration soll es allerdings auch in der Region des unteren Don und dem Nordkaukasus geben. Von der anderen Seite her kam aus dem Gebiet des Mittleren Dons die expandierende Repin-Kultur nach Südwesten in die Dnjepr-Region. Außerdem migrierte sie in südlicher Richtung zum unteren Don und der unteren Wolga (Rassamakin 1999: 123-125). Die beobachteten F_{ST} -Werte könnten auch, zumindest teilweise, eine Folge dieser Ereignisse gegen Ende des Äneolithikums sein. Rassamakin schreibt außerdem, dass die Jamnaja-Kultur Elemente von allen lokalen spätäneolithischen Formationen absorbierte, aber möglicherweise wären die Kvitanska- und die Životilovka-Volčansk-Gruppe davon ausgenommen (1999: 125 f). Das Kerngebiet der Kvitanska-Kultur befand sich in etwa in der Region um den Dnjepr und nördlich des Asowschen Meeres.

Importfunde in der nordpontischen Steppe aus der etwa zeitgleichen, nordwestlich des Untersuchungsraumes angesiedelten Kugelamphorenkultur (Dergačev 1998, Rassamakin 1999: 127) könnten auch mit einem (regional beschränkten) genetischen Influx einhergegangen sein, der zu den Distanzen zwischen dem Gebiet des Dnjepr und den anderen Regionen führte. Allerdings gibt es bisher, inklusive der vier im Rahmen dieser Studie sequenzierten Individuen, nur fünf DNA-Sequenzen aus der Kugelamphorenkultur (Kozłowski *et al.* 2014: 1 Individuum mit Haplogruppe K2a*). Es bedarf mehr Proben dieser Kultur, um diesen Erklärungsansatz zu überprüfen.

Eine alternative oder weitere mögliche Ursache für die Beobachtungen könnten in der geographischen Lage der/des etwaigen Ursprungsgebiete/s und der Art der Ausbreitung der Jamnaja zu finden sein. Eine direkte Untersuchung diesbezüglich ist aufgrund der fehlenden Datengrundlage in der Don- und Wolga-Region nicht möglich, aber für indirekte Indizien können die sechs Proben der Poltavka-Kultur verwendet werden. Bei dieser Kultur handelt es sich archäologisch um die Nachfolger der Jamnaja im Wolga-Gebiet. Anders als in der nordpontischen Steppe wurden nördlich des

Kaspischen Meeres keine Katakomben angelegt. Sämtliche paarweisen F_{ST} -Werte zwischen der Poltavka-Kultur und den Jamnaja-Gruppen aus den verschiedenen Regionen sind entweder negativ oder maximal 0,001. Die „Ausnahme“ bilden kleinere Gruppierungen, die die Fundplatz-Einheiten 6 und 7 enthielten, mit F_{ST} -Werten zur Poltavka von 0,01461 bis 0,02061, die nicht signifikant sind. Ein ähnliches Bild ergibt der Vergleich regionaler Jamnaja-Gruppen mit der zeitlich überlappenden frühen Phase der Katakombengrabkultur, die hauptsächlich aus Kalmückien stammen (siehe unten). Das Phänomen „*isolation by distance*“ wird demnach nicht beobachtet. Damit fehlt ein wichtiger Hinweis auf einen reinen Ideentransfer; ein Modell, bei dem der Aktionsradius der regionalen Gemeinschaften eher begrenzt wäre und die neue Bestattungssitte von einem Ursprung aus sukzessive von benachbarten, aber relativ eigenständigen Verbänden übernommen wurde. Man beachte allerdings, dass es sich hier nur um indirekte Vergleiche handelt, da aus der Jamnaja-Kultur an sich kaum Proben östlich des Dons vorliegen. Insgesamt sind jedoch die Werte der Diversitätsstatistiken in allen Gebieten hoch, was für eine große Populationsgröße und eine überregionale Vermischung spricht.

Das große Verbreitungsgebiet der Jamnaja-Kultur – im Westen scheinen einige Gruppen bis in die Ungarische Tiefebene vorgedrungen zu sein – spricht für Migrationen von Jamnaja-Gruppen. Die genetischen Distanzen zwischen den küstennahen und küstenfernen Fundorten könnten vielleicht aus Migrationsrouten resultieren. So haben sich einige Gruppen möglicherweise entlang der Küste bewegt, andere weiter nördlich in der Steppe und Waldsteppe, während in manchen Gebieten keine Bewegungen stattfanden.

Basierend auf dem momentanen Datensatz kann noch kein Modell zu den bevölkerungsbiologischen Hintergründen zur Verbreitung des Grubengrabritus entwickelt werden. Entweder ist die Jamnaja-Kultur insgesamt sehr homogen und die Abweichung über dem Asowschen Meer hat eine externe Ursache: eine Stichprobenverzerrung oder eine Signatur der Kulturverhältnisse im Äneolithikum. Oder es gab Strukturen innerhalb der Jamnaja: eine Ost-West Unterteilung, wie sie z. B. von Anthony (2007: 304), allerdings mit dem Don als Grenze, beschrieben wurde, oder Verhältnisse, die zu einer Sonderstellung der Asowschen Subkultur(en) führten.

Weitere aDNA-Daten der Jamnaja, insbesondere aus dem Gebiet zwischen Dnjepr und Wolga, sowie zusätzliche neolithische und äneolithische Proben könnten zu einer Aufschlüsselung der Prozesse beitragen. Die vermuteten Migrationen vor und während der Jamnaja-Phase könnten allerdings viele Spuren verwischt oder überdeckt haben.

10.3.3 Die Katakombengrabkultur

Während im gesamten Steppenraum und weit darüber hinaus die Jamnaja-Kultur mit ihren schachtartigen Grubengräbern verbreitet war, tauchten die ersten frühen Katakombengrabstrukturen ab etwa 2 700 v. Chr. zunächst zwischen Don und Donez, dann auch im nördlichen Kaukasus und Kalmückien auf. In ihrer entwickelten Phase dehnte sich die Katakombengrabkultur auch in die Steppengebiete westlich des Dnjepr aus und löste schließlich die Jamnaja-Kultur ab (Kaiser 2003: 65-68, 2007, Chernykh 2008).

Eine der bedeutendsten Fragen ist die nach der Identität der Menschen, die die Katakombengräber bauten. Der vorliegende Datensatz liefert hierzu einige Anhaltspunkte. Vorab allerdings eine kleine Anmerkung: Dem fachkundigen Leser dürfte die bereits erwähnte Problematik von aDNA-Studien bezüglich der kleinen Individuenzahlen und deren ungleichmäßigen Verteilung hinlänglich bekannt sein. Dieses Phänomen ist bei den Proben der Katakombengrabkultur besonders prägnant, da 10 der 23 Individuen aus zwei nur wenige Kilometer voneinander entfernten Fundorten (Pesčanij V und Temrta III und V) und nur vier Kurganen stammen. Damit ist die Wahrscheinlichkeit sehr hoch, dass hier direkte Verwandtschaftsbeziehungen die genetische Zusammensetzung der Stichprobe prägen. Ein Kind (PEJ 4) aus einer Doppelbestattung wurde nicht in die Analysen mit einbezogen, da es sich bei PEJ 3 offensichtlich um dessen Mutter handelte. Im Folgenden werden die genetischen Befunde im Kontext der prähistorischen Entwicklung diskutiert, doch sollte insbesondere in dieser Zeitscheibe des Datensatzes ein *sampling bias* als Erklärungsmöglichkeit nicht außer Acht gelassen werden.

In Bezug auf den Ursprung oder die Identität der Erbauer der Katakombengräber unterstützen die niedrigen, nicht signifikanten F_{ST} -Werte prinzipiell eine Entwicklung aus der lokalen Jamnaja-Bevölkerung heraus. Die Distanz zwischen den beiden Kulturkomplexen liegt bei 0,02354. Dabei beträgt die Distanz der Jamnaja- zur synchronen frühen Katakombengrabkultur lediglich 0,00723, zur nachfolgenden entwickelten Katakombengrabkultur 0,01323. Allerdings beschränkt sich die genetische Affinität zwischen den beiden Kulturen auf die Jamnaja-Gruppen westlich des Dnjeprs, wobei über 90 % der Proben der Katakombengrabkultur aus Fundplätzen östlich des Flusses stammen. Die aus dem gleichen Gebiet östlich des Dnjepr stammenden Proben (Einheiten 6 bis 10) weisen einen F_{ST} -Wert von 0,15020 ($p < 0,01$) zwischen den beiden Kulturen auf. Die Jamnaja-Gruppen nördlich des Asowschen Meeres scheinen demnach nicht die Katakombengräber konstruiert zu haben. Die Distanz zwischen der Katakombengrabkultur, insbesondere der frühen Phase, und der äneolithischen Gruppe sind mit F_{ST} -Werten von 0,05435 bzw. 0,05953 etwas größer und marginal signifikant ($p < 0,1$). Dies ist durch den zeitlichen und räumlichen Abstand erklärbar.

Das auffälligste Merkmal der Katakombengrabkultur, neben der geringen Diversität, ist die hohe Frequenz der Haplogruppe U4 (30 %) im Datensatz. Diese ist wahrscheinlich nicht auf direkte Verwandtschaften innerhalb der Stichprobe zurückzuführen. Die sieben Individuen, die U4-Linien tragen, stammen aus sechs Kurganen von fünf verschiedenen Fundplätzen. Stattdessen ist die Häufigkeit von U4 entweder die Folge lokaler Entwicklungen oder von Einflüssen von außerhalb des Untersuchungsgebiet:

Die sehr geringe Diversität der frühen Katakombengrabkultur ist, wie bereits erwähnt, höchstwahrscheinlich von artifizieller Natur. Man könnte sie aber auch als Folge eines Gründereffekts interpretieren. In diesem Szenario hätte eine kleine lokale Gruppe innerhalb der Jamnaja begonnen, Katakombengräber zu bauen und sich möglicherweise zeitweilig sozial und genetisch isoliert. Dies würde das Fehlen der bei den Jamnaja-Individuen noch häufigen Haplogruppe T, sowie die plötzliche Zunahme der Frequenz der Haplogruppe U4 erklären. Allerdings wäre in diesem Fall auch eine genetische Distanz der frühen Katakombengrabkultur zu den westlichen und nicht nur

zu den östlichen Jamnaja-Gruppen zu erwarten. Letztlich sprechen auch die leicht negativen, wenn auch nicht signifikanten Werte der Tajima's D - und Fu's F_s -Statistik gegen einen Gründereffekt.

Basierend auf den Beobachtungen in den veröffentlichten Jäger/Sammler-Populationen Westeurasiens ist es denkbar, dass bereits in der ursprünglichen Jäger/Sammler-Population der nordpontischen Steppe U4-Linien verbreitet waren und bis mindestens in die mittlere Bronzezeit noch nachzuweisen sind. Allerdings würde man entsprechende Anteile auch in der Jamnaja-Kultur erwarten. Entweder wurden diese nicht erfasst, weil sie z. B. auf die Jamnaja-Gruppen in der Don-Wolga-Region beschränkt gewesen sein könnten, oder die Beobachtungen lassen sich durch ein Modell mit externen Einflüssen erklären.

Nach diesem Modell wäre der rapide Anstieg des Anteils an U4-Linien auf eine Zuwanderung in die Steppe zurückzuführen sein. Entsprechend den Ausführungen in Abschnitt 10.2.5 könnte deren Ausgangspunkt in den Jäger/Sammler-Refugien im östlichen Baltikum liegen. Weitere aDNA-Daten, z. B. von östlichen Jamnaja-Gruppen oder aus den Gebieten nördlich des Untersuchungsraums, könnten Aufschluss geben, ob ein „steppeninternes“ oder „-externes“ Modell zu favorisieren ist.

Archäologischen Erkenntnissen zufolge erschien die Katakombengrabkultur in ihrer frühen Phase als uniform. In der entwickelten Phase lassen sich zwei Gruppen mit Unterschieden im Grabbrauch ausmachen: die westliche Ingulgruppe um den Dnjepr und westlich davon und eine östliche Gruppe am Donez und unterem Don, mit einem Übergangsgebiet nördlich des Asowschen Meeres zwischen Dnjepr und Donez, in dem beide Bestattungstypen nachgewiesen sind (Kaiser 2003: 55). Anhand der F_{ST} -Werte lassen sich genetisch weder temporale noch spatiale Populationsstrukturen innerhalb der Katakombengrabkultur feststellen. Ähnlich wie bei der Jamnaja-Kultur fehlen allerdings auch hier für eine Identifizierung etwaiger Strukturen die entsprechenden Proben für eine sinnvoll geographisch basierte Analyse, in diesem Fall Proben westlich des Dnjepr.

10.3.4 Die Poltavka-Kultur

In der Wolga-Region etablierten sich keine Katakombengräber, hier wurde der Grubengrabbau durch die Poltavka-Kultur weiter betrieben (Mallory & Adams 1997: 439 f, Häusler 1998). Wie bereits gezeigt, ist diese genetisch eng mit der Jamnaja-Kultur verwandt (siehe S. 219). Die voneinander unabhängige Weiterentwicklung der Bestattungsriten zwischen dem nordkaspischen und dem nordpontischen Raum ist auch genetisch sowohl durch einen hohen F_{ST} -Wert (0,10078; $p < 0,05$) als auch einem geringen Anteil identischer Haplotypen (0 % – 4,4 %) zwischen der Poltavka- und der frühen Katakombengrabbkultur fassbar (zum Vergleich: Jamnaja – Poltavka 1,3 % – 10,3 %). Die geringere genetische Distanz in der darauffolgenden Phase ($F_{ST} = 0,03680$; $p > 0,1$) lässt auf eine Intensivierung der Kontakte zwischen den beiden Kulturen schließen. Bei der populationsgenetischen Interpretation der Statistiken zur Poltavka-Kultur ist allerdings die geringe Stichprobengröße von sechs Individuen zu berücksichtigen.

10.4 Marker unter Selektionsverdacht

Im Kapitel zu Selektion in der Humanevolution wurde bereits eine kurze Übersicht zu den im Rahmen dieser Arbeit betrachteten *loci* gegeben (Abschnitt 5.4). Hier folgt nun eine genauere Beschreibung des Wissenstands zu den Funktionen der Gene, den funktionellen bzw. phänotypischen Folgen der untersuchten Punktmutationen und der zum Teil sehr spekulativen Selektionshypothesen. Darin eingebettet ist die Diskussion der Genotypisierungsergebnisse für die Kulturen der nordpontischen Steppe und ihre wissenschaftliche Bedeutung. Zunächst werden die SNPs besprochen, für die eine Selektion aufgrund ihrer Bedeutung im Vitamin D-Haushalt angenommen wird. Dann folgen *loci* mit Einfluss auf den Wasser- bzw. Elektrolythaushalt. Den Abschluss der klimatischen Faktoren bildet die Anpassung an Kälte. Wie der Organismus mit Stoffen aus der Umwelt umgeht, zeigen die Gene in den Kategorien Ernährung und Pathogene. Die Aufteilung der *loci* in diese Gruppen ist keineswegs als starre

Abtrennung zu sehen. Wie sich zeigen wird, spielen manche Gene in mehreren dieser Bereiche eine Rolle.

10.4.1 Pigmentierung & Vitamin D-Haushalt

10.4.1.1 Pigmentierungs-Pathway

Das für die Haut-, Haar- und Augenfarbe verantwortliche Pigment ist Melanin. Es kommt in zwei Formen vor, dem hellen, orange-gelben Phäomelanin und dem dunklen, schwarz-braunen Eumelanin. Synthetisiert wird es in Melanosomen, Organellen in speziellen Pigmentzellen, den Melanocyten. Diese kommen in der basalen Schicht der Epidermis, in Haarfollikeln und der mittleren Schicht des Auges vor. Die Pigmentierung hängt nicht von der Anzahl der Melanocyten, sondern von der Anzahl, Größe und Verteilung der Phäo- und Eumelanosomen ab (Jablonski & Chaplin 2000). Melanosomen reifen in vier Stadien. Die mit Lysosomen verwandten Organellen werden vom Golgi-Apparat und dem Endoplasmatischen Reticulum produziert. Phase I-Prämelanosomen werden, nachdem Enzyme wie Tyrosinase oder andere strukturelle Proteine eingebracht sind, als Vesikel vom *trans*-Golgi-Netzwerk abgeschnürt. Daraufhin beginnt das zweite Stadium mit der Melaninsynthese. Phäomelanosomen bleiben in dieser Phase II. Eumelanosomen benötigen zusätzliche Enzyme und eine fibrilläre Matrix. Daher durchlaufen sie noch zwei weitere Phasen, die mit dem Import der Enzyme TYRP1 (tyrosinase-related protein 1) und DCT (DOPAchrome tautomerase oder auch TYRP2) beginnen. Reife Melanosomen werden durch die dendritischen Ausläufer der Melanocyten transportiert und an benachbarte Keratinocyten transferiert. Diese machen den größten Anteil der Zellen in allen Epidermisschichten aus. Die Melanosomen mit dem darin gelagerten Melanin werden in den Keratinocyten als Schutz vor der schädlichen UV-Strahlung verteilt. Ausnahme ist das Auge, wo die Melanosomen in den Melanocyten der Iris verbleiben (Sturm *et al.* 2001, Scherer & Kumar 2010).

An der in Phase II-Prämelanosomen beginnenden Melanogenese ist eine Reihe von Enzymen direkt oder indirekt beteiligt. Das Enzym Tyrosinase katalysiert die ersten

beiden Schritte der Umwandlung der Aminosäure Tyrosin in Phäo- oder Eumelanin. Tyrosin wird zunächst zu DOPA (3,4-Dihydroxyphenylalanin), und weiter zu DOPACHINON oxidiert. Der erste Schritt ist dabei geschwindigkeitslimitierend. In Anwesenheit von Cystein oder Glutathion wird DOPACHINON zu 5-S-Cysteinyl-DOPA umgewandelt. Aus diesem entsteht durch Polymerisation Phäomelanin über das Zwischenprodukt 1,4-Benzothiazinylalanin.

Ist kein Cystein vorhanden, so autooxidiert DOPACHINON zu LeukoDOPACHROM (CycloDOPA) dann DOPACHROM. Dieses instabile Molekül decarboxyliert zu DHI (5,6-Dihydroxyindol). Mithilfe von DCT kann DOPACHROM aber auch zu DHICA (5,6-Dihydroxyindol-2-Carbonsäure) umgewandelt werden. Diese beiden Produkte polymerisieren zu Eumelanin (Sturm *et al.* 2001, Ito 2003, Sturm & Frudakis 2004, Scherer & Kumar 2010).

Das Mengenverhältnis zwischen Phäo- und Eumelanin wird zu Beginn des Signalwegs zur Eumelanin-Synthese bestimmt (Scherer & Kumar 2010). MC1R (melanocortin-1 receptor) an der Oberfläche der Melanocyten wird durch Bindung von α -MSH (α -Melanocyten-stimulierendes Hormon) stimuliert. Hierdurch wird cAMP (cyclisches Adenosinmonophosphat), ein *second messenger*, aktiviert und die Signaltransduktion eingeleitet. Durch die erhöhten cAMP-Level wird Proteinkinase A stimuliert, die im Zellkern Transkriptionsfaktoren der CREB-Familie (cAMP responsive element binding protein) phosphoryliert. Diese verstärken daraufhin die Transkription von *MITF* (*microphthalmia-associated transcription factor*), welcher wiederum die Transkription von *TYR*, *TYRP1* und *DCT* aktiviert. Die Proteine werden in die Melanosomen transportiert, wo die oben beschriebene Melaninsynthese stattfindet. Ohne diese Aktivierungskette ist nur eine gewisse Grundmenge an Tyrosinase vorhanden, und es wird Phäomelanin produziert. Erst die Stimulierung von MC1R führt zur Synthese von Eumelanin. Durch Bindung von ASIP (agouti signaling protein), einem Antagonist von MC1R, wird die cAMP-Konzentration reduziert und Phäomelanin produziert (Scherer & Kumar 2010).

Wie diese Ausführungen andeuten, ist das System der Pigmentierung äußerst komplex und auf der einen Seite genetisch konstituiert, aber auch den akuten

Umweltbedingungen gegenüber in gewissem Maße flexibel. Die Mechanismen sind noch nicht vollständig entschlüsselt. Die beteiligten Gene liegen in verschiedenen Varianten vor, deren Vorkommen teilweise global stark variieren und dabei mit der Sonnenintensität korrelieren. Für einige dieser Gene werden diese Unterschiede auf Selektion zurückgeführt (siehe z. B. Lao *et al.* 2007, Sabeti *et al.* 2007, Rees & Harding 2012). Vier an der Pigmentierung beteiligten Gene wurden im Rahmen dieser Arbeit untersucht: *TYR*, *SLC45A2*, *HERC2* und *SLC24A5*. In jedem dieser Gene wurde ein funktionaler, unter Selektionsverdacht stehender SNP betrachtet und die Häufigkeiten der Allele in den kupfer- und bronzezeitlichen Populationen ermittelt. Die Ergebnisse zu den ersten drei Genen bzw. SNPs wurden bereits in Wilde *et al.* (2014) veröffentlicht. Dennoch sollen sie hier der Vollständigkeit halber kurz besprochen werden. Die Selektionshypothese, die mindestens dreien der vier Gene zugrunde liegt, wird im Anschluss der Betrachtung des *SLC24A5*-Gens diskutiert (Abschnitt 10.4.1.6).

10.4.1.2 TYR

Die Funktion von *TYR* (*tyrosinase*) in der Melaninsynthese wurde bereits beschrieben. Bei dem untersuchten SNP (rs1042602) handelt es sich um eine Transversion (C > A), die einen Aminosäureaustausch (Ser192Tyr) zur Folge hat. Das Alter der Mutation wurde mit der *rho*-Statistik auf etwa 6 100 Jahre, mit Bayesscher Koaleszenz auf ca. 15 600 Jahre geschätzt. Die abgeleitete *Tyrosinase*-Variante kommt hauptsächlich in nordafrikanischen (ca. 46 %), europäischen (ca. 36 %) und südasiatischen (ca. 20 %) Populationen vor und geht mit einer verminderten Enzymaktivität einher (Hudjashov *et al.* 2013). In Zellkulturen wurden positive Zusammenhänge zwischen der katalytischen Aktivität der Tyrosinase, dem pH-Wert in den Melanosomen und der Hautfarbe nachgewiesen (Ancans *et al.* 2001, Fuller *et al.* 2001).

Bezüglich der phänotypischen, die Pigmentierung betreffenden Ausprägung des abgeleiteten Allels in menschlichen Populationen gibt es verschiedene Beobachtungen. In südasiatischen Population stellten Stokowski *et al.* (2007) eine signifikante Assoziation dieses SNPs mit dem Melanin-Gehalt in der Haut fest.

Untersuchungen an einer afrikanisch-europäischen Mischpopulation kamen zu einem ähnlichen Ergebnis (Beleza *et al.* 2013). Obwohl *TYR* in europäischen Populationen hochpolymorph ist, wurde dort, mit Ausnahme der Studie von Shriver *et al.* (2003), kein signifikanter Einfluss des SNPs auf die Hautfarbe beobachtet (Norton *et al.* 2007, Sulem *et al.* 2007). Sulem und Kollegen (2007) stellten in Isländern und Dänen einen Zusammenhang zwischen dem abgeleiteten Allel und dem Fehlen von Sommersprossen fest, während laut Frudakis *et al.* (2003) eine schwache Assoziation des SNPs mit der Augenfarbe in Europäern besteht. Selektion konnte anhand von modernen Daten nicht eindeutig nachgewiesen oder abgelehnt werden (Norton *et al.* 2007, Sulem *et al.* 2007, Sturm 2009, Beleza *et al.* 2013, Hudjashov *et al.* 2013).

Mit den vorliegenden Genotypisierungen der prähistorischen Individuen konnte eine positive Selektion des abgeleiteten Allels in Osteuropa innerhalb der letzten fünftausend Jahre aufgezeigt werden. Bezüglich der beiden Allele des *TYR locus* (Genotypen: 40x CC, 4x CA) befindet sich die untersuchte Gesamtpopulation nicht im Hardy-Weinberg-Gleichgewicht. Die Allelfrequenzen zwischen den kupfer- und bronzezeitlichen Proben und einer modernen ukrainischen Stichprobe (Hollfelder 2013) sind signifikant unterschiedlich. Mithilfe einer *forward-in-time*-Simulation wurde in Wilde *et al.* (2014) für eine große Spannweite an Parametern gezeigt, dass die beobachtete Zunahme des abgeleiteten Allels nur unter der Wirkung von Selektion, mit einem Vorteil von etwa 2,6 % bei Kodominanz (1,6 % – 10,4 % bei dominanter/rezessiver Vererbung), zu erklären ist.

10.4.1.3 *SLC45A2*

SLC45A2 (*solute carrier family 45, member 2*), in der Literatur z. T. auch als *MATP* (*membrane associated transporter protein*) oder *AIM1* (*absent in melanoma 1* (Antigen)) bezeichnet, kodiert ein Membranprotein, das – nachgewiesen am Mausmodell – am intrazellulären Transport der Tyrosinase und anderen Pigmentierungsenzymen (z. B. *TYRP1*) zu den Melanosomen beteiligt ist (Costin *et al.* 2003). Aufgrund eines hohen Homologiegrades zu Saccharose/Protonen-Symportern in Pflanzen vermuten Newton *et al.* (2001) eine ähnliche Funktion bei *SLC45A2*.

Eventuell könnte der Transporter das osmotische Potential und damit das Volumen der Melanosomen oder den pH-Wert beeinflussen, der sich wiederum auf die Enzymaktivität der Tyrosinase auswirkt. Auch wenn über die genaue Funktion von *SLC45A2* spekuliert wird, so ist eine Assoziation mit Pigmentierung eindeutig nachgewiesen. Eine nicht-synonyme Mutation, C1122G / Leu374Phe (rs16891982) beeinflusst die Haut-, Haar- und Augenfarbe. Insbesondere für die ancestrale Leucin-Variante ist die signifikante Kopplung mit dunkler Pigmentierung und hoher Tyrosinaseaktivität bewiesen (z. B. Graf *et al.* 2005, Stokowski *et al.* 2007, Cook *et al.* 2009, Nan *et al.* 2009, Valenzuela *et al.* 2010). Das abgeleitete Allel ist nach Sabeti *et al.* (2006) in Europa (beinahe) fixiert. Den Untersuchungen von Norton *et al.* (2007) zufolge tritt es in Europäern mit einer Häufigkeit von 87 %, in Nicht-Europäern von 17 % auf. Dabei ist ein Süd-Nord-Gradient zunehmenden Vorkommens zu beobachten. In der gleichen Studie wurde anhand von drei Neutralitäts- und Diversitätstests (Tajima's *D*, *LSBL* (*locus-specific branch length*) und *InRH* (*logarithmically transformed ratio of expected heterozygosity*)) ein starkes Selektionssignal für das mit heller Pigmentierung assoziierte Phenylalanin-Allel in der europäischen Population nachgewiesen. Zu diesem Ergebnis kamen auch Soejima *et al.* (2006), die mithilfe von vier Neutralitätstests (Tajima's *D*, *Fu* und *Li's F*, *Fu* und *Li's D*, sowie *Fay* und *Wu's H*) und eine Haplotypenanalyse einen *selective sweep* in der Region des *SLC45A2*-Gens nachwies. Auch in einer genomweiten Studie wurde *SLC45A2* als Ziel von Selektion identifiziert (Sabeti *et al.* 2006). Die Mutation entstand vermutlich nach der Aufspaltung von Asiaten und Europäern (Norton *et al.* 2007). Der Beginn des *sweeps* wurde auf 11 000 bis 19 000 Jahre vor heute datiert, der Selektionskoeffizient auf $S = 0,04 - 0,05$ geschätzt (Beleza *et al.* 2013).

Die vorliegende Studie konnte, wie in Wilde *et al.* (2014) bereits veröffentlicht, diese Ergebnisse mit aDNA-Daten bestätigen. Am besten wurden die Veränderung der Allelfrequenzen über die Zeit mit einem Selektionskoeffizienten von $S = 0,03$ in einem kodominanten, $S = 0,022$ in einem rezessiven und $S = 0,088$ in einem dominanten Vererbungsmodell des abgeleiteten Allels erklärt. Während am *TYR locus* in der Kupfer- und Bronzezeit noch das ancestrale Allel vorherrschte, trugen über 43 % der Individuen (20 von 46) aus dieser Zeit bereits das *SLC45A2**G-Allel. Demnach hat hier

die Selektion in den letzten 5 000 Jahren auf *standing variation* gewirkt, bzw. der *sweep* war zur Lebenszeit der Proben bereits mitten im Gange, wie auch von Beleza *et al.* (2013) berechnet. Überraschenderweise fällt der Anteil der GG-tragenden Individuen von der Jamnaja-Kultur zur Katakombengrabkultur um 33,3 % auf 12,5 % ab, während die Heterozygoten nur um 4 % weniger werden. Möglicherweise besteht ein Zusammenhang mit der mtDNA. Im Gesamtdatensatz tragen die zwei bzw. drei für *SLC45A2* genotypisierten U4- und U5-Typen CC oder CG. Lediglich das Individuum mit der mtDNA-Haplogruppe U2 trägt GG. U-Linien (insbesondere U4) sind besonders häufig in der Katakombengrabkultur. Die „neolithischen“ (nicht U-) Linien verteilen sich auf alle drei Genotypen.

10.4.1.4 HERC2

Der dritte *locus*, zu dem bereits die Ergebnisse veröffentlicht wurden, ist der SNP rs12913832 in *HERC2* (*hect domain and RCC1-like domain 2*). *HERC2* gehört nicht zu den Pigmentierungsgenen, sondern ist ubiquitär an diversen Zellprozessen beteiligt: an Proteintransporten, intrazellulären Degradationsprozessen und der Regulation des Zellzyklus (Ji *et al.* 1999, Nicholls & Knepper 2001, Kuhnle *et al.* 2011, Cubillos-Rojas *et al.* 2014). 11,7 kb (Kilobasen) *downstream* von *HERC2* liegt das *OCA2*-Gen (*oculocutaneous albinism II*). Dieses kodiert das P-Protein (melanocyte-specific transporter protein oder pink-eyed dilution protein homolog), welches am Transport von kleinen Molekülen in die Melanosomen beteiligt ist und somit die Melanogenese mit einleitet. Zu seinen Substraten zählen auch Anionen. Auf diese Weise trägt das P-Protein zur Neutralisation des melanosomalen pH-Wertes bei. Wie bereits angedeutet, wird dadurch die Produktion von Eumelanin gefördert (Puri *et al.* 2000, Ancans *et al.* 2001, Fuller *et al.* 2001).

Einige Polymorphismen in *OCA2* wurden mit Unterschieden in der Pigmentierung in Verbindung gebracht. Die stärkste Assoziation, insbesondere mit der Augenfarbe (Eiberg *et al.* 2008, Sturm *et al.* 2008), aber auch mit der Haarfarbe und Bräunungsfähigkeit (fakultative Pigmentierung) (Han *et al.* 2008), wurde für einen SNP in Intron 86 des *HERC2* Gens festgestellt. Rs12913832 liegt in einer hochkonservierten

Region von *HERC2*, die vermutlich als Regulator von *OCA2* fungiert (Eiberg *et al.* 2008, Sturm *et al.* 2008). Die Mutation (A > G) führt zu einer schwächeren Chromatin-Schleife zwischen dem *enhancer*-Element und dem *OCA2-promotor*, so dass die Expression des Pigmentiergens reduziert wird (Visser *et al.* 2012). Homozygote Träger des abgeleiteten G-Allels haben mit fast hundertprozentiger Wahrscheinlichkeit blaue oder grau-/grün-blaue Augen (Eiberg *et al.* 2008, Sturm *et al.* 2008, Branicki *et al.* 2009). Über 75 % der heterozygoten und für das ancestrale Allel homozygoten Testpersonen in der Studie von Branicki *et al.* (2009) hatten braune Augen, lediglich 15 % hatten eine blaue oder blau-graue Iris (die restlichen Individuen hatten grüne Augen). Die Autoren testeten Modelle der verschiedenen Erbgänge mit dem Ergebnis, dass das mit blauer Augenfarbe assoziierte G-Allel rezessiv vererbt wird.

Eiberg *et al.* (2008) schätzten das Alter der Mutation auf 6 – 11 000 Jahre, mit einem möglichen Ursprung in der nordwestlichen Schwarzmeerregion. Auch vermuteten sie, dass eventuell Selektion zu den hohen Frequenzen des G-Allels in Skandinavien und dem Baltikum geführt habe, allerdings wurde diese Hypothese nicht getestet. Die bisherigen Studien beschränkten sich auf die Untersuchung von Assoziationen zwischen Geno- und Phänotyp. In Wilde *et al.* (2014) wurde rs12913832 erstmals auf Selektion untersucht. Mit einem Selektionskoeffizient von $S = 0,036$ scheint das rezessive G-Allel in den letzten 5 000 Jahren positiv selektiert worden zu sein.

Im vorliegenden Datensatz sind nur 6 der 52⁴⁰ Individuen (11,54 %) homozygot für das G-Allel und damit höchstwahrscheinlich blauäugig gewesen. Interessanterweise zeigen etwas jüngere Proben der Andronovo-Kultur (AND; 1800 – 1400 v. Chr.) und der Karasuk-Kultur (1400 – 800 v. Chr.) aus Zentralasien in vier von sieben bzw. zwei von zwei Fällen den Genotypen GG (Bouakaze *et al.* 2009, Keyser *et al.* 2009). Dies könnte durch Migrationsbewegungen aus dem Westen und *allelic surfing* zu erklären sein. Verbindungen zwischen der Andronovo-Kultur und den nordpontischen

⁴⁰ In Wilde *et al.* (2014) sind 4 von 47 Individuen als GG-Träger aufgeführt. Dort nicht enthalten sind ZAU 1 (AG) und ZAU 2 (GG) der Kugelamphorenkultur, außerdem die Probe TET 2 (AA), die zum Zeitpunkt der informatischen Analysen für den Artikel noch nicht vollständig reproduziert war, sowie PES 4 (GG) und ZAG 3 (AG), von denen nur drei PCR-Produkte vorliegen.

Steppenpopulationen wurden bereits auf der maternalen Vererbungslinie aufgewiesen (siehe Abschnitt 10.2.7).

10.4.1.5 *SLC24A5*

Der Marker *SLC24A5* (*solute carrier family 24, member 5* bzw. *NCKX5*, $\text{Na}^+/\text{Ca}^{2+}+\text{K}^+$ *exchanger 5*) fand erstmals Beachtung durch die Arbeit von Lamason *et al.* (2005), die anhand des orthologen Gens in Zebrafischen seine Bedeutung in der menschlichen Pigmentierung aufzeigten. Bei dem Genprodukt handelt es sich um einen Kalium-abhängigen Natrium-Calcium-Austauscher in der melanosomalen Membran. Durch die Regulierung der Ca^{2+} -Konzentration in den Organellen trägt er möglicherweise zu ihrer Reifung bei (Scherer & Kumar 2010). Für *slc24a5* im Zebrafisch wurde als zusätzliche oder alternative Funktion eine Kopplung mit einer V-ATPase vorgeschlagen. Hierdurch würde *slc24a5* indirekt an der Regulierung des pH-Wertes durch aktiven Protonentransport beitragen (Lamason *et al.* 2005).

In der gleichen Studie wurde mit rs1426654 ein Polymorphismus identifiziert, der bereits in der ersten HapMap-Veröffentlichung (2003) durch extreme Unterschiede in Allelfrequenzen zwischen europäischen und afrikanischen Populationen aufgefallen war. Dieser SNP (G814A) korreliert mit einem Aminosäureaustausch in Exon 3, Ala111Thre, wobei das abgeleitete Threonin-Allel mit heller Hautfarbe assoziiert ist (Lamason *et al.* 2005). Der SNP erklärt zwischen 22 und 38 % der Variation der Hautfarbe, wobei der Erbgang vermutlich intermediär ist (Lamason *et al.* 2005, Stokowski *et al.* 2007, Mallick *et al.* 2013).

Das abgeleitete Allel ist in Europa und im Nahen Osten (nahezu) fixiert, während es in asiatischen und afrikanischen Populationen äußerst selten vorkommt (Sabeti *et al.* 2006). In Südasien kommen beide Allele in mittleren Frequenzen vor (Norton *et al.* 2007). Genomweite Analysen haben Hinweise auf Selektion in der Umgebung des *SLC24A5* locus gefunden. Sabeti *et al.* (2007) grenzten die Anzahl möglicher selektierter SNPs durch den *XP-EHH*-Test (*cross population extended haplotype homozygosity*) ein und identifizierten den funktionalen rs1426654 als Ziel der

Selektion. Voight und Kollegen (2006) stellten in der Region um *SLC24A5* ein *iHS*-Signal (*integrated haplotype score*) fest. Auch Lamason *et al.* (2005) und Norton *et al.* (2007) konnten Selektion nachweisen.

Das Alter der Mutation wird auf 22 000 bis 28 000 (5 000 – 58 000) Jahre geschätzt (Mallick *et al.* 2013). Der Beginn des Selektionsereignisses wird unter einem additiven Modell auf 11 368 (1 036 – 55 832) Jahre vor heute datiert, mit einem Selektionskoeffizienten von 0,08 (0,02 – 0,55) (Beleza *et al.* 2013).

Das Datierungsintervall des *selective sweeps* lässt sich insofern weiter eingrenzen, dass das abgeleitete, in Europäern selektierte Allel in der aus 41 Individuen bestehenden Stichprobe des vorliegenden 5 000 Jahre alten Datensatzes mit einer Häufigkeit von 100 % auftritt. Dementsprechend dürfte der Selektionskoeffizient tatsächlich noch etwas höher als 8 % sein. Dies wäre nicht überraschend, wenn man den großen Einfluss des Gens auf die Hautpigmentierung bedenkt.

Laut der auf südasiatischen Populationen basierten Studie von Stokowski *et al.* (2007) hat der SNP in *SLC24A5* einen weitaus größeren Effekt auf die Hautfarbe als die in *SLC45A2* und *TYR*. Betrachtet man rein diese drei Polymorphismen, ohne Berücksichtigung von Variationen in anderen, hier nicht untersuchten Genen, so dürften die Steppenvölker der Kupfer- und Bronzezeit bereits relativ hellhäutig gewesen sein (aber im Durchschnitt vermutlich etwas dunkler als heutige Populationen dieser Regionen), mit vorwiegend dunkler Haar- und Augenfarbe. Im komplexen Zusammenspiel aller direkt und indirekt an dem Pigmentierungs-*pathway* beteiligten Gene könnten aber auch noch ganz andere Faktoren bei der Determination der phänotypischen Ausprägungen eine Rolle spielen.

10.4.1.6 Selektionsfaktor Sonneneinstrahlung und Vitamin D-Homöostasis

Für die untersuchten SNPs in den Pigmentiergenen *TYR*, *SLC45A2* und *SLC24A5* sind mit unterschiedlicher Sicherheit Assoziationen mit der Hautfarbe nachgewiesen. Für die beiden ersten SNPs konnte mithilfe von aDNA-Daten rezente Selektion nachgewiesen werden. Bei *SLC24A5* war das Ereignis vor 5 000 Jahren mit der

Fixierung des selektierten Allels anscheinend bereits abgeschlossen. Einhergehend mit der geographischen Verteilung des Phänotyps „helle Hautpigmentierung“ wird allgemein die Sonneneinstrahlung als Selektionsfaktor für die Pigmentiergene anerkannt (Jablonski & Chaplin 2000, Jablonski 2012).

Melanin absorbiert sichtbares Licht, Ultraviolett- (UV-) und ionisierende Strahlung. Außerdem neutralisiert es Moleküle der reaktiven Sauerstoffspezies, die bei Kontakt dieser Strahlen mit der Haut entstehen (Ito 2003, Rees 2004). Dabei ist Eumelanin wesentlich stabiler und effektiver (Jablonski 2012). In tropischen Regionen, bzw. Gebieten mit starker Sonneneinstrahlung schützt somit die dunkle Haut vor der schädlichen UV-Strahlung. Eine verbreitete Hypothese ist dabei insbesondere der Schutz der Folsäure vor Photolyse durch die bis in die Dermis eindringenden UV-A-Strahlen⁴¹. Folsäure wird für Prozesse wie die RNA- und DNA-Synthese, DNA-Reparatur, Zellteilung, Spermatogenese, Produktion von Neurotransmittern und Formation von Myelin benötigt (Jablonski 2004, 2012 und Referenzen darin).

Mit zunehmenden Breitengraden lässt die Sonnenintensität und damit der Selektionsdruck auf die dunkle Pigmentierung nach. Allerdings ist die helle Hautfarbe nicht rein darauf zurückzuführen (Norton *et al.* 2007). Vielmehr nimmt die Notwendigkeit der Aufnahme von Sonnenstrahlen zu. Vitamin D₃ ist ein Prohormon, das in Keratinocyten und Fibroblasten mithilfe von UV-B-Strahlung durch die Photolyse von 7-Dehydrocholesterol produziert wird (Abbildung 10.2) (DeLuca 2004 und Referenzen darin). Die weiteren Umwandlungsschritte finden in verschiedenen Zelltypen statt, zunächst in der Leber, dann vor allem in der Niere und anderen Geweben. Über das Zwischenprodukt 25-Hydroxyvitamin D₃ wird Vitamin D₃ in das physiologisch aktive, lebensnotwendige Hormon 1 α ,25-Dihydroxyvitamin D₃ (1,25(OH)₂D₃, Calcitriol) umgewandelt (Schuster 2011). Starke Pigmentierung verhindert zwar nicht die Produktion von Vitamin D₃, aber reduziert durch Absorption des Lichts die Effizienz und Rate, so dass bei geringer UV-Strahlung nicht genug Vitamin D₃ synthetisiert wird, und es folglich zu einem Vitamin D-, genauer gesagt,

⁴¹ Ultraviolettes Licht wird nach Wellenlängenbereichen eingeteilt: UV-A = 380 – 315 nm, UV-B = 315 – 280 nm, UV-C = 280 – 100 nm (Quelle: <http://de.wikipedia.org/wiki/Ultraviolettstrahlung>, Stand: 01.10.2014).

1,25(OH)₂D₃-Mangel mit einer Reihe gesundheitlicher Folgen kommt. Eine Überproduktion von Vitamin D₃ ist hingegen kein Problem, da UV-A- und UV-B-Strahlung das Molekül in eine Reihe anderer inerter Produkte umwandeln kann (Holick *et al.* 1981).

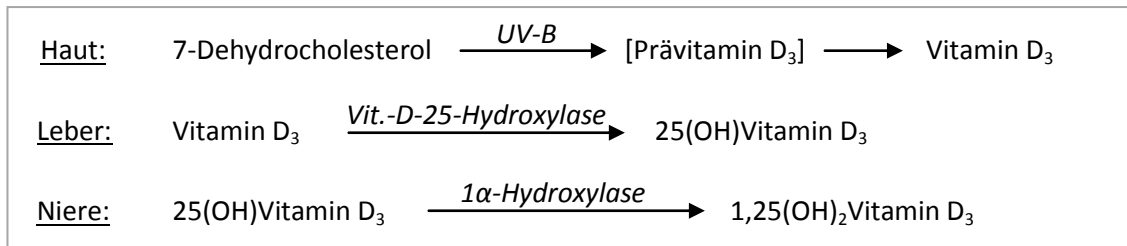


Abbildung 10.2: Produktion und Aktivierung von Vitamin D₃

1,25(OH)₂D₃ hat vielfältige Aufgaben im Körper, insbesondere in der Calcium-Homöostase. Dieser Mineralstoff ist an fast allen physiologischen Prozessen beteiligt, z. B. als *second messenger*, im Stoffwechsel, bei Zellteilungen, bei muskulärer und neuraler Aktivität etc., und nicht zuletzt im Knochenaufbau. Folglich muss dieser Haushalt ständig ausbalanciert sein. Insbesondere der Calcium-Phosphat-Spiegel im Plasma muss konstant gehalten werden. Sinkt der Spiegel, wird die Aufnahme der Ionen aus der Nahrung im Dünndarm, durch Resorption aus der Niere und, falls nötig, durch Demineralisierung von Knochen eingeleitet. An allen diesen Prozessen ist Vitamin D essentiell beteiligt (Jones *et al.* 1998). Vitamin D-Mangel führt zu Rachitis in Kindern bzw. Osteomalazie in Erwachsenen: Calcium wird im Darm nicht ausreichend aufgenommen. Es kommt zur Resorption aus dem Knochen. Da 1,25(OH)₂D₃ vermutlich auch stimulierend auf den Knochenauf- und -umbau wirkt, aber die Mineralien fehlen, kann dieser kaum noch stattfinden (Holick *et al.* 1981, Jones *et al.* 1998, Wharton & Bishop 2003). Unzureichende Vitamin D-Zufuhr in schwangeren Frauen kann schon im Embryo zu Rachitis führen (Wharton & Bishop 2003 und Referenzen darin).

1,25(OH)₂D₃ hat noch weitere Aufgaben im Organismus. Vitamin D-Mangel führt zu Osteoporose, Muskelschwäche, Multipler Sklerose, Diabetes, Herz-Kreislauf-

Erkrankungen, Krebs und mentalen Störungen, um nur einige zu nennen (Pittas *et al.* 2007, Kulie *et al.* 2009). Doch allein schon seine essentielle Bedeutung im Calcium-Haushalt und die damit verbundenen Einflüssen auf physiologische Vorgänge im ganzen Körper sind hinreichend für die Annahme, dass eine ausreichende Vitamin D-Produktion ein starker Selektionsfaktor für die helle Pigmentierung ist. Die prähistorischen Fundplätze im Untersuchungsgebiet liegen etwa zwischen 42° N und 54° N. Jablonski & Chaplin (2000) zeigten, dass in diesem Gebiet die jahresdurchschnittliche UV-Strahlung insuffizient für eine ausreichende Vitamin D₃-Synthese in moderat (Typ V) oder stark pigmentierter Haut (Typ VI) ist. In einer weiteren Studie (2010) unterschieden sie zwischen mittleren Breitengraden von 23° bis 46°, wo ein abgeschwächt depigmentierter Phänotyp mit der Fähigkeit zur fakultativen Sonnenbräunung bevorzugt wird, und höheren Breitengraden, wo helle Haut Voraussetzung für eine ausreichende Vitamin D₃-Synthese ist. 46 der 55 Proben, von denen Genotypen der Pigmentier-SNPs vorliegen (42 von 50 ohne *HERC2*), stammen aus Fundplätzen nördlich des 46. Breitengrades. Damit ist die allgemein für Europa aufgestellte Selektionshypothese auch durchaus für den Untersuchungsraum gerechtfertigt.

Für den regulatorischen SNP des *OCA2*-Gens, der hauptsächlich mit der Augenfarbe assoziiert ist, trifft diese Selektionshypothese eher nicht zu. Hier könnte die Erklärung für die deutliche Zunahme des Allels in sexueller Selektion / assortativer Partnerwahl liegen. Möglicherweise hat der SNP neben der Determination der Augenfarbe noch einen anderen, bisher nicht bekannten Einfluss auf den Organismus. Vielleicht steht er nur indirekt durch eine nicht identifizierte Kopplung mit einem anderen Gen unter Selektion oder unterliegt einem *hitchhiking*-Effekt. Eine detaillierte Diskussion der verschiedenen Selektionshypothesen für die Pigmentier-SNPs mit Bezug auf den Untersuchungsraum können in Wilde *et al.* (2014) nachgelesen werden.

10.4.1.7 CYP3A4

CYP3A4 (*cytochrome P450, subfamily IIIA, polypeptide 4*) gehört zu einer Familie von Enzymen (CYP3A), die aufgrund ihrer Variabilität für den Metabolismus einer großen Bandbreite endogener und exogener Substanzen verantwortlich ist, unter anderem Steroidhormone, zu denen Vitamin D gehört (Smith *et al.* 1998, Lamba *et al.* 2002, Schirmer *et al.* 2006). Die genetischen Variationen, die der interindividuellen phänotypischen Variabilität bezüglich der Verstoffwechslung z. B. von Medikamenten zugrunde liegen, sind bei weitem noch nicht aufgedeckt (Lamba *et al.* 2002). Das Gen *CYP3A4* liegt zusammen mit *CYP3A5*, *CYP3A7* und *CYP3A43* in einem Cluster auf Chromosom 7. In vielen Populationen macht CYP3A4 den Hauptanteil an produziertem CYP3A-Protein aus (Thompson *et al.* 2004).

Schirmer *et al.* (2006) genotypisierten 26 SNPs auf diesem *CYP3A*-Cluster in Afro-Amerikanern, Europäern und Asiaten, und untersuchten den *locus* auf Haplotypenstrukturen und Selektionssignale. Der größte *linkage disequilibrium*-Block wurde in der europäischen Population gefunden. Er enthielt über 80 % der analysierten DNA-Sequenzen. Darin enthalten war der SNP rs2740574 (G > A). Er liegt in der 5'-regulatorischen Region von *CYP3A4*. Mithilfe der P_{excess} -Statistik wurden zwei Bereiche identifiziert, von denen einer in Europäern, der andere in Asiaten ein Selektionssignal aufwies. Die Autoren schlossen aus diesen und einigen weiteren Indizien, dass rs2740574, an dem sich die beiden Sequenz-Regionen überlappten, möglicherweise das eigentliche Ziel der Selektion sei. Sie argumentierten für eine Selektion gegen das G-Allel (*CYP3A4**1B) in nicht-afrikanischen Populationen.

Funktional unterscheiden sich die beiden Allele durch eine höhere Expression und Enzymaktivität des ancestralen G-Allels. Neben seinen vielen anderen metabolischen Aufgaben ist *CYP3A4* am Vitamin D-Stoffwechsel beteiligt. Das Enzym ist eines der Kandidaten, die möglicherweise am ersten Schritt der Aktivierung des aus der Nahrung stammenden Prohormons Vitamin D₂ (weniger D₃) in Leber und Dünndarm mitwirken (Gupta *et al.* 2004, Schuster 2011, Wang *et al.* 2013). Wie auch bei Vitamin D₃ ist das aktivierte Produkt der Vitamin D₂-Umwandlung Calcitriol, 1,25(OH)₂D₃. Dieses und verschiedene Metabolite aktivieren durch Bindung an einen Vitamin D-

Rezeptor eine große Anzahl unterschiedlichster Zielgene (bis zu 5 % des Genoms) (Schuster 2011). $1,25(\text{OH})_2\text{D}_3$ reguliert seinen eigenen Katabolismus durch Induktion einer 24-Hydroxylase. Somit unterliegt der Vitamin D-Haushalt einer internen Kontrolle, durch die die Homöostase gewährleistet bleibt. Neben CYP24A1 kann auch CYP3A4 das aktive Vitamin D abbauen (Gupta *et al.* 2005, Wang *et al.* 2013).

Diese kurze Beschreibung kann nicht die höchst komplexen Vorgänge, die bei weitem noch nicht vollständig aufgeklärt sind und gewebespezifisch variieren, erfassen. Sie sind insofern interessant, dass, wie schon bei den Pigmentgenen, auch für CYP3A4 der Vitamin D-Haushalt bzw. Rachitis / Osteomalazie eine tragende Rolle in der Selektionshypothese spielen. Allerdings gibt es neben Schirmer *et al.* (2006), deren Argumente für eine negative Selektion speziell des CYP3A4*1B-Allels nicht hundertprozentig überzeugen können, nur die Studie von Thompson *et al.* (2004) (siehe Abschnitt 10.4.2.1 zu CYP3A5), die den CYP3A-Cluster gezielt auf selektierte Polymorphismen untersuchten. Genomweite Studien konnten Selektionssignale nur allgemein für die Region des CYP3A-Clusters feststellen, ohne das genaue Ziel einzugrenzen (z. B. Sabeti *et al.* 2006, Voight *et al.* 2006, Tang *et al.* 2007). Möglicherweise ist das abgeleitete A-Allel durch einen *hitchhiking*-Effekt fixiert worden. Polymorphismen in der CYP3A-Region sind vor allem das Objekt klinischer Studien zur Wirkung und Verträglichkeit von Pharmazeutika.

Da der betreffende SNP sowohl in der Region mit europäischen als auch in der mit asiatischen Selektionssignal liegt, müsste das Selektionsereignis wahrscheinlich vor der Aufspaltung der Populationen stattgefunden haben, oder es fand parallele Selektion statt. Tendenziell ist dabei die erste Variante zu bevorzugen. Gegen eine rezente Selektion spricht auch, dass das negativ selektierte, ancestrale G-Allel (CYP3A4*1) in den prähistorischen Steppenkulturen nicht vorkommt. Alle 32 an rs2740574 genotypisierten Individuen tragen das abgeleitete A-Allel. Aufgrund der großen Bandbreite an Substraten von CYP3A4 könnte statt verminderter Sonnenintensität auch der Kontakt mit neuartigen Substanzen mit dem Verlassen Afrikas eine selektierende Rolle gespielt haben. Stattdessen oder parallel dazu kommen auch demographische Faktoren im Zuge der Migrationen in Frage, zur Fixierung des abgeleiteten Allels beigetragen zu haben (siehe auch 10.4.2.1).

10.4.2 Wasser- und Elektrolythaushalt

10.4.2.1 CYP3A5

CYP3A5 (cytochrome P450, subfamily IIIA, polypeptide 5) liegt etwa 77 kb von dem oben diskutierten *CYP3A4* entfernt und gehört zur gleichen Enzymfamilie mit ihren vielfältigen Substraten. Während *CYP3A4* in der Leber das vorrangig vorliegende Enzym ist, so ist *CYP3A5* das mengenmäßig dominante Enzym in der Niere (Haehner *et al.* 1996). Eine Punktmutation (A > G; *CYP3A5**1/*3) in Intron 3 von *CYP3A5* (rs776746) generiert eine *splicing site*. Durch das Spleißen wird in das Transkript ein Stopp-Codon integriert und somit die Translation an Aminosäure 109 vorzeitig beendet. Es entsteht ein nicht-funktionelles Protein (Kuehl *et al.* 2001, Bolbrinker *et al.* 2012).

Zu den Aktivitäten von *CYP3A5* gehört unter anderem die Umwandlung des Stresshormons Cortisol zu 6 β -Hydroxycortisol in der Niere (Wrighton *et al.* 1990). Verschiedene Tier- und *in vitro*-Studien zeigten, dass 6 β -Hydroxycortisol den aktiven Transport von und damit die Reabsorption von Na⁺ induziert (Duncan *et al.* 1988, Watlington *et al.* 1988, Grogan *et al.* 1990, Ghosh *et al.* 1995). Dem Verlust von Elektrolyten wie Natriumchlorid entgegenzuwirken ist in tropischen Regionen mit heißem, feuchtem Klima und salzarmer Umgebung lebensnotwendig. Durch Schwitzen wird eine Überhitzung des Körpers verhindert. Thompson *et al.* (2004) stellten fest, dass die Häufigkeit des nicht exprimierten *CYP3A5**3-Allels mit der Entfernung zum Äquator und damit den Tropen positiv korreliert (Allelfrequenzen von 0,06 – 0,07 in Subsahara-Afrika bis 0,95 – 0,96 in Europa und Teilen Chinas).

Kuehl *et al.* (2001) hatten bereits vorgeschlagen, dass das funktionelle *CYP3A5**1-Allel in afrikanischen Populationen, zum Beispiel in Regionen mit Wassermangel, positiv selektiert sein könnte. Thompson und Kollegen (2004) griffen diese Idee auf und wiesen im DNA-Abschnitt um *CYP3A4* und *CYP3A5* Selektionssignale (ungewöhnliche Haplotypenstruktur, negatives Tajima's *D* und große Allelfrequenzunterschiede) nach. Während Schirmer *et al.* (2006) einen SNP des *CYP3A4*-Gens als Ursprung der Selektionssignale sehen, argumentieren Thompson und

Kollegen (2004) dafür, dass mit der Auswanderung aus Afrika in kühlere und salzreichere Regionen die Retention von Natrium den Salzhaushalt ungünstig beeinflusse. Dies zeige sich darin, dass in Afroamerikanern eine Prädisposition für Bluthochdruck stärker verbreitet ist als in Amerikanern europäischer Abstammung, die vorwiegend das nicht-funktionelle Allel tragen (Thompson *et al.* 2004, Kurian & Cardarelli 2007). In einer Gruppe von afroamerikanischen Probanden korrelierte der *CYP3A5*-Genotyp mit dem Blutdruck. Homozygote Träger des *CYP3A5*1*-Allels hatten signifikant höhere Blutdruckwerte als Heterozygote. *CYP3A5*3/CYP3A5*3*-Träger waren nicht unter den Testpersonen (Givens *et al.* 2003). Während Zhang *et al.* (2010) ebenfalls einen signifikanten Zusammenhang zwischen dem *CYP3A5*-Genotyp und dem Blutdruck feststellten, konnten andere Studien keine solche Assoziation oder lediglich eine schwache nachweisen (Langae *et al.* 2007, Xi *et al.* 2011). Auch für das *AGT*- (*angiotensinogen*-) Gen im Renin-Angiotensin-Aldosteron-System (RAAS), für das eine Untersuchung in der vorliegenden Studie vorgesehen war, aber verworfen werden musste, gibt es Selektionshypothesen bezüglich Wasser- und Elektrolythaushalt in Verbindung mit Bluthochdruck (Nakajima *et al.* 2004).

Sowohl die Sonneneinstrahlung / Vitamin D-Hypothese für *CYP3A4*, als auch die Salzhaushalt / Bluthochdruck-Hypothese für *CYP3A5* sind plausible Erklärungen für die beobachteten Selektionssignale. Die beiden Allele *CYP3A4*1B* und *CYP3A5*3* sind stark, jedoch nicht vollständig, miteinander gekoppelt (Thompson *et al.* 2004, Schirmer *et al.* 2006, Semiz *et al.* 2011). In den prähistorischen Steppenkulturen ist, wie gesagt, das *CYP3A4*1B*-Allel anscheinend fixiert. Für *CYP3A5*3* sind 29 der 30 gewonnenen Genotypen von rs776746 homozygot, während ein Individuum heterozygot ist. Dabei handelt es sich um TEM 1, einen Vertreter der Katakombengrabbkultur aus dem Fundort Temrta in Kalmückien. Im Rahmen der 454-Sequenzierungen konnte für dieses Individuum der Genotyp für den *CYP3A4*-SNP leider nicht bestimmt werden. Insgesamt scheint das auf die *CYP3A*-Region wirkende Selektionsereignis zum Zeitpunkt der Stichprobe bereits abgeschlossen gewesen zu sein, welcher Faktor auch immer den entsprechenden Druck ausgeübt hat. Eine alternative Erklärung für die großen Frequenzunterschiede zwischen Afrikanern und Nicht-Afrikanern könnte auch genetische Drift sein, indem es im Rahmen der

Migrationen und räumlichen Expansionen aus Afrika heraus zu dem bereits erwähnten *allelic surfing*-Phänomen kam (Currat *et al.* 2006, Hofer *et al.* 2009).

10.4.2.2 SLC12A3

Ein weiteres Gen, das an der Regulierung des Elektrolythaushalts und damit des Blutdrucks beteiligt ist, ist *SLC12A3* (*solute carrier family 12, member 3*), auch unter den Bezeichnungen *NCC* (*sodium-chloride cotransporter*) oder *TSC* (*thiazide-sensitive Na-Cl cotransporter*) zu finden. *SLC12A3* kodiert einen $\text{Na}^+\text{-Cl}^-$ -Symporter in den Nierenkanälchen (genauer gesagt, im Distalen Konvolut), der evolutionär hoch konserviert ist (Mastroianni *et al.* 1996, Plotkin *et al.* 1996).

Zu den Hauptfunktionen der Niere gehören das Herausfiltern von Stoffwechselendprodukten aus dem Blut, die dann über den Urin ausgeschieden werden, sowie der Wasser-Elektrolyt-Haushalt, der auf den Blutdruck wirkt. Im Primärharn landen bei der Erstfilterung auch Stoffe, die vom Körper benötigt werden, z. B. Elektrolyte, Aminosäuren oder Glucose. Diese werden dann von verschiedenen Membrantransportern im Tubulussystem rückresorbiert. Es gibt mehrere Membranproteine für den Rücktransport von Natrium in der Niere. Ein Großteil der Na^+ -Ionen wird bereits im Proximalen Konvolut resorbiert. Den Transportern im Distalen Tubulus fällt eher die Aufgabe der Feinregulation zu (Giebisch & Windhager 1964, Breyer & Jacobson 1990, Knepper & Brooks 2001).

Das *SLC12A3*-Protein befindet sich in der dem Urin zugewandten (luminalen) Membran der Tubuluszellen. Eine Natrium-Kalium-ATPase auf der basolateralen Seite pumpt aktiv Na^+ aus der Zelle ins Blut, wodurch ein Konzentrationsgefälle zwischen dem Harn und dem Inneren der Tubuluszelle entsteht bzw. aufrechterhalten wird. Der Import von Natriumionen (zusammen mit Chlorid) durch den *SLC12A3*-Transporter wirkt dem entgegen (Warnock & Eveloff 1982, Breyer & Jacobson 1990). Es sollte noch angemerkt werden, dass ein $\text{Na}^+/\text{Ca}^{2+}$ -Antiporter Calcium, das wahrscheinlich passiv aus dem Urin in die Epithelzelle gelangt, aus der Zelle in das Blut pumpt (Ramachandran & Brunette 1989, Friedman & Gesek 1993).

SLC12A3 wird vermutlich durch Hormone wie Aldosteron, Adiuretin oder Angiotensin II stimuliert, die an der Wasserresorption und Blutdruckregulierung beteiligt sind (Kim *et al.* 1998, Ecelbarger *et al.* 2001, Moes *et al.* 2014). Gehemmt wird es durch Thiaziddiuretika, eine Gruppe harntreibender Substanzen, die bei Bluthochdruck oder Herzschwäche eingesetzt werden (Ellison *et al.* 1987). *Loss-of-function*-Mutationen in *SLC12A3* führen zum Gitelman-Syndrom, das sich unter anderem durch Hypokali- und Hypomagnesiämie auszeichnet (Simon *et al.* 1996). Moreno *et al.* (2004) untersuchten durch in vitro-Mutagenese und Expression in Oocyten des *Xenopus laevis* (Glatter Krallenfrosch) die funktionellen Eigenschaften von rs1529927. Die nichtsynonyme Transversion (G > C; Gly264Ala) in Exon 6 von *SLC12A3* führte zu einer signifikant niedrigeren Ionentransportrate des Proteins. Vormfelde *et al.* (2007) bestätigten diese Beobachtung in einer klinischen Studie an Männern europäischer Abstammung. Träger des Alanin-Allels schieden ein höheres Volumen an Urin und größere Mengen an Elektrolyten bzw. deren Ionen nach Einnahme von Schleifendiuretika (harntreibende Mittel, die an Transportern der Henle'schen Schleife wirken) aus. Signifikant waren die Unterschiede bei Kalium und Chlorid, etwas schwächer war der Effekt auf Natrium und Calcium.

Anhand von HapMap-Daten (2003) zeigte sich, dass das C- bzw. Ala-Allel genau einen Haplotypen definiert, der eine Frequenz von 7 % hat, und signifikant im *extended haplotype homozygosity*-Test (EHH) auffiel (Vormfelde *et al.* 2007). Die Variante tritt lediglich in Individuen europäischer Abstammung auf, und auch dort nur im heterozygoten Genotyp. Vormfelde und Kollegen vermuteten, dass der weniger aktive SLC12A3-Transporter positiv selektiert sein könnte. Als Selektionsgrund gaben sie die Stärkung des Calciumhaushaltes an. Ein verlangsamter Natrium-Import durch SLC12A3, führt zu einer verminderten Konzentration in der Zelle. Dies fördere die Calcium-Resorption, da diese, wie oben erwähnt, mit einem Transport von Na⁺ aus dem Blut in die Tubuluszelle gekoppelt ist (Gesek & Friedman 1995, Vormfelde *et al.* 2007). Der Einfluss auf den Salzhaushalt bzw. den Blutdruck, den Thompson und Kollegen (2004) als Selektionsfaktor für *CYP3A5* sahen, wurde von Vormfelde *et al.* (2007) nicht in Erwägung gezogen. In Anbetracht dessen, dass Bluthochdruck gerade in modernen Populationen ein weitverbreitetes Phänomen und Problem ist, könnte

der etwaige positive Selektionsdruck durchaus direkt den Na^+ -Transport und nicht indirekt den Calciumhaushalt betreffen. Falls rs1529927 tatsächlich unter positiver Selektion steht, so ist diese entweder sehr schwach und/oder extrem rezent, da selbst in Europa die Frequenz bei nur 7,4 % Heterozygoten (1000 Genomes) liegt, mit 0 % homozygoten C-Allel-Trägern. Allerdings könnte der CC-Genotyp, der bisher noch in keiner Population beobachtet wurde, auch wiederum schädlich (oder gar nicht lebensfähig) sein, so dass negative und positive Selektion miteinander konkurrieren (Ji *et al.* 2008).

Zwei Individuen (6,67 %) im vorliegenden Datensatz sind heterozygot für rs1529927. Sie stammen beide aus der jüngsten Zeitscheibe, NEV 1 von der entwickelten Katakombengrabkultur und POD 1 aus der Poltavka-Kultur. Alle anderen sind homozygot für das ancestrale Allel. Im modernen Datensatz der Ukraine liegt die Heterozygotenfrequenz bei 11,9 %, und damit leicht unter dem Durchschnitt des mitteleuropäischen Datensatzes des HapMap-Panels (16,7 %; NCBI - dbSNP). Innerhalb der letzten 5 000 Jahre hat keine signifikante Änderung der Allelfrequenzen für den betroffenen SNP stattgefunden. Diese Beobachtung, das Nicht-Auffallen des *locus* in genomweiten Analysen (Voight *et al.* 2006, Sabeti *et al.* 2007, Tang *et al.* 2007) und der Umstand, dass die Rolle von SLC12A3 in der Feinregulation des Elektrolythaushalts liegt und ihn daher nur in geringem Maße beeinflusst, sprechen gegen ein Selektionsereignis.

10.4.3 Temperaturadaptation

Die beiden Gene *EDAR* und *ABCC11* werden unter der Rubrik „Temperaturadaptation“ zusammengefasst, da für beide der Schutz vor kaltem, trockenem Klima als Selektionshypothese vorgeschlagen wurde. Allerdings kommen für beide Fälle noch weitere mögliche Selektionsfaktoren in Frage.

10.4.3.1 EDAR

Das erste Gen, *EDAR* (*ectodysplasin A receptor*), ist ein Mitglied der *tumor necrosis factor receptor family*, die eine wichtige Rolle in der ektodermalen Differenzierung spielt. *EDAR* ist Bestandteil verschiedener Signalwege und an der Entwicklung von Haarfollikeln, Zähnen und exokrinen Drüsen (z. B. Schweißdrüsen) beteiligt (Clarke 1987, Monreal *et al.* 1999, Kumar *et al.* 2001, Kamberov *et al.* 2013). Intensiv erforscht ist vor allem die embryonale Bildung von Haarfollikeln. Zu dieser gehört ein komplexes Interaktions- und Signalsystem zwischen dem Neuroektoderm und dem darunterliegenden Mesenchym (Schmidt-Ullrich & Paus 2005, Rishikaysh *et al.* 2014). Einer von mehreren Signalwegen bei der Morphogenese von Haarfollikeln ist der EDA-A1/EDAR/NF- κ B-Weg, der mit einem weiteren Signalweg, dem Wnt/ β -Catenin-Weg in Wechselbeziehung steht (Rishikaysh *et al.* 2014). Der Rezeptor EDAR, der in der sich entwickelnden Epidermis (in Keratinocyten) liegt, bindet den Liganden EDA-A1 (Yan *et al.* 2000, Kumar *et al.* 2001). An seiner intrazellulären Todesdomäne (*death domain*) rekrutiert EDAR daraufhin den Adapter EDARADD, wodurch ein Signalweg über verschiedene TRAF-Proteine eingeleitet wird. Dies führt schließlich zur Aktivierung von NF- κ B-Transkriptionsfaktoren und deren Transport in den Nukleus (Kumar *et al.* 2001, Botchkarev & Fessing 2005). NF- κ B-Transkriptionsfaktoren kommen in allen Zelltypen und Geweben vor und spielen eine wichtige Rolle in Prozessen wie z. B. Immunantwort, Zellproliferation und Apoptose (Basak *et al.* 2012). Vermutlich regulieren die Transkriptionsfaktoren bei der Haarfollikelentwicklung ihre Zielgene dahingehend, dass die Zellproliferation eingestellt und die Differenzierung eingeleitet wird (Seitz *et al.* 1998, Kumar *et al.* 2001). Möglicherweise induziert NF- κ B aber auch das Wachstum der Plakode (Epidermisverdickung), die die Basis für die Follikelentwicklung bildet (Schmidt-Ullrich & Paus 2005).

Rs3827760 (T1540C) führt zu einem Aminosäureaustausch (Val370Ala) in der hochkonservierten Todesdomäne von EDAR. Dies wirkt sich auf die Bindung von EDARADD aus und beeinflusst in der Folge die Aktivität des Transkriptionsfaktors NF- κ B. Mit einem Luciferase-*assay* zeigten Fujimoto *et al.* (2008a), dass die abgeleitete Alanin-Variante mit einer niedrigeren Aktivität von NF- κ B einhergeht. Bryk *et al.* (2008) kamen zu einem gegenteiligen Ergebnis, nämlich einer erhöhten NF- κ B-

Aktivität. Mou *et al.* (2008) bestätigten deren Befunde und vermuteten, wie auch Fujimoto *et al.* (2008b), dass die anscheinend widersprüchlichen Beobachtungen an unterschiedlichen Parametern im Versuchsaufbau lagen, z. B. mit welchem Zeitabstand die Messungen durchgeführt wurden. Einig waren sich die drei Gruppen, dass 370Ala phänotypisch zu dickeren Haaren führt, wahrscheinlich aufgrund einer höheren Aktivität der Haarentwicklung. Außerdem ist das Haar glatt und hat einen runden Querschnitt (Mou *et al.* 2008).

Dieser Haartyp ist vor allem in Asien verbreitet. Damit eng korrelierend sind die hohen Frequenzen zwischen 70 und 100 % des abgeleiteten *EDAR*-Allels in den meisten asiatischen Populationen und indigenen Populationen Amerikas (Sabeti *et al.* 2007, Bryk *et al.* 2008, Fujimoto *et al.* 2008a, Fujimoto *et al.* 2008b). Im HapMap-Panel von 2005 ist rs3827760 der am drittstärksten differenzierte nicht-synonyme SNP. Für die genomische Region um den *EDAR locus* wurden mithilfe verschiedener Tests (*LRH*, *iHS*, *XP-EHH*, Tajima's *D*, Fay & Wu's *H*, *F_{ST}* etc.) deutliche Selektionssignale in Ostasiaten (hauptsächlich basierend auf den HCB- und JPT-Datensätzen des HapMap-Projekts) nachgewiesen (u. a. Carlson *et al.* 2005, Sabeti *et al.* 2007, Fujimoto *et al.* 2008a, Xue *et al.* 2009). Mithilfe von *approximate Bayesian computation*-Simulationen (*ABC*) und mit einer *maximum likelihood*-Analyse (*ML*) berechneten Kamberov *et al.* (2013) den Zeitpunkt der Mutation und den Selektionskoeffizienten. Demnach liegt der Ursprung des mit dicken Haaren assoziierten Allels vor über 30 000 Jahren in Zentralchina und sein Fitnessvorteil zwischen 0,030 und 0,186 ($S = 0,114$ mit *ABC* bzw. 0,0744 mit *ML*).

Aufgrund der pleiotropen Effekte von *EDAR* ist unklar, welche phänotypische Ausprägung die Ursache des Selektionsdrucks auf die Alanin-Variante ist. Zunächst wurde vermutet, dass die dicken Haare insbesondere in den nördlicheren Regionen Ostasiens vor einem Verlust von Körperwärme über den Kopf schützen (Fujimoto *et al.* 2008a). Möglicherweise bot auch eine veränderte Drüsenstruktur, z. B. vergrößerte Talg- und Meibom⁴²-Drüsen Schutz vor einer Austrocknung im vor 10 000 bis 25 000 Jahren besonders trocken-kaltem Klima (Chang *et al.* 2009). Kamberov *et al.* (2013)

⁴² Talgdrüsen an den Augenlidern

argumentieren hingegen, dass das Klima im prähistorischen China eher warm und humid war und die größere Anzahl ekkriner Schweißdrüsen in Trägern der abgeleiteten *EDAR*-Variante von Vorteil gewesen sein könnte. Weitere Selektionsspekulationen beinhalten Resistenz gegen Ektoparasiten, Ernährung (Zahnentwicklung), sexuelle Selektion oder ein noch unentdecktes Merkmal (Bryk *et al.* 2008, Mou *et al.* 2008, Xue *et al.* 2009, Kamberov *et al.* 2013).

In 44 genotypisierten Proben der osteuropäischen Steppe ist das abgeleitete Allel von *EDAR* lediglich in einem heterozygoten Individuum der Jamnaja-Kultur (POP 4) zu finden (auch heterozygot für *ABCC11*, siehe unten). Die phänotypische Ausprägung in Heterozygoten für rs3827760 ist wahrscheinlich intermediär (Fujimoto *et al.* 2008a, Fujimoto *et al.* 2008b, Mou *et al.* 2008, Kimura *et al.* 2009). Somit hatten die bronzezeitlichen Steppenbewohner im Großen und Ganzen nicht die dicken und glatten Haare, wie sie heute in Asien verbreitet sind, sondern dürften mit denen in Europa vergleichbar sind. Direkt aus der heutigen Ukraine liegen für den SNP keine Daten vor, aber in den angrenzenden Regionen Ungarn und Nordkaukasus (Adygeja) liegt die Häufigkeit des abgeleiteten C-Allels bei 1,7 bzw. 0,9 %. Während in Kasachstan die Frequenz bereits bei über 50 % liegt, herrscht im westliche Russland ebenfalls die ancestrale Variante vor (ALFRED, <http://alfred.med.yale.edu>; Osier *et al.* 2002, Rajeevan *et al.* 2003).

10.4.3.2 *ABCC11*

ABCC11 (*ATP-binding cassette, subfamily C, member 11*) kodiert einen ATP-abhängigen Membrantransporter (selten unter dem Namen *MRP8, multidrug resistance-associated protein 8*, veröffentlicht), der als Efflux-Pumpe verschiedener Substanzen agiert. Hierzu gehören z. B. zyklische Nukleotide wie cGMP und cAMP, Gallensäuren und weitere Steroide (Guo *et al.* 2003, Chen *et al.* 2005). Das Gen wird in verschiedenen Geweben exprimiert, u. a. in den apokrinen⁴³ Duft-/Schweißdrüsen und

⁴³ Apokrine Drüsen geben bei der Sekretion einen Teil des Cytoplasmas und der Membran durch Abschnürung mit ab. Ekkrine Drüsen hingegen sezernieren ohne Verlust von Cytoplasma. Ekkrine

den Cerumen-produzierenden apokrinen Drüsen im Hörkanal (Yabuuchi *et al.* 2001, Bortfeld *et al.* 2006, Yoshiura *et al.* 2006, Martin *et al.* 2010). Dort wurde das Protein sowohl in der Plasmamembran sezernierender Zellen als auch in intrazellulären Granula und großen Vakuolen nachgewiesen (Toyoda *et al.* 2009).

Yoshiura *et al.* (2006) identifizierten mit rs17822931 einen SNP, der die Zusammensetzung von Ohrenschmalz (Cerumen) determiniert. Der Aminosäureaustausch von Glycin zu Arginin an Proteinposition 180, hervorgerufen durch eine C > T-Transition, führt im homozygoten Zustand zu trockenem, flockigem, weiß-gräulichem Schmalz, während die dominante Wildtyp-Variante gelb-bräunlich und wachsartig-klebrig ist. Dies ist vermutlich auf eine *loss-of-function* zurückzuführen. Toyoda *et al.* (2009) zeigten, dass die Arginin-Variante von ABCC11 bei der posttranslationalen Modifikation im Endoplasmatischen Retikulum nicht N-glykosyliert wird. Dies führt zu einer fehlerhaften Faltung des Protein, woraufhin es proteosomal degradiert wird.

Neben dem Ohrenschmalz beeinflusst rs17822931 außerdem den Achselgeruch (Nakano *et al.* 2009, Harker *et al.* 2014). Viele der Substrate von ABCC11 kommen in Achselschweiß vor oder sind Vorläufer solcher Substanzen (Martin *et al.* 2010, Baumann *et al.* 2014, Prokop-Prigge *et al.* 2014). Möglicherweise werden die Komponenten über ABCC11 in Vesikel transportiert und dann von der apokrinen Zelle sezerniert. Der frische Schweiß an sich ist geruchslos, doch infolge des Kontaktes mit der Hautflora kommt es zur Zersetzung des Sekrets (insbesondere der Lipide und Proteine), und es entstehen unangenehm riechende Produkte. Fehlender oder geringer Schweißgeruch korreliert vollständig mit trockenem Cerumen und somit mit dem T/T-Genotyp, was auf einen nicht stattfindenden Transport diverser, für den Schweißgeruch notwendiger Substanzen zurückzuführen sein dürfte (Martin *et al.* 2010).

In Ostasien liegt die Häufigkeit des abgeleiteten Allels bei 80 – 95 %, in Europa durchschnittlich bei 14 % und Afrika hingegen bei 0 % (1000 Genomes, Yoshiura *et al.*

Schweißdrüsen sind fast über die gesamte Haut verteilt, während apokrine auf die Achselhöhlen, Brustwarzen, den Anal- und Genitalbereich beschränkt sind (Folk *et al.* 1991).

2006). Die Frequenz nimmt entlang eines Ost-West- und eines Nord-Süd-Gradienten ab. Yoshiura *et al.* (2006) vermuten den Ursprung der Mutation in Nordosturasien, von wo aus sie sich im Zuge von Migrationen und aufgrund von Selektion ausbreitete. Wang *et al.* (2007b) hingegen nehmen einen europäischen Ursprung an, mit einem rapiden Anstieg der Häufigkeit in asiatischen Populationen. Rs17822931 gehört zu den am stärksten zwischen Populationen differenzierten nicht-synonymen SNPs (The International HapMap Consortium 2005). Die Region um *ABCC11* weist eine geringe Nukleotiddiversität auf (Diller *et al.* 2002). Diese Beobachtungen sowie ein statistischer Nachweis durch einen *LRH*-Test (Wang *et al.* 2007b) sprechen für eine Selektion des Arginin-Allels in ostasiatischen Populationen. Mithilfe von Koaleszenzsimulationen ermittelten Ohashi *et al.* (2011) einen Selektionskoeffizienten *S* von ungefähr 0,01. Das Alter der Mutation schätzten sie auf 2 006 Generationen (95 % CI: 1 023 – 3 901 Generationen), etwa 50 000 Jahre bei einer Generationszeit von 25 Jahren.

Als Selektionshypothese wird eine Anpassung an das trocken-kalte Klima in Nordosturasien vorgeschlagen, möglicherweise durch reduziertes Schwitzen und damit einen geringeren Verlust an Körperwärme (Yoshiura *et al.* 2006, Ohashi *et al.* 2011). Allerdings fanden Martin *et al.* (2010) keine signifikanten Unterschiede in der Menge an produziertem Schweiß zwischen den drei Genotypen, lediglich der Intensität des Achselschweißgeruchs. Sexuelle Selektion könnte demnach eine Rolle gespielt haben.

Interessanterweise ist im vorliegenden aDNA-Datensatz die Allelfrequenz des „asiatischen“ Arginin-Allels mit 28,8 % (19,99 – 39,5 %) deutlich höher als in den von Nina Hollfelder (2013) untersuchten modernen Ukrainern (11,4 %; 6,3 – 19,7 %). Für diese Beobachtung gibt es mehrere Erklärungsmöglichkeiten. Das Mutationsereignis hat möglicherweise bereits kurz nach einem „*out of Africa*“-Event stattgefunden (Ohashi *et al.* 2011). Während der Migration könnte das „neue“ Allel durch *allelic surfing* an Häufigkeit zugenommen haben. In kalten Regionen wie Nordosturasien wurde das Allel durch Selektion weiter gefördert, während es in Europa aufgrund von genetischer Drift fast vollständig wieder verloren ging. Der *F_{ST}*-Wert von 0,0804 für *ABCC11* zwischen der bronzezeitlichen und modernen Bevölkerung ist zwar signifikant,

aber bedeutend niedriger als bei SNPs wie in *HERC2*, *SLC45A2* und *TYR*, für die Selektion in europäischen Populationen nachgewiesen wurde. Alternativ könnte die vergleichsweise hohe Frequenz des Arginin-Allels vor allem in der äneolithischen und der Jamnaja-Stichprobe ein Hinweis auf östliche Einflüsse sein. Diese müssten aber in einer früheren Zeitscheibe stattgefunden haben. Mitochondrial sind im Neolithikum noch „asiatische“ Linien in der Steppe zu finden (Nikitin *et al.* 2012), aber weder die mtDNA noch die Gene *ADH1B* und *EDAR* bestätigen Einflüsse aus Osten auf die Steppenulturen des dritten vorchristlichen Jahrtausends. Betrachtet man statt der Allel- die Genotypenhäufigkeit, so fällt auf, dass der Abfall der T-Frequenz im Verlauf der Bronzezeit vor allem auf einem Wegfall der Homozygoten und einer Zunahme der Heterozygoten beruht. Möglicherweise wurde das ancestrale Allel bei der Genese der Katakombengrabkultur durch die in Kapitel 10.2.5 vorgeschlagenen genetischen Einflüsse aus nördlicher Richtung, dem Jäger/Sammler-Refugium im östlichen Baltikum eingetragen. Allerdings ist keine Korrelation zwischen den maternalen Haplotypen und den Genotypen von *ABCC11* festzustellen. Genetische Drift ist daher als Erklärung der Entwicklungen im Untersuchungsraum zu bevorzugen.

10.4.4 Ernährung

10.4.4.1 *LCT*

Das Paradebeispiel für positive Selektion in der rezenten Humanevolution ist *LCT* (*lactase*; auch *lactase-phlorizin hydrolase*, *LPH*). Das in den Mikrovilli des Dünndarms vorkommende Enzym Laktase katalysiert hauptsächlich die Spaltung von Laktose in die Monosaccharide Glukose und Galaktose (Abbildung 10.3). An einem zweiten aktiven Zentrum wirkt sie außerdem als Glykosylceramidase (Arribas *et al.* 2000).

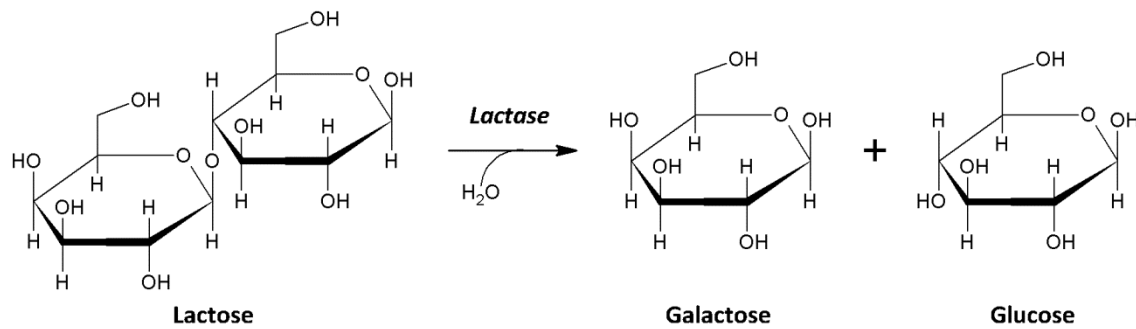


Abbildung 10.3: Hydrolytische Spaltung von Laktose

Nach der Geburt ist für Säugetiere die Muttermilch zunächst die einzige Nahrungsquelle, wobei die Kohlenhydratzufuhr durch die Laktose gesichert wird. Der Milchzucker muss erst in seine beiden Bestandteile zerlegt werden, bevor diese über die Dünndarmschleimhaut vom Körper aufgenommen werden können. Natürlicherweise endet die Ernährung von Milch mit dem Abstillen. Wie bei allen Säugetieren wird bei einem Großteil der menschlichen Bevölkerung anschließend die Laktase-Produktion um etwa 90 bis 95 % herunterreguliert (Swagerty *et al.* 2002). Konsum von laktosehaltigen Milchprodukten führt dann (ab gewissen Mengen, die individuell variieren) zu unangenehmen Symptomen. Durch den nicht absorbierten Zucker kommt es zu einem osmotischen Flüssigkeitsinflux in die Därme. Gelangt die Laktose in den Dickdarm, so wird sie von den Darmbakterien gespalten und als Nährstoff verarbeitet. Bei diesen Gärprozessen entstehen unter anderem Laktat (Milchsäure), Methan, Wasserstoff, Kohlendioxid und kurzkettige Fettsäuren. Die Folge sind Aufblähungen, Krämpfe, Flatulenz und unter Umständen Diarrhö (Hove *et al.* 1999, Swagerty *et al.* 2002, Lomer *et al.* 2008). Man bezeichnet diesen ursprünglichen Phänotyp als Laktoseintoleranz, Laktosemalabsorption oder Hypolaktasie.

Es gibt allerdings Menschen, die auch im Erwachsenenalter problemlos größere Mengen an Milch verdauen können. Bei ihnen herrscht Laktasepersistenz vor, d. h. das Gen wird dauerhaft exprimiert. Diese Eigenschaft wird vor allem in Nord- und Mitteleuropa beobachtet, sowie in Populationen mit pastoraler Subsistenz in Afrika

und im Nahen und Mittleren Osten (Itan *et al.* 2010). Der Fokus soll hier auf der in Europäern verbreiteten Variante liegen.

Die molekulare Grundlage für die Laktasepersistenz in Europäern wurde von Enattah *et al.* (2002) entdeckt. Sie identifizierten in Nordeuropäern den SNP C-13 910T (rs4988235) *upstream* des *LCT locus* in Intron 13 von *MCM6* (*minichromosome maintenance complex component 6*), der vollständig mit dem Phänotyp assoziiert ist. Das abgeleitete T-Allel ist dominant über das ancestrale Allel und geht mit Laktasepersistenz und damit Laktosetoleranz einher. Ein weitere SNP 22 018 bp (G > A, rs182549) oberhalb des Startcodons von *LCT*, in Intron 9 von *MCM6* korrelierte in 97 % der Fälle mit der entsprechenden phänotypischen Ausprägung. Nach Olds & Sibley (2003) ist vor allem der C-13 910T-SNP für die unterschiedliche Transkriptionsaktivität von *LCT* verantwortlich. Er liegt in einem regulatorischen Element.

Die Region um den *Laktase locus* weist deutliche Selektionssignale auf, unter anderem eine hohe genetische Differenzierung der europäischen zu anderen Populationen, wie Bersaglieri *et al.* (2004) durch die F_{ST} - und p_{excess} -Statistiken zeigten. Ferner wiesen sie durch die *REHH*-Statistik (*relative extended haplotype homozygosity*) nach, dass der Haplotyp, der -13 910*T und -22 018*A trägt, in Europa weit verbreitet und überdurchschnittlich lang ist. Auch in genomweiten Analysen fiel diese Region auf (z. B. The International HapMap Consortium 2005, Voight *et al.* 2006, Sabeti *et al.* 2007). Das Alter des mit Laktasepersistenz assoziierten Haplotyps wurde auf 2 188 bis 20 650 Jahre und der Selektionskoeffizient auf $S = 0,014$ bis $0,19$ geschätzt (Bersaglieri *et al.* 2004).

Ein Selektionsvorteil der Laktasepersistenz ist nur gegeben, wenn Milch als Nahrungsmittel zur Verfügung steht. Zu diesem Umstand kam es mit der neolithischen Transition, als sich Vieh- und Milchwirtschaft betreibende Gesellschaften entwickelten. Itan *et al.* (2009) grenzten mithilfe von Simulationanalysen von Modellen, in die die Verbreitung von Laktasepersistenz, Milchwirtschaft und die Ausbreitung von Bauern im Neolithikum integriert waren, den Zeitpunkt des Selektionsbeginns auf etwa 7 500 Jahre vor heute ein. Man spricht hier von einem

Gen-Kultur-Koevolutionsmodell (Aoki 1986, Beja-Pereira *et al.* 2003, Itan *et al.* 2009, Laland *et al.* 2010).

Die Geschichte der *Laktase* in Europa ist multidisziplinär untersucht. Die frühesten archäologischen und isotopenchemischen Hinweise auf Milchwirtschaft stammen aus Westanatolien und Südosteuropa und sind älter als 8 000 Jahre (Evershed *et al.* 2008). Im fünften Jahrtausend v. Chr. erreichte sie auch die Britischen Inseln (Copley *et al.* 2003). Die Frequenzen des abgeleiteten Allels nehmen in Europa entlang eines Südost-Nordwest-Gradienten zu und korrelieren dabei mit der Variation von Milchproteinen in europäischen Rindern (Beja-Pereira *et al.* 2003). Itan *et al.* (2009) erklärten dieses Muster mit der Ausbreitung der Bauern, wobei das Persistenz-Allel a) innerhalb der Population selektiert wurde und b) auf der demographischen Ausbreitungswelle durch Gründereffekte „mitwanderte“ und an Häufigkeit zunahm (*allelic surfing*). Der Ursprung der Mutation lag den Simulationen zufolge zwischen dem Karpatenbecken und Mitteleuropa, von wo aus auch die Ausbreitung der Linearbandkeramik begann.

In einer palaeogenetischen Studie von Burger *et al.* (2007) wurde in der aDNA von einem mesolithischen und acht neolithischen Skelettfunden das abgeleitete T-Allel der Laktase nicht gefunden. Unter Vorbehalt, aufgrund der kleinen Stichprobengröße, muss es im frühen Neolithikum demnach noch sehr selten gewesen sein. Dies spricht, ebenso wie das junge Alter der Mutation, gegen die sogenannte „reverse cause“-Hypothese, nach der Populationen, in denen Laktosetoleranz vorherrschte, begannen, Milchwirtschaft zu betreiben. Wahrscheinlich war die Entwicklung der Laktasepersistenz komplex, mit regional und zeitlich unterschiedlich starken Einflüssen von Selektionsdruck und demographischen Faktoren wie Migration, Expansion, Kulturkontakten und Gendrift (Gerbault *et al.* 2009, Itan *et al.* 2009, Gerbault *et al.* 2011).

Nach den obigen Ausführungen wäre zu erwarten, dass in den etwa 5 000 Jahre alten Populationen, die in der vorliegenden Studie untersucht wurden, bereits Laktasepersistenz zu beobachten wäre, insbesondere, da die Subsistenz der Jamnaja- und der Katakombengrabkultur hauptsächlich auf der Haltung von Rindern, Ziegen und Schafen beruhte. Überraschenderweise sind alle 43 für rs4988235 genotypisierten

Individuen homozygot für das ancestrale C-Allel. In modernen Ukrainern liegt die Allelfrequenz von -13 910*T bei 35,9 % (Hollfelder 2013). Für den SNP G-22 018A waren immerhin 3 von 37 Individuen heterozygot. Allerdings ist für diesen *locus* kein signifikanter Einfluss auf die *LCT*-Aktivität bekannt (Olds & Sibley 2003). Dementsprechend hoch sind die F_{ST} -Werte zwischen den alten und den modernen Proben: 0,34373 für rs4988235 und 0,23672 für rs182549 (jeweils $p < 10^{-5}$). Vermutlich haben die Viehhirten der osteuropäischen Steppe vor allem fermentierte Milchprodukte konsumiert. Die Selektion des Laktasepersistenz-Allels hat entweder später eingesetzt, und der Selektionsdruck muss besonders hoch, möglicherweise höher als bisher vermutet gewesen sein, oder demographische Faktoren haben eine entscheidende Rolle bei dem starken Anstieg der Allelfrequenz gespielt. Auf jeden Fall kann ausgeschlossen werden, dass die nordpontischen Steppenpopulationen zur raschen Zunahme der Laktasepersistenz in Mitteleuropa beigetragen haben. In der ebenfalls etwa 5 000 Jahre alten mittelneolithischen Grübchenkeramischen Kultur in Skandinavien konnte -13 910*T immerhin in einem vom 20 Chromosomen (10 Individuen) nachgewiesen werden (Malmström *et al.* 2009). Deren Subsistenz beruhte auf Jagen und Sammeln. Auch in acht spätneolithischen, etwa 3 750 Jahre alten Individuen von der Iberischen Halbinsel wurde lediglich das ancestrale Allel authentifiziert (Sverrisdóttir *et al.* 2014). Es bleibt abzuwarten, ab welchem Zeitpunkt in der Prähistorie das T-Allel nachweislich an Häufigkeit zunahm.

Über den zugrundeliegenden Selektionsdruck wird viel spekuliert. Milch an sich ist ein Calcium- und in geringen Maßen Vitamin D-Lieferant. 1973 stellten Flatz & Rotthauwe die Calcium-Assimilations-Hypothese auf. Im sonnenarmen Nordeuropa sei die Vitamin D-Versorgung und damit die Calciumabsorption nicht ausreichend gewährleistet (siehe 10.4.1). Laktasepersistenz sei eine Art Ersatzmechanismus zum Schutz vor Rachitis bzw. Osteomalazie, da Laktose nachweislich die Calciumabsorption genauso gut wie Vitamin D fördere. Allerdings ist dies nicht unumstritten (siehe Diskussion in Ziegler & Fomon 1983, Zittermann *et al.* 2000). Als weiterer Vorteil wurde Milch als Nahrungsergänzung vorgeschlagen, da Ernten wahrscheinlich oft fehlschlugen und außerdem saisonal beschränkt sind (Itan *et al.* 2009). Auf der sozialen Seite könnten Kühe (und damit der Zugang zu Milch) eine Art Statussymbol

gewesen sein. In Europa wahrscheinlich weniger, aber dafür z. B. in afrikanischen Populationen ist Milch als „saubere“ Alternative zu kontaminierten Wasserquellen. Keine der Hypothesen allein scheint den starken Selektionsdruck auf das *Laktase*-Gen erklären zu können. Wahrscheinlich akkumulierten viele Faktoren zu einem insgesamt großen Vorteil.

Wie bereits erwähnt, kommt Laktasepersistenz auch außerhalb Europas vor. Beispielsweise in nordafrikanischen Berber-Gruppen (Myles *et al.* 2005), in Zentralasien (Heyer *et al.* 2011) und Indien (Gallego Romero *et al.* 2011) findet sich das 13 910*T-Allel. Einige weitere Populationen in Afrika zeigen ebenfalls Laktosetoleranz (Myles *et al.* 2005, Ingram *et al.* 2007, Tishkoff *et al.* 2007, Enattah *et al.* 2008, Ranciaro *et al.* 2014). Dort scheinen andere Polymorphismen im *Laktase*-Gen als die beiden hier untersuchten SNPs die genetische Grundlage des Phänotyps zu sein. Bei all diesen Populationen spielt die Vieh- und Milchwirtschaft eine große Rolle in der Subsistenzwirtschaft. Laktasepersistenz scheint also mehrfach im Rahmen kultureller Entwicklungen entstanden zu sein.

10.4.4.2 TRPV6

Ebenfalls in Zusammenhang mit Calcium in der Ernährung steht *TRPV6* (*transient receptor potential cation channel, subfamily V (vanilloid), member 6*). Andere Bezeichnungen sind *ECAC2* (*epithial calcium channel 2*), *CAT1* (*calcium transporter 1*) oder *CATL* (*calcium transport-like protein*). *TRPV6* wird in Epithelgeweben unter anderem des Magen-Darm-Traktes, der Niere und der Plazenta exprimiert (Müller *et al.* 2000). Der spannungsabhängige Ionenkanal ist auf den transzellulären Transport von Ca^{2+} spezialisiert. Er spielt eine Schlüsselrolle in der Aufnahme von Calcium aus der Nahrung im Darm, aber auch in anderen Geweben (Peng *et al.* 1999, Peng *et al.* 2000, Bödding 2005). Calcium gelangt passiv und aktiv, im letzteren Fall z. B. durch *TRPV6*, durch die apikale (auf der Seite zum Darmlumen) Membran in die Epithelzelle. Dort wird es an Calbindin, ein Transportermolekül gebunden und zur basolateralen Membran gebracht, wo es durch andere Membrankanäle in den extrazellulären Raum und schließlich ins Blut gelangt (van Abel *et al.* 2005). *TRPV6* wird durch Hormone wie

aktives Vitamin D und Östrogen, aber auch Ca^{2+} aus der Nahrung reguliert und spielt dadurch, dass der Calcium-Transport durch diesen Kanal der geschwindigkeitsbestimmende Schritt bei der aktiven Absorption spielt, eine wichtige Rolle im Calcium-Haushalt (van Abel *et al.* 2005). Die Aktivität von TRPV6 wird unter anderem durch intrazelluläres Calcium beeinflusst. Hohe Konzentrationen führen zur zeitweiligen Inaktivierung des Membrankanals (Nilius *et al.* 2001). Während bei ausreichender Calcium-Zufuhr mit der Nahrung die Aufrechterhaltung der Konzentration im Körper durch passive Absorption gewährleistet zu sein scheint, gewinnt der aktive Transport durch TRPV6 bei Calcium-armer Ernährung an Bedeutung (Lieben *et al.* 2010).

In zwei Studien wurden jeweils über einhundert *loci* in Amerikanern afrikanischer und europäischer Abstammung untersucht, und beide Male wurden in der zweiten Population in der chromosomalen Region um *TRPV6* Selektionssignale durch *summary statistics* zur Allelfrequenzverteilung und durch Modellierung demographischer Modelle identifiziert (Akey *et al.* 2004, Stajich & Hahn 2005). Eine d_N/d_S -Analyse (Analyse des Verhältnisses der Häufigkeit nicht-synonymer zu synonymen SNPs) bestätigte die Hypothese von positiver Selektion, und Neutralitätstests an weiteren Populationen weiteten sie allgemein auf nicht-afrikanische Bevölkerungen aus (Akey *et al.* 2006, Soejima *et al.* 2009).

Dabei scheint ein Haplotyp mit drei Mutationen in *TRPV6* die selektierte Variante zu sein. Alle drei SNP führen zu Aminosäuresubstitutionen: Arg157Cys durch C > T in Exon 4 (rs4987657), Val378Met durch G > A in Exon 9 (rs4987667) und Thr681Met durch C > T in Exon 15 (rs4987682). Der abgeleitete Haplotyp (Cystein-Methionin-Methionin) ist in nicht-afrikanischen Populationen beinahe fixiert, während in afrikanischen Populationen das Verhältnis der abgeleiteten und ancestralen Allele der SNPs in etwa ausgeglichen ist (Akey *et al.* 2004). Es wird von einem *soft sweep* basierend auf *standing variation* ausgegangen. Akey *et al.* (2006) schlugen vor, dass *TRPV6* in Afrika möglicherweise unter ausgleichender Selektion stehe. Außerhalb Afrikas scheint der Selektionsdruck geographisch und vielleicht zeitlich variiert zu haben. Hughes *et al.* (2008) verwarfen mit ihren auf *extended homozygosity* basierenden Simulationsanalysen ein Modell mit einem einzigen Selektionsereignis in

einer ancestralen nicht-afrikanischen Population. Stattdessen favorisieren sie unabhängige, parallele Events in Europäern und Asiaten, mit einer rezenteren und intensiveren Selektion in Asien.

Über den Zeitpunkt des oder der *sweeps* herrscht Unklarheit. In ihrer ursprünglichen Studie (2004) vermuteten Akey und Kollegen eine rezente Selektion in Europäern vor etwa 10 000 Jahren im Rahmen der Domestikation milchproduzierender Tiere. Sie zogen dabei Parallelen zum *Laktase*-Gen und der Calcium-Absorptionshypothese von Flatz & Rotthauwe (1973). Die Erweiterung ihres Datensatzes in ihrer Folgestudie (2006), führte zu einer Revidierung dieser Datierung. Da alle nicht-afrikanischen Populationen für den abgeleiteten Haplotypen fast fixiert sind, wurde der *sweep* auf etwa 60 – 100 000 im Zuge der *out-of-Africa*-Migrationen datiert. Hughes *et al.* (2008) schließen dies, wie oben bereits beschrieben aus. Ihren Analysen zufolge waren die zwei oder mehreren parallelen *sweeps* vor etwa 7 000 Jahren durch die Fixierung des Cys-Met-Met-Haplotyps abgeschlossen.

Wie auch in den modernen nicht-afrikanischen Populationen sind die beiden Methionin-Allele mit einem Vorkommen von jeweils 95 % in der nordpontischen Steppe vor 5 000 Jahren fast fixiert. Dies steht im kompletten Gegensatz zur *Laktase* und spricht für einen (Akey *et al.* 2006) oder mehrere (Hughes *et al.* 2008) ältere *sweeps*, bei denen spätestens vor 7 000 bis 5 000 Jahren die heutigen Frequenzen von fast 100 % erreicht wurden. Die Genotypisierung des vermeintlichen Arg157Cys-SNPs musste leider verworfen werden, da sich später herausstellte, dass das Primersystem nicht spezifisch war und gleichzeitig mit TRPV6 ein Abschnitt von *TRPV5*⁴⁴ amplifiziert wurde. Die beiden Gene *TRPV5* und *TRPV6* sind homolog und gehen wahrscheinlich auf eine Genduplikation zurück. TRPV5 ist für Calciumtransport vor allem in der Niere zuständig (Peng *et al.* 2000, van Abel *et al.* 2005). Die Sequenzierung ergab in allen Fällen einen scheinbar heterozygoten Genotyp für diesen SNP. Vermutlich wäre das C-Allel *TRPV5* und das T-Allel *TRPV6* zuzuordnen.

⁴⁴ Zum Zeitpunkt der Entwicklung der Primer war die DNA-Sequenz von *TRPV5* noch nicht in der NCBI-Nukleotid-Datenbank enthalten.

Die funktionellen Auswirkungen der drei Mutationen sind noch ungeklärt. Suzuki *et al.* (2008) befanden, dass die ancestrale Variante von TRPV6 eine höhere Ca^{2+} -Transportaktivität aufweist (bei heterologer Expression in *Xenopus*-Oocyten). Heterozygote zeigten eine intermediäre Aktivität. Der kausale SNP schien dabei Val378Met zu sein, dessen Aminosäure extrazellulär scheinbar nahe einer Bindungsstelle für Regulatoren liegt. Hughes *et al.* (2008) konnten hingegen keine signifikanten Aktivitätsunterschiede feststellen (mit der Patch-Clamp-Methode an humanen Zellkulturen). Sie mutmaßen, dass möglicherweise durch den mutierten Kanal aufgrund einer – allerdings nicht signifikanten – schnelleren Ca^{2+} -abhängigen Inaktivierung, weniger Calcium absorbiert werden könne. Es könnte aber auch regulatorische Unterschiede auf der Protein- oder der Genexpressionsebene geben.

Da der Phänotyp des selektierten Haplotypen noch nicht sicher identifiziert ist, kann über den Selektionshintergrund nur spekuliert werden. Wie auch Akey *et al.* (2004) schlagen Hughes und Kollegen (Hughes *et al.* 2008) Ernährungsumstellungen im Zuge der Einführung von Agrikultur vor, wobei die Selektion parallel in den unterschiedlichen Regionen stattfand. Zwar scheint die Selektion auf einer *standing variation* von mittleren Frequenzen zu beruhen, dennoch muss ein deutlicher Selektionsvorteil vorgelegen haben, damit ein Haplotyp bereits 3 000 bis 5 000 Jahre später fixiert ist. Suzuki *et al.* (2008) spekulierten, dass mit der heller werdenden Haut (siehe Abschnitt 10.4.1) die Menschen mehr Vitamin D_3 und folglich $1\alpha,25$ -Dihydroxyvitamin D_3 produzierten, was wiederum die Aufnahme von Calcium im Darm erhöhte. Um das Risiko einer Hypercalcurie (vermehrte Ausscheidung von Calcium über den Urin) – ausgelöst durch Hypercalcämie (erhöhter Calciumspiegel im Serum) – und damit einer Bildung von Harnsteinen zu reduzieren, hätte sich der abgeleitete Haplotyp unter hellhäutigen Menschen verbreitet.

TRPV6 ist auch an der Aktivierung von T-Lymphocyten beteiligt (Cui *et al.* 2002). Daher könnte die Selektion sich auch auf das Immunsystem beziehen, möglicherweise im Rahmen neuer Krankheiten oder Pathogene im Zuge von Kolonisierungen, durch die zunehmenden Populationsdichten oder Kontakten zu Tieren im Neolithikum (Akey *et al.* 2006, Hughes *et al.* 2008). Da Ca^{2+} in allen Zellen, Geweben und physiologischen

Vorgängen, z. B. als *second messenger* eine wichtige Rolle spielt, könnte der Selektionshintergrund auch in anderen Bereichen liegen.

10.4.4.3 NAT2

NAT2 (*N-acetyltransferase 2*, auch *AAC2*, *arylamide acetylase 2*) kodiert eine von zwei N-Acetyltransferasen, die am Metabolismus und der Detoxifikation von Xenobiotika beteiligt sind. Ein drittes Gen dieser Familie, *NATP*, ist ein Pseudogen (Blum *et al.* 1990). Die Hauptsubstrate von *NAT2* sind Arylamine und Arylhydrazine. Das sind Derivate von Ammoniak (NH_3) bzw. Hydrazin (N_2H_4), bei denen mindestens ein Wasserstoffatom durch einen Rest mit einem aromatischen Grundgerüst ersetzt ist. Auch heterocyclische Amine (im Kohlenstoffring ist mindestens ein Stickstoffatom enthalten) zählen dazu. Sie werden von *NAT2* mithilfe von Acetyl-Coenzym A acetyliert und dadurch deaktiviert (N-Acetylierung) oder aktiviert (O-Acetylierung) (Abbildung 10.4).

Diese aromatischen Stickstoffverbindungen finden sich in vielen Stoffen. Sie sind Bestandteile von Pharmazeutika, finden sich in Zigarettenrauch und entstehen bei Verbrennung fossiler Brennstoffe. Heterocyclische Amine bilden sich beim Braten von Fleisch und Fisch durch die Pyrolyse von Aminosäuren. Zum Teil sind diese Verbindungen karzinogen oder werden während ihrer Metabolisierung, z. B. durch die Bildung instabiler, mutagener Zwischenprodukte. Daher ist *NAT2* vor allem Untersuchungsgegenstand in Krebsstudien (Hein 1988, Nagao *et al.* 1996).

Phänotypisch unterscheidet man zwei Typen von *NAT2*, die „schnellen Acetylierer“ und die „langsamen Acetylierer“. Zu den ersteren gehört die ancestrale Variante *NAT2**4. Allele der Gruppe *5 produzieren die Enzyme mit der geringsten Umsatzrate. Der definierende SNP für den *NAT2**5-Allel-Cluster ist die Mutation 341T > C (rs1801280), die zu einem Aminosäureaustausch Ile114Thr führt (Vatsis *et al.* 1991, Hein 2002, Hein *et al.* 2008). Der „langsame“ Phänotyp ist möglicherweise die Folge einer erhöhten oder schnelleren Degradierung des Proteins (Zang *et al.* 2004).

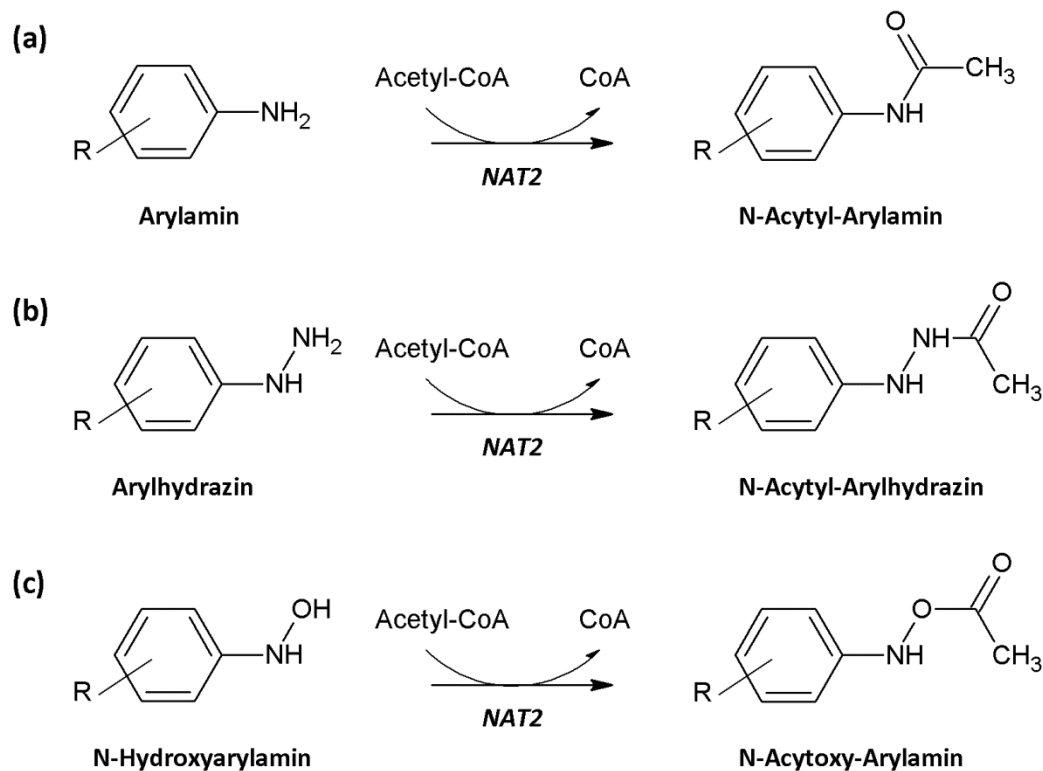


Abbildung 10.4: Beispiele für Acetylierungen durch NAT2

Durch Transfer einer Acetylgruppe von Acetyl-Coenzym A (CoA) auf seine Substrate werden diese aktiviert oder deaktiviert. (a) Beispiel einer N-Acetylierung eines primären Arylamins, (b) N-Acetylierung eines Arylhydrazins, (c) O-Acetylisierung eines N-Hydroxyarylamins. R = Rest (repräsentiert die verschiedenen Typen innerhalb der Substratfamilien).

Patin *et al.* fanden 2006 mithilfe von Sequenz-basierten Neutralitätstests und dem *LRH*-Test heraus, dass der Haplotyp *NAT2**5B (341T>C / Ile114Thr, 481C>T / Leu161Leu, 803A>G / Lys268Arg) in West- und Zentraleurasien unter rezenter Selektion stand oder steht. Dort lag die Häufigkeit des Haplotypen bei über 50 % während sie in osteurasischen Repräsentanten zwischen 6 und 30 % betrug. *Maximum-likelihood*- und koaleszenzbasierte Abschätzungen ergaben, dass in West- und Zentraleurasien das 341C-Allel vor 6 315 (95 % CI: 5 797 – 7 005) Jahren begann, an Häufigkeit zuzunehmen. Als Selektionskoeffizient wurde $S = 0,032$ (0,0124 – 0,0913) ermittelt. Der Zeitrahmen des Selektionsereignisses korreliert mit der Verbreitung von Agrikultur in Westeurasien. Mit der Einführung der produzierenden Wirtschaftsweise wurde der Mensch nach und nach mit neuen Nahrungsquellen,

neuen Verarbeitungstechniken von Lebensmitteln und veränderten Umweltbedingungen konfrontiert (Ferguson 2002, Cordain *et al.* 2005, Luca *et al.* 2010). Daher könnte die Selektion der langsamen Acetylator-Variante in Zusammenhang mit dieser Entwicklung, z. B. in der Nahrungszubereitung, stehen (Patin *et al.* 2006).

Andere Arbeiten bestätigten diese Ergebnisse. Luca *et al.* (2008) und Magalon *et al.* (2008) beobachteten ebenfalls Selektionssignale, befanden aber, dass nicht sicher sei, ob es sich dabei um einen gerichteten oder ausgleichenden Modus handele. Außerdem wurde in drei verschiedenen Studien festgestellt, dass, wie schon von Patin *et al.* (2006) vermutet, die Allelfrequenzen mit der Wirtschaftsweise von Populationen zusammenhängen. In heutigen Jäger/Sammler-Populationen überwiegen deutlich die „schnellen Acetylierer“, in Pastoralismus-betreibenden Gruppierungen und noch stärker in landwirtschaftlichen Gesellschaften die „langsamen Acetylierer“ (Luca *et al.* 2008, Magalon *et al.* 2008, Sabbagh *et al.* 2011). Luca *et al.* (2008) spezifizierten die Selektionshypothese auf den Folsäurehaushalt, der bereits im Rahmen der Pigmentiergene als möglicher Selektionshintergrund diskutiert wurde. Sie schlugen vor, dass es im Neolithikum zu einer Umstellung der Ernährung von Folsäure-reichen grünblättrigen Gemüsen, Leber und Fisch auf Folsäure-arme Getreide, die durch Lagerung und Kochen weiter degradiert wurden, kam. Die *NAT*-Gene spielen eine Rolle im Katabolismus von Folat. Die langsame Variante würde demnach die Erhaltung oder Retention von Folat fördern. Allerdings ist nicht eindeutig geklärt, ob *NAT2* daran beteiligt ist, oder lediglich *NAT1* (Cao *et al.* 2010, Laurieri *et al.* 2014).

Eine Reihe von SNPs spielen eine Rolle in der Acetylierungs-Effizienz von *NAT2*. Im Rahmen dieser Dissertation wurde der in allen *NAT2*5*-Allelen abgeleitet vorkommende SNP T341C untersucht. Das ancestrale Allel überwiegt in der prähistorischen als auch der heutigen Steppe mit etwa 60 %. Dabei sind 62,8 % der kupfer- und bronzezeitlichen Individuen heterozygot, was aufgrund von Kodominanz zu einem intermediären Phänotyp führt (Hein 1988, Zang *et al.* 2004). Der Anteil der Heterozygoten nimmt in der Phase der Katakombengrab- und der Poltavka-Kultur im Vergleich zur Jamnaja-Kultur zu, während der homozygot ancestrale Typ abnimmt. Die Häufigkeit der homozygoten *NAT2*5*-Träger bleibt durchweg niedrig zwischen 0 und 14 %.

Ein weiterer Aspekt der Ernährung ist der Konsum von Alkohol. Dieser wird vorrangig in der Leber durch Alkoholdehydrogenasen (ADH) und Aldehyddehydrogenasen (ALDH) in zwei aufeinanderfolgenden oxidativen Enzymreaktionen metabolisiert (Abbildung 10.5).

Alkoholdehydrogenasen sind Homo- oder Heterodimere, deren Untereinheiten von mehreren Genen kodiert werden. Im Menschen wurden bisher sieben Gene bzw. Untereinheiten identifiziert, die in fünf Klassen unterteilt werden. Die Bezeichnungen haben sich im Verlauf der Forschungsgeschichte mehrfach geändert. Hier wird die offiziell anerkannte Nomenklatur verwendet, wie sie in Edenberg *et al.* (2007) beschrieben ist. Die Ausführungen beschränken sich auf die Enzyme der Klasse I, insbesondere das Gen *ADH1B* (*alcohol dehydrogenase 1B, class I, beta polypeptide*), das früher unter dem Namen *ADH2* (*alcohol dehydrogenase 2*) publiziert wurde. Klasse I setzt sich zusammen aus *ADH1A*, *ADH1B* und *ADH1C* mit den entsprechenden Untereinheiten α , β und γ , die zum fertigen Protein – $[\alpha\alpha]$, $[\beta\beta]$, $[\gamma\gamma]$, $[\alpha\beta]$, $[\alpha\gamma]$ oder $[\beta\gamma]$ – kombiniert werden (Ramchandani *et al.* 2001). Die drei Gene sind Paraloge und zu 94 % identisch zueinander (Carrigan *et al.* 2012). *ADH1A* scheint monomorph zu sein, während man bei *ADH1B* und *C* drei (β_1 , β_2 , β_3) bzw. zwei (γ_1 , γ_2) Varianten charakterisiert, die sich in ihren kinetischen Eigenschaften unterscheiden. Bei *ADH1B* weisen β_2 und β_3 wesentlich höhere Enzymaktivitäten als β_1 auf (Edenberg 2007).

*ADH1B*2* (β_2) ist bis auf einen Aminosäureaustausch identisch mit *ADH1B*1* (β_1). An Position 48⁴⁵ trägt es ein Histidin anstatt Arginin aufgrund einer Punktmutation (G > A; rs1229984) an der zweiten Codonposition (Jörnvall *et al.* 1984, Matsuo *et al.* 1989). Die 48. Aminosäure befindet sich an der Bindungsstelle für das Pyrophosphat des Coenzym NAD(H) (Nicotinamidadenindinukleotid). Histidin führt zu einer schwächeren Bindung, so dass das Coenzym nach der Reaktion vermutlich schneller von der Dehydrogenase dissoziiert (Jörnvall *et al.* 1984, Hurley *et al.* 1990).

Viele verschiedene Studien haben eine Assoziation zwischen der aktiveren β_2 -Variante von *ADH1B* und einem reduzierten Risiko der Alkoholabhängigkeit

⁴⁵ In der älteren Nomenklatur und einigen Publikationen ist die Position als Arg47His angegeben. Mittlerweile wird die vom Startcodon kodierte Aminosäure Methionin mitgezählt, auch wenn diese bei den meisten Proteinen in der posttranslationalen Prozessierung abgespalten wird.

festgestellt (z. B. Thomasson *et al.* 1991, Chen *et al.* 1999, Borràs *et al.* 2000, Bierut *et al.* 2011, Li *et al.* 2011a). Durch eine höhere Umsatzrate von Ethanol zu Ethanal (Acetaldehyd), oder auch eine niedrige Aktivität der Aldehyddehydrogenase im zweiten Oxidationsschritt, kommt es zu einer Akkumulation des giftigen Zwischenprodukts (Crabb *et al.* 2004). Dies führt zu Symptomen wie starkes Erröten (*flush*-Syndrom), Herzrasen, Kopfschmerzen, Schwindelgefühl und Übelkeit (Wolff 1972, Mizoi *et al.* 1979), und folglich zu einer Aversion gegen übermäßigen Alkoholkonsum. Hinzu kommt möglicherweise noch eine Produktinhibition von ADH durch angehäuften Acetaldehyd, wodurch die Wirkung des Alkohols weiter verstärkt wird (Thomasson *et al.* 1993, Kang *et al.* 2014).

Die Histidin-Variante von *ADH1B* kommt vor allem in Asiaten, insbesondere Ostasiaten vor, während es in Europäern selten ist (u. a. Goedde *et al.* 1992, Borràs *et al.* 2000, Li *et al.* 2007, Xue *et al.* 2009). Rs1229984 gehört zu den am stärksten differenzierten nicht-synonymen SNPs im HapMap-Datensatz (2005). Nachdem der *ADH*-Cluster in genomweiten Analysen auffällig war (Voight *et al.* 2006), untersuchten Han und Kollegen (2007) den Klasse I-*locus* auf Selektionssignale. Mithilfe der F_{ST} -Statistik und dem *LRH*-Test wiesen sie die Selektion eines Haplotypen mit der *ADH1B**48His-Variante in mehreren ostasiatischen Populationen nach. Unter einem semidominanten Modell ermittelten sie einen Selektionskoeffizienten von $S = 0,03$. Außerdem identifizierten sie einen weiteren SNP 1 761 bp *upstream* von *ADHB1*, rs3811801 (C > T), der vermutlich einem regulatorischen Element des Gens zuzuordnen ist und das primäre Ziel der Selektion sein könnte (Han *et al.* 2007, Li *et al.* 2008, Xue *et al.* 2009). Li *et al.* (2008) spekulieren, dass eine Kombination von erhöhter Enzymaktivität durch *48His und einer größeren Menge an Enzym aufgrund von *-1 761T in der Promotor-Region die phänotypischen Grundlage für die Selektion bildete. Auch sie fanden Selektionssignale nur in bestimmten asiatischen Populationen. Sie machten Andeutungen, dass hier die subsistenzwirtschaftliche Geschichte eine Rolle spielen könnte, aber es war keine eindeutige Korrelation zwischen den genetischen und historischen oder ökologischen Mustern zu erkennen.

Peng *et al.* (2010) stellten einen Zusammenhang zwischen den Frequenzen des *ADH1B**His-Allels und der Domestikation von Reis fest. Bereits 7 000 v. Chr. wurde in

Südchina Reis fermentiert bzw. in Gemischen mit anderen Lebensmitteln zu alkoholischen Getränken vergoren (McGovern *et al.* 2004). Fermentation macht Lebensmittel haltbar, tötet Pathogene ab, bietet eine Alternative zu verunreinigtem Wasser, erhöht die Bioverfügbarkeit von Spurenelementen etc. (Vallee 1998, McGovern *et al.* 2004, Hotz & Gibson 2007). Peng *et al.* (2010) schlugen vor, dass mit zunehmendem Konsum von Alkohol durch die Fermentierung von Reis im Neolithikum das abgeleitete Enzym vorteilhaft wurde. Da es eine höhere Umsatzrate hat, ist die Leber dem Ethanol weniger bzw. kürzer ausgesetzt. Auf der anderen Seite kommt es jedoch durch die Acetaldehyd-Anhäufung zu unangenehmen Nebenwirkungen. Die Selektion scheint wohl auf einem schützenden Effekt vor übermäßigem Alkoholkonsum und seinen gesundheitlichen Folgen zu beruhen. Als weiterer Faktor könnte noch Pilzbefall von Reis (ein allgemein in der Agrikultur großes Problem) hinzukommen als Ursache für die Selektion ist. Schimmeliger Reis enthält Mykotoxine, deren toxikologischen Effekte durch Alkoholkonsum noch verstärkt werden (Bennett & Klich 2003, Han *et al.* 2007). Alternativ kämen vielleicht noch unbekannt infektiöse Krankheiten durch Alcedaldehyd-empfindliche Mikroorganismen in Frage (Han *et al.* 2007).

Das vermutliche Alter des *ADH1B**2-Allels von etwa 7 000 bis 10 000 Jahren stützt die Hypothese zur Gen-Kultur-Koevolution im Zuge der Reisdomeestikation (Peng *et al.* 2010). Allerdings war die Datierung nach Slatkin & Rannala (2000) eher unsicher. Li *et al.* (2011b) schätzten den Expansionsbeginn der *48His tragenden Haplotypen auf vor etwa 50 – 70 Tausend Jahre im Nahen Osten, also in die Zeit der Migrationen aus Afrika. Die Mutation in der Promotorregion entstand vor etwa 4 500 – 2 000 Jahren auf einem dieser Haplotypen und nahm in Ostasien rasch an Häufigkeit zu. Das Team ging aber davon aus, das Alter unterschätzt zu haben, da sie eine Generationszeit von nur 20 Jahren annahmen.

Das Ergebnis der Genotypisierung der Steppenpopulationen war eindeutig. Alle 29 bzw. 32 Individuen waren homozygot für die ancestralen Allele von rs3811801 und rs1229984. Dies ist ein weiteres Indiz für eine eher westliche Orientierung der Beziehungen dieser Kulturen. Hier liegt demnach, zumindest was die Alkoholdehydrogenase betrifft, keine Hypersensitivität gegen Alkohol vor. In den

modernen Ukrainern kommen vereinzelt heterozygote Individuen vor, tendenziell eher für den Aminosäureaustausch (4 von 23 Individuen, 1 von 46 für den regulatorischen SNP). Borinskaya *et al.* (2009) kamen mit einer Histidinfrequenz von 8,3 % (91 Arg/Arg, 18 Arg/His) auf die gleiche Allelhäufigkeit in Ukrainern wie Hollfelder (2013). Interessanterweise führt das Vorkommen von Heterozygoten im modernen Datensatz zu einem F_{ST} -Wert von 0,08519 ($p < 0,05$) zu den alten Proben, obwohl sich die Allelfrequenzen nicht signifikant unterscheiden. Im Vergleich zu den nachweislich selektierten Pigmentier-*loci* ist der F_{ST} -Wert jedoch eher niedrig.

10.4.5 Pathogene

10.4.5.1 CASP12

Die Caspasen (CASP) bzw. Cyteinyalaspartasen sind eine Familie von Proteasen mit einem Cystein im aktiven Zentrum, die Proteine gezielt hinter Aspartatresten spalten. Sie werden entsprechend ihrer Funktionen in drei Gruppen eingeteilt. Zur Gruppe I gehören die an der unspezifischen Immunabwehr beteiligten (inflammatorischen) Caspasen. Die Gruppen II und III setzen sich aus Enzymen zusammen, die für die Initiierung und Durchführung des programmierten Zelltods (Apoptose) zuständig sind.

CASP1, 4, 5 und *12* (Gruppe I) liegen in einem Cluster auf Chromosom 11 und gingen höchstwahrscheinlich aus dem gleichen ancestralen Gen hervor (Fischer *et al.* 2002). Sie spielen eine bedeutende Rolle bei der Regulierung von Entzündungsreaktionen auf Pathogene. Deren Vorhandensein wird durch Rezeptoren des Immunsystems anhand molekularer Merkmale, sogenannter *pathogen-associated molecular patterns* (z. B. bakterielle Membranbestandteile wie Lipopolysaccharide, bakterielle oder virale DNA/RNA etc.) erkannt, woraufhin verschiedene rezeptorspezifische Schutzreaktionswege eingeleitet werden. So wird zum einen über den nach dem beteiligten Transkriptionsfaktor benannten *Nf- κ B pathway* Pro-Interleukin-1 β produziert (Goto *et al.* 1999, Ogura *et al.* 2006). Zum anderen werden in den Makrophagen die Inflammasomen stimuliert. Diese Proteinkomplexe enthalten die Caspase 1. Sie katalysiert unter anderem die Reifung der proinflammatorischen

Botenstoffe (Pro-)Interleukin-1 β und -18, die eine Reihe von Mechanismen aktivieren, welche zur Entzündung des umliegenden Gewebes, d. h. seiner Rekrutierung für die Abwehrreaktion, führen (Keller *et al.* 2008). Im Zuge der Immunantwort werden noch weitere Inflammationsmediatoren ausgeschüttet (Cytokine, Chemokine, Lipide, Radikale etc.), die im komplexen Zusammenspiel zum Teil antagonistische Funktionen haben, so dass ein regulierter Ablauf der Entzündungsreaktion gewährleistet ist (Cohen 2002, Schulte *et al.* 2013). Ein weiterer Regulationsmechanismus ist die Caspase 12 mit ihrer inhibitorischen Wirkung sowohl auf Nf- κ B als auch auf CASP1. Damit dämpft CASP12 die Antwort auf bakterielle Infektionen. Dies scheint interessanterweise ohne eine proteolytische Aktivität des Enzyms zu geschehen (Saleh *et al.* 2006, Labbé *et al.* 2010).

Allerdings tragen die meisten Menschen eine inaktive Form der *Caspase 12*. Durch eine *nonsense*-Mutation an Nukleotidposition 629 (rs497116) wird das für Arginin kodierende Codon CGA im Exon 4 durch das Stopp-Codon TGA ersetzt (Fischer *et al.* 2002). Die mRNA wird dann über durch den *nonsense-mediated mRNA decay* degradiert, so dass das verkürzte Protein nicht synthetisiert wird (Yeretssian *et al.* 2009, Yngvadottir *et al.* 2009). Die Immunantwort auf pathogene Reize hängt stark vom Genotyp eines Individuums ab. Bei Menschen, die homozygot für das Pseudogen (*CASP12-S*) sind, ist die Reaktion am stärksten, bei Homozygoten für das funktionelle Enzym (*CASP12-L*) am schwächsten, und bei Heterozygoten dazwischenliegend (Saleh *et al.* 2004).

Das aneztrale *CASP12-L* ist außerhalb Afrikas äußerst selten (Saleh *et al.* 2004, Wang *et al.* 2006, Xue *et al.* 2006). Es ist nur noch in vereinzelt Regionen in Israel, Pakistan, China und Mexiko mit einer Häufigkeit von weniger als einem Prozent zu finden. In Europa ist das abgeleitete *CASP12-S*-Allel fixiert. Lediglich im subsaharischen Afrika tragen nach Saleh *et al.* (2004) ungefähr 20 % der Menschen mindestens ein *CASP12-L*-Allel, wobei dessen Frequenz bei etwa 11 % liegt. Xue und Kollegen (2006) stellten basierend auf dem HGDP-CEPH-Panel⁴⁶ mit ca. 28 % eine weitaus höhere

⁴⁶ HGDP-CEPH ist eine Zellbank mit 1063 menschlichen lymphoblastoiden Zelllinien, die im Rahmen des HUMAN GENOME DIVERSITY PROJECT (HGDP) aus verschiedenen Laboratorien gesammelt wurden und im CENTRE D'ÉTUDE DU POLYMORPHISME HUMAIN (CEPH) in Paris lagern bzw. kultiviert werden. Die Zelllinien,

Frequenz des C-Allels in Afrika fest. Die Häufigkeiten variieren dabei deutlich zwischen einzelnen Populationen. Bei den kenianischen Bantu liegt die Allelfrequenz von *CASP12-L* bei unter 4 %, bei den Bambuti und den San ist sie mit 60 bzw. 57 % am höchsten.

In zwei unabhängigen Studien wurden starke Hinweise darauf gefunden, dass die beinahe vollständige Verdrängung des funktionellen Allels in weiten Teilen der Welt die Folge einer positiven Selektion der *nonsense*-Mutation ist. Xue *et al.* (2006) untersuchten eine resequenzierte Stichprobe des HapMap-Panels (2003). In der fast gleichzeitig publizierten Studie von Wang und Kollegen (2006) wurden Proben der HGDP-CEPH-Zellbank (hauptsächlich afrikanischer Herkunft) analysiert. Beide Arbeitsgruppen basierten ihre Selektionshypothese auf sequenzbasierten Neutralitätstests und Koaleszenzsimulationen, Wang *et al.* (2006) außerdem noch auf der Haplotypenverteilung.

Der Zeitpunkt der *nonsense*-Mutation wurde mithilfe einer auf der Phylogenie des Gens basierten Methode (Bandelt *et al.* 1999), je nach Wurzelung des Baums, auf vor 552 ± 276 bzw. 397 ± 223 Tausend Jahre geschätzt (Xue *et al.* 2006). Der Zeitpunkt der beginnenden Ausbreitung des inaktiven Allels wurde dabei auf 61 ± 16 000 Jahre vor heute festgelegt. Der Selektionskoeffizient wurde durch eine *composite likelihood ratio*-Analyse (CLR) (Kim & Stephan 2002) und eine *full likelihood*-Methode nach Coop und Griffiths (2004) abgeschätzt. Der Selektionsvorteil des inaktiven Gens liegt demnach zwischen 1,7 und 0,8 % (Xue *et al.* 2006).

Bei der Suche nach dem selektiven Vorteil berufen sich die Autoren der beiden populationsgenetischen Studien auf die Ergebnisse von Saleh und Kollegen von 2004 und 2006. Diese stellten anhand von Mausmodellen fest, dass Träger des Pseudogens eine höhere Resistenz gegen bakterielle Sepsis haben als Individuen mit der ursprünglichen Form der *Caspase 12*. Unter einer Sepsis versteht man ein äußerst komplexes klinisches Syndrom, das durch eine systemische Entzündungsreaktion des Körpers auf eine Infektion hervorgerufen wird (Bone *et al.* 1992, Levy *et al.* 2003). Sie

deren DNA zu Studienzwecken an Forschungsinstitutionen verteilt wird, stammen von 1050 Individuen aus 52 Populationen zur Repräsentation der weltweiten Diversität (Cann *et al.* 2002).

kann entstehen, wenn es der Körper nicht schafft, eine lokale Infektion einzudämmen, so dass die Keime in den Blutkreislauf gelangen und im gesamten Organismus zu Abwehrreaktionen führen. Während eine räumlich begrenzte Entzündung durchaus effektiv für die Bekämpfung eines lokal eingedrungenen Erregers ist, so hat ihre Ausbreitung schwerwiegende Folgen. Zu den Symptomen der Sepsis zählen je nach Schweregrad Fieber, Hypothermie, beschleunigter Herzschlag, beschleunigte Atmung, Verwirrtheit, erhöhte oder verminderte Anzahl an Leukocyten. Kommt es zu Organdysfunktionen, spricht man von schwerer Sepsis. Ein septischer Schock ist ein akutes Kreislaufversagen durch anhaltend starke arterielle Hypotonie. Da all diese Symptome sehr unspezifisch sind und auch bei anderen Krankheitsbildern auftreten, ist eine genaue Definition von Sepsis problematisch. Entsprechend schwierig gestaltet sich die Suche nach Ursachen und Behandlungsmethoden der oft tödlich verlaufenden Krankheit (siehe u. a. Bone *et al.* 1992, Levy *et al.* 2003, Riedemann *et al.* 2003).

Wahrscheinlich ist die Sepsis auf eine Dysregulation bei der Immunantwort zurückzuführen. Dabei ist das Gleichgewicht zwischen den verschiedenen Mediatoren gestört – wie genau, wird noch erforscht. Zum einen ist womöglich die Ausschüttung proinflammatorischer Substanzen nicht pulsartig am Anfang der Reaktionskaskade, sondern andauernd. Zum anderen kommen außerdem antiinflammatorische Prozesse in Gang, wobei es unter anderem zur Apoptose von Immunzellen kommt. Dieses Wechselspiel führt zu Gewebeschädigungen, diese zur Beeinträchtigung des Kreislaufsystems und schließlich Organversagen (Cohen 2002, Pinsky 2004, Giamarellos-Bourboulis *et al.* 2011, Schulte *et al.* 2013). Da eine funktionelle Caspase 12 durch Inhibition von Nf- κ B und Caspase 1 die Reifung bzw. Aktivierung der Cytokine Interleukin-1 β und -18 reduziert, wird durch sie möglicherweise das Verhältnis der Antagonisten weiter in Richtung Antiinflammation verschoben. Vielleicht ist im Vorfeld schon das Risiko größer, dass ein Infektionsherd nicht lokal eingedämmt wird, wenn die initiale Immunantwort gedämpft ist. Saleh *et al.* (2004, 2006) zeigten entsprechend, dass bakterielle Angriffe in *CASP12-S*-Modellen effizienter abgewehrt werden. Außerdem stellten sie in Humanstudien fest, dass in afrikanischen Populationen schwere Sepsis signifikant häufiger bei Trägern des *CASP12-L*-Allels auftritt (7,8-fach bei Homozygoten). Auch die Sterberate in Folge von

Sepsis ist bei dieser Gruppe wesentlich höher (54 %) als bei Homozygoten für *CASP12-S* (17%).

Basierend auf dem Auftreten von schwerer Sepsis sowie der Sterberate in Abhängigkeit vom Genotyp errechneten Wang und Kollegen (2006) einen Selektionskoeffizienten von 0,9 %, der damit den unteren Bereich des Intervalls aus der Studie von Xue *et al.* (2006) bestätigt (siehe S. 267). In letzterer Studie wird als möglicher Auslöser der Selektion die Populationsexpansion in Afrika im Palaeolithikum diskutiert. Da eine hohe Bevölkerungsgröße und -dichte eine Grundlage für die Verbreitung von Krankheitserregern darstellt, spielten Infektionen wahrscheinlich eine zunehmende Rolle. Unklar ist, warum *CASP12-L* heutzutage in einigen afrikanischen Populationen immer noch vergleichsweise häufig auftritt (McCall *et al.* 2010, Hermel & Klapstein 2011).

Den Ergebnissen der hier beschriebenen Studien entsprechend, ist die das Stopp-Codon hervorrufende abgeleitete Variante von rs497116 im vorliegenden prähistorischen Datensatz fixiert. Der auf *standing variation* beruhende *soft sweep*, der wahrscheinlich in Afrika begann, war vor über 5 000 Jahren zumindest in Osteuropa wohl bereits „complete“, also abgeschlossen.

10.4.5.2 ABCB1

ABCB1 (*ATP-binding cassette, subfamily B, member 1*) kodiert einen Transmembrantransporter mit breitem Substratspektrum. Verbreitet ist auch die Bezeichnung *MDR1* (*multidrug resistance 1*), seltener *PGY1* (*P-glycoprotein 1*). Das Gen wird hauptsächlich in Epithelzellen von Darm, Niere, Leber, Pankreas, Plazenta und Zentralnervensystem exprimiert (Thiebaut *et al.* 1987, Sugawara *et al.* 1988, Cordon-Cardo *et al.* 1989). Das Membranprotein transportiert aktiv cytotoxische Substanzen unter ATP-Verbrauch aus Zellen heraus. Hierzu zählen viele Arzneimittel, Giftstoffe, schädliche Nahrungsmittelbestandteile, Hormone, Peptide etc. (Croop *et al.* 1988, Schinkel 1997). *ABCB1* fungiert dabei vermutlich als „Flippase“ oder „Staubsauger“ (Higgins & Gottesman 1992, Sharom 2014). Die schädliche oder fremde Substanz

dringt von außen (z. B. aus dem Darmlumen) in die Zellmembran ein und gelangt zur cytoplasmatischen Lipidschicht. Dort docken sie an die Substratbindungsstelle von ABCB1 an. Nach dem „Staubsauger“-Modell werden sie direkt in die extrazelluläre Phase ausgeströmt. Beim „Flippase“-Modell geschieht der Transport über den „Flipflop“-Mechanismus zur äußeren Membranschicht, von wo aus die Substanz sich in den extrazellulären Raum absetzt. Die Beförderung durch die Membran erfolgt im Rahmen einer Konformitätsänderung des Proteins.

ABCB1 schützt zum einen den Körper im Allgemeinen vor schädlichen Substanzen, indem er z. B. im Darm deren Aufnahme verhindert. Zum anderen ist der Transporter Teil der Schutzbarrieren von sensitiven Geweben, beispielsweise der Blut-Hirn-, der Blut-Hoden-, der Plazenta-Schranke zum Schutz des Fötus oder des Endokards, das unter anderem als eine Art Blut-Herz-Schranke fungiert (Schinkel *et al.* 1994, Lankas *et al.* 1998, Smit *et al.* 1999, Melaine *et al.* 2002, Sissung *et al.* 2011). Es gibt Hinweise, dass *ABCB1* mit den bereits in 10.4.1 und 10.4.2 diskutierten *CYP3A4*- und *CYP3A5*-Genen interagiert, um den Schutz des Körpers vor xenobiotischen und die Regulation des Transports endogener Substanzen zu gewährleisten (Schuetz *et al.* 2000, Kivistö *et al.* 2004, Eap *et al.* 2007).

Es wurde eine Reihe von Polymorphismen in *ABCB1* identifiziert, die die Funktion des Transporters beeinflussen (Fromm 2003). Dazu gehören auch drei SNPs, die unter Selektionsverdacht stehen: C1236T in Exon 12 (rs1128503), G2677A/T (rs2032582) in Exon 21 und C3435T (rs1045642) in Exon 26 (Tang *et al.* 2004, Wang *et al.* 2007a). Die Mutation in Exon 21 hat einen Aminosäureaustausch Ala893Thr/Ser zur Folge, die beiden anderen sind synonym. 2677G bzw. A befinden sich meist auf einem Haplotypen mit 3435C, während 2677T eher mit 3435T korreliert (Horinouchi *et al.* 2002). Tang *et al.* (2002, Wang *et al.* 2007b) zeigten starkes *linkage disequilibrium* zwischen allen drei SNPs, aber unterschiedliche Varianten in asiatischen Populationen. Dabei identifizierten sie über das etwa 200 kb lange Gen einen Haplotypen (mh7), der den Subhaplotypen CGC (die Basen repräsentieren die Allele der drei hier untersuchten SNPs) enthielt, und 34 % der afrikanischen Chromosomen ausmachte, während in nicht-afrikanischen Populationen ein Haplotyp mit dem TTT-Subhaplotyp (mh5) mit 25 bis 50 % weit verbreitet ist. In der Stichprobe „Caucasians“ war mh5 mit

33,80 % und mh7 mit 18,83 % vertreten (Tang *et al.* 2004). Durch eine Adaption des *LRH*-Tests auf einzelne SNPs identifizierten Tang und Kollegen (2004) eine positive Selektion von 3435C in afrikanischen Populationen und von 2677T in Malaien. 2677T und 3435T wiesen ein signifikantes Selektionssignal in Chinesen auf; in Indern und Europäern nur unter bestimmten Populationsmodellen und Parametern. Wang *et al.* (2007b) beobachteten eine Selektion der abgeleiteten Allele aller drei SNPs in Chinesen. Außerdem schien 1236T in Europäern selektiert. Zusätzlich fanden sie Selektionssignale in der regulatorischen Region. Die Selektion der ancestralen Variante 3435C in Afrika konnten Wang *et al.* (Wang *et al.* 2007b) nicht bestätigen.

Aufgrund seiner Funktion ist es naheliegend, dass die Konfrontation mit neuen Umweltbedingungen bei der Besiedelung verschiedener Erdteile zu Selektionsdrücken auf *ABCB1* führte. Allerdings sind die Ergebnisse zur funktionellen Bedeutung der drei Mutationen nicht eindeutig. Für 2677T und/oder 3435T wurden sowohl verminderte (Hoffmeyer *et al.* 2000) als auch erhöhte (Kim *et al.* 2001, Nakamura *et al.* 2002) Expression bzw. Transportaktivität, als auch keine signifikante Veränderung festgestellt (Kroetz *et al.* 2003). Während C2677A/T durch einen nicht-konservativen Aminosäureaustausch zu einer direkten Veränderung des Proteins führt, handelt es sich, wie bereits erwähnt, bei C1236T und C3435T um stille Mutationen. Kimchi-Sarfaty *et al.* (2007) demonstrierten, dass durch die T-Allele zwei seltene Codons für die entsprechenden Aminosäuren Glycin und Isoleucin zum Einsatz kommen. Dies beeinflusst das *timing* der kotranslationalen Faltung des Proteins und seine Insertion in die Membran. Dadurch komme es zu strukturellen und folglich funktionellen Veränderungen in den Substrat- und Modulatorbindestellen (Kimchi-Sarfaty *et al.* 2007, Fung *et al.* 2014).

Lee *et al.* (2004) sowie Tan *et al.* (2005) stellten eine Assoziation zwischen C1236T, G2677A/T und C3435T und dem Parkinson-Syndrom in Chinesen fest, wobei Träger des abgeleiteten TTT-Haplotypen ein geringeres Erkrankungsrisiko haben. Tang *et al.* (2004) vermuteten daraufhin, dass die Selektion einen erhöhten Schutz des Gehirns vor xenobiotischen Angriffen an der Blut-Hirn-Schranke betrifft. Parkinson an sich dürfte wohl kaum ein Selektionsgrund sein, da die Krankheit erst im Alter auftritt und somit nicht die reproduktive Fitness beeinträchtigt. Da *ABCB1* am Schutz vieler

wichtiger Gewebe beteiligt ist, lässt sich die Selektionsursache nicht sicher bestimmen. Neben dem Schutz vor schädigen Substanzen könnte auch die Assoziation des 3435C-Allels mit Bluthochdruck eine Rolle spielen (Eap *et al.* 2007, Lacchini *et al.* 2012).

Die prähistorischen Daten dieser Arbeit bestätigen die unvollständige Kopplung der drei SNPs. C1236T, G2677A/T stimmen in allen 19 Individuen, für die Genotypen an allen drei Positionen vorliegen im Genotyp überein, d. h. sie sind immer beide homozygot ancestral, abgeleitet oder heterozygot. In einem einzigen Fall treten sie heterozygot als CT – GA auf. Dies deutet auf eine vollständige Kopplung hin. Allerdings ist dies gesichert, da die gametischen Phasen nicht bekannt sind. Die Häufigkeit des 3435T-Allels ist höher als die der beiden anderen abgeleiteten Varianten. In fünf Fällen tritt es auf einem der beiden Chromosomen zusammen mit den ancestralen Allelen der beiden anderen SNPs auf, d. h. CT – GT – TT bzw. CC – GG – CT, allerdings nie auf beiden Chromosomen. Auch tritt das 3435C-Allel nicht zusammen mit 1236T oder 2677T auf. Dies könnte für 3435T als primäres Ziel einer etwaigen Selektion sprechen oder allgemein gegen Selektion. Insgesamt stimmen die Allelfrequenzen der alten und modernen Proben in etwa überein, so dass, falls Selektion stattgefunden hat, dieses Ereignis älter als 5 000 Jahre sein muss. Dies wird gestützt durch Signale in verschiedenen nicht-afrikanischen Populationen, so dass es sich wahrscheinlich um eine Selektion im Zuge der *out-of-Africa*-Migrationen handelt. Für eine Wirkung von Selektion auf *ABCB1* spricht seine wichtige Rolle im Organismus. Die Tatsache, dass zwei der drei Mutationen still sind und bisher kein eindeutiger Einfluss der SNPs auf die Funktion des Gens nachgewiesen werden konnte, indiziert einen vorwiegend demographischen Einfluss auf die Entwicklung der Allelfrequenzen. In den prähistorischen Steppenkulturen deutet der exakte Hardy-Weinberg-Test auf eine neutrale Evolution der *ABCB1*-SNPs hin. Interessanterweise ist die Häufigkeit der abgeleiteten Allele in der Katakombengrabkultur 41 bzw. 14 % niedriger als in der Jamnaja-Kultur bzw. der modernen Population der Ukraine. Allerdings sind die Unterschiede nicht signifikant. Es wurden keine Korrelationen zwischen den Genotypen und bestimmten mtDNA-Haplotypen, wie z. B. Haplogruppe U4, festgestellt. Es könnte sich möglicherweise um noch nicht aufgedeckte Strukturen im Datensatz oder einfach um einen *sampling bias* handeln.

11 Synthese und Ausblick

Mit der vorliegenden Arbeit wurden erstmals prähistorische Bevölkerungsstrukturen in der osteuropäischen Steppe von der Oberthrakischen Tiefebene bis zur Wolga populationsgenetisch untersucht. Hierfür wurden in Zusammenarbeit mit M. U. Protokolle für Multiplex-PCRs entwickelt, mit denen von der mitochondrialen DNA neben der HVR I noch 30 kurze Abschnitte in der *coding region* mit 32 kladenspezifischen Polymorphismen, sowie 20 putativ selektierte autosomale SNPs und ein geschlechtsbestimmender Locus amplifiziert wurden. Für die Parallelsequenzierung der Proben auf der 454-Plattform wurde die Methode des Sequenz-tagging von Meyer *et al.* (2008) für die Anwendung auf alte DNA modifiziert. Von 66 der 180 kupfer- und bronzezeitlichen Individuen wurden mtDNA-Sequenzen gewonnen und reproduziert, wobei die HVR I von 65 für die Datenanalysen verwendet wurden. Sie wurden vier Kulturkomplexen zugeordnet: Äneolithikum (10 Sequenzen), Jamnaja-Kultur (22), Katakombengrabkultur (23) und Poltavka-Kultur (6). Von 61 Individuen liegen Genotypen vor. Zu Vergleichszwecken wurden bereits veröffentlichte aDNA-Daten aus Mittel- und Osteuropa, Zentralasien, sowie moderne DNA-Sequenzen herangezogen.

Der mitochondriale Genpool der untersuchten Steppenpopulation setzt sich aus Linien der Haplogruppe U, die mit Jägern/Sammlern assoziiert sind, und Linien, die in Mitteleuropa ab dem Neolithikum nachgewiesen sind, zusammen. Direkt aus dem Untersuchungsgebiet gibt es keine prä-neolithischen Sequenzdaten, aber basierend auf Beobachtungen im Jungpalaeolithikum und Mesolithikum Mittel- und Osteuropas ist anzunehmen, dass auch in der Steppe ursprünglich typische „Jäger/Sammler-Linien“ vorherrschten. Demnach stützt die Haplogruppenzusammensetzung der Steppenpopulation die archäologisch basierte Hypothese, dass die Criş-Kultur, eine der vermutlichen Vorläufer-Kulturen der LBK, der erste Impulsgeber zur sukzessiven Etablierung der Viehwirtschaft in der Steppe war. Dabei sind Bauern in die Steppe eingewandert (demische Diffusion). In der frühen Bronzezeit ist die Viehhaltung als Subsistenzwirtschaft voll etabliert, und U- und andere Linien sind in gemeinsamen

Fundplätzen nachgewiesen. Der SNP in *NAT2*, der in heutigen Populationen mit deren Wirtschaftsweise assoziiert ist, tritt vor allem als heterozygoter Genotyp in der prähistorischen Steppenbevölkerung auf. Diese Beobachtungen deuten auf ein *admixture* zwischen Bauern und ursprünglichen Steppenbewohnern hin.

Die zunächst mittleren genetischen Distanzen zwischen den äneolithischen Steppenbewohnern zur LBK und der niedrige F_{ST} -Wert (0,02807) zwischen dieser und der Jamnaja-Kultur weist auf zunehmende Beziehungen der beiden Kulturkomplexe hin. Von den Jäger/Sammler-Gruppen ist die Jamnaja-Population distanziert, während zu den der LBK nachfolgenden Kulturen bis zu Bronzezeit praktisch keine genetischen Distanzen festzustellen sind. Dies spricht für andauernde oder wiederkehrende Kontakte. Aufgrund der durchweg um Null liegenden F_{ST} -Werte konnte nicht festgestellt werden, ob im Speziellen die Jamnaja-Kultur aktiv an der Genese der Schnurkeramik beteiligt war, oder lediglich kulturelle Beziehungen zwischen den ungefähr kontemporären Kulturen bestanden. Die hohe genetische Diversität der Jamnaja-Kultur, sowie die signifikant negativen Werte von Tajima's D und Fu's F_S sprechen für eine Populationsexpansion, die auch archäologisch, z. B. durch die zehntausende von Kurganen über den nordpontischen Raum verteilt bis in die Ungarische Tiefebene, nachgewiesen ist.

In der Katakombengrabkultur gibt es diese Hinweise nicht mehr. Mit F_{ST} -Werten von 0,03 bis 0,08 ist die Kultur signifikant von den mitteleuropäischen Bauern distanziert (Ausnahme: zur Glockenbecherkultur, $F_{ST} = 0,018$). Im Gegensatz zu den anderen nordpontischen Steppenkulturen zeigt die Katakombengrabkultur (insbesondere in der frühen Phase) eine geringe genetische Distanz zu den nordeuropäischen Jägern/Sammlern der Grübchenkeramik. Beide Populationen, sowie Mesolithiker des Baltikums und Nordwestrusslands weisen sich durch einen hohen Anteil an U4-Haplotypen aus. Es ist wahrscheinlich, dass bereits in der ursprünglichen Jäger/Sammler-Population der nordpontischen Steppe U4-Linien verbreitet waren. Bei einer steppeninternen Entwicklung würde man allerdings auch entsprechende Anteile der Haplogruppe in der Jamnaja-Kultur erwarten. Entweder wurden diese nicht erfasst, weil sie womöglich auf die Gruppen in der Don-Wolga-Region beschränkt waren, oder Immigranten aus den ostbaltischen Jäger/Sammler-Refugien waren an

der Genese der Katakombengrabkultur beteiligt. Die frühe Katakombengrabkultur zeigt einen hohen F_{ST} -Wert zu der kleinen Stichprobe der Poltavka-Kultur in der Wolga-Region, die sich weder kulturell noch genetisch von ihrer Vorläuferkultur, der Jamnaja, unterscheidet.

Während im Neolithikum noch zentralasiatische Haplogruppen gefunden wurden, so ist der vorliegende Datensatz ausschließlich von europäischen mtDNA-Linien geprägt. Auch bei den autosomalen SNPs sind asiatische Einflüsse nahezu absent. Lediglich das in Asien hochfrequente, in Europa seltene Allel von *ABCC11* ist in der nordpontischen Bronzezeit häufiger als in der modernen Stichprobe.

Die genetischen Distanzen zwischen den nordpontischen Steppenkulturen und den Vergleichspopulationen aus der mittleren Bronzezeit Zentralasiens sind insgesamt sehr hoch. Keine Distanz ist allerdings zwischen der Jamnaja-Kultur und einer bronzezeitlichen Kultur in Kasachstan festzustellen. Dies könnte ein Indiz für die von Archäologen vermuteten Migrationen von Jamnaja- oder Repin- (Vorläufer der östlichen Jamnaja-Kultur) Gruppen nach Osten sein. Ein weiterer Sonderfall sind die niedrigen F_{ST} -Werte der nordpontischen Kulturen zur jüngeren Phase der Andronovo-Kultur. Im archäologischen Fundgut gibt es Hinweise auf Kontakte zwischen der spätbronzezeitlichen Srubna-Kultur im Nordschwarzmeerraum und der Andronovo-Kultur östlich des Kaspischen Meeres.

Phänotypisch ist die kupfer- und bronzezeitliche Bevölkerung der nordpontischen Steppe den heutigen Ukrainern bezüglich der untersuchten *loci* sehr ähnlich, in der Pigmentierung im Durchschnitt vermutlich etwas dunkler. Die einzige Ausnahme ist das vollständige Fehlen der Laktasepersistenz, die heute in der ukrainischen Population bei etwa 35 % liegt. Dies widerlegt die Vermutung, dass das abgeleitete *LCT*-Allel durch die Jamnaja-Kultur in das europäische Erbgut gelangte.

Die niedrigen F_{ST} -Werte zwischen den prähistorischen Steppenpopulationen und der heutigen Bevölkerung Mittel- und Osteuropas weisen auf genetische Kontinuität hin. Nach dem aktuellen Kenntnisstand lässt sich der moderne europäische Genpool auf drei Wurzeln zurückführen: die westeuropäischen mesolithischen Jäger/Sammler, die aus dem Nahen Osten eingewanderten frühen Bauern und eine nordeurasische

Komponente jungpalaeolithischen Ursprungs. Die nordpontische Population könnte womöglich ein Kandidat sein, über den letztere in das Erbgut der spätneolithischen Europäer gelangt ist.

Der vorliegende Datensatz und die mithilfe von *summary statistics* gewonnenen Erkenntnisse bieten eine Grundlage für weiterführende Untersuchungen. Auf Modellsimulationen basierende Analysen könnten Aufschluss über den Grad der Vermischung zwischen den Jägern/Sammlern der Steppe und frühneolithischen Bauern geben, oder den Einfluss der Steppe auf die Entwicklungen im mitteleuropäischen Neolithikum quantifizieren.

Zukünftig könnten aDNA-Daten aus Ost- und Südosteuropa die Lücke zwischen Mitteleuropa und der Steppe schließen. Die Binnenstruktur des nordpontischen Steppenraums betreffend, könnte die Analyse weiterer Proben, insbesondere von äneolithischen Gruppen östlich des Dnjeprs, von Jamnaja-Stichproben aus der Region zwischen Don und Wolga und von Funden aus Katakombengräbern westlich des Dnjepr, einige der hier beschriebenen Ergebnisse und Hypothesen festigen oder widerlegen, bzw. weitere Erkenntnisse zu den steppeninternen Strukturen liefern. Zukünftige palaeogenetische Studien zum Nordschwarzmeerraum im Meso- und Neolithikum oder zu Entwicklungen ab dem zweiten Jahrtausend v. Chr. werden von dem bisher erarbeiteten Datensatz profitieren.

Mithilfe von *DNA capture* könnten die vorliegenden Proben noch genauer auf der Ebene des nukleären Genoms charakterisiert werden und Hinweise liefern, ob die nordpontische Steppe als Korridor für den Einfluss der palaeolithischen nordeurasischen Komponente in den modernen europäischen Genpool gedient haben könnte.

12 Referenzen

- 1000 Genomes: <http://browser.1000genomes.org/> (*abgerufen am 01.10.2014*).
- HaploGrep: <https://haplogrep.uibk.ac.at/> (*abgerufen am 01.10.2014*).
- Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, Gragert L, Babrzadeh F, Gharizadeh B, Luo M, Plummer FA, Kimani J, Carrington M, Middleton D, Rajalingam R, Beksac M, Marsh SG, Maiers M, Guethlein LA, Tavoularis S, Little AM, Green RE, Norman PJ and Parham P (2011) *The shaping of modern human immune systems by multiregional admixture with archaic humans*. *Science* 334: 89-94.
- Adessi C, Matton G, Ayala G, Turcatti G, Mermod JJ, Mayer P and Kawashima E (2000) *Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms*. *Nucleic Acids Res* 28: E87.
- Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA and Kruglyak L (2004) *Population history and natural selection shape patterns of genetic variation in 132 genes*. *PLoS Biol* 2: e286.
- Akey JM, Swanson WJ, Madeoy J, Eberle M and Shriver MD (2006) *TRPV6 exhibits unusual patterns of polymorphism and divergence in worldwide populations*. *Hum Mol Genet* 15: 2106-2113.
- Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, Richmond TA, Middle CM, Rodesch MJ, Packard CJ, Weinstock GM and Gibbs RA (2007) *Direct selection of human genomic loci by microarray hybridization*. *Nat Methods* 4: 903-905.
- Allen JA (1877) *The influence of physical conditions in the genesis of species*. *Radic Rev* 1: 108-140.
- Allentoft M, Schuster SC, Holdaway R, Hale M, McLay E, Oskam C, Gilbert MT, Spencer P, Willerslev E and Bunce M (2009) *Identification of microsatellites from an extinct moa species using high-throughput (454) sequence data*. *Bioinformatics* 24: 195-200.
- Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) *Basic local alignment search tool*. *J Mol Biol* 215: 403-410.
- Ammerman AJ and Cavalli-Sforza LL (1984) *The neolithic transition and the genetics of populations in Europe*. Princeton Univ. Pr., Princeton, NJ.
- Ancans J, Tobin DJ, Hoogduijn MJ, Smit NP, Wakamatsu K and Thody AJ (2001) *Melanosomal pH controls rate of melanogenesis, eumelanin/phaeomelanin ratio and melanosome maturation in melanocytes and melanoma cells*. *Exp Cell Res* 268: 26-35.
- Anderson CN, Ramakrishnan U, Chan YL and Hadly EA (2005) *Serial SimCoal: a population genetics model for data from multiple populations and points in time*. *Bioinformatics* 21: 1733-1734.
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM and Howell N (1999) *Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA*. *Nat Genet* 23: 147.
- Anthony DW and Brown DR (2000) *Eneolithic horse exploitation in the Eurasian steppes: Diet, ritual and riding*. *Antiquity* 74: 75-86.

- Anthony DW (2007) *The Horse, the Wheel, and Language*. Princeton University Press, Princeton.
- Aoki K (1986) *A stochastic model of gene-culture coevolution suggested by the "culture historical hypothesis" for the evolution of adult lactose absorption in humans*. Proc Natl Acad Sci U S A 83: 2929-2933.
- Arribas JC, Herrero AG, Martin-Lomas M, Canada FJ, He S and Withers SG (2000) *Differential mechanism-based labeling and unequivocal activity assignment of the two active sites of intestinal lactase/phlorizin hydrolase*. Eur J Biochem 267: 6996-7005.
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA and Saunders NC (1987) *Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics*. Ann Rev Ecol Syst 18: 489-522.
- Balaresque PL, Ballereau SJ and Jobling MA (2007) *Challenges in human genetic diversity: demographic history and adaptation*. Hum Mol Genet 16 Spec No. 2: R134-139.
- Bandelt HJ, Forster P and Rohlf A (1999) *Median-joining networks for inferring intraspecific phylogenies*. Mol Biol Evol 16: 37-48.
- Bánffy E, Brandt G and Alt KW (2012) *'Early Neolithic' graves of the Carpathian Basin are in fact 6000 years younger—appeal for real interdisciplinarity between archaeology and ancient DNA research*. J Hum Genet 57: 467-469; author reply 470-461.
- Bar-Yosef O (1998) *The Natufian culture in the Levant, threshold to the origins of agriculture*. Evol Anthropol 6: 159-177.
- Barbujani G, Magagni A, Minchi E and Cavalli-Sforza LL (1997) *An apportionment of human DNA diversity*. Proc Natl Acad Sci U S A 94: 4516-4519.
- Basak S, Behar M and Hoffmann A (2012) *Lessons from mathematically modeling the NF-kappaB pathway*. Immunol Rev 246: 221-238.
- Baumann T, Bergmann S, Schmidt-Rose T, Max H, Martin A, Enthaler B, Terstegen L, Schweiger D, Kalbacher H, Wenck H, Jedlitschky G and Jovanovic Z (2014) *Glutathione-conjugated sulfanylalkanols are substrates for ABCC11 and gamma-glutamyl transferase 1: a potential new pathway for the formation of odorant precursors in the apocrine sweat gland*. Exp Dermatol 23: 247-252.
- Beall CM, Cavalleri GL, Deng L, Elston RC, Gao Y, Knight J, Li C, Li JC, Liang Y, McCormack M, Montgomery HE, Pan H, Robbins PA, Shianna KV, Tam SC, Tsering N, Veeramah KR, Wang W, Wangdi P, Weale ME, Xu Y, Xu Z, Yang L, Zaman MJ, Zeng C, Zhang L, Zhang X, Zhaxi P and Zheng YT (2010) *Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders*. Proc Natl Acad Sci U S A 107: 11459-11464.
- Beaumont MA (2010) *Approximate Bayesian Computation in Evolution and Ecology*. Annu Rev Ecol Syst 41: 379-406.
- Beja-Pereira A, Luikart G, England PR, Bradley DG, Jann OC, Bertorelle G, Chamberlain AT, Nunes TP, Metodiev S, Ferrand N and Erhardt G (2003) *Gene-culture coevolution between cattle milk protein genes and human lactase genes*. Nat Genet 35: 311-313.
- Beleza S, Santos AM, McEvoy B, Alves I, Martinho C, Cameron E, Shriver MD, Parra EJ and Rocha J (2013) *The timing of pigmentation lightening in Europeans*. Mol Biol Evol 30: 24-35.

- Belle EM, Ramakrishnan U, Mountain JL and Barbujani G (2006) *Serial coalescent simulations suggest a weak genealogical relationship between Etruscans and modern Tuscans*. Proc Natl Acad Sci U S A 103: 8012-8017.
- Benecke N (1998) *Diachroner Vergleich der Pferdehaltung im Karpatenbecken und in der osteuropäischen Steppe während der vorchristlichen Metallzeiten*. In: Hänsel B and Machnik J (Hrsg), Das Karpatenbecken und die osteuropäische Steppe. Südosteuropa-Schriften, Band 20. VML Verlag Marie Leidorf, Rahden/Westf. 91-98.
- Bennett JW and Klich M (2003) *Mycotoxins*. Clin Microbiol Rev 16: 497-516.
- Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovskiy Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Mullikin JC, Hurler ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R and Smith AJ (2008) *Accurate whole human genome sequencing using reversible terminator chemistry*. Nature 456: 53-59.
- Bergmann C (1847) *Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse*. Göttinger Studien Abt. 1: 595-708.
- Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE and Hirschhorn JN (2004) *Genetic signatures of strong recent positive selection at the lactase gene*. Am J Hum Genet 74: 1111-1120.
- Bierut LJ, Goate AM, Breslau N, Johnson EO, Bertelsen S, Fox L, Agrawal A, Buchholz KK, Grucza R, Hesselbrock V, Kramer J, Kuperman S, Nurnberger J, Porjesz B,

- Saccone NL, Schuckit M, Tischfield J, Wang JC, Foroud T, Rice JP and Edenberg HJ (2011) *ADH1B is associated with alcohol dependence and alcohol consumption in populations of European and African ancestry*. Mol Psychiatry
- Binladen J, Gilbert MT, Bollback JP, Panitz F, Bendixen C, Nielsen R and Willerslev E (2007) *The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing*. PLoS One 2: e197.
- Biswas S and Akey JM (2006) *Genomic insights into positive selection*. Trends Genet 22: 437-446.
- Blum M, Grant DM, McBride W, Heim M and Meyer UA (1990) *Human arylamine N-acetyltransferase genes: isolation, chromosomal localization, and functional expression*. DNA Cell Biol 9: 193-203.
- Bödding M (2005) *Voltage-dependent changes of TRPV6-mediated Ca²⁺ currents*. J Biol Chem 280: 7022-7029.
- Bolbrinker J, Seeberg S, Schostak M, Kempkensteffen C, Baelde H, de Heer E and Kreutz R (2012) *CYP3A5 genotype-phenotype analysis in the human kidney reveals a strong site-specific expression of CYP3A5 in the proximal tubule in carriers of the CYP3A5*1 allele*. Drug Metab Dispos 40: 639-641.
- Bollongino R (2006) *Die Herkunft der Hausrinder in Europa. Eine aDNA-Studie an neolithischen Knochenfunden*. Habelt, Bonn. UPA Band 130 (zugleich Dissertation Johannes Gutenberg-Universität Mainz 2005).
- Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkosova Z, Powell A and Burger J (2013) *2000 years of parallel societies in Stone Age Central Europe*. Science 342: 479-481.
- Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, Schein RM and Sibbald WJ (1992) *Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine*. Chest 101: 1644-1655.
- Borinskaya S, Kal'ina N, Marusin A, Faskhutdinova G, Morozova I, Kutuev I, Koshechkin V, Khusnutdinova E, Stepanov V, Puzyrev V, Yankovsky N and Rogaev E (2009) *Distribution of the alcohol dehydrogenase ADH1B*47His allele in Eurasia*. Am J Hum Genet 84: 89-92; author reply 92-84.
- Borràs E, Coutelle C, Rosell A, Fernandez-Muixi F, Broch M, Crosas B, Hjelmqvist L, Lorenzo A, Gutierrez C, Santos M, Szczepanek M, Heilig M, Quattrocchi P, Farres J, Vidal F, Richart C, Mach T, Bogdal J, Jornvall H, Seitz HK, Couzigou P and Pares X (2000) *Genetic polymorphism of alcohol dehydrogenase in europeans: the ADH2*2 allele decreases the risk for alcoholism and is associated with ADH3*1*. Hepatology 31: 984-989.
- Bortfeld M, Rius M, König J, Herold-Mende C, Nies AT and Keppler D (2006) *Human multidrug resistance protein 8 (MRP8/ABCC11), an apical efflux pump for steroid sulfates, is an axonal protein of the CNS and peripheral nervous system*. Neuroscience 137: 1247-1257.
- Botchkarev VA and Fessing MY (2005) *Edar signaling in the control of hair follicle development*. J Investig Dermatol Symp Proc 10: 247-251.
- Bouakaze C, Keyser C, Crubezy E, Montagnon D and Ludes B (2009) *Pigment phenotype and biogeographical ancestry from ancient skeletal remains:*

- inferences from multiplexed autosomal SNP analysis*. Int J Legal Med 123: 315-325.
- Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind CJ, Lueth F, Terberger T, Hiller J, Matsumura S, Forster P and Burger J (2009) *Genetic discontinuity between local hunter-gatherers and central Europe's first farmers*. Science 326: 137-140.
- Brandt G, Haak W, Adler CJ, Roth C, Szecsenyi-Nagy A, Karimnia S, Moller-Rieker S, Meller H, Ganslmeier R, Friederich S, Dresely V, Nicklisch N, Pickrell JK, Sirocko F, Reich D, Cooper A and Alt KW (2013) *Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity*. Science 342: 257-261.
- Branicki W, Brudnik U and Wojas-Pelc A (2009) *Interactions between HERC2, OCA2 and MC1R may influence human pigmentation phenotype*. Ann Hum Genet 73: 160-170.
- Breasted JH (1914) *Earliest Man, the Orient, Greece, and Rome*. In: Robinson JH and Breasted JH (Hrsg), *Outlines of European History - Part 1*. Atheneum Press/Ginn & Co, Boston. 1-314.
- Breyer J and Jacobson HR (1990) *Molecular mechanisms of diuretic agents*. Annu Rev Med 41: 265-275.
- Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prufer K, Meyer M, Krause J, Ronan MT, Lachmann M and Paabo S (2007) *Patterns of damage in genomic DNA sequences from a Neandertal*. Proc Natl Acad Sci U S A 104: 14616-14621.
- Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Lalueza-Fox C, Rudan P, Brajkovic D, Kucan Z, Gusic I, Schmitz R, Doronichev VB, Golovanova LV, de la Rasilla M, Fortea J, Rosas A and Pääbo S (2009) *Targeted retrieval and analysis of five Neandertal mtDNA genomes*. Science 325: 318-321.
- Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, Richards SM, Sarkissian CD, Ganslmeier R, Friederich S, Dresely V, van Oven M, Kenyon R, Van der Hoek MB, Korfach J, Luong K, Ho SY, Quintana-Murci L, Behar DM, Meller H, Alt KW, Cooper A, Adhikarla S, Ganesh Prasad AK, Pitchappan R, Varatharajan Santhakumari A, Balanovska E, Balanovsky O, Bertranpetit J, Comas D, Martinez-Cruz B, Mele M, Clarke AC, Matisoo-Smith EA, Dulik MC, Gaieski JB, Owings AC, Schurr TG, Vilar MG, Hobbs A, Soodyall H, Javed A, Parida L, Platt DE, Royyuru AK, Jin L, Li S, Kaplan ME, Merchant NC, John Mitchell R, Renfrew C, Lacerda DR, Santos FR, Soria Hernanz DF, Spencer Wells R, Swamikrishnan P, Tyler-Smith C, Paulo Vieira P and Ziegler JS (2013) *Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans*. Nat Commun 4: 1764.
- Bryk J, Hardouin E, Pugach I, Hughes D, Strotmann R, Stoneking M and Myles S (2008) *Positive selection in East Asians for an EDAR allele that enhances NF-kappaB activation*. PLoS ONE 3: e2209.
- Buhler S and Sanchez-Mazas A (2011) *HLA DNA sequence variation among human populations: molecular signatures of demographic and selective events*. PLoS One 6: e14643.
- Burbano HA, Hodges E, Green RE, Briggs AW, Krause J, Meyer M, Good JM, Maricic T, Johnson PL, Xuan Z, Rooks M, Bhattacharjee A, Brizuela L, Albert FW, de la Rasilla M, Fortea J, Rosas A, Lachmann M, Hannon GJ and Pääbo S (2010)

- Targeted investigation of the Neandertal genome by array-based sequence capture.* Science 328: 723-725.
- Burger J, Hummel S, Hermann B and Henke W (1999) *DNA preservation: a microsatellite-DNA study on ancient skeletal remains.* Electrophoresis 20: 1722-1728.
- Burger J (2007) *Palaeogenetik.* In: Wagner G (Hrsg), Einführung in die Archäometrie. Springer, Berlin, Heidelberg. 279-298.
- Burger J, Kirchner M, Bramanti B, Haak W and Thomas MG (2007) *Absence of the lactase-persistence-associated allele in early Neolithic Europeans.* Proc Natl Acad Sci U S A 104: 3736-3741.
- Calafell F, Underhill P, Tolun A, Angelicheva D and Kalaydjieva L (1996) *From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks.* Ann Hum Genet 60: 35-49.
- Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, Piouffre L, Bodmer J, Bodmer WF, Bonne-Tamir B, Cambon-Thomsen A, Chen Z, Chu J, Carcassi C, Contu L, Du R, Excoffier L, Ferrara GB, Friedlaender JS, Groot H, Gurwitz D, Jenkins T, Herrera RJ, Huang X, Kidd J, Kidd KK, Langaney A, Lin AA, Mehdi SQ, Parham P, Piazza A, Pistillo MP, Qian Y, Shu Q, Xu J, Zhu S, Weber JL, Greely HT, Feldman MW, Thomas G, Dausset J and Cavalli-Sforza LL (2002) *A human genome diversity cell line panel.* Science 296: 261-262.
- Cano RJ, Poinar HN, Pieniazek NJ, Acra A and Poinar GO, Jr. (1993) *Amplification and sequencing of DNA from a 120-135-million-year-old weevil.* Nature 363: 536-538.
- Cao W, Strnatka D, McQueen CA, Hunter RJ and Erickson RP (2010) *N-acetyltransferase 2 activity and folate levels.* Life Sci 86: 103-106.
- Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ and Nickerson DA (2005) *Genomic regions exhibiting positive selection identified from dense genotype data.* Genome Res 15: 1553-1565.
- Carrigan MA, Uryasev O, Davis RP, Zhai L, Hurley TD and Benner SA (2012) *The natural history of class I primate alcohol dehydrogenases includes gene duplication, gene loss, and gene conversion.* PLoS One 7: e41175.
- Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, Kaslow R, Buchbinder S, Hoots K and O'Brien SJ (1999) *HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage.* Science 283: 1748-1752.
- Cavalli-Sforza LL, Menozzi P and Piazza A (1993) *Demic expansions and human evolution.* Science 259: 639-646.
- Chang SH, Jobling S, Brennan K and Headon DJ (2009) *Enhanced Edar signalling has pleiotropic effects on craniofacial and cutaneous glands.* PLoS One 4: e7591.
- Charlesworth B, Morgan MT and Charlesworth D (1993) *The effect of deleterious mutations on neutral molecular variation.* Genetics 134: 1289-1303.
- Charlesworth D (2006) *Balancing selection and its effects on sequences in nearby genome regions.* PLoS Genet 2: e64.
- Chen CC, Lu RB, Chen YC, Wang MF, Chang YC, Li TK and Yin SJ (1999) *Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism.* Am J Hum Genet 65: 795-807.

- Chen ZS, Guo Y, Belinsky MG, Kotova E and Kruh GD (2005) *Transport of bile acids, sulfated steroids, estradiol 17-beta-D-glucuronide, and leukotriene C4 by human multidrug resistance protein 8 (ABCC11)*. *Mol Pharmacol* 67: 545-557.
- Chernykh E (2008) *The "Steppe Belt" of stockbreeding cultures in Eurasia during the Early Metal Age*. *Trabajos de Prehistoria* 65: 73-93.
- Childe VG (1936) *Man Makes Himself*. Watts & Co, London.
- Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW and McCabe AM (2009) *The Last Glacial Maximum*. *Science* 325: 710-714.
- Clarke A (1987) *Hypohidrotic ectodermal dysplasia*. *J Med Genet* 24: 659-663.
- Cohen J (2002) *The immunopathogenesis of sepsis*. *Nature* 420: 885-891.
- Cook AL, Chen W, Thurber AE, Smit DJ, Smith AG, Bladen TG, Brown DL, Duffy DL, Pastorino L, Bianchi-Scarra G, Leonard JH, Stow JL and Sturm RA (2009) *Analysis of cultured human melanocytes based on polymorphisms within the SLC45A2/MATP, SLC24A5/NCKX5, and OCA2/P loci*. *J Invest Dermatol* 129: 392-405.
- Cooke GS and Hill AV (2001) *Genetics of susceptibility to human infectious disease*. *Nat Rev Genet* 2: 967-977.
- Coop G and Griffiths RC (2004) *Ancestral inference on gene trees under selection*. *Theor Popul Biol* 66: 219-232.
- Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, Myers RM, Cavalli-Sforza LL, Feldman MW and Pritchard JK (2009) *The role of geography in human adaptation*. *PLoS Genet* 5: e1000500.
- Cooper A and Poinar HN (2000) *Ancient DNA: Do It Right or Not at All*. *Science* 289: 1139.
- Copley MS, Berstan R, Dudd SN, Docherty G, Mukherjee AJ, Straker V, Payne S and Evershed RP (2003) *Direct chemical evidence for widespread dairying in prehistoric Britain*. *Proc Natl Acad Sci U S A* 100: 1524-1529.
- Cordain L, Eaton SB, Sebastian A, Mann N, Lindeberg S, Watkins BA, O'Keefe JH and Brand-Miller J (2005) *Origins and evolution of the Western diet: health implications for the 21st century*. *Am J Clin Nutr* 81: 341-354.
- Cordon-Cardo C, O'Brien JP, Casals D, Rittman-Grauer L, Biedler JL, Melamed MR and Bertino JR (1989) *Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood-brain barrier sites*. *Proc Natl Acad Sci U S A* 86: 695-698.
- Costin GE, Valencia JC, Vieira WD, Lamoreux ML and Hearing VJ (2003) *Tyrosinase processing and intracellular trafficking is disrupted in mouse primary melanocytes carrying the underwhite (uw) mutation. A model for oculocutaneous albinism (OCA) type 4*. *J Cell Sci* 116: 3203-3212.
- Crabb DW, Matsumoto M, Chang D and You M (2004) *Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology*. *Proc Nutr Soc* 63: 49-63.
- Croop JM, Gros P and Housman DE (1988) *Genetics of multidrug resistance*. *J Clin Invest* 81: 1303-1309.
- Cubillos-Rojas M, Amair-Pinedo F, Peiro-Jordan R, Bartrons R, Ventura F and Rosa JL (2014) *The E3 Ubiquitin Protein Ligase HERC2 Modulates Activity of Tumor Protein p53 by Regulating its Oligomerization*. *J Biol Chem*

- Cui J, Bian JS, Kagan A and McDonald TV (2002) *CaT1 contributes to the store-operated calcium current in Jurkat T-lymphocytes*. J Biol Chem 277: 47175-47183.
- Currat M and Excoffier L (2005) *The effect of the Neolithic expansion on European molecular diversity*. Proc Biol Sci 272: 679-688.
- Currat M, Excoffier L, Maddison W, Otto SP, Ray N, Whitlock MC and Yeaman S (2006) *Comment on "Ongoing adaptive evolution of ASPM, a brain size determinant in Homo sapiens" and "Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans"*. Science 313: 172; author reply 172.
- Currat M and Excoffier L (2011) *Strong reproductive isolation between humans and Neanderthals inferred from observed patterns of introgression*. Proc Natl Acad Sci U S A 108: 15129-15134.
- Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paabo S, Arsuaga JL and Meyer M (2013) *Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments*. Proc Natl Acad Sci U S A 110: 15758-15763.
- DeGiorgio M, Lohmueller KE and Nielsen R (2014) *A model-based approach for identifying signatures of ancient balancing selection in genetic data*. PLoS Genet 10: e1004561.
- DeLuca HF (2004) *Overview of general physiologic features and functions of vitamin D*. Am J Clin Nutr 80: 1689S-1696S.
- Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J, Daley GQ, Eggan K, Hochedlinger K, Thomson J, Wang W, Gao Y and Zhang K (2009) *Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming*. Nat Biotechnol 27: 353-360.
- Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, Koshelev S, Zaporozhchenko V, Gronenborn D, Moiseyev V, Kolpakov E, Shumkin V, Alt KW, Balanovska E, Cooper A and Haak W (2013) *Ancient DNA reveals prehistoric gene-flow from siberia in the complex human population history of North East Europe*. PLoS Genet 9: e1003296.
- Dergačev VA (1998) *Kulturelle und historische Entwicklungen im Raum zwischen Karpaten und Dnepr. Zu den Beziehungen zwischen frühen Gesellschaften im nördlichen Südost- und Osteuropa*. In: Hänsel B and Machnik J (Hrsg), Das Karpatenbecken und die osteuropäische Steppe. Südosteuropa-Schriften, Band 20. VML Verlag Marie Leidorf, Rahden/Westf. 27-64.
- Dillenberger S (2013) *Entwicklung einer Y-chromosomalen Multiplex-PCR zur palaeogenetischen Analyse humaner Populationen*. Magisterarbeit an der Johannes Gutenberg-Universität Mainz.
- Diller KC, Gilbert WA and Kocher TD (2002) *Selective sweeps in the human genome: a starting point for identifying genetic differences between modern humans and chimpanzees*. Mol Biol Evol 19: 2342-2345.
- Dressman D, Yan H, Traverso G, Kinzler KW and Vogelstein B (2003) *Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations*. Proc Natl Acad Sci U S A 100: 8817-8822.

- Duncan RL, Grogan WM, Kramer LB and Watlington CO (1988) *Corticosterone's metabolite is an agonist for Na⁺ transport stimulation in A6 cells*. *Am J Physiol* 255: F736-748.
- Eap CB, Bochud M, Elston RC, Bovet P, Maillard MP, Nussberger J, Schild L, Shamlaye C and Burnier M (2007) *CYP3A5 and ABCB1 genes influence blood pressure and response to treatment, and their effect is modified by salt*. *Hypertension* 49: 1007-1014.
- Ecelbarger CA, Knepper MA and Verbalis JG (2001) *Increased abundance of distal sodium transporters in rat kidney during vasopressin escape*. *J Am Soc Nephrol* 12: 207-217.
- Edenberg HJ (2007) *The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants*. *Alcohol Res Health* 30: 5-13.
- Edmonds CA, Lillie AS and Cavalli-Sforza LL (2004) *Mutations arising in the wave front of an expanding population*. *Proc Natl Acad Sci U S A* 101: 975-979.
- Edwards CJ, Bollongino R, Scheu A, Chamberlain A, Tresset A, Vigne J-D, Baird JF, Larson G, Ho SYW, Heupink TH, Shapiro B, Freeman AR, Thomas MG, Arbogast R-M, Arndt B, Bartosiewicz L, Benecke N, Budja M, Chaix L, Choyke AM, Coqueugniot E, Döhle H-J, Göldner H, Hartz S, Helmer D, Herzig B, Hongo H, Mashkour M, Özdoğan M, Pucher E, Roth G, Schade-Lindig S, Schmölcke U, Schulting RJ, Stephan E, Uerpmann H-P, Vörös I, Voytek B, Bradley DG and Burger J (2007) *Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs*. *Proc Biol Sci* 274: 1377-1385.
- Edwards CJ, Ginja C, Kantanen J, Perez-Pardal L, Tresset A, Stock F, Gama LT, Penedo MC, Bradley DG, Lenstra JA and Nijman IJ (2011) *Dual origins of dairy cattle farming--evidence from a comprehensive survey of European Y-chromosomal variation*. *PLoS ONE* 6: e15922.
- Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, Kjaer KW and Hansen L (2008) *Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression*. *Hum Genet* 123: 177-187.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J and Turner S (2009) *Real-time DNA sequencing from single polymerase molecules*. *Science* 323: 133-138.
- Elhaik E (2012) *Empirical distributions of F(ST) from large-scale human polymorphism data*. *PLoS One* 7: e49837.
- Ellison DH, Velazquez H and Wright FS (1987) *Thiazide-sensitive sodium chloride cotransport in early distal tubule*. *Am J Physiol* 253: F546-554.
- Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP and Pääbo S (2002) *Molecular evolution of FOXP2, a gene involved in speech and language*. *Nature* 418: 869-872.

- Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L and Jarvela I (2002) *Identification of a variant associated with adult-type hypolactasia*. *Nat Genet* 30: 233-237.
- Enattah NS, Jensen TG, Nielsen M, Lewinski R, Kuokkanen M, Rasinpera H, El-Shanti H, Seo JK, Alifrangis M, Khalil IF, Natah A, Ali A, Natah S, Comas D, Mehdi SQ, Groop L, Vestergaard EM, Imtiaz F, Rashed MS, Meyer B, Troelsen J and Peltonen L (2008) *Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture*. *Am J Hum Genet* 82: 57-72.
- Endicott P, Gilbert MT, Stringer C, Lalueza-Fox C, Willerslev E, Hansen AJ and Cooper A (2003) *The genetic origins of the Andaman Islanders*. *Am J Hum Genet* 72: 178-184.
- Ermini L, Olivieri C, Rizzi E, Corti G, Bonnal R, Soares P, Luciani S, Marota I, De Bellis G, Richards MB and Rollo F (2008) *Complete mitochondrial genome sequence of the Tyrolean Iceman*. *Curr Biol* 18: 1687-1693.
- Evershed RP, Payne S, Sherratt AG, Copley MS, Coolidge J, Urem-Kotsu D, Kotsakis K, Ozdogan M, Ozdogan AE, Nieuwenhuys O, Akkermans PM, Bailey D, Andeescu RR, Campbell S, Farid S, Hodder I, Yalman N, Ozbasaran M, Bicaçci E, Garfinkel Y, Levy T and Burton MM (2008) *Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding*. *Nature* 455: 528-531.
- Excoffier L, Smouse PE and Quattro JM (1992) *Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data*. *Genetics* 131: 479-491.
- Excoffier L and Yang Z (1999) *Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees*. *Mol Biol Evol* 16: 1357-1368.
- Excoffier L, Novembre J and Schneider S (2000) *SIMCOAL: a general coalescent program for the simulation of molecular data in interconnected populations with arbitrary demography*. *J Hered* 91: 506-509.
- Excoffier L and Hamilton G (2003) *Comment on "Genetic structure of human populations"*. *Science* 300: 1877; author reply 1877.
- Excoffier L and Lischer HE (2010) *Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows*. *Mol Ecol Resour* 10: 564-567.
- Excoffier L and Foll M (2011) *fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios*. *Bioinformatics* 27: 1332-1334.
- Fay JC and Wu CI (2000) *Hitchhiking under positive Darwinian selection*. *Genetics* 155: 1405-1413.
- Fehren-Schmitz L, Reindel M, Cagigao ET, Hummel S and Herrmann B (2010) *Pre-Columbian population dynamics in coastal southern Peru: A diachronic investigation of mtDNA patterns in the Palpa region by ancient DNA analysis*. *Am J Phys Anthropol* 141: 208-221.
- Fehren-Schmitz L (2012) *Population dynamics, cultural evolution and climate change in pre-Columbian western South America*. In: Kaiser E, Burger J and Schier W (Hrsg), *Population Dynamics in Prehistory and Early History, New Approaches Using Stable Isotopes and Genetics*. De Gruyter, Berlin, Boston. 55-73.

- Ferguson LR (2002) *Natural and human-made mutagens and carcinogens in the human diet*. Toxicology 181-182: 79-82.
- Fischer H, Koenig U, Eckhart L and Tschachler E (2002) *Human caspase 12 has acquired deleterious mutations*. Biochem Biophys Res Commun 293: 722-726.
- Flatz G and Rotthauwe HW (1973) *Lactose nutrition and natural selection*. Lancet 2: 76-77.
- Folk GE, Jr. and Semken HA, Jr. (1991) *The evolution of sweat glands*. Int J Biometeorol 35: 180-186.
- Francois O, Currat M, Ray N, Han E, Excoffier L and Novembre J (2010) *Principal component analysis under population genetic models of range expansion and admixture*. Mol Biol Evol 27: 1257-1268.
- Friedman PA and Gesek FA (1993) *Calcium transport in renal epithelial cells*. Am J Physiol 264: F181-198.
- Frudakis T, Thomas M, Gaskin Z, Venkateswarlu K, Chandra KS, Ginjupalli S, Gunturi S, Natrajan S, Ponnuswamy VK and Ponnuswamy KN (2003) *Sequences associated with human iris pigmentation*. Genetics 165: 2071-2083.
- Fu Q, Mitnik A, Johnson PL, Bos K, Lari M, Bollongino R, Sun C, Giemisch L, Schmitz R, Burger J, Ronchitelli AM, Martini F, Cremonesi RG, Svoboda J, Bauer P, Caramelli D, Castellano S, Reich D, Pääbo S and Krause J (2013) *A revised timescale for human evolution based on ancient mitochondrial genomes*. Curr Biol 23: 553-559.
- Fu YX (1997) *Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection*. Genetics 147: 915-925.
- Fujimoto A, Kimura R, Ohashi J, Omi K, Yuliwulandari R, Batubara L, Mustofa MS, Samakkarn U, Settheetham-Ishida W, Ishida T, Morishita Y, Furusawa T, Nakazawa M, Ohtsuka R and Tokunaga K (2008a) *A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness*. Hum Mol Genet 17: 835-843.
- Fujimoto A, Ohashi J, Nishida N, Miyagawa T, Morishita Y, Tsunoda T, Kimura R and Tokunaga K (2008b) *A replication study confirmed the EDAR gene to be a major contributor to population differentiation regarding head hair thickness in Asia*. Hum Genet 124: 179-185.
- Fuller BB, Spaulding DT and Smith DR (2001) *Regulation of the catalytic activity of preexisting tyrosinase in black and Caucasian human melanocyte cell cultures*. Exp Cell Res 262: 197-208.
- Fung KL, Pan J, Ohnuma S, Lund PE, Pixley JN, Kimchi-Sarfaty C, Ambudkar SV and Gottesman MM (2014) *MDR1 synonymous polymorphisms alter transporter specificity and protein stability in a stable epithelial monolayer*. Cancer Res 74: 598-608.
- Gallego Romero I, Basu Mallick C, Liebert A, Crivellaro F, Chaubey G, Itan Y, Metspalu M, Easwarkhanth M, Pitchappan R, Villems R, Reich D, Singh L, Thangaraj K, Thomas MG, Swallow DM, Mirazon Lahr M and Kivisild T (2011) *Herders of Indian and European Cattle Share Their Predominant Allele for Lactase Persistence*. Mol Biol Evol
- García-Garcerà M, Gigli E, Sanchez-Quinto F, Ramirez O, Calafell F, Civit S and Lalueza-Fox C (2011) *Fragmentation of contaminant and endogenous DNA in ancient*

- samples determined by shotgun sequencing; prospects for human palaeogenomics*. PLoS One 6: e24161.
- Geörg C (2012) *Paläopopulationsgenetik von Schwein und Schaf in Südosteuropa und Transkaukasien*. VML Verlag Marie Leidorf, Rahden/Westf. Menschen - Kulturen - Traditionen, Studien aus den Forschungsclustern des Deutschen Archäologischen Instituts Band 9 (zugleich Dissertation Johannes Gutenberg-Universität Mainz 2011).
- Gerbault P, Moret C, Currat M and Sanchez-Mazas A (2009) *Impact of selection and demography on the diffusion of lactase persistence*. PLoS One 4: e6369.
- Gerbault P, Liebert A, Itan Y, Powell A, Currat M, Burger J, Swallow DM and Thomas MG (2011) *Evolution of lactase persistence: an example of human niche construction*. Philos Trans R Soc Lond B Biol Sci 366: 863-877.
- Gerling C, Heyd V, Pike AC, Bánffy E, Dani J, Köhler K, Kulcár G, Kaiser E and Schier W (2012) *Identifying kurgan graves in Eastern Hungary: A burial mound in the light of strontium and oxygen isotope analysis*. In: Kaiser E, Burger J and Schier W (Hrsg), *Population Dynamics in Prehistory and Early History, New Approaches Using Stable Isotopes and Genetics*. De Gruyter, Berlin, Boston. 165-176.
- Gerling C (im Druck) *Prähistorische Mobilität und Ernährung im westlichen Eurasien. Stabile Isotopenanalysen an Mensch und Tier zwischen 3500 und 300 v.Chr.* Dissertationsresümee, Archäologisches Nachrichtenblatt.
- Gesek FA and Friedman PA (1995) *Sodium entry mechanisms in distal convoluted tubule cells*. Am J Physiol 268: F89-98.
- Ghosh SS, Basu AK, Ghosh S, Hagley R, Kramer L, Schuetz J, Grogan WM, Guzelian P and Watlington CO (1995) *Renal and hepatic family 3A cytochromes P450 (CYP3A) in spontaneously hypertensive rats*. Biochem Pharmacol 50: 49-54.
- Giamarellos-Bourboulis EJ, van de Veerdonk FL, Mouktaroudi M, Raftogiannis M, Antonopoulou A, Joosten LA, Pickkers P, Savva A, Georgitsi M, van der Meer JW and Netea MG (2011) *Inhibition of caspase-1 activation in Gram-negative sepsis and experimental endotoxemia*. Crit Care 15: R27.
- Giebisch G and Windhager EE (1964) *Renal Tubular Transfer of Sodium, Chloride and Potassium*. Am J Med 36: 643-669.
- Gilbert MT, Hansen AJ, Willerslev E, Rudbeck L, Barnes I, Lynnerup N and Cooper A (2003a) *Characterization of genetic miscoding lesions caused by postmortem damage*. Am J Hum Genet 72: 48-61.
- Gilbert MT, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N and Cooper A (2003b) *Distribution patterns of postmortem damage in human mitochondrial DNA*. Am J Hum Genet 72: 32-47.
- Gilbert MT, Kivisild T, Gronnow B, Andersen PK, Metspalu E, Reidla M, Tamm E, Axelsson E, Gotherstrom A, Campos PF, Rasmussen M, Metspalu M, Higham TF, Schwenninger JL, Nathan R, De Hoog CJ, Koch A, Moller LN, Andreasen C, Meldgaard M, Villems R, Bendixen C and Willerslev E (2008) *Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland*. Science 320: 1787-1789.
- Gimbutas M (1956) *The Prehistory of Eastern Europe, Part I: Mesolithic, Neolithic and Copper Age Cultures in Russia and the Baltic Area*. Peabody Museum, Cambridge, Mass.

- Gimbutas M (1994) *Das Ende Alteuropas: Der Einfall von Steppennomaden aus Südrussland und die Indogermanisierung Mitteleuropas*. Institut für Sprachwissenschaft der Universität Innsbruck ; Archaeolingua, Innsbruck; Budapest.
- Givens RC, Lin YS, Dowling AL, Thummel KE, Lamba JK, Schuetz EG, Stewart PW and Watkins PB (2003) *CYP3A5 genotype predicts renal CYP3A activity and blood pressure in healthy adults*. *J Appl Physiol* 95: 1297-1300.
- Glenn TC (2011) *Field guide to next-generation DNA sequencers*. *Mol Ecol Resour* 11: 759-769.
- Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, Gabriel S, Jaffe DB, Lander ES and Nusbaum C (2009) *Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing*. *Nat Biotechnol* 27: 182-189.
- Goedde HW, Agarwal DP, Fritze G, Meier-Tackmann D, Singh S, Beckmann G, Bhatia K, Chen LZ, Fang B, Lisker R and et al. (1992) *Distribution of ADH2 and ALDH2 genotypes in different populations*. *Hum Genet* 88: 344-346.
- Goto M, Katayama KI, Shirakawa F and Tanaka I (1999) *Involvement of NF-kappaB p50/p65 heterodimer in activation of the human pro-interleukin-1beta gene at two subregions of the upstream enhancer element*. *Cytokine* 11: 16-28.
- Graefen A (2010) Development of a multiplex PCR for the assessment of ancient DNA preservation levels in prehistoric human remains. Magisterarbeit an der Johannes Gutenberg Universität Mainz.
- Graf J, Hodgson R and van Daal A (2005) *Single nucleotide polymorphisms in the MATP gene are associated with normal human pigmentation variation*. *Hum Mutat* 25: 278-284.
- Graffelman J (2013) *HardyWeinberg: Graphical tests for Hardy-Weinberg equilibrium*. R package version 1.5.2, <http://CRAN.R-project.org/package=HardyWeinberg>.
- Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, Du L, Egholm M, Rothberg JM, Paunovic M and Pääbo S (2006) *Analysis of one million base pairs of Neanderthal DNA*. *Nature* 444: 330-336.
- Green RE, Malaspinas AS, Krause J, Briggs AW, Johnson PL, Uhler C, Meyer M, Good JM, Maricic T, Stenzel U, Prufer K, Siebauer M, Burbano HA, Ronan M, Rothberg JM, Egholm M, Rudan P, Brajkovic D, Kucan Z, Gusic I, Wikstrom M, Laakkonen L, Kelso J, Slatkin M and Pääbo S (2008) *A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing*. *Cell* 134: 416-426.
- Green RE, Briggs AW, Krause J, Prufer K, Burbano HA, Siebauer M, Lachmann M and Paabo S (2009) *The Neandertal genome and ancient DNA authenticity*. *Embo J* 28: 2494-2502.
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T, Alkan C, Prufer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Hober B, Hoffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Z, Gusic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PL, Eichler EE, Falush D, Birney E,

- Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D and Pääbo S (2010) *A draft sequence of the Neandertal genome*. Science 328: 710-722.
- Grogan WM, Phillips VM, Schuetz EG, Guzelian PS and Watlington CO (1990) *Corticosterone 6 beta-hydroxylase in A6 epithelia: a steroid-inducible cytochrome P-450*. Am J Physiol 258: C480-488.
- Gronenborn D (1999) *A Variation on a Basic Theme: The Transition to Farming in Southern Central Europe*. J World Prehist 13: 123-210.
- Guba Z, Hadadi E, Major A, Furka T, Juhasz E, Koos J, Nagy K and Zeke T (2011) *HVS-I polymorphism screening of ancient human mitochondrial DNA provides evidence for N9a discontinuity and East Asian haplogroups in the Neolithic Hungary*. J Hum Genet
- Guilaine J and Manen C (2007) *From Mesolithic to Early Neolithic in the western Mediterranean*. In: Whittle A and Cummings V (Hrsg), Going Over: The Mesolithic-Neolithic Transition in North-West Europe. Proceedings of the British Academy, 144. Oxford University Press, Oxford. 21-51.
- Guo SW and Thompson EA (1992) *Performing the exact test of Hardy-Weinberg proportion for multiple alleles*. Biometrics 48: 361-372.
- Guo Y, Kotova E, Chen ZS, Lee K, Hopper-Borge E, Belinsky MG and Kruh GD (2003) *MRP8, ATP-binding cassette C11 (ABCC11), is a cyclic nucleotide efflux pump and a resistance factor for fluoropyrimidines 2',3'-dideoxycytidine and 9'-(2'-phosphonylmethoxyethyl)adenine*. J Biol Chem 278: 29509-29514.
- Gupta RP, Hollis BW, Patel SB, Patrick KS and Bell NH (2004) *CYP3A4 is a human microsomal vitamin D 25-hydroxylase*. J Bone Miner Res 19: 680-688.
- Gupta RP, He YA, Patrick KS, Halpert JR and Bell NH (2005) *CYP3A4 is a vitamin D-24- and 25-hydroxylase: analysis of structure function by site-directed mutagenesis*. J Clin Endocrinol Metab 90: 1210-1219.
- Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, Vilems R, Renfrew C, Gronenborn D, Alt KW and Burger J (2005) *Ancient DNA from the first European farmers in 7500-year-old Neolithic sites*. Science 310: 1016-1018.
- Haak W (2006) *Populationsgenetik der ersten Bauern Mitteleuropas: Eine aDNA-Studie an neolithischem Skelettmaterial*. Dissertation an der Johannes Gutenberg-Universität Mainz.
- Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V, Hawkesworth C, Pike AW, Meller H and Alt KW (2008) *Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age*. Proc Natl Acad Sci U S A 105: 18226-18231.
- Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, Der Sarkissian CS, Brandt G, Schwarz C, Nicklisch N, Dresely V, Fritsch B, Balanovska E, Vilems R, Meller H, Alt KW and Cooper A (2010) *Ancient DNA from European early neolithic farmers reveals their near eastern affinities*. PLoS Biol 8: e1000536.
- Haehner BD, Gorski JC, Vandenbranden M, Wrighton SA, Janardan SK, Watkins PB and Hall SD (1996) *Bimodal distribution of renal cytochrome P450 3A activity in humans*. Mol Pharmacol 50: 52-59.
- Haensch S, Bianucci R, Signoli M, Rajerison M, Schultz M, Kacki S, Vermunt M, Weston DA, Hurst D, Achtman M, Carniel E and Bramanti B (2010) *Distinct clones of Yersinia pestis caused the black death*. PLoS Pathogens 6: e1001134.

- Hamblin MT and Di Rienzo A (2000) *Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus*. Am J Hum Genet 66: 1669-1679.
- Hamblin MT, Thompson EE and Di Rienzo A (2002) *Complex signatures of natural selection at the Duffy blood group locus*. Am J Hum Genet 70: 369-383.
- Hamilton MB (2009) *Population Genetics*. Wiley-Blackwell, Hoboken, NJ.
- Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, Hankinson SE, Hu FB, Duffy DL, Zhao ZZ, Martin NG, Montgomery GW, Hayward NK, Thomas G, Hoover RN, Chanock S and Hunter DJ (2008) *A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation*. PLoS Genet 4: e1000074.
- Han Y, Gu S, Oota H, Osier MV, Pakstis AJ, Speed WC, Kidd JR and Kidd KK (2007) *Evidence of positive selection on a class I ADH locus*. Am J Hum Genet 80: 441-456.
- Handt O, Hoss M, Krings M and Pääbo S (1994) *Ancient DNA: methodological challenges*. Experientia 50: 524-529.
- Handt O, Krings M, Ward RH and Paabo S (1996) *The retrieval of ancient human DNA sequences*. Am J Hum Genet 59: 368-376.
- Hardenbol P, Yu F, Belmont J, Mackenzie J, Bruckner C, Brundage T, Boudreau A, Chow S, Eberle J, Erbilgin A, Falkowski M, Fitzgerald R, Ghose S, Iartchouk O, Jain M, Karlin-Neumann G, Lu X, Miao X, Moore B, Moorhead M, Namsaraev E, Pasternak S, Prakash E, Tran K, Wang Z, Jones HB, Davis RW, Willis TD and Gibbs RA (2005) *Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay*. Genome Res 15: 269-275.
- Hardy GH (1908) *Mendelian Proportions in a Mixed Population*. Science 28: 49-50.
- Harker M, Carvell AM, Marti VP, Riazanskaia S, Kelso H, Taylor D, Grimshaw S, Arnold DS, Zillmer R, Shaw J, Kirk JM, Alcasid ZM, Gonzales-Tanon S, Chan GP, Rosing EA and Smith AM (2014) *Functional characterisation of a SNP in the ABCC11 allele - effects on axillary skin metabolism, odour generation and associated behaviours*. J Dermatol Sci 73: 23-30.
- Harris EE and Meyer D (2006) *The molecular signature of selection underlying human adaptations*. Am J Phys Anthropol Suppl 43: 89-130.
- Häusler A (1998) *Struktur und Evolution der Bestattungsriten zwischen Wolga und Karpatenbecken vom Äneolithikum bis zur frühen Bronzezeit. Ein diachroner Vergleich*. In: Hänsel B and Machnik J (Hrsg), Das Karpatenbecken und die osteuropäische Steppe. Südosteuropa-Schriften, Band 20. VML Verlag Marie Leidorf, Rahden/Westf. 135-161.
- Häusler A (2003) *Nomaden, Indogermanen, Invasionen. Zur Entstehung eines Mythos*. Orientwissenschaftliches Zentrum, Halle. Orientwissenschaftliche Hefte 5; Mitteilungen des SFB „Differenz und Integration“ 3.
- Hein DW (1988) *Acetylator genotype and arylamine-induced carcinogenesis*. Biochim Biophys Acta 948: 37-66.
- Hein DW (2002) *Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis*. Mutat Res 506-507: 65-77.
- Hein DW, Boukouvala S, Grant DM, Minchin RF and Sim E (2008) *Changes in consensus arylamine N-acetyltransferase gene nomenclature*. Pharmacogenet Genomics 18: 367-368.

- Hermel E and Klapstein KD (2011) *A possible mechanism for maintenance of the deleterious allele of human CASPASE-12*. *Med Hypotheses* 77: 803-806.
- Hermisson J and Pennings PS (2005) *Soft sweeps: molecular population genetics of adaptation from standing genetic variation*. *Genetics* 169: 2335-2352.
- Hershkovitz I, Donoghue HD, Minnikin DE, Besra GS, Lee OY, Gernaey AM, Galili E, Eshed V, Greenblatt CL, Lemma E, Bar-Gal GK and Spigelman M (2008) *Detection and molecular characterization of 9,000-year-old Mycobacterium tuberculosis from a Neolithic settlement in the Eastern Mediterranean*. *PLoS One* 3: e3426.
- Hervella M, Plantinga TS, Alonso S, Ferwerda B, Izagirre N, Fontecha L, Fregel R, van der Meer JW, de-la-Rua C and Netea MG (2012) *The loss of functional caspase-12 in europe is a pre-neolithic event*. *PLoS ONE* 7: e37022.
- Heyer E, Brazier L, Segurel L, Hegay T, Austerlitz F, Quintana-Murci L, Georges M, Pasquet P and Veuille M (2011) *Lactase persistence in central Asia: phenotype, genotype, and evolution*. *Hum Biol* 83: 379-392.
- Higgins CF and Gottesman MM (1992) *Is the multidrug transporter a flippase?* *Trends Biochem Sci* 17: 18-21.
- Higuchi R, Bowman B, Freiberger M, Ryder OA and Wilson AC (1984) *DNA sequences from the quagga, an extinct member of the horse family*. *Nature* 312: 282-284.
- Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA and Cox DR (2005) *Whole-genome patterns of common DNA variation in three human populations*. *Science* 307: 1072-1079.
- Ho SY and Gilbert MT (2010) *Ancient mitogenomics*. *Mitochondrion* 10: 1-11.
- Hoban S, Bertorelle G and Gaggiotti OE (2012) *Computer simulations: tools for population and evolutionary genetics*. *Nat Rev Genet* 13: 110-122.
- Hofer T, Ray N, Wegmann D and Excoffier L (2009) *Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection*. *Ann Hum Genet* 73: 95-108.
- Hoffmeyer S, Burk O, von Richter O, Arnold HP, Brockmoller J, John A, Cascorbi I, Gerloff T, Roots I, Eichelbaum M and Brinkmann U (2000) *Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo*. *Proc Natl Acad Sci U S A* 97: 3473-3478.
- Hofreiter M, Serre D, Poinar HN, Kuch M and Pääbo S (2001) *Ancient DNA*. *Nat Rev Genet* 2: 353-359.
- Holick MF, MacLaughlin JA and Doppelt SH (1981) *Regulation of cutaneous previtamin D3 photosynthesis in man: skin pigment is not an essential regulator*. *Science* 211: 590-593.
- Hollfelder N (2013) *Frequenzen von rezent positiv selektierten Allelen in der ukrainischen Population*. Magisterarbeit an der Johannes Gutenberg-Universität Mainz.
- Holliday TW and Hilton CE (2010) *Body proportions of circumpolar peoples as evidenced from skeletal data: Ipiutak and Tigara (Point Hope) versus Kodiak Island Inuit*. *Am J Phys Anthropol* 142: 287-302.
- Holsinger KE and Weir BS (2009) *Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}* . *Nat Rev Genet* 10: 639-650.

- Horinouchi M, Sakaeda T, Nakamura T, Morita Y, Tamura T, Aoyama N, Kasuga M and Okumura K (2002) *Significant genetic linkage of MDR1 polymorphisms at positions 3435 and 2677: functional relevance to pharmacokinetics of digoxin*. *Pharm Res* 19: 1581-1585.
- Hotz C and Gibson RS (2007) *Traditional food-processing and preparation practices to enhance the bioavailability of micronutrients in plant-based diets*. *J Nutr* 137: 1097-1100.
- Hove H, Norgaard H and Mortensen PB (1999) *Lactic acid bacteria and the human gastrointestinal tract*. *Eur J Clin Nutr* 53: 339-350.
- Howes RE, Patil AP, Piel FB, Nyangiri OA, Kabaria CW, Gething PW, Zimmerman PA, Barnadas C, Beall CM, Gebremedhin A, Menard D, Williams TN, Weatherall DJ and Hay SI (2011) *The global distribution of the Duffy blood group*. *Nat Commun* 2: 266.
- Hudjashov G, VILLEMS R and KIVISILD T (2013) *Global patterns of diversity and selection in human tyrosinase gene*. *PLoS One* 8: e74307.
- Hudson RR, Kreitman M and Aguade M (1987) *A test of neutral molecular evolution based on nucleotide data*. *Genetics* 116: 153-159.
- Hudson RR (1990) *Gene genealogies and the coalescent process*. In: Futuyama D and Antonovics J (Hrsg), *Oxford Surveys in Evolutionary Biology*, Vol. 7. Oxford University Press, New York. 1-44.
- Hughes AL and Nei M (1988) *Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection*. *Nature* 335: 167-170.
- Hughes AL and Nei M (1989) *Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection*. *Proc Natl Acad Sci U S A* 86: 958-962.
- Hughes DA, Tang K, Strotmann R, Schoneberg T, Prenen J, Nilius B and Stoneking M (2008) *Parallel selection on TRPV6 in human populations*. *PLoS One* 3: e1686.
- Hummel S, Schmidt D, Kremeyer B, Herrmann B and Oppermann M (2005) *Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons*. *Genes Immun* 6: 371-374.
- Hurley TD, Edenberg HJ and Bosron WF (1990) *Expression and kinetic characterization of variants of human beta 1 beta 1 alcohol dehydrogenase containing substitutions at amino acid 47*. *J Biol Chem* 265: 16366-16372.
- Ingram CJ, Elamin MF, Mulcare CA, Weale ME, Tarekegn A, Raga TO, Bekele E, Elamin FM, Thomas MG, Bradman N and Swallow DM (2007) *A novel polymorphism associated with lactose tolerance in Africa: multiple causes for lactase persistence?* *Hum Genet* 120: 779-788.
- International Human Genome Sequencing Consortium (2001) *Initial sequencing and analysis of the human genome*. *Nature* 409: 860-921.
- Itan Y, Powell A, Beaumont MA, Burger J and Thomas MG (2009) *The origins of lactase persistence in Europe*. *PLoS Comput Biol* 5: e1000491.
- Itan Y, Jones BL, Ingram CJ, Swallow DM and Thomas MG (2010) *A worldwide correlation of lactase persistence phenotype and genotypes*. *BMC Evol Biol* 10: 36.
- Ito S (2003) *The IFPCS presidential lecture: a chemist's view of melanogenesis*. *Pigment Cell Res* 16: 230-236.

- IUPAC-IUB Joint Commission on Biochemical Nomenclature (1993) *Nomenclature and symbolism for amino acids and peptides. Corrections to recommendations 1983*. Eur J Biochem 213: 2.
- Jablonski NG and Chaplin G (2000) *The evolution of human skin coloration*. J Hum Evol 39: 57-106.
- Jablonski NG (2004) *The evolution of human skin and skin color*. Annu Rev Anthropol 33: 585-623.
- Jablonski NG and Chaplin G (2010) *Colloquium paper: human skin pigmentation as an adaptation to UV radiation*. Proc Natl Acad Sci U S A 107 Suppl 2: 8962-8968.
- Jablonski NG (2012) *The evolution of human skin colouration and its relevance to health in the modern world*. J R Coll Physicians Edinb 42: 58-63.
- Jaenicke-Després V, Buckler ES, Smith BD, Gilbert MT, Cooper A, Doebley J and Pääbo S (2003) *Early allelic selection in maize as revealed by ancient DNA*. Science 302: 1206-1208.
- Jeunesse C and Willigen Sv (2010) *Westmediterranes Frühneolithikum und westliche Linearbandkeramik: Impulse, Interaktionen, Mischkulturen*. In: Gronenborn D and Petrasch J (Hrsg), *Die Neolithisierung Mitteleuropas - The Spread of the Neolithic to Central Europe (Internationale Tagung, Mainz 24. bis 26. Juni 2005)*. Schnell & Steiner, Regensburg. 569-605.
- Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D and Lifton RP (2008) *Rare independent mutations in renal salt handling genes contribute to blood pressure variation*. Nat Genet 40: 592-599.
- Ji Y, Walkowicz MJ, Buiting K, Johnson DK, Tarvin RE, Rinchik EM, Horsthemke B, Stubbs L and Nicholls RD (1999) *The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities*. Hum Mol Genet 8: 533-542.
- Jobling MA, Hurles ME and Tyler-Smith C (2004) *Human Evolutionary Genetics*. Garland Science, New York.
- Jones G, Strugnell SA and DeLuca HF (1998) *Current understanding of the molecular actions of vitamin D*. Physiol Rev 78: 1193-1231.
- Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT and Batzer MA (2000) *The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data*. Am J Hum Genet 66: 979-988.
- Jörnvall H, Hempel J, Vallee BL, Bosron WF and Li TK (1984) *Human liver alcohol dehydrogenase: amino acid substitution in the beta 2 beta 2 Oriental isozyme explains functional properties, establishes an active site structure, and parallels mutational exchanges in the yeast enzyme*. Proc Natl Acad Sci U S A 81: 3024-3028.
- Kaiser E (2003) *Studien zur Katakombengrabkultur zwischen Dnepr und Prut*. Philipp von Zabern, Mainz. Archäologie in Eurasien 14 (zugleich Dissertation Freie Universität Berlin 2000).
- Kaiser E (2007) *Die Bestattungen der Katakombengrabkultur*. In: Bunjatjan KP, Kaiser E and Nikolova AV (Hrsg), *Bronzezeitliche Bestattungen aus dem unteren Dneprgebiet*. Schriften des Zentrums für Archäologie und Kulturgeschichte des Schwarzmeerraums 8. Beier & Beran, Langenweißbach. 33-45.

- Kaiser E (2010) *Der Übergang zur Rinderzucht im nördlichen Schwarzmeerraum*. Godišnjak Centar za balkanološka ispitivanja 23-34.
- Kalicz N (1998) *Östliche Beziehungen während der Kupferzeit in Ungarn*. In: Hänsel B (Hrsg), Mensch und Umwelt in der Bronzezeit Europas. Oetker-Voges, Kiel. 163-177.
- Kamberov YG, Wang S, Tan J, Gerbault P, Wark A, Tan L, Yang Y, Li S, Tang K, Chen H, Powell A, Itan Y, Fuller D, Lohmueller J, Mao J, Schachar A, Paymer M, Hostetter E, Byrne E, Burnett M, McMahon AP, Thomas MG, Lieberman DE, Jin L, Tabin CJ, Morgan BA and Sabeti PC (2013) *Modeling recent human evolution in mice by expression of a selected EDAR variant*. Cell 152: 691-702.
- Kang G, Bae KY, Kim SW, Kim J, Shin HY, Kim JM, Shin IS, Yoon JS and Kim JK (2014) *Effect of the Allelic Variant of Alcohol Dehydrogenase ADH1B*2 on Ethanol Metabolism*. Alcohol Clin Exp Res 38: 1502-1509.
- Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B, Franke A, Mayer J, Spangler J, McLaughlin S, Shah M, Lee C, Harkins TT, Sartori A, Moreno-Estrada A, Henn B, Sikora M, Semino O, Chiaroni J, Rootsi S, Myres NM, Cabrera VM, Underhill PA, Bustamante CD, Vigl EE, Samadelli M, Cipollini G, Haas J, Katus H, O'Connor BD, Carlson MR, Meder B, Blin N, Meese E, Pusch CM and Zink A (2012) *New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing*. Nat Commun 3: 698.
- Keller M, Ruegg A, Werner S and Beer HD (2008) *Active caspase-1 is a regulator of unconventional protein secretion*. Cell 132: 818-831.
- Keyser-Tracqui C, Crubezy E, Pamzav H, Varga T and Ludes B (2006) *Population origins in Mongolia: genetic structure analysis of ancient and modern DNA*. Am J Phys Anthropol 131: 272-281.
- Keyser C, Bouakaze C, Crubezy E, Nikolaev VG, Montagnon D, Reis T and Ludes B (2009) *Ancient DNA provides new insights into the history of south Siberian Kurgan people*. Hum Genet 126: 395-410.
- Kim GH, Masilamani S, Turner R, Mitchell C, Wade JB and Knepper MA (1998) *The thiazide-sensitive Na-Cl cotransporter is an aldosterone-induced protein*. Proc Natl Acad Sci U S A 95: 14552-14557.
- Kim RB, Leake BF, Choo EF, Dresser GK, Kubba SV, Schwarz UI, Taylor A, Xie HG, McKinsey J, Zhou S, Lan LB, Schuetz JD, Schuetz EG and Wilkinson GR (2001) *Identification of functionally variant MDR1 alleles among European Americans and African Americans*. Clin Pharmacol Ther 70: 189-199.
- Kim Y and Stephan W (2002) *Detecting a local signature of genetic hitchhiking along a recombining chromosome*. Genetics 160: 765-777.
- Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV and Gottesman MM (2007) *A "silent" polymorphism in the MDR1 gene changes substrate specificity*. Science 315: 525-528.
- Kimura M (1968) *Evolutionary rate at the molecular level*. Nature 217: 624-626.
- Kimura R, Yamaguchi T, Takeda M, Kondo O, Toma T, Haneji K, Hanihara T, Matsukusa H, Kawamura S, Maki K, Osawa M, Ishida H and Oota H (2009) *A common variation in EDAR is a genetic determinant of shovel-shaped incisors*. Am J Hum Genet 85: 528-535.
- Kingman JFC (1982) *The coalescent*. 13: 235-248.

- Kirsanow K and Burger J (2012) *Ancient human DNA*. *Ann Anat* 194: 121-132.
- Kistler L and Shapiro B (2011) *Ancient DNA confirms a local origin of domesticated chenopod in eastern North America*. *J Archaeol Sci* 38: 3549-3554.
- Kivistö KT, Niemi M and Fromm MF (2004) *Functional interaction of intestinal CYP3A4 and P-glycoprotein*. *Fundam Clin Pharmacol* 18: 621-626.
- Klopfstein S, Currat M and Excoffier L (2006) *The fate of mutations surfing on the wave of a range expansion*. *Mol Biol Evol* 23: 482-490.
- Kloss-Brandstätter A, Pacher D, Schonherr S, Weissensteiner H, Binna R, Specht G and Kronenberg F (2011) *HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups*. *Hum Mutat* 32: 25-32.
- Knapp M and Hofreiter M (2010) *Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives*. *Genes (Basel)* 1: 227-243.
- Knepper MA and Brooks HL (2001) *Regulation of the sodium transporters NHE3, NKCC2 and NCC in the kidney*. *Curr Opin Nephrol Hypertens* 10: 655-659.
- Kozłowski T, Stepańczak B, Reitsema LJ, Osipowicz G, Szostek K, Płoszaj T, Jędrychowska-Dańska K, Pawłyta J, Paluszkiwicz C and Witas HW (2014) *Osteological, chemical and genetic analyses of the human skeleton from a Neolithic site representing the Globular Amphora Culture (Kowal, Kuyavia region, Poland)*. *Anthropologie (Brno)* 52: 91-111.
- Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, Burbano HA, Hublin JJ, Hanni C, Fortea J, de la Rasilla M, Bertranpetit J, Rosas A and Pääbo S (2007) *The derived FOXP2 variant of modern humans was shared with Neandertals*. *Curr Biol* 17: 1908-1912.
- Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A and Pääbo S (2010) *A complete mtDNA genome of an early modern human from Kostenki, Russia*. *Curr Biol* 20: 231-236.
- Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M and Pääbo S (1997) *Neandertal DNA sequences and the origin of modern humans*. *Cell* 90: 19-30.
- Kroetz DL, Pauli-Magnus C, Hodges LM, Huang CC, Kawamoto M, Johns SJ, Stryke D, Ferrin TE, DeYoung J, Taylor T, Carlson EJ, Herskowitz I, Giacomini KM and Clark AG (2003) *Sequence diversity and haplotype structure in the human ABCB1 (MDR1, multidrug resistance transporter) gene*. *Pharmacogenetics* 13: 481-494.
- Kruskal JB (1964) *Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis*. *Psychometrika* 29: 1-27.
- Kubijovyč V (2001a) *Forest-steppe*. In: Internet Encyclopedia of Ukraine. Studies CloU, Toronto.
<http://www.encyclopediaofukraine.com/display.asp?linkpath=pages\F\O\Forest6steppe.htm> (abgerufen am 27.06.2014).
- Kubijovyč V (2001b) *Dnieper Rapids*. In: Internet Encyclopedia of Ukraine. Studies CloU, Toronto.
<http://www.encyclopediaofukraine.com/display.asp?linkpath=pages\D\N\DnieperRapids.htm> (abgerufen am 27.06.2014).
- Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, Watkins PB, Daly A, Wrighton SA, Hall SD, Maurel P, Relling M, Brimer C, Yasuda K, Venkataramanan R, Strom S, Thummel K, Boguski MS and Schuetz E (2001) *Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression*. *Nat Genet* 27: 383-391.

- Kuhnle S, Kogel U, Glockzin S, Marquardt A, Ciechanover A, Matentzoglou K and Scheffner M (2011) *Physical and functional interaction of the HECT ubiquitin-protein ligases E6AP and HERC2*. J Biol Chem 286: 19410-19416.
- Kulie T, Groff A, Redmer J, Hounshell J and Schrager S (2009) *Vitamin D: an evidence-based review*. J Am Board Fam Med 22: 698-706.
- Kumar A, Eby MT, Sinha S, Jasmin A and Chaudhary PM (2001) *The ectodermal dysplasia receptor activates the nuclear factor-kappaB, JNK, and cell death pathways and binds to ectodysplasin A*. J Biol Chem 276: 2668-2677.
- Kurian AK and Cardarelli KM (2007) *Racial and ethnic differences in cardiovascular disease risk factors: a systematic review*. Ethn Dis 17: 143-152.
- Labbé K, Miu J, Yeretsian G, Serghides L, Tam M, Finney CA, Erdman LK, Goulet ML, Kain KC, Stevenson MM and Saleh M (2010) *Caspase-12 dampens the immune response to malaria independently of the inflammasome by targeting NF-kappaB signaling*. J Immunol 185: 5495-5502.
- Lacchini R, Figueiredo VN, Demacq C, Coeli-Lacchini FB, Martins LC, Yugar-Toledo J, Coca A, Tanus-Santos JE and Moreno H, Jr. (2012) *MDR-1 C3435T polymorphism may affect blood pressure in resistant hypertensive patients independently of its effects on aldosterone release*. J Renin Angiotensin Aldosterone Syst 15: 170-176.
- Laland KN, Odling-Smee J and Myles S (2010) *How culture shaped the human genome: bringing genetics and the human sciences together*. Nat Rev Genet 11: 137-148.
- Lalueza-Fox C, Sampietro ML, Gilbert MT, Castri L, Facchini F, Pettener D and Bertranpetit J (2004) *Unravelling migrations in the steppe: mitochondrial DNA sequences from ancient central Asians*. Proc Biol Sci 271: 941-947.
- Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynec MJ, Mao X, Humphreville VR, Humbert JE, Sinha S, Moore JL, Jagadeeswaran P, Zhao W, Ning G, Makalowska I, McKeigue PM, O'Donnell D, Kittles R, Parra EJ, Mangini NJ, Grunwald DJ, Shriver MD, Canfield VA and Cheng KC (2005) *SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans*. Science 310: 1782-1786.
- Lamba JK, Lin YS, Schuetz EG and Thummel KE (2002) *Genetic contribution to variable human CYP3A-mediated metabolism*. Adv Drug Deliv Rev 54: 1271-1294.
- Langaee TY, Gong Y, Yarandi HN, Katz DA, Cooper-DeHoff RM, Pepine CJ and Johnson JA (2007) *Association of CYP3A5 polymorphisms with hypertension and antihypertensive response to verapamil*. Clin Pharmacol Ther 81: 386-391.
- Lankas GR, Wise LD, Cartwright ME, Pippert T and Umbenhauer DR (1998) *Placental P-glycoprotein deficiency enhances susceptibility to chemically induced birth defects in mice*. Reprod Toxicol 12: 457-463.
- Lao O, de Gruijter JM, van Duijn K, Navarro A and Kayser M (2007) *Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms*. Ann Hum Genet 71: 354-369.
- Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A, Vigne JD, Edwards CJ, Schlumbaum A, Dinu A, Balacescu A, Dolman G, Tagliacozzo A, Manaseryan N, Miracle P, Van Wijngaarden-Bakker L, Masseti M, Bradley DG

- and Cooper A (2007a) *Ancient DNA, pig domestication, and the spread of the Neolithic into Europe*. Proc Natl Acad Sci U S A 104: 15276-15281.
- Larson G, Cucchi T, Fujita M, Matisoo-Smith E, Robins J, Anderson A, Rolett B, Spriggs M, Dolman G, Kim TH, Thuy NT, Randi E, Doherty M, Due RA, Boltt R, Djubiantono T, Griffin B, Intoh M, Keane E, Kirch P, Li KT, Morwood M, Pedrina LM, Piper PJ, Rabett RJ, Shooter P, Van den Bergh G, West E, Wickler S, Yuan J, Cooper A and Dobney K (2007b) *Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania*. Proc Natl Acad Sci U S A 104: 4834-4839.
- Laurieri N, Dairou J, Egleton JE, Stanley LA, Russell AJ, Dupret JM, Sim E and Rodrigues-Lima F (2014) *From arylamine N-acetyltransferase to folate-dependent acetyl CoA hydrolase: impact of folic acid on the activity of (HUMAN)NAT1 and its homologue (MOUSE)NAT2*. PLoS One 9: e96370.
- Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI, Nordenfelt S, Li H, de Filippo C, Prufer K, Sawyer S, Posth C, Haak W, Hallgren F, Fornander E, Rohland N, Delsate D, Francken M, Guinet JM, Wahl J, Ayodo G, Babiker HA, Bailliet G, Balanovska E, Balanovsky O, Barrantes R, Bedoya G, Ben-Ami H, Bene J, Berrada F, Bravi CM, Brisighelli F, Busby GB, Cali F, Churnosov M, Cole DE, Corach D, Damba L, van Driem G, Dryomov S, Dugoujon JM, Fedorova SA, Gallego Romero I, Gubina M, Hammer M, Henn BM, Hervig T, Hodoqlugil U, Jha AR, Karachanak-Yankova S, Khusainova R, Khusnutdinova E, Kittles R, Kivisild T, Klitz W, Kucinskis V, Kushniarevich A, Laredj L, Litvinov S, Loukidis T, Mahley RW, Melegh B, Metspalu E, Molina J, Mountain J, Nakkalajarvi K, Nesheva D, Nyambo T, Osipova L, Parik J, Platonov F, Posukh O, Romano V, Rothhammer F, Rudan I, Ruizbakiev R, Sahakyan H, Sajantila A, Salas A, Starikovskaya EB, Tarekegn A, Toncheva D, Turdikulova S, Uktveryte I, Utevska O, Vasquez R, Villena M, Voevoda M, Winkler CA, Yepiskoposyan L, Zalloua P, Zemunik T, Cooper A, Capelli C, Thomas MG, Ruiz-Linares A, Tishkoff SA, Singh L, Thangaraj K, VILLEMS R, Comas D, Sukernik R, Metspalu M, Meyer M, Eichler EE, Burger J, Slatkin M, Paabo S, Kelso J, Reich D and Krause J (2014) *Ancient human genomes suggest three ancestral populations for present-day Europeans*. Nature 513: 409-413.
- Lee CG, Tang K, Cheung YB, Wong LP, Tan C, Shen H, Zhao Y, Pavanni R, Lee EJ, Wong MC, Chong SS and Tan EK (2004) *MDR1, the blood-brain barrier transporter, is associated with Parkinson's disease in ethnic Chinese*. J Med Genet 41: e60.
- Lee EJ, Makarewicz C, Renneberg R, Harder M, Krause-Kyora B, Muller S, Ostritz S, Fehren-Schmitz L, Schreiber S, Muller J, von Wurmb-Schwark N and Nebel A (2012) *Emerging genetic patterns of the European Neolithic: perspectives from a late Neolithic Bell Beaker burial site in Germany*. Am J Phys Anthropol 148: 571-579.
- Leonard JA, Wayne RK, Wheeler J, Valadez R, Guillen S and Vila C (2002) *Ancient DNA evidence for Old World origin of New World dogs*. Science 298: 1613-1616.
- Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, Cohen J, Opal SM, Vincent JL and Ramsay G (2003) *2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference*. Intensive Care Med 29: 530-538.

- Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL and Venter JC (2007) *The diploid genome sequence of an individual human*. PLoS Biol 5: e254.
- Lewontin RC (1972) *The Apportionment of Human Diversity*. Evol Biol 6: 381-398.
- Li D, Zhao H and Gelernter J (2011a) *Strong association of the alcohol dehydrogenase 1B gene (ADH1B) with alcohol dependence and alcohol-induced medical diseases*. Biol Psychiatry 70: 504-512.
- Li H, Mukherjee N, Soundararajan U, Tarnok Z, Barta C, Khaliq S, Mohyuddin A, Kajuna SL, Mehdi SQ, Kidd JR and Kidd KK (2007) *Geographically separate increases in the frequency of the derived ADH1B*47His allele in eastern and western Asia*. Am J Hum Genet 81: 842-846.
- Li H, Gu S, Cai X, Speed WC, Pakstis AJ, Golub EI, Kidd JR and Kidd KK (2008) *Ethnic related selection for an ADH Class I variant within East Asia*. PLoS ONE 3: e1881.
- Li H, Gu S, Han Y, Xu Z, Pakstis AJ, Jin L, Kidd JR and Kidd KK (2011b) *Diversification of the ADH1B gene during expansion of modern humans*. Ann Hum Genet 75: 497-507.
- Lichardus J and Lichardus-Itten M (1998) *Nordpontische Gruppen und ihre westlichen Nachbarn. Ein Beitrag zur Entstehung der frühen Kupferzeit Alteuropas*. In: Hänsel B and Machnik J (Hrsg), *Das Karpatenbecken und die osteuropäische Steppe. Südosteuropa-Schriften, Band 20*. VML Verlag Marie Leidorf, Rahden/Westf. 99-133.
- Lichter C (2005) *Introduction to the workshop*. In: Lichter C (Hrsg), *How Did Farming Reach Europe? Byzas 2, Istanbul*. 1-11.
- Lieben L, Benn BS, Ajibade D, Stockmans I, Moermans K, Hediger MA, Peng JB, Christakos S, Bouillon R and Carmeliet G (2010) *Trpv6 mediates intestinal calcium absorption during calcium restriction and contributes to bone homeostasis*. Bone 47: 301-308.
- Lindahl T (1993) *Instability and decay of the primary structure of DNA*. Nature 362: 709-715.
- Lindqvist C, Schuster SC, Sun Y, Talbot SL, Qi J, Ratan A, Tomsho LP, Kasson L, Zeyl E, Aars J, Miller W, Ingolfsson O, Bachmann L and Wiig O (2010) *Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear*. Proc Natl Acad Sci U S A 107: 5053-5057.
- Lira J, Linderholm A, Olaria C, Brandstrom Durling M, Gilbert MT, Ellegren H, Willerslev E, Liden K, Arsuaga JL and Gotherstrom A (2010) *Ancient DNA reveals traces of Iberian Neolithic and Bronze Age lineages in modern Iberian horses*. Mol Ecol 19: 64-78.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L and Law M (2012) *Comparison of next-generation sequencing systems*. J Biomed Biotechnol 2012: 251364.
- Livingstone FB (1984) *The Duffy blood groups, vivax malaria, and malaria selection in human populations: a review*. Hum Biol 56: 413-425.
- Lomer MC, Parkes GC and Sanderson JD (2008) *Review article: lactose intolerance in clinical practice--myths and realities*. Aliment Pharmacol Ther 27: 93-103.

- Lubbock J (1865) *Pre-historic times as illustrated by ancient remains, and the manners and customs of modern savages*. Williams & Norgate, London.
- Luca F, Bubba G, Basile M, Brdicka R, Michalodimitrakis E, Rickards O, Vershubsky G, Quintana-Murci L, Kozlov AI and Novelletto A (2008) *Multiple advantageous amino acid variants in the NAT2 gene in human populations*. PLoS ONE 3: e3136.
- Luca F, Perry GH and Di Rienzo A (2010) *Evolutionary adaptations to dietary changes*. Annu Rev Nutr 30: 291-314.
- Magalon H, Patin E, Austerlitz F, Hegay T, Aldashev A, Quintana-Murci L and Heyer E (2008) *Population genetic diversity of the NAT2 gene supports a role of acetylation in human adaptation to farming in Central Asia*. Eur J Hum Genet 16: 243-251.
- Mallick CB, Iliescu FM, Mols M, Hill S, Tamang R, Chaubey G, Goto R, Ho SY, Gallego Romero I, Crivellaro F, Hudjashov G, Rai N, Metspalu M, Mascie-Taylor CG, Pitchappan R, Singh L, Mirazon-Lahr M, Thangaraj K, Villems R and Kivisild T (2013) *The light skin allele of SLC24A5 in South Asians and Europeans shares identity by descent*. PLoS Genet 9: e1003912.
- Mallory JP and Adams DQ (Hrsg) (1997) *Encyclopedia of Indo-European Culture*. Fitzroy Dearborn, London, Chicago.
- Malmström H, Gilbert MT, Thomas MG, Brandström M, Stora J, Molnar P, Andersen PK, Bendixen C, Holmlund G, Götherström A and Willerslev E (2009) *Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians*. Curr Biol 19: 1758-1762.
- Malyarchuk BA and Derenko MV (2001) *Mitochondrial DNA variability in Russians and Ukrainians: implication to the origin of the Eastern Slavs*. Ann Hum Genet 65: 63-78.
- Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Wozniak M and Miscicka-Sliwka D (2002) *Mitochondrial DNA variability in Poles and Russians*. Ann Hum Genet 66: 261-283.
- Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J and Turner DJ (2010) *Target-enrichment strategies for next-generation sequencing*. Nat Methods 7: 111-118.
- Manzura I (2005) *Steps to the Steppe: Or, How the North Pontic Region Was Colonised*. Oxf J Archaeol 24: 313-338.
- Mardis ER (2008) *Next-generation DNA sequencing methods*. Annu Rev Genomics Hum Genet 9: 387-402.
- Mardis ER (2013) *Next-generation sequencing platforms*. Annu Rev Anal Chem 6: 287-303.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF and Rothberg JM (2005) *Genome sequencing in microfabricated high-density picolitre reactors*. Nature 437: 376-380.

- Maricic T and Pääbo S (2009) *Optimization of 454 sequencing library preparation from small amounts of DNA permits sequence determination of both DNA strands*. *Biotechniques* 46: 51-52, 54-57.
- Martin A, Saathoff M, Kuhn F, Max H, Terstegen L and Natsch A (2010) *A functional ABCC11 allele is essential in the biochemical formation of human axillary odor*. *J Invest Dermatol* 130: 529-540.
- Masson VM (1998) *Die europäischen Steppen in den frühen Metallzeiten. Kultur und Sozialfortschritt im Lichte der neuen archäologischen Entdeckungen*. In: Hänsel B and Machnik J (Hrsg), *Das Karpatenbecken und die osteuropäische Steppe*. Südosteuropa-Schriften, Band 20. VML Verlag Marie Leidorf, Rahden/Westf. 19-26.
- Mastroianni N, De Fusco M, Zollo M, Arrigo G, Zuffardi O, Bettinelli A, Ballabio A and Casari G (1996) *Molecular cloning, expression pattern, and chromosomal localization of the human Na-Cl thiazide-sensitive cotransporter (SLC12A3)*. *Genomics* 35: 486-493.
- Matsuo Y, Yokoyama R and Yokoyama S (1989) *The genes for human alcohol dehydrogenases beta 1 and beta 2 differ by only one nucleotide*. *Eur J Biochem* 183: 317-320.
- McCall MB, Ferwerda B, Hopman J, Ploemen I, Maiga B, Daou M, Dolo A, Hermsen CC, Doumbo OK, Bedu-Addo G, van der Meer JW, Troye-Blomberg M, van der Ven AJ, Schumann RR, Sauerwein RW, Mockenhaupt FP and Netea MG (2010) *Persistence of full-length caspase-12 and its relation to malaria in West and Central African populations*. *Eur Cytokine Netw* 21: 77-83.
- McDonald JH and Kreitman M (1991) *Adaptive protein evolution at the Adh locus in Drosophila*. *Nature* 351: 652-654.
- McGovern PE, Zhang J, Tang J, Zhang Z, Hall GR, Moreau RA, Nunez A, Butrym ED, Richards MP, Wang CS, Cheng G, Zhao Z and Wang C (2004) *Fermented beverages of pre- and proto-historic China*. *Proc Natl Acad Sci U S A* 101: 17593-17598.
- Melaine N, Lienard MO, Dorval I, Le Goascogne C, Lejeune H and Jegou B (2002) *Multidrug resistance genes and p-glycoprotein in the testis of the rat, mouse, Guinea pig, and human*. *Biol Reprod* 67: 1699-1707.
- Metzker ML (2010) *Sequencing technologies - the next generation*. *Nat Rev Genet* 11: 31-46.
- Meyer M, Stenzel U, Myles S, Prufer K and Hofreiter M (2007) *Targeted high-throughput sequencing of tagged nucleic acid samples*. *Nucleic Acids Res* 35: e97.
- Meyer M, Stenzel U and Hofreiter M (2008) *Parallel tagged sequencing on the 454 platform*. *Nat Protoc* 3: 267-278.
- Meyer M, Fu Q, Aximu-Petri A, Glocke I, Nickel B, Arsuaga JL, Martinez I, Gracia A, de Castro JM, Carbonell E and Pääbo S (2014) *A mitochondrial genome sequence of a hominin from Sima de los Huesos*. *Nature* 505: 403-406.
- Miller LH, Mason SJ, Clyde DF and McGinniss MH (1976) *The resistance factor to Plasmodium vivax in blacks. The Duffy-blood-group genotype, FyFy*. *N Engl J Med* 295: 302-304.
- Miller W, Drautz DI, Janecka JE, Lesk AM, Ratan A, Tomsho LP, Packard M, Zhang Y, McClellan LR, Qi J, Zhao F, Gilbert MT, Dalen L, Arsuaga JL, Ericson PG, Huson

- DH, Helgen KM, Murphy WJ, Gotherstrom A and Schuster SC (2009) *The mitochondrial genome sequence of the Tasmanian tiger (Thylacinus cynocephalus)*. *Genome Res* 19: 213-220.
- Mizoi Y, Ijiri I, Tatsuno Y, Kijima T, Fujiwara S, Adachi J and Hishida S (1979) *Relationship between facial flushing and blood acetalddehyde levels after alcohol intake*. *Pharmacol Biochem Behav* 10: 303-311.
- Moes AD, van der Lubbe N, Zietse R, Loffing J and Hoorn EJ (2014) *The sodium chloride cotransporter SLC12A3: new roles in sodium, potassium, and blood pressure regulation*. *Pflugers Arch* 466: 107-118.
- Molodin VI, Pilipenko AS, Romaschenko AG, Zhuravlev AA, Trapezov RO, Chikisheva TA and Pozdnyakov DV (2012) *Human migrations in the southern region of the West Siberian Plain during the Bronze Age: Archaeological, palaeogenetic and anthropological data*. In: Kaiser E, Burger J and Schier W (Hrsg), *Population Dynamics in Prehistory and Early History, New Approaches Using Stable Isotopes and Genetics*. De Gruyter, Berlin, Boston. 93-111.
- Monreal AW, Ferguson BM, Headon DJ, Street SL, Overbeek PA and Zonana J (1999) *Mutations in the human homologue of mouse dl cause autosomal recessive and dominant hypohidrotic ectodermal dysplasia*. *Nat Genet* 22: 366-369.
- Moreno E, Tovar-Palacio C, de los Heros P, Guzman B, Bobadilla NA, Vazquez N, Riccardi D, Poch E and Gamba G (2004) *A single nucleotide polymorphism alters the activity of the renal Na⁺:Cl⁻ cotransporter and reveals a role for transmembrane segment 4 in chloride and thiazide affinity*. *J Biol Chem* 279: 16553-16560.
- Motuzaitė-Matuzevičiūtė G, Telizhenko S and Jones MK (2013) *The earliest evidence of domesticated wheat in the Crimea at Chalcolithic Ardych-Burun*. *J Field Archaeol* 38: 120-128.
- Mou C, Thomason HA, Willan PM, Clowes C, Harris WE, Drew CF, Dixon J, Dixon MJ and Headon DJ (2008) *Enhanced ectodysplasin-A receptor (EDAR) signaling alters multiple fiber characteristics to produce the East Asian hair form*. *Hum Mutat* 29: 1405-1411.
- Müller D, Hoenderop JG, Meij IC, van den Heuvel LP, Knoers NV, den Hollander AI, Eggert P, Garcia-Nieto V, Claverie-Martin F and Bindels RJ (2000) *Molecular cloning, tissue distribution, and chromosomal mapping of the human epithelial Ca²⁺ channel (ECAC1)*. *Genomics* 67: 48-53.
- Müller J (2009a) *Kupfer, Megalithen und neue Technologien*. In: Jockenhövel A (Hrsg), *WBG Weltgeschichte. Eine globale Geschichte von den Anfängen bis ins 21. Jahrhundert, Band I: Grundlagen der globalen Welt*. Wissenschaftliche Buchgesellschaft, Darmstadt. 297-328.
- Müller J (2009b) *Die Jungsteinzeit (6000 - 2000 v. Chr.)*. In: Schnurbein Sv (Hrsg), *Atlas der Vorgeschichte: Europa von den ersten Menschen bis Christi Geburt* Konrad Theiss Verlag, Stuttgart. 58-105.
- Mullis K, Faloona F, Scharf S, Saiki R, Horn G and Erlich H (1986) *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction*. *Cold Spring Harb Symp Quant Biol* 51 Pt 1: 263-273.
- Myles S, Bouzekri N, Haverfield E, Cherkaoui M, Dugoujon JM and Ward R (2005) *Genetic evidence in support of a shared Eurasian-North African dairying origin*. *Hum Genet* 117: 34-42.

- Nagao M, Wakabayashi K, Ushijima T, Toyota M, Totsuka Y and Sugimura T (1996) *Human exposure to carcinogenic heterocyclic amines and their mutational fingerprints in experimental animals*. Environ Health Perspect 104 Suppl 3: 497-501.
- Nakajima T, Wooding S, Sakagami T, Emi M, Tokunaga K, Tamiya G, Ishigami T, Umemura S, Munkhbat B, Jin F, Guan-Jun J, Hayasaka I, Ishida T, Saitou N, Pavelka K, Lalouel JM, Jorde LB and Inoue I (2004) *Natural selection and population history in the human angiotensinogen gene (AGT): 736 complete AGT sequences in chromosomes from around the world*. Am J Hum Genet 74: 898-916.
- Nakamura T, Sakaeda T, Horinouchi M, Tamura T, Aoyama N, Shirakawa T, Matsuo M, Kasuga M and Okumura K (2002) *Effect of the mutation (C3435T) at exon 26 of the MDR1 gene on expression level of MDR1 messenger ribonucleic acid in duodenal enterocytes of healthy Japanese subjects*. Clin Pharmacol Ther 71: 297-303.
- Nakano M, Miwa N, Hirano A, Yoshiura K and Niikawa N (2009) *A strong association of axillary osmidrosis with the wet earwax type determined by genotyping of the ABCC11 gene*. BMC Genet 10: 42.
- Nan H, Kraft P, Hunter DJ and Han J (2009) *Genetic variants in pigmentation genes, pigmentary phenotypes, and risk of skin cancer in Caucasians*. Int J Cancer 125: 909-917.
- NCBI - BLAST - Basic Local Alignment Search Tool. Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
- NCBI - dbSNP - Database of Single Nucleotide Polymorphisms. Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine (dbSNP Build ID: 129-141): <http://www.ncbi.nlm.nih.gov/SNP/>.
- NCBI - OMIM - Online Mendelian Inheritance in Man, OMIM[®]. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) <http://omim.org/>.
- Nei M and Li WH (1979) *Mathematical model for studying genetic variation in terms of restriction endonucleases*. Proc Natl Acad Sci U S A 76: 5269-5273.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.
- Newton JM, Cohen-Barak O, Hagiwara N, Gardner JM, Davisson MT, King RA and Brilliant MH (2001) *Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4*. Am J Hum Genet 69: 981-988.
- Nicholls RD and Knepper JL (2001) *Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes*. Annu Rev Genomics Hum Genet 2: 153-175.
- Nielsen R, Hellmann I, Hubisz M, Bustamante C and Clark AG (2007) *Recent and ongoing selection in the human genome*. Nat Rev Genet 8: 857-868.
- Nielsen R and Beaumont MA (2009) *Statistical inferences in phylogeography*. Mol Ecol 18: 1034-1047.

- Nikitin AG, Sokhatsky MP, Kovaliukh MM and Videiko MY (2010) *Comprehensive Site Chronology and Ancient Mitochondrial DNA Analysis from Verteba Cave – a Trypillian Culture Site of Eneolithic Ukraine*. IANSA 1: 9-18.
- Nikitin AG, Newton JR and Potekhina ID (2012) *Mitochondrial haplogroup C in ancient mitochondrial DNA from Ukraine extends the presence of East Eurasian genetic lineages in Neolithic Central and Eastern Europe*. J Hum Genet 57: 610-612.
- Nikolova AV (2007) *Die Bestattungen der Jamnaja-Kultur*. In: Bunjatjan KP, Kaiser E and Nikolova AV (Hrsg), *Bronzezeitliche Bestattungen aus dem unteren Dneprgebiet*. Schriften des Zentrums für Archäologie und Kulturgeschichte des Schwarzmeerraums 8. Beier & Beran, Langenweißbach. 13-32.
- Nilius B, Prenen J, Vennekens R, Hoenderop JG, Bindels RJ and Droogmans G (2001) *Modulation of the epithelial calcium channel, ECaC, by intracellular Ca²⁺*. Cell Calcium 29: 417-428.
- Nomenclature Committee of the International Union of Biochemistry (1985) *Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984*. Biochem J 229: 281-286.
- Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, Chen F, Platt D, Paabo S, Pritchard JK and Rubin EM (2006) *Sequencing and analysis of Neanderthal genomic DNA*. Science 314: 1113-1118.
- Norton HL, Kittles RA, Parra E, McKeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, McEvoy B and Shriver MD (2007) *Genetic evidence for the convergent evolution of light skin in Europeans and East Asians*. Mol Biol Evol 24: 710-722.
- Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M and Bustamante CD (2008) *Genes mirror geography within Europe*. Nature 456: 98-101.
- Novembre J and Stephens M (2008) *Interpreting principal component analyses of spatial population genetic variation*. Nat Genet 40: 646-649.
- Novembre J and Di Rienzo A (2009) *Spatial patterns of variation due to natural selection in humans*. Nat Rev Genet 10: 745-755.
- Ogura Y, Sutterwala FS and Flavell RA (2006) *The inflammasome: first line of the immune response to cell stress*. Cell 126: 659-662.
- Ohashi J, Naka I and Tsuchiya N (2011) *The impact of natural selection on an ABCC11 SNP determining earwax type*. Mol Biol Evol 28: 849-857.
- Olalde I, Allentoft ME, Sanchez-Quinto F, Santpere G, Chiang CW, DeGiorgio M, Prado-Martinez J, Rodriguez JA, Rasmussen S, Quilez J, Ramirez O, Marigorta UM, Fernandez-Callejo M, Prada ME, Encinas JM, Nielsen R, Netea MG, Novembre J, Sturm RA, Sabeti P, Marques-Bonet T, Navarro A, Willerslev E and Lalueza-Fox C (2014) *Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European*. Nature 507: 225-228.
- Olds LC and Sibley E (2003) *Lactase persistence DNA variant enhances lactase promoter activity in vitro: functional role as a cis regulatory element*. Hum Mol Genet 12: 2333-2340.
- Oleksyk TK, Smith MW and O'Brien SJ (2010) *Genome-wide scans for footprints of natural selection*. Philos Trans R Soc Lond B Biol Sci 365: 185-205.
- Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, Johnson PL, Fumagalli M, Vilstrup JT, Raghavan M, Korneliusen T, Malaspina AS, Vogt J, Szklarczyk D, Kelstrup CD,

- Vinther J, Dolocan A, Stenderup J, Velazquez AM, Cahill J, Rasmussen M, Wang X, Min J, Zazula GD, Seguin-Orlando A, Mortensen C, Magnussen K, Thompson JF, Weinstock J, Gregersen K, Roed KH, Eisenmann V, Rubin CJ, Miller DC, Antczak DF, Bertelsen MF, Brunak S, Al-Rasheid KA, Ryder O, Andersson L, Mundy J, Krogh A, Gilbert MT, Kjaer K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wang J and Willerslev E (2013) *Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse*. *Nature* 499: 74-78.
- Osier MV, Cheung KH, Kidd JR, Pakstis AJ, Miller PL and Kidd KK (2002) *ALFRED: An allele frequency database for anthropology*. *Am J Phys Anthropol* 119: 77-83.
- Otoni C, Flink LG, Evin A, Georg C, De Cupere B, Van Neer W, Bartosiewicz L, Linderholm A, Barnett R, Peters J, Decorte R, Waelkens M, Vanderheyden N, Ricaut FX, Cakirlar C, Cevik O, Hoelzel AR, Mashkour M, Karimlu AF, Seno SS, Daujat J, Brock F, Pinhasi R, Hongo H, Perez-Enciso M, Rasmussen M, Frantz L, Megens HJ, Crooijmans R, Groenen M, Arbuckle B, Benecke N, Vidarsdottir US, Burger J, Cucchi T, Dobney K and Larson G (2013) *Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics*. *Mol Biol Evol* 30: 824-832.
- Ovchinnikov IV, Gotherstrom A, Romanova GP, Kharitonov VM, Liden K and Goodwin W (2000) *Molecular analysis of Neanderthal DNA from the northern Caucasus*. *Nature* 404: 490-493.
- Özdoğan M (2005) *Westward expansion of the Neolithic way of life: what we know and what we do not know*. In: Lichter C (Hrsg), *How Did Farming Reach Europe?* Byzas 2, Istanbul. 13-27.
- Özdoğan M (2007) *Von Zentralanatolien nach Europa - Die Ausbreitung der neolithischen Lebensweise*. In: Karlsruhe BL (Hrsg), *Die ältesten Monumente der Menschheit*. Theiss, Stuttgart. 150-165.
- Özdoğan M (2011) *Archaeological Evidence on the Westward Expansion of Farming Communities from Eastern Anatolia to the Aegean and the Balkans*. *Curr Anthropol* 52: S415-S430.
- Pääbo S (1985) *Molecular cloning of Ancient Egyptian mummy DNA*. *Nature* 314: 644-645.
- Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L and Hofreiter M (2004) *Genetic analyses from ancient DNA*. *Annu Rev Genet* 38: 645-679.
- Panchal M and Beaumont MA (2007) *The automation and evaluation of nested clade phylogeographic analysis*. *Evolution* 61: 1466-1480.
- Parpola A (2008) *Proto-Indo-European speakers of the Late Tripolye culture as the inventors of wheeled vehicles: Linguistic and archaeological considerations of the PIE homeland problem*. In: Jones-Bley K, Huld ME, Volpe AD and Dexter MR (Hrsg), *Proceedings of the 19th annual UCLA Indo-European conference*. *Journal of Indo-European Studies Monograph* 54. Institute for the Study of Man, Washington, D. C. 1-59.
- Parzinger H (1998a) *Kulturverhältnisse in der eurasischen Steppe während der Bronzezeit*. In: Hänsel B (Hrsg), *Mensch und Umwelt in der Bronzezeit Europas*. Oetker-Voges, Kiel. 457-479.

- Parzinger H (1998b) *Der nordpontische Raum und das untere Donaugebiet in der späten Kupferzeit: das Ende des Kodžadermen-Gumelnița-Karanovo VI-Verbandes und die Cernavodă I-Kultur*. In: Hänsel B and Machnik J (Hrsg), *Das Karpatenbecken und die osteuropäische Steppe*. Südosteuropa-Schriften, Band 20. VML Verlag Marie Leidorf, Rahden/Westf. 123-134.
- Parzinger H (2006) *Die frühen Völker Eurasiens. Vom Neolithikum bis zum Mittelalter*. C.H. Beck, München.
- Patin E, Barreiro LB, Sabeti PC, Austerlitz F, Luca F, Sajantila A, Behar DM, Semino O, Sakuntabhai A, Guiso N, Gicquel B, McElreavey K, Harding RM, Heyer E and Quintana-Murci L (2006) *Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes*. *Am J Hum Genet* 78: 423-436.
- Pavúk J (2004) *Early Linear Pottery Culture in Slovakia and the Neolithisation of Central Europe*. In: Lukes A and Zvelebil M (Hrsg), *LBK Dialogues: Studies in the formation of the Linear Pottery Culture*. BAR Internat Ser 1304, Oxford. 71-82.
- Peng JB, Chen XZ, Berger UV, Vassilev PM, Tsukaguchi H, Brown EM and Hediger MA (1999) *Molecular cloning and characterization of a channel-like transporter mediating intestinal calcium absorption*. *J Biol Chem* 274: 22739-22746.
- Peng JB, Chen XZ, Berger UV, Weremowicz S, Morton CC, Vassilev PM, Brown EM and Hediger MA (2000) *Human calcium transport protein CaT1*. *Biochem Biophys Res Commun* 278: 326-332.
- Peng Y, Shi H, Qi XB, Xiao CJ, Zhong H, Ma RL and Su B (2010) *The ADH1B Arg47His polymorphism in east Asian populations and expansion of rice domestication in history*. *BMC Evol Biol* 10: 15.
- Pheiffer BH and Zimmerman SB (1983) *Polymer-stimulated ligation: enhanced blunt- or cohesive-end ligation of DNA or deoxyribonucleotides by T4 DNA ligase in polymer solutions*. *Nucleic Acids Res* 11: 7853-7871.
- Pinsky MR (2004) *Dysregulation of the immune response in severe sepsis*. *Am J Med Sci* 328: 220-229.
- Pittas AG, Lau J, Hu FB and Dawson-Hughes B (2007) *The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis*. *J Clin Endocrinol Metab* 92: 2017-2029.
- Plotkin MD, Kaplan MR, Verlander JW, Lee WS, Brown D, Poch E, Gullans SR and Hebert SC (1996) *Localization of the thiazide sensitive Na-Cl cotransporter, rTSC1 in the rat kidney*. *Kidney Int* 50: 174-183.
- Poinar HN, Schwarz C, Qi J, Shapiro B, Macphee RD, Buigues B, Tikhonov A, Huson DH, Tomsho LP, Auch A, Rampp M, Miller W and Schuster SC (2006) *Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA*. *Science* 311: 392-394.
- Price TD (2000) *Europe's first farmers: an introduction*. In: Price TD (Hrsg), *Europe's First Farmers*. Cambridge University Press, Cambridge. 1-18.
- Price TD, Bentley RA, Lüning J, Gronenborn D and Wahl J (2001) *Prehistoric human migration in the Linearbandkeramik of Central Europe*. *Antiquity* 75: 593-603.
- Prokop-Prigge KA, Thaler E, Wysocki CJ and Preti G (2014) *Identification of volatile organic compounds in human cerumen*. *J Chromatogr B Analyt Technol Biomed Life Sci* 953-954: 48-52.

- Prugnolle F, Manica A, Charpentier M, Guegan JF, Guernier V and Balloux F (2005) *Pathogen-driven selection and worldwide HLA class I diversity*. *Curr Biol* 15: 1022-1027.
- Puri N, Gardner JM and Brilliant MH (2000) *Aberrant pH of melanosomes in pink-eyed dilution (p) mutant melanocytes*. *J Invest Dermatol* 115: 607-613.
- Quitta H (1960a) *Zur Frage der ältesten Bandkeramik in Mitteleuropa (Teil II)*. *Prähistorische Zeitschrift* 38: 153-188.
- Quitta H (1960b) *Zur Frage der ältesten Bandkeramik in Mitteleuropa (Teil I)*. *Prähistorische Zeitschrift* 38: 1-38.
- R Core Team (2013) *R: A language and environment for statistical computing*. In: Wien, Österreich. <http://www.R-project.org/> (abgerufen am
- Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford TW, Jr., Orlando L, Metspalu E, Karmin M, Tambets K, Rootsi S, Magi R, Campos PF, Balanovska E, Balanovsky O, Khusnutdinova E, Litvinov S, Osipova LP, Fedorova SA, Voevoda MI, DeGiorgio M, Sicheritz-Ponten T, Brunak S, Demeshchenko S, Kivisild T, Villems R, Nielsen R, Jakobsson M and Willerslev E (2014) *Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans*. *Nature* 505: 87-91.
- Rajeevan H, Osier MV, Cheung KH, Deng H, Druskin L, Heinzen R, Kidd JR, Stein S, Pakstis AJ, Tosches NP, Yeh CC, Miller PL and Kidd KK (2003) *ALFRED: the ALlele FREquency Database. Update*. *Nucleic Acids Res* 31: 270-271.
- Ramachandran C and Brunette MG (1989) *The renal Na⁺/Ca²⁺ exchange system is located exclusively in the distal tubule*. *Biochem J* 257: 259-264.
- Ramakrishnan U and Hadly EA (2009) *Using phylochronology to reveal cryptic population histories: review and synthesis of 29 ancient DNA studies*. *Mol Ecol* 18: 1310-1330.
- Ramchandani VA, Bosron WF and Li TK (2001) *Research advances in ethanol metabolism*. *Pathol Biol (Paris)* 49: 676-682.
- Ranciaro A, Campbell MC, Hirbo JB, Ko WY, Froment A, Anagnostou P, Kotze MJ, Ibrahim M, Nyambo T, Omar SA and Tishkoff SA (2014) *Genetic origins of lactase persistence and the spread of pastoralism in Africa*. *Am J Hum Genet* 94: 496-510.
- Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, Bertalan M, Nielsen K, Gilbert MT, Wang Y, Raghavan M, Campos PF, Kamp HM, Wilson AS, Gledhill A, Tridico S, Bunce M, Lorenzen ED, Binladen J, Guo X, Zhao J, Zhang X, Zhang H, Li Z, Chen M, Orlando L, Kristiansen K, Bak M, Tommerup N, Bendixen C, Pierre TL, Gronnow B, Meldgaard M, Andreasen C, Fedorova SA, Osipova LP, Higham TF, Ramsey CB, Hansen TV, Nielsen FC, Crawford MH, Brunak S, Sicheritz-Ponten T, Villems R, Nielsen R, Krogh A, Wang J and Willerslev E (2010) *Ancient human genome sequence of an extinct Palaeo-Eskimo*. *Nature* 463: 757-762.
- Rassamakin YY (1994) *The Main Directions of the Development of Early Pastoral Societies of Northern Pontic Zone: 4500-2450 BC (Pre-Yamnaya Cultures and Yamnaya Culture)*. In: Kadrow S, Klochko VI, Kosko A, S.Z. P, Rassamakin YY and Videiko MY (Hrsg), *Nomadism and Pastoralism in the Circle of Baltic-Pontic Early Agrarian Cultures: 5000 - 1650 BC*. *Baltic-Pontic Studies* 2, Poznań. 29-70.

- Rassamakin YY (1999) *The Eneolithic of the Black Sea Steppe: Dynamics of Cultural and Economic Development 4500-2300 BC*. In: Levine M, Rassamakin YY, Kislenko A and Tatarintseva N (Hrsg), *Late Prehistoric Exploitation of the Eurasian Steppe*. McDonald Institute for Archaeological Research, Cambridge, Mass. 59-182.
- Rees JL (2004) *The genetics of sun sensitivity in humans*. *Am J Hum Genet* 75: 739-751.
- Rees JL and Harding RM (2012) *Understanding the evolution of human pigmentation: recent contributions from population genetics*. *J Invest Dermatol* 132: 846-853.
- Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, Maricic T, Good JM, Marques-Bonet T, Alkan C, Fu Q, Mallick S, Li H, Meyer M, Eichler EE, Stoneking M, Richards M, Talamo S, Shunkov MV, Derevianko AP, Hublin JJ, Kelso J, Slatkin M and Pääbo S (2010) *Genetic history of an archaic hominin group from Denisova Cave in Siberia*. *Nature* 468: 1053-1060.
- Reingruber A (2011) *Early Neolithic settlement patterns and exchange networks in the Aegean*. *Documenta Praehistorica* 38: 291-305.
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Golge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Di Rienzo A, Novelletto A, Oppenheim A, Norby S, Al-Zaheri N, Santachiara-Benerecetti S, Scozari R, Torroni A and Bandelt HJ (2000) *Tracing European founder lineages in the Near Eastern mtDNA pool*. *Am J Hum Genet* 67: 1251-1276.
- Riedemann NC, Guo RF and Ward PA (2003) *The enigma of sepsis*. *J Clin Invest* 112: 460-467.
- Rishikaysh P, Dev K, Diaz D, Qureshi WM, Filip S and Mokry J (2014) *Signaling involved in hair follicle morphogenesis and development*. *Int J Mol Sci* 15: 1647-1670.
- Romualdi C, Balding D, Nasidze IS, Risch G, Robichaux M, Sherry ST, Stoneking M, Batzer MA and Barbujani G (2002) *Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms*. *Genome Res* 12: 602-612.
- Ronaghi M, Karamohamed S, Pettersson B, Uhlen M and Nyren P (1996) *Real-time DNA sequencing using detection of pyrophosphate release*. *Anal Biochem* 242: 84-89.
- Ronaghi M (2001) *Pyrosequencing sheds light on DNA sequencing*. *Genome Res* 11: 3-11.
- Rosenberg NA and Nordborg M (2002) *Genealogical trees, coalescent theory and the analysis of genetic polymorphisms*. *Nat Rev Genet* 3: 380-390.
- Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT and Bustillo J (2011) *An integrated semiconductor device enabling non-optical genome sequencing*. *Nature* 475: 348-352.
- Ruff C (2002) *Variation in human body size and shape*. *Annu Rev Anthropol* 31: 211-232.

- Ruff CB (1994) *Morphological adaptation to climate in modern and fossil hominids*. *Am J Phys Anthropol* 37: 65-107.
- Sabbagh A, Darlu P, Crouau-Roy B and Poloni ES (2011) *Arylamine N-acetyltransferase 2 (NAT2) genetic diversity and traditional subsistence: a worldwide population survey*. *PLoS ONE* 6: e18507.
- Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R and Lander ES (2002) *Detecting recent positive selection in the human genome from haplotype structure*. *Nature* 419: 832-837.
- Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D and Lander ES (2006) *Positive natural selection in the human lineage*. *Science* 312: 1614-1620.
- Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Sham PC, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Johnson TA, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Yakub I, Birren BW, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R and Stewart J (2007) *Genome-wide detection and characterization of positive selection in human populations*. *Nature* 449: 913-918.

- Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA and Arnheim N (1985) *Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia*. Science 230: 1350-1354.
- Saleh M, Vaillancourt JP, Graham RK, Huyck M, Srinivasula SM, Alnemri ES, Steinberg MH, Nolan V, Baldwin CT, Hotchkiss RS, Buchman TG, Zehnbauser BA, Hayden MR, Farrer LA, Roy S and Nicholson DW (2004) *Differential modulation of endotoxin responsiveness by human caspase-12 polymorphisms*. Nature 429: 75-79.
- Saleh M, Mathison JC, Wolinski MK, Bensinger SJ, Fitzgerald P, Droin N, Ulevitch RJ, Green DR and Nicholson DW (2006) *Enhanced bacterial clearance and sepsis resistance in caspase-12-deficient mice*. Nature 440: 1064-1068.
- Salo WL, Aufderheide AC, Buikstra J and Holcomb TA (1994) *Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy*. Proc Natl Acad Sci U S A 91: 2091-2094.
- Sanger F, Nicklen S and Coulson AR (1977) *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A 74: 5463-5467.
- Sawyer S, Krause J, Guschanski K, Savolainen V and Pääbo S (2012) *Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA*. PLoS ONE 7: e34131.
- Scherer D and Kumar R (2010) *Genetics of pigmentation in skin cancer--a review*. Mutat Res 705: 141-153.
- Scheu A (2012) *Palaeogenetische Studien zur Populationsgeschichte von Rind und Ziege mit einem Schwerpunkt auf dem Neolithikum in Südosteuropa*. VML Verlag Marie Leidorf, Rahden/Westf. Menschen - Kulturen - Traditionen, Studien aus den Forschungsclustern des Deutschen Archäologischen Instituts Band 4 (zugleich Dissertation Johannes Gutenberg-Universität Mainz 2011).
- Schinkel AH, Smit JJ, van Tellingen O, Beijnen JH, Wagenaar E, van Deemter L, Mol CA, van der Valk MA, Robanus-Maandag EC, te Riele HP and et al. (1994) *Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs*. Cell 77: 491-502.
- Schinkel AH (1997) *The physiological function of drug-transporting P-glycoproteins*. Semin Cancer Biol 8: 161-170.
- Schirmer M, Toliat MR, Haberl M, Suk A, Kamdem LK, Klein K, Brockmoller J, Nurnberg P, Zanger UM and Wojnowski L (2006) *Genetic signature consistent with selection against the CYP3A4*1B allele in non-African populations*. Pharmacogenet Genomics 16: 59-71.
- Schmidt-Ullrich R and Paus R (2005) *Molecular principles of hair follicle induction and morphogenesis*. Bioessays 27: 247-261.
- Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jager G, Bos KI, Herbig A, Economou C, Benjak A, Busso P, Nebel A, Boldsen JL, Kjellstrom A, Wu H, Stewart GR, Taylor GM, Bauer P, Lee OY, Wu HH, Minnikin DE, Besra GS, Tucker K, Roffey S, Sow SO, Cole ST, Nieselt K and Krause J (2013) *Genome-wide comparison of medieval and modern Mycobacterium leprae*. Science 341: 179-183.
- Schuetz EG, Umbenhauer DR, Yasuda K, Brimer C, Nguyen L, Relling MV, Schuetz JD and Schinkel AH (2000) *Altered expression of hepatic cytochromes P-450 in mice deficient in one or more mdr1 genes*. Mol Pharmacol 57: 188-197.

- Schulte W, Bernhagen J and Bucala R (2013) *Cytokines in sepsis: potent immunoregulators and potential therapeutic targets--an updated view*. *Mediators Inflamm* 2013: 165974.
- Schuster I (2011) *Cytochromes P450 are essential players in the vitamin D signaling system*. *Biochim Biophys Acta* 1814: 186-199.
- Seitz CS, Lin Q, Deng H and Khavari PA (1998) *Alterations in NF-kappaB function in transgenic epithelial tissue demonstrate a growth inhibitory role for NF-kappaB*. *Proc Natl Acad Sci U S A* 95: 2307-2312.
- Semiz S, Dujic T, Ostanek B, Prnjavorac B, Bego T, Malenica M, Mlinar B, Marc J and Causevic A (2011) *Analysis of CYP3A4*1B and CYP3A5*3 polymorphisms in population of Bosnia and Herzegovina*. *Med Glas (Zenica)* 8: 84-89.
- Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Mennecier P, Hofreiter M, Possnert G and Pääbo S (2004) *No evidence of Neandertal mtDNA contribution to early modern humans*. *PLoS Biol* 2: E57.
- Shapiro B and Hofreiter M (2014) *A paleogenomic perspective on evolution and gene function: new insights from ancient DNA*. *Science* 343: 1236573.
- Sharom FJ (2014) *Complex Interplay between the P-Glycoprotein Multidrug Efflux Pump and the Membrane: Its Role in Modulating Protein Function*. *Front Oncol* 4: 41.
- Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD and Church GM (2005) *Accurate multiplex polony sequencing of an evolved bacterial genome*. *Science* 309: 1728-1732.
- Shendure J and Ji H (2008) *Next-generation DNA sequencing*. *Nat Biotechnol* 26: 1135-1145.
- Shishlina NI (2001) *The Seasonal Cycle of Grassland Use in the Caspian Sea Steppe During the Bronze Age: A New Approach to an Old Problem*. *Eur J Archaeol* 4: 346-366.
- Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart CJ, McKeigue PM and Kittles RA (2003) *Skin pigmentation, biogeographical ancestry and admixture mapping*. *Hum Genet* 112: 387-399.
- Simon DB, Nelson-Williams C, Bia MJ, Ellison D, Karet FE, Molina AM, Vaara I, Iwata F, Cushner HM, Koolen M, Gainza FJ, Gittleman HJ and Lifton RP (1996) *Gitelman's variant of Bartter's syndrome, inherited hypokalaemic alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl cotransporter*. *Nat Genet* 12: 24-30.
- Sissung TM, Gardner ER, Piekarczyk RL, Howden R, Chen X, Woo S, Franke R, Clark JA, Miller-DeGraff L, Steinberg SM, Venzon D, Liewehr D, Kleeberger SR, Bates SE, Price DK, Rosing DR, Cabell C, Sparreboom A and Figg WD (2011) *Impact of ABCB1 allelic variants on QTc interval prolongation*. *Clin Cancer Res* 17: 937-946.
- Skoglund P, Malmstrom H, Raghavan M, Stora J, Hall P, Willerslev E, Gilbert MT, Gotherstrom A and Jakobsson M (2012) *Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe*. *Science* 336: 466-469.
- Slatkin M and Rannala B (2000) *Estimating allele age*. *Annu Rev Genomics Hum Genet* 1: 225-249.

- Smit JW, Huisman MT, van Tellingen O, Wiltshire HR and Schinkel AH (1999) *Absence or pharmacological blocking of placental P-glycoprotein profoundly increases fetal drug exposure*. *J Clin Invest* 104: 1441-1447.
- Smith G, Stubbins MJ, Harries LW and Wolf CR (1998) *Molecular genetics of the human cytochrome P450 monooxygenase superfamily*. *Xenobiotica* 28: 1129-1165.
- Smith JM and Haigh J (1974) *The hitch-hiking effect of a favourable gene*. *Genet Res* 23: 23-35.
- Soejima M, Tachida H, Ishida T, Sano A and Koda Y (2006) *Evidence for recent positive selection at the human AIM1 locus in a European population*. *Mol Biol Evol* 23: 179-188.
- Soejima M, Tachida H and Koda Y (2009) *Sequence analysis of human TRPV6 suggests positive selection outside Africa*. *Biochem Genet* 47: 147-153.
- Sokal RR, Oden NL and Wilson C (1991) *Genetic evidence for the spread of agriculture in Europe by demic diffusion*. *Nature* 351: 143-145.
- Sommer S (2005) *The importance of immune gene variability (MHC) in evolutionary ecology and conservation*. *Front Zool* 2: 16.
- Stajich JE and Hahn MW (2005) *Disentangling the effects of demography and selection in human history*. *Mol Biol Evol* 22: 63-73.
- Stebelsky I (2001) *Steppe*. In: Internet Encyclopedia of Ukraine. Studies CloU, Toronto. <http://www.encyclopediaofukraine.com/display.asp?linkpath=pages\S\T\Steppe.htm> (abgerufen am 27.06.2014).
- Stiller M, Green RE, Ronan M, Simons JF, Du L, He W, Egholm M, Rothberg JM, Keates SG, Ovodov ND, Antipina EE, Baryshnikov GF, Kuzmin YV, Vasilevski AA, Wuenschell GE, Termini J, Hofreiter M, Jaenicke-Despres V and Pääbo S (2006) *Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA*. *Proc Natl Acad Sci U S A* 103: 13578-13584.
- Stokowski RP, Pant PV, Dadd T, Fereday A, Hinds DA, Jarman C, Filsell W, Ginger RS, Green MR, van der Ouderaa FJ and Cox DR (2007) *A genomewide association study of skin pigmentation in a South Asian population*. *Am J Hum Genet* 81: 1119-1132.
- Stone AC and Stoneking M (1998) *mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World*. *Am J Hum Genet* 62: 1153-1170.
- Stone AC, Wilbur AK, Buikstra JE and Roberts CA (2009) *Tuberculosis and leprosy in perspective*. *Am J Phys Anthropol* 140 Suppl 49: 66-94.
- Sturm RA, Teasdale RD and Box NF (2001) *Human pigmentation genes: identification, structure and consequences of polymorphic variation*. *Gene* 277: 49-62.
- Sturm RA and Frudakis TN (2004) *Eye colour: portals into pigmentation genes and ancestry*. *Trends Genet* 20: 327-332.
- Sturm RA, Duffy DL, Zhao ZZ, Leite FP, Stark MS, Hayward NK, Martin NG and Montgomery GW (2008) *A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color*. *Am J Hum Genet* 82: 424-431.
- Sturm RA (2009) *Molecular genetics of human pigmentation diversity*. *Hum Mol Genet* 18: R9-17.

- Sugawara I, Kataoka I, Morishita Y, Hamada H, Tsuruo T, Itoyama S and Mori S (1988) *Tissue distribution of P-glycoprotein encoded by a multidrug-resistant gene as revealed by a monoclonal antibody, MRK 16*. *Cancer Res* 48: 1926-1929.
- Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, Jakobsdottir M, Steinberg S, Palsson S, Jonasson F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Kiemenev LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U and Stefansson K (2007) *Genetic determinants of hair, eye and skin pigmentation in Europeans*. *Nat Genet* 39: 1443-1452.
- Suzuki Y, Pasch A, Bonny O, Mohaupt MG, Hediger MA and Frey FJ (2008) *Gain-of-function haplotype in the epithelial calcium channel TRPV6 is a risk factor for renal calcium stone formation*. *Hum Mol Genet* 17: 1613-1618.
- Sverrisdóttir OO, Timpson A, Toombs J, Lecoœur C, Froguel P, Carretero JM, Arsuaga Ferreras JL, Gotherstrom A and Thomas MG (2014) *Direct estimates of natural selection in iberia indicate calcium absorption was not the only driver of lactase persistence in Europe*. *Mol Biol Evol* 31: 975-983.
- Swagerty DL, Jr., Walling AD and Klein RM (2002) *Lactose intolerance*. *Am Fam Physician* 65: 1845-1850.
- Szmyt M (2001) *The Absolute (Radiocarbon) Chronology of the Central and Eastern Groups of the Globular Amphora Culture*. In: Czebreszuk J and Müller J (Hrsg), *Die absolute Chronologie in Mitteleuropa 3000-2000 v. Chr. Studien zur Archäologie in Ostmitteleuropa 1*. VML Verlag Marie Leidorf, Rahden/Westf. 25-80.
- Tajima F (1983) *Evolutionary relationship of DNA sequences in finite populations*. *Genetics* 105: 437-460.
- Tajima F (1989) *Statistical method for testing the neutral mutation hypothesis by DNA polymorphism*. *Genetics* 123: 585-595.
- Tamura K and Nei M (1993) *Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees*. *Mol Biol Evol* 10: 512-526.
- Tan EK, Chan DK, Ng PW, Woo J, Teo YY, Tang K, Wong LP, Chong SS, Tan C, Shen H, Zhao Y and Lee CG (2005) *Effect of MDR1 haplotype on risk of Parkinson disease*. *Arch Neurol* 62: 460-464.
- Tang K, Ngoi SM, Gwee PC, Chua JM, Lee EJ, Chong SS and Lee CG (2002) *Distinct haplotype profiles and strong linkage disequilibrium at the MDR1 multidrug transporter gene locus in three ethnic Asian populations*. *Pharmacogenetics* 12: 437-450.
- Tang K, Wong LP, Lee EJ, Chong SS and Lee CG (2004) *Genomic evidence for recent positive selection at the human MDR1 gene locus*. *Hum Mol Genet* 13: 783-797.
- Tang K, Thornton KR and Stoneking M (2007) *A new approach for using genome scans to detect recent positive selection in the human genome*. *PLoS Biol* 5: e171.
- Templeton AR, Routman E and Phillips CA (1995) *Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum**. *Genetics* 140: 767-782.
- Templeton AR (1998) *Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history*. *Mol Ecol* 7: 381-397.

- Terberger T (2009) *Die Alt- und Mittelsteinzeit (1,3 Millionen Jahre - 4000 v. Chr.)*. In: Schnurbein Sv (Hrsg), Atlas der Vorgeschichte: Europa von den ersten Menschen bis Christi Geburt Konrad Theiss Verlag, Stuttgart. 10-57.
- Tewhey R, Warner JB, Nakano M, Libby B, Medkova M, David PH, Kotsopoulos SK, Samuels ML, Hutchison JB, Larson JW, Topol EJ, Weiner MP, Harismendy O, Olson J, Link DR and Frazer KA (2009) *Microdroplet-based PCR enrichment for large-scale targeted sequencing*. Nat Biotechnol 27: 1025-1031.
- Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL, Germonpre MB, Sablin MV, Lopez-Giraldez F, Domingo-Roura X, Napierala H, Uerpmann HP, Loponte DM, Acosta AA, Giemsch L, Schmitz RW, Worthington B, Buikstra JE, Druzhkova A, Graphodatsky AS, Ovodov ND, Wahlberg N, Freedman AH, Schweizer RM, Koepfli KP, Leonard JA, Meyer M, Krause J, Pääbo S, Green RE and Wayne RK (2013) *Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs*. Science 342: 871-874.
- The 1000 Genomes Project Consortium (2010) *A map of human genome variation from population-scale sequencing*. Nature 467: 1061-1073.
- The International HapMap 3 Consortium (2010) *Integrating common and rare genetic variation in diverse human populations*. Nature 467: 52-58.
- The International HapMap Consortium (2003) *The International HapMap Project*. Nature 426: 789-796.
- The International HapMap Consortium (2005) *A haplotype map of the human genome*. Nature 437: 1299-1320.
- The International HapMap Consortium (2007) *A second generation human haplotype map of over 3.1 million SNPs*. Nature 449: 851-861.
- The International SNP Map Working Group (2001) *A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms*. Nature 409: 928-933.
- Thiebaut F, Tsuruo T, Hamada H, Gottesman MM, Pastan I and Willingham MC (1987) *Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues*. Proc Natl Acad Sci U S A 84: 7735-7738.
- Thomas MG, Kivisild T, Chikhi L and Burger J (2013) *Europe and western Asia: genetics and population history*. In: Ness I and Bellwood P (Hrsg), The Encyclopedia of Global Human Migration. Wiley-Blackwell, Hoboken, NJ. Chapter 18.
- Thomasson HR, Edenberg HJ, Crabb DW, Mai XL, Jerome RE, Li TK, Wang SP, Lin YT, Lu RB and Yin SJ (1991) *Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men*. Am J Hum Genet 48: 677-681.
- Thomasson HR, Crabb DW, Edenberg HJ and Li TK (1993) *Alcohol and aldehyde dehydrogenase polymorphisms and alcoholism*. Behav Genet 23: 131-136.
- Thompson EE, Kuttub-Boulos H, Witonsky D, Yang L, Roe BA and Di Rienzo A (2004) *CYP3A variation and the evolution of salt-sensitivity variants*. Am J Hum Genet 75: 1059-1069.
- Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghorji J, Bumpstead S, Pritchard JK, Wray GA and Deloukas P (2007) *Convergent adaptation of human lactase persistence in Africa and Europe*. Nat Genet 39: 31-40.

- Tournamille C, Colin Y, Cartron JP and Le Van Kim C (1995a) *Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals*. Nat Genet 10: 224-228.
- Tournamille C, Le Van Kim C, Gane P, Cartron JP and Colin Y (1995b) *Molecular basis and PCR-DNA typing of the Fya/fyb blood group polymorphism*. Hum Genet 95: 407-410.
- Toyoda Y, Sakurai A, Mitani Y, Nakashima M, Yoshiura K, Nakagawa H, Sakai Y, Ota I, Lezhava A, Hayashizaki Y, Niikawa N and Ishikawa T (2009) *Earwax, osmidrosis, and breast cancer: why does one SNP (538G>A) in the human ABC transporter ABCB11 gene determine earwax type?* Faseb J 23: 2001-2013.
- Trinks A (2010) *Populationsgenetische Untersuchung spätbronze- und früheisenzeitlicher Skelettfunde von Camelus bactrianus aus West-Sibirien und Süd-Uzbekistan*. Magisterarbeit an der Johannes Gutenberg-Universität Mainz.
- Underhill PA and Kivisild T (2007) *Use of y chromosome and mitochondrial DNA population structure in tracing human migrations*. Annu Rev Genet 41: 539-564.
- Unterländer M (2014) *Populationsgenetik eisenzeitlicher Reiternomaden der Eurasischen Steppe*. Dissertation an der Johannes Gutenberg-Universität Mainz.
- Valenzuela RK, Henderson MS, Walsh MH, Garrison NA, Kelch JT, Cohen-Barak O, Erickson DT, John Meaney F, Bruce Walsh J, Cheng KC, Ito S, Wakamatsu K, Frudakis T, Thomas M and Brilliant MH (2010) *Predicting phenotype from genotype: normal pigmentation*. J Forensic Sci 55: 315-322.
- Vallee BL (1998) *Alcohol in the western world*. Sci Am 278: 80-85.
- Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A and Johnson SM (2008) *A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning*. Genome Res 18: 1051-1063.
- van Abel M, Hoenderop JG and Bindels RJ (2005) *The epithelial calcium channels TRPV5 and TRPV6: regulation and implications for disease*. Naunyn Schmiedebergs Arch Pharmacol 371: 295-306.
- van Oven M and Kayser M (2009) *Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation*. Hum Mutat 30: E386-394.
- Vatsis KP, Martell KJ and Weber WW (1991) *Diverse point mutations in the human gene for polymorphic N-acetyltransferase*. Proc Natl Acad Sci U S A 88: 6333-6337.
- Veeramah KR and Hammer MF (2014) *The impact of whole-genome sequencing on the reconstruction of human population history*. Nat Rev Genet 15: 149-162.
- Venables WN and Ripley BD (2002) *Modern Applied Statistics with S. Fourth Edition*. Springer, New York.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-

- Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A and Zhu X (2001) *The sequence of the human genome*. Science 291: 1304-1351.
- Verginelli F, Capelli C, Coia V, Musiani M, Falchetti M, Ottini L, Palmirota R, Tagliacozzo A, De Grossi Mazzorin I and Mariani-Costantini R (2005) *Mitochondrial DNA from prehistoric canids highlights relationships between dogs and South-East European wolves*. Mol Biol Evol 22: 2541-2551.
- Visser M, Kayser M and Palstra RJ (2012) *HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter*. Genome Res 22: 446-455.
- Voelkerding KV, Dames SA and Durtschi JD (2009) *Next-generation sequencing: from basic research to diagnostics*. Clin Chem 55: 641-658.
- Voight BF, Kudaravalli S, Wen X and Pritchard JK (2006) *A map of recent positive selection in the human genome*. PLoS Biol 4: e72.
- Vormfelde SV, Sehr D, Toliat MR, Schirmer M, Meineke I, Tzvetkov M, Nurnberg P and Brockmoller J (2007) *Genetic variation in the renal sodium transporters NKCC2, NCC, and ENaC in relation to the effects of loop diuretic drugs*. Clin Pharmacol Ther 82: 300-309.

- Wagner DM, Klunk J, Harbeck M, Devault A, Waglechner N, Sahl JW, Enk J, Birdsell DN, Kuch M, Lumibao C, Poinar D, Pearson T, Fourment M, Golding B, Riehm JM, Earn DJ, Dewitte S, Rouillard JM, Grupe G, Wiechmann I, Bliska JB, Keim PS, Scholz HC, Holmes EC and Poinar H (2014) *Yersinia pestis and the plague of Justinian 541-543 AD: a genomic analysis*. *Lancet Infect Dis* 14: 319-326.
- Wang H, Ding K, Zhang Y, Jin L, Kullo IJ and He F (2007a) *Comparative and evolutionary pharmacogenetics of ABCB1: complex signatures of positive selection on coding and regulatory regions*. *Pharmacogenet Genomics* 17: 667-678.
- Wang X, Grus WE and Zhang J (2006) *Gene losses during human origins*. *PLoS Biol* 4: e52.
- Wang Z, Wang J, Tantoso E, Wang B, Tai AY, Ooi LL, Chong SS and Lee CG (2007b) *Signatures of recent positive selection at the ATP-binding cassette drug transporter superfamily gene loci*. *Hum Mol Genet* 16: 1367-1380.
- Wang Z, Schuetz EG, Xu Y and Thummel KE (2013) *Interplay between vitamin D and the drug metabolizing enzyme CYP3A4*. *J Steroid Biochem Mol Biol* 136: 54-58.
- Warnock DG and Eveloff J (1982) *NaCl entry mechanisms in the luminal membrane of the renal tubule*. *Am J Physiol* 242: F561-574.
- Watlington CO, Atkins JL, McNeil JS, Grogan WM and Johnson JP (1988) *Corticosterone is converted to 6 beta-hydroxycorticosterone in rat: effects of the metabolite on urinary electrolyte excretion*. *J Steroid Biochem* 31: 947-954.
- Watson CL and Lockwood DN (2009) *Single nucleotide polymorphism analysis of European archaeological *M. leprae* DNA*. *PLoS One* 4: e7547.
- Watterson GA (1975) *On the number of segregating sites in genetical models without recombination*. *Theor Popul Biol* 7: 256-276.
- Wechler K-P (1998) *Zum Neolithikum der osteuropäischen Steppe und Waldsteppe (Dnestr-Donetz-Gebiet)*. In: Hänsel B and Machnik J (Hrsg), *Das Karpatenbecken und die osteuropäische Steppe*. Südosteuropa-Schriften, Band 20. VML Verlag Marie Leidorf, Rahden/Westf. 71-89.
- Wechler K-P (2001) *Studien zum Neolithikum der osteuropäischen Steppe*. Philipp von Zabern, Mainz. DAI. Archäologie in Eurasien Band 12.
- Weinberg W (1908) *Über Vererbungsgesetze beim Menschen*. *Zeitschrift für Induktive Abstammungs- und Vererbungslehre* 1: 440-460.
- Weinstock J, Willerslev E, Sher A, Tong W, Ho SY, Rubenstein D, Storer J, Burns J, Martin L, Bravi C, Prieto A, Froese D, Scott E, Xulong L and Cooper A (2005) *Evolution, systematics, and phylogeography of pleistocene horses in the new world: a molecular perspective*. *PLoS Biol* 3: e241.
- Weir BS and Cockerham CC (1984) *Estimating F-Statistics for the Analysis of Population Structure*. *Evolution* 38: pp. 1358-1370.
- Weisdorf JL (2005) *From Foraging To Farming: Explaining The Neolithic Revolution*. *J Econ Surv* 19: 561-586.
- Weiss G and von Haeseler A (1998) *Inference of population history using a likelihood approach*. *Genetics* 149: 1539-1546.
- Westropp HM (1872) *Pre-historic phases; or, Introductory essays on pre-historic archaeology*. Bell & Daldy, London.
- Wharton B and Bishop N (2003) *Rickets*. *Lancet* 362: 1389-1400.
- Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski

- JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA and Rothberg JM (2008) *The complete genome of an individual by massively parallel DNA sequencing*. Nature 452: 872-876.
- Wigginton JE, Cutler DJ and Abecasis GR (2005) *A note on exact tests of Hardy-Weinberg equilibrium*. Am J Hum Genet 76: 887-893.
- Wilde S, Timpson A, Kirsanow K, Kaiser E, Kayser M, Unterlander M, Hollfelder N, Potekhina ID, Schier W, Thomas MG and Burger J (2014) *Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y*. Proc Natl Acad Sci U S A 111: 4832-4837.
- Willerslev E, Gilbert MT, Binladen J, Ho SY, Campos PF, Ratan A, Tomsho LP, da Fonseca RR, Sher A, Kuznetsova TV, Nowak-Kemp M, Roth TL, Miller W and Schuster SC (2009) *Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution*. BMC Evol Biol 9: 95.
- Wolff PH (1972) *Ethnic differences in alcohol sensitivity*. Science 175: 449-450.
- Woodward SR, Weyand NJ and Bunnell M (1994) *DNA sequence from Cretaceous period bone fragments*. Science 266: 1229-1232.
- Wright S (1951) *The genetical structure of populations*. Ann Eugen 15: 323-354.
- Wrighton SA, Brian WR, Sari MA, Iwasaki M, Guengerich FP, Raucy JL, Molowa DT and Vandenbranden M (1990) *Studies on the expression and metabolic capabilities of human liver cytochrome P450III_{A5} (HLP₃)*. Mol Pharmacol 38: 207-213.
- Wuren T, Simonson TS, Qin G, Xing J, Huff CD, Witherspoon DJ, Jorde LB and Ge RL (2014) *Shared and unique signals of high-altitude adaptation in geographically distinct Tibetan populations*. PLoS One 9: e88252.
- Xi B, Wang C, Liu L, Zeng T, Liang Y, Li J and Mi J (2011) *Association of the CYP3A5 polymorphism (6986G>A) with blood pressure and hypertension*. Hypertens Res 34: 1216-1220.
- Xue Y, Daly A, Yngvadottir B, Liu M, Coop G, Kim Y, Sabeti P, Chen Y, Stalker J, Huckle E, Burton J, Leonard S, Rogers J and Tyler-Smith C (2006) *Spread of an inactive form of caspase-12 in humans is due to recent positive selection*. Am J Hum Genet 78: 659-670.
- Xue Y, Zhang X, Huang N, Daly A, Gillson CJ, Macarthur DG, Yngvadottir B, Nica AC, Woodruff C, Chen Y, Conrad DF, Ayub Q, Mehdi SQ, Li P and Tyler-Smith C (2009) *Population differentiation as an indicator of recent positive selection in humans: an empirical evaluation*. Genetics 183: 1065-1077.
- Yabuuchi H, Shimizu H, Takayanagi S and Ishikawa T (2001) *Multiple splicing variants of two new human ATP-binding cassette transporters, ABCC11 and ABCC12*. Biochem Biophys Res Commun 288: 933-939.
- Yan M, Wang LC, Hymowitz SG, Schilbach S, Lee J, Goddard A, de Vos AM, Gao WQ and Dixit VM (2000) *Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors*. Science 290: 523-527.
- Yeretssian G, Doiron K, Shao W, Leavitt BR, Hayden MR, Nicholson DW and Saleh M (2009) *Gender differences in expression of the human caspase-12 long variant determines susceptibility to Listeria monocytogenes infection*. Proc Natl Acad Sci U S A 106: 9016-9020.

- Yngvadottir B, Xue Y, Searle S, Hunt S, Delgado M, Morrison J, Whittaker P, Deloukas P and Tyler-Smith C (2009) *A genome-wide survey of the prevalence and evolutionary forces acting on human nonsense SNPs*. *Am J Hum Genet* 84: 224-234.
- Yoshiura K, Kinoshita A, Ishida T, Ninokata A, Ishikawa T, Kaname T, Bannai M, Tokunaga K, Sonoda S, Komaki R, Ihara M, Saenko VA, Alipov GK, Sekine I, Komatsu K, Takahashi H, Nakashima M, Sosonkina N, Mapendano CK, Ghadami M, Nomura M, Liang DS, Miwa N, Kim DK, Garidkhuu A, Natsume N, Ohta T, Tomita H, Kaneko A, Kikuchi M, Russomando G, Hirayama K, Ishibashi M, Takahashi A, Saitou N, Murray JC, Saito S, Nakamura Y and Niikawa N (2006) *A SNP in the ABCC11 gene is the determinant of human earwax type*. *Nat Genet* 38: 324-330.
- Zang Y, Zhao S, Doll MA, States JC and Hein DW (2004) *The T341C (Ile114Thr) polymorphism of N-acetyltransferase 2 yields slow acetylator phenotype by enhanced protein degradation*. *Pharmacogenetics* 14: 717-723.
- Zeder MA (2008) *Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact*. *Proc Natl Acad Sci U S A* 105: 11597-11604.
- Zeder MA (2011) *The Origins of Agriculture in the Near East*. *Curr Anthropol* 52: S221-S235.
- Zhang L, Miyaki K, Wang W and Muramatsu M (2010) *CYP3A5 polymorphism and sensitivity of blood pressure to dietary salt in Japanese men*. *J Hum Hypertens* 24: 345-350.
- Ziegler EE and Fomon SJ (1983) *Lactose enhances mineral absorption in infancy*. *J Pediatr Gastroenterol Nutr* 2: 288-294.
- Zittermann A, Bock P, Drummer C, Scheld K, Heer M and Stehle P (2000) *Lactose does not enhance calcium bioavailability in lactose-tolerant, healthy adults*. *Am J Clin Nutr* 71: 931-936.
- Zvelebil M (2000) *The social context of the agricultural transition in Europe*. In: Renfrew C and Boyle K (Hrsg), *Archaeogenetics: DNA and the population prehistory of Europe*. McDonald Institute for Archaeological Research, Cambridge. 57-79.

13 Anhang

13.1 Labormaterialien

Tabelle 13.1: Verwendete Reagenzien und sonstige Chemikalien

Chemikalie	Hersteller
50bp Ladder Gene Ruler	Fermentas - Thermo Fisher Scientific
Agarose UltraPure™	Invitrogen Corp. - Life Technologies Corp.
Alconox	Sigma-Aldrich
AmpliAq Gold DNA Polymerase	Applied Biosystems Inc. (Roche)
ATP (Adenosin 5'-triphosphat)	Fermentas - Thermo Fisher Scientific
Big Dye™ Terminator	Applied Biosystems Inc. (Roche)
5x Sequencing Buffer	
Big Dye™ Terminator	Applied Biosystems Inc. (Roche)
Cycle Sequencing 3.1 Kit	
Bovine Serum Albumin (BSA)	Roche Diagnostics GmbH, Mannheim
Bromphenolblau	Fermentas - Thermo Fisher Scientific
Bst DNA Polymerase Large Fragment	New England BioLabs Inc., Ipswich, MA (USA)
Buffer Tango	Fermentas - Thermo Fisher Scientific
Calf Intestine Phosphatase (CIP)	New England BioLabs Inc., Ipswich, MA (USA)
DanKlorix Hygienereiniger	Colgate-Palmolive AG
DNA-ExitusPlus™ IF	AppliChem GmbH, Darmstadt
dNTP-Mix, PCR-Grade	Qiagen GmbH, Hilden
EDTA pH 8	Ambion - Life Technologies
Ethanol (≥ 99,8 % p.a.)	Carl Roth GmbH & Co. KG, Karlsruhe
Ethanol (≥ 99,8 % vergällt)	Carl Roth GmbH & Co. KG, Karlsruhe
Ethidiumbromid	Carl Roth GmbH & Co. KG, Karlsruhe
Exonuclease I (ExoI)	Fermentas - Thermo Fisher Scientific
FastAP Thermosensitive Alkaline Phosphatase	Fermentas - Thermo Fisher Scientific
Gene Ruler™ 50 bp DNA Ladder	Fermentas - Thermo Fisher Scientific
GeneAmp 10X PCR Gold Buffer	Applied Biosystems Inc. (Roche)
Hi-Di™ Formamide	Applied Biosystems Inc. (Roche)
HPLC-Wasser	Acros Organics - Thermo Fisher Scientific
Hydroxylapatit C13-17 MV 1792	Chemische Fabrik Budenheim KG, Budenheim
Invisorb Spin Swab Kit	Invitex GmbH (Stratec), Berlin
Loading Dye Solution	Fermentas - Thermo Fisher Scientific
MgCl ₂ Solution 25 mM	Applied Biosystems Inc. (Roche)
MinElute® PCR Purification Kit	Qiagen GmbH, Hilden
MSB Spin PCRapace Kit	Invitex GmbH (Stratec), Berlin
N-Laurylsarcosin	Merck KGaA, Darmstadt
Oligonukleotide (Primer, Meyer-tags)	BioSpring GmbH, Frankfurt/Main
PEG 4000	Fermentas - Thermo Fisher Scientific

Chemikalie	Hersteller
POP-6™ Polymer	Applied Biosystems Inc. (Roche)
Proteinase K recombinant, PCR Grade	Roche Diagnostics GmbH, Mannheim
Puffer für ABI PRISM® 3130 Genetic Analyzer	Applied Biosystems Inc. (Roche)
Qiagen Multiplex PCR Kit	Qiagen GmbH, Hilden
Qubit® dsDNA HS Assay Kit	Invitrogen Corp. - Life Technologies Corp.
Roti®-Phenol/Chloroform/Isoamylalkohol	Carl Roth GmbH & Co. KG, Karlsruhe
Rotipuran® ≥ 99 %, p.a. Chloroform	Carl Roth GmbH & Co. KG, Karlsruhe
Seesand p.a., gereinigt und geglüht	Carl Roth GmbH & Co. KG, Karlsruhe
Sephadex G-50 FINE	GE Healthcare Europe GmbH: Life Sciences
Shrimp Alkaline Phosphatase (SAP)	Fermentas - Thermo Fisher Scientific
SrfI Restriction Enzyme	Stratagene - Agilent Technologies Inc.
T4 DNA Polymerase	Fermentas - Thermo Fisher Scientific
T4 Ligase	Fermentas - Thermo Fisher Scientific
T4 Polynukleotidkinase	Fermentas - Thermo Fisher Scientific
TE-Buffer	Sigma-Aldrich
Tween 20	Carl Roth GmbH & Co. KG, Karlsruhe

Table 13.2: Sonstige verwendete Verbrauchsmaterialien

Material	Hersteller
Amicon® Ultra-15 (50 kDa)	Millipore - Merck Millipore
Astronautenhauben	Hansa Medical, Kirchhoff Group
chirurgischer Mundschutz	Hansa Medical, Kirchhoff Group
Diamanttrennscheiben	Dremel Corp., Horico Dental GmbH & Cie
Einweg-Pipettenspitzen	Sarstedt AG & Co., Nümbrecht Thermo Fisher Scientific
Einweg-Pipettenspitzen mit Filter	Abimed GmbH, Langenfeld Thermo Fisher Scientific
Einweg-Untersuchungshandschuhe (Latex, Vinyl, Nitril)	Hansa Medical, Kirchhoff Group MaiMed GmbH, Neuenkirchen
Falcon-Gefäße: 15 ml, 50 ml	Sarstedt AG & Co., Nümbrecht
Helme und Visiere	Carl Roth GmbH & Co. KG, Karlsruhe
Klebefolienmatten	Purus International Inc. (über Basan GmbH, VWR International GmbH)
Kosmetiktücher	Kimberly-Clark Inc.
Küchenrollen	Verschiedene Händler
MultiScreen-Platten	Millipore - Merck Millipore
Reaktionsgefäße: 0,5 ml, 1,5 ml (DNA LoBind Tubes)	Eppendorf AG, Hamburg
Reaktionsgefäße: 0,5 ml, 1,5 ml, 2 ml (Safe-Lock Tubes)	Eppendorf AG, Hamburg
Reaktionsgefäße: 200 µl	Biozym Scientific GmbH, Hessisch Oldendorf

Material	Hersteller
Spezial-Edelkorund, 250µ	Harnisch+Rieth GmbH & Co. KG, Winterbach
Spezial-Edelkorund, 50µ	Harnisch+Rieth GmbH & Co. KG, Winterbach
Thermo-fast 96 PCR Detection AB-1100	ABgene – Thermo Fisher Scientific
Tyvek® Comfort Einweg-Schuhschutz	DuPont AG (über Basan GmbH, VWR International GmbH)
Tyvek® Overall mit Kapuze, Modell Classic	DuPont AG (über Basan GmbH, VWR International GmbH)

Tabelle 13.3: Verwendete Laborgeräte

Gerät	Hersteller
ABI PRISM® 3130 Genetic Analyzer	Applied Biosystems Inc.
Brutschrank Heraeus/ Typ B 5042	Heraeus Holding GmbH, Hanau (über Carl Roth GmbH & Co. KG, Karlsruhe)
Column Loader	Millipore - Merck Millipore
Einplatz-Absauganlage D-LE 260S	Harnisch+Rieth GmbH & Co. KG, Winterbach
Eppendorf 5415/R Zentrifuge	Eppendorf AG, Hamburg
Geldokumentationsanlage	INTAS Science Imaging Instruments GmbH, Göttingen
Gelkammern	Carl Roth GmbH & Co. KG, Karlsruhe
Helme und Visiere	Carl Roth GmbH & Co. KG, Karlsruhe
Jouan GR.142 Zentrifuge	Jouan SA, Saint Herblain, Frankreich
Kugelschwingmühle MM200	Retsch GmbH, Haan
Mastercycler	Eppendorf AG, Hamburg
Mastercycler (gradient)	Eppendorf AG, Hamburg
Mikromotor Elekter Emax NK-261	MAFRA Maile + Hass GmbH, Schutterwald
Osmoseanlage	IEM GmbH Mainz
Pipetten	Abimed GmbH, Langenfeld Eppendorf AG, Hamburg
Präzisionswaage EMB 2000	Kern & Sohn GmbH, Balingen
Punktstrahlgerät P-G 400	Harnisch+Rieth GmbH & Co. KG, Winterbach
Qubit® Fluorometer	Invitrogen Corp. - Life Technologies Corp.
Rotationssäge KaVo Typ 10 KWL	KaVo Dental GmbH, Biberach/Riß
Schnellspannhandstück IH-300	MAFRA Maile + Hass GmbH, Schutterwald
Spannungsgerät Consort / E 143, S/N 60441	Consort bvba, Turnhout, Belgien (über Carl Roth GmbH & Co. KG, Karlsruhe)
TPersonal Thermocycler	Biometra GmbH, Göttingen
UV-Boxen (Bestrahlung, Arbeiten)	Eigenbau
UV-Tauchlampe,UV -Handlampe	UV Consulting Peschl e.K., Mainz
Vortexer	Carl Roth GmbH & Co. KG, Karlsruhe
Wippschüttler	Carl Roth GmbH & Co. KG, Karlsruhe
Zentrifuge Rotina 380	Hettich GmbH & Co. KG, Tuttlingen

13.2 F_{ST} -Werte

Aufgrund der zum Teil sehr kleinen Individuenzahlen und Datenlücken, z. B. der Jamnaja-Gruppen in den östlichen Gebieten, wurden die Jamnaja- und die Katakombengrabbkultur zunächst zu einer Population zusammengefasst und verschiedene, archäologisch festgestellte Strukturen überprüft.

In Tabelle 13.4 wird die bronzezeitliche Bevölkerung auf Spuren der kulturell beobachteten Strukturen des Neolithikums untersucht, in Tabelle 13.5 und Tabelle 13.6 auf Spuren aus dem frühen bzw. mittleren Äneolithikum.

In rein geographische Gruppen wurde die zusammengefasste Population der Bronzezeit in den folgenden Matrizen unterteilt. In Tabelle 13.7 beinhalten die Gruppen Fundplätze aus den nebeneinander angeordneten Gebieten zwischen den großen Flüssen, die den Untersuchungsraum unterteilen. Tabelle 13.8 fasst die entlang der einzelnen Flüsse gelegenen Fundplätze zusammen. Tabelle 13.9 unterteilt den Untersuchungsraum jeweils in ein West- und ein Ost-Gebiet, wobei drei verschiedene Grenzen (Südlicher Bug, Dnjepr und Don) getestet werden.

Die Matrix in Tabelle 13.10 zeigt die paarweisen genetischen Distanzen zwischen allen steppeninternen Gruppierungen mit kultureller Differenzierung. Die Gruppennbildung basierte auf geographischen, landschaftlichen und populationsgeschichtlichen Aspekten

Tabelle 13.11 ist eine Ergänzung zu Tabelle 9.17 und enthält die paarweisen genetischen Distanzen zwischen allen betrachteten DNA-Datensätzen Westeurasiens.



Abbildung 13.1: Karte der Fundplatz-Einheiten

Es handelt sich hier um ein Duplikat der Abbildung 9.9 zur Zuordnung und Lokalisation der Fundplatz-Einheiten in den folgenden F_{ST} -Matrizen (Tabelle 13.10 f).

Tabelle 13.4: F_{ST} -Werte zwischen den Verbreitungsgebieten neolithischer Kulturen

Die Individuen der Jamnaja- und der Katakombengrabkultur wurden zu einer bronzzeitlichen Population zusammengefasst, und eine Aufteilung in Gruppen entsprechend der Verbreitungsgebiete neolithischer Kulturen nach Rassamakin (1999) und Kapitel 4.2.1. vorgenommen. N = Anzahl der Individuen, FPE = Fundplatz-Einheiten nach Abbildung 13.1. Allen Zahlenwerten (außer den Individuenzahlen) in der Matrix ist der Zusatz „ $\times 10^{-5}$ “ anzuhängen. Die Werte unterhalb der Diagonale der Matrix geben die paarweisen F_{ST} -Werte, basierend auf paarweisen Unterschieden, an. In grauer Schrift sind Vergleiche zwischen Gruppen angegeben, bei denen ein Teil der Individuen in beiden Populationen vorkommen. Fettgedruckt sind signifikante F_{ST} -Werte ($p < 0,05$). In der oberen Diagonale sind die p -Werte, basierend auf 10 000 Permutationen, aufgeführt. Die Bug-Dnjestr- und die Sursker-Kultur sind lediglich der Vollständigkeit halber aufgeführt. Ihre Stichprobengrößen sind zu klein für eine Auswertung.

Gruppen	N	FPE	Südost-europa	Bug-Dnjestr	Dnjepr-Donetz	Sursk	Seroglazov
			2	3,4	5,6,8	7	9,10,11
Neolithikum Südosteuropa	6	2		64033 ± 460	27284 ± 460	8009 ± 290	57668 ± 450
Bug-Dnjestr-Kultur	5	3,4	-3448		51817 ± 500	8593 ± 290	70399 ± 470
Dnjepr-Donetz-Kultur	16	5,6,8	2168	-989		129 ± 30	58351 ± 490
Sursker Kultur	5	7	13413	16667	27800		5118 ± 230
Seroglazov-Kultur	19	9,10,11	-1280	-2988	-834	10862	

Tabelle 13.5: F_{ST} -Werte zwischen den Verbreitungsgebieten der Kulturen des frühen Äneolithikums

Die Individuen der Jamnaja- und der Katakombengrabkultur wurden zu einer bronzezeitlichen Population zusammengefasst, und eine Aufteilung in Gruppen entsprechend der Verbreitungsgebiete neolithischer Kulturen nach Rassamakin (1999) und Kapitel 4.2.2. Für weitere Informationen siehe Tabelle 13.4. Die Suvorovo-Gruppe ist lediglich der Vollständigkeit halber aufgeführt. Ihre Stichprobengröße ist zu klein für eine Auswertung.

Gruppen			Varna	Suvorovo	Skelanska	Svbodnoe	Khvalynsk
	N	FPE	2	3,4	5,6,7,8	9,10	11
Varna-Kultur	6	2		63499 ± 450	40422 ± 520	59756 ± 440	59450 ± 510
Suvorovo-Gruppe	5	3,4	-3448		50292 ± 420	56826 ± 510	79151 ± 470
Skelanska-Kultur	21	5,6,7,8	335	-866		96010 ± 220	31878 ± 430
Svbodnoe-Kultur	13	9,10	-1486	-1328	-3867		25295 ± 350
Khvalynsk-Kultur	6	11	-556	-4666	1327	2492	

Tabelle 13.6: F_{ST} -Werte zwischen den Verbreitungsgebieten der Kulturen des mittleren Äneolithikums

Die Individuen der Jamnaja- und der Katakombengrabkultur wurden zu einer bronzezeitlichen Population zusammengefasst, und eine Aufteilung in Gruppen entsprechend der Verbreitungsgebiete neolithischer Kulturen nach Rassamakin (1999) und Kapitel 4.2.2. Für weitere Informationen siehe Tabelle 13.4.

Gruppen			Cernavodă	U. Mikhail.	Kvitanska	Majkop	Khvalynsk
	N	FPE	2	4,7	3,5,6,8	9,10	11
Cernavodă-Kultur	8	2		31066 ± 450	15672 ± 350	60261 ± 470	59875 ± 440
U. Mikhailovka-Kultur	8	4,7	2013		495 ± 80	18949 ± 390	77537 ± 430
Kvitanska-Kultur	18	3,5,6,8	3586	13062		78052 ± 380	3247 ± 170
Majkop-Kultur	13	9,10	-1486	3828	-2233		24116 ± 430
Khvalynsk-Kultur	6	11	-556	-5803	8870	2492	

Tabelle 13.7: F_{ST} -Werte zwischen den Zwischenstromgebieten

Die Individuen der Jamnaja- und der Katakombengrabkultur wurden zu einer bronzezeitlichen Population zusammengefasst, und eine Aufteilung in geographische Gruppen mit den großen Flüssen des Untersuchungsgebiets als Grenzen. Für weitere Informationen siehe Tabelle 13.4.

Gruppen			Donau	Pruth-Dnjepr	Dnjepr-Don	Kuban-Manych	Wolga
	N	FPE	2	3,4,5	6,7,8	9,10	11
Donau	6	2		49975 ± 470	31343 ± 440	59073 ± 490	5947 9 ± 490
Pruth-Dnjepr	10	3,4,5	-553		12286 ± 330	39927 ± 510	24829 ± 460
Dnjepr-Don	16	6,7,8	1383	3696		87882 ± 310	42144 ± 530
Kuban-Manych	13	9,10	-1486	292	-3607		25245 ± 440
Wolga	6	11	-556	2807	55	2492	

Tabelle 13.8: F_{ST} -Werte zwischen Gruppen entlang der Flüsse

Die Individuen der Jamnaja- und der Katakombengrabkultur wurden zu einer bronzezeitlichen Population zusammengefasst, und eine Aufteilung in geographische Gruppen entlang der großen Flüsse des Untersuchungsgebiets. In grauer Schrift sind Vergleiche zwischen Gruppen angegeben, bei denen ein Teil der Individuen in beiden Populationen vorkommen. Für weitere Informationen siehe Tabelle 13.4. Die Dnjestr/Bug-Gruppe ist lediglich der Vollständigkeit halber aufgeführt. Ihre Stichprobengröße ist zu klein für eine Auswertung.

Gruppen			Donau	Dnjestr/Pruth	Dnjepr	Donez	Kuban/Manych	Manych	Wolga
	N	FPE	2	3,4	5,6,7	8	9,10	10	11
Donau	6	2		64360 ± 480	38313 ± 510	21929 ± 400	59618 ± 530	58430 ± 440	60192 ± 430
Dnjestr/Pruth	5	3,4	-3448		31343 ± 470	44035 ± 460	57390 ± 530	69389 ± 440	78943 ± 430
Dnjepr	13	5,6,7	624	2053		4336 ± 170	58014 ± 500	28918 ± 450	51886 ± 490
Donez	8	8	4664	-890	8637		52569 ± 520	83635 ± 350	12365 ± 330
Kuban, Manych	13	9,10	-1486	-1328	-913	-2085		99673 ± 50	24889 ± 410
Manych	11	10	-1294	-3071	1834	-5374	-7442		10940 ± 270
Wolga	6	11	-556	-4666	-1512	8767	2492	5489	

Tabelle 13.9: F_{ST} -Werte zwischen östlichen und westlichen Gruppen

Die Individuen der Jamnaja- und der Katakombengrabbkultur wurden zu einer bronzezeitlichen Population zusammengefasst, und Aufteilungen in West- und Ostgruppen mit verschiedenen Grenzflüssen. In grauer Schrift sind Vergleiche zwischen Gruppen angegeben, bei denen ein Teil der Individuen in beiden Populationen vorkommen. Für weitere Informationen siehe Tabelle 13.4.

Gruppen			Südl. Bug West	Südl. Bug Ost	Dnjepr West	Dnjepr Ost	Don West	Don Ost
	N	FPE	2-4	5-11	2-5	6-11	2-8	9-11
Südl. Bug West	11	2-4		40036 ± 450	99990 ± 0	32918 ± 480	96703 ± 170	69785 ± 430
Südl. Bug Ost	40	5-11	23		28393 ± 430	99990 ± 0	99743 ± 50	99950 ± 20
Dnjepr West	16	2-5	-7214	717		12949 ± 310	94476 ± 260	45065 ± 530
Dnejr Ost	35	6-11	602	-2542	1981		94614 ± 230	99970 ± 10
Don West	32	2-8	-3439	-2003	-2394	-1613		96000 ± 190
Don Ost	19	9-11	-1537	-3384	-114	-3567	-2232	

Tabelle 13.10: Steppeninterne F_{ST} -Werte

Die Tabelle befindet sich auf den folgenden vier Seiten. Allen Zahlenwerten (außer den Individuenzahlen) in der Matrix ist der Zusatz „ $\times 10^{-5}$ “ anzuhängen. Die Werte unterhalb der Diagonale der Matrix geben die paarweisen F_{ST} -Werte, basierend auf paarweisen Unterschieden, an. In grauer Schrift sind Vergleiche zwischen Gruppen angegeben, bei denen ein Teil der Individuen in beiden Populationen vorkommen. Fettgedruckt sind signifikante F_{ST} -Werte ($p < 0,05$). In der oberen Diagonale sind die p -Werte, basierend auf 10 000 Permutationen, aufgeführt.

Es sind Distanzen zwischen temporalen und regionalen Gruppierungen angegeben. Zeile 1 und Spalte 1 geben die Zeitschichten an. Ä1-Ä2: äneolithische Gruppierungen (ÄNL); J1-J25: Jamnaja-Gruppierungen (JAM); K1-K18: Gruppierungen der (frühen/entwickelten) Katakombengrabbkultur ((f/e)KGK); P1: Poltavka-Kultur (POL). Spalte 2 gibt die Fundplätze an, aus der die Proben der jeweiligen Gruppe stammen, gemäß der Nummerierung der Fundplatz-Einheiten (FPE) in Abbildung 13.1. Spalte 3 gibt die grobe regionale Einordnung der Gruppen, sowie die Anzahl der Individuen (N) an. S.Bug = Südlicher Bug; W, NW, N, NO = Gebiete basierend auf den entsprechenden Himmelsrichtungen.

Tabelle 13.10: Steppeninterne F_{ST}-Werte (S. 1 von 4) Nähere Informationen auf voriger Seite.

Ä1	ÄNL	gesamt 10		93456 ± 240	52094 ± 460	52807 ± 440	56905 ± 500	28552 ± 470	33729 ± 420	24611 ± 420	25661 ± 440	46055 ± 450	30225 ± 470	25968 ± 440
Ä2	2,4,5,7	Steppe 6	-10278		23730 ± 440	44807 ± 510	42521 ± 490	15315 ± 350	12157 ± 300	15236 ± 420	10801 ± 310	24235 ± 400	16751 ± 400	13048 ± 320
J1	JAM	gesamt 22	-794	2574		94416 ± 230	97307 ± 170	94050 ± 230	99990 ± 0	84883 ± 360	99901 ± 30	99980 ± 10	85477 ± 350	99891 ± 30
J2	2	Donau 6	-1659	645	-5630		99990 ± 0	99772 ± 50	95080 ± 220	75240 ± 370	49540 ± 470	59172 ± 470	77982 ± 410	50619 ± 490
J3	2,3,4	Donau-Dnjestr 10	-1389	-47	-4242	-12367		99990 ± 0	97099 ± 170	98079 ± 160	65637 ± 460	68795 ± 470	97317 ± 170	58697 ± 450
J4	2,3,4,5	Donau-Dnjepr 14	1306	4725	-3093	-9161	-8296		97119 ± 160	99812 ± 50	68706 ± 510	63093 ± 420	99881 ± 30	60350 ± 400
J5	2,3,4,5, 6,7	Donau-Don 19	641	4980	-4722	-6077	-4786	-4021		91793 ± 260	99960 ± 20	99762 ± 50	91377 ± 270	99911 ± 30
J6	3,4,5	Pruth-Dnjepr 8	2193	5958	-3241	-3610	-7432	-8307	-4370		82705 ± 390	75785 ± 390	99990 ± 0	75022 ± 450
J7	3,4,5,6, 7	Pruth-Don 13	1789	6765	-4711	-785	-1912	-1806	-5617	-4515		99990 ± 0	79537 ± 400	99990 ± 0
J8	3,4,5,6, 7,9,10	Pruth-Wolga 16	-534	3006	-4869	-1444	-1894	-1138	-4560	-3018	-6543		74478 ± 460	99990 ± 0
J9	4,5	Dnjestr-Dnjepr 7	1599	4706	-3646	-4616	-7624	-8761	-4700	-14264	-4531	-3221		77735 ± 400
J10	4,5,6,7	Dnjestr-Don 12	1723	6456	-4788	-808	-1336	-1268	-5621	-3517	-8352	-6660	-4582	
J11	4,5,6,7, 8,10	Dnjestr-Wolga 15	-652	2597	-4862	-1419	-1370	-607	-4420	-2109	-6368	-6688	-3093	-7026
J12	5,6,7	S.Bug-Don 9	5589	12596	-3445	2497	3576	3017	-4338	2280	-7866	-5853	1875	-8698
J13	5,6,7,9	S.Bug-Kuban 11	5553	12227	-2580	3462	5138	4972	-2164	4794	-5118	-5162	4515	-5857
J14	5,6,7, 9,10	S.Bug-Wolga 12	1133	5839	-4340	335	1774	2192	-3654	1748	-6018	-6595	1202	-6727
J15	6,7,9	Dnjepr-Kuban 7	12452	17856	4218	12634	14564	16841	5936	18860	2334	539	18957	1190
J16	6,7,9, 10	Dnjepr-Wolga 8	4281	7139	-225	6099	7945	10752	1685	11933	-1299	-3509	11614	-2394
J17	2,4,7,9	Küste 14	383	2763	-4451	-5627	-2732	-176	-3845	1242	-3776	-4710	299	-4508
J18	2,4	Küste W 9	-1748	-1146	-4236	-13621	-11209	-8186	-4651	-6364	-1251	-1519	-7967	-1411
J19	2,4,7	Küste NW 12	82	1922	-4865	-7427	-4781	-1935	-5112	-453	-4767	-4593	-1615	-5552
J20	4,7,9	Küste NO 8	5984	7904	-313	5561	6149	9200	1263	9106	-1982	-3649	8188	-3282
J21	4,7	Küste N 6	6897	7647	-106	5714	5130	8871	231	8337	-3765	-3452	7048	-5400
J22	3,5,6,10	Landesinneres 8	-102	4893	-3197	-3304	-5264	-6430	-3984	-10276	-3785	-2850	-9220	-2514
J23	3,5,6	Landesinneres NW 7	5706	14399	-1764	-97	-1957	-4589	-3246	-8761	-3657	-1572	-6827	-2064
J24	5,6,10	Landesinneres NO 7	-1510	2960	-3952	-4719	-5534	-7015	-4631	-10124	-4086	-3393	-11240	-3842
J25	5,6	Landesinneres N 6	4865	13659	-2611	-1477	-1967	-5197	-4104	-8541	-4228	-2215	-9096	-3830
K1	KGK	gesamt 23	5438	11350	2287	3496	1026	848	2714	-1057	3732	2849	825	5350
K2	6,8,10	fKGK gesamt 11	5953	12950	544	1876	-1698	-1124	941	-1901	2790	1679	1079	4869
K3	3,5,7,8, 10	eKGK gesamt 12	2696	7192	1167	1588	351	230	1672	-2171	2029	1168	-1122	3157
K4	3,5,6,7, 8	Pruth-Don 13	2053	5544	935	1802	-468	82	1844	-1961	2632	1105	-434	4030
K5	5,6,7,8	S.Bug-Don 12	2404	5677	1486	2021	-77	1111	2633	-76	3776	1843	1718	5263
K6	5,6,7,8, 10	S.Bug-Wolga 22	5757	11665	2628	3665	1289	1450	3175	60	4417	3319	2115	6096
K7	6,7,8	Dnjepr-Don 11	1999	4538	1749	2386	50	1706	3100	501	4150	1958	2380	5665
K8	6,7,8,10	Dnjepr-Wolga 21	5489	11091	2681	3774	1265	1654	3314	237	4494	3290	2344	6190
K9	8,10	Don-Wolge 18	7389	14248	3355	4995	1171	1288	3459	-177	5072	4499	2887	7272
K10	8	Donez 8	4606	8131	3131	4664	-265	1467	3690	529	5541	4225	4224	7939
K11	10	Manych 10	5639	13776	-231	184	-1806	-2229	-457	-3716	755	677	-1425	2501
K12	3,5,6,8, 10	Landesinneres 21	6830	13930	2348	4203	1042	532	2525	-1412	3724	3138	1085	5662
K13	3,5,6,8	Landesinneres NW 11	3697	8691	807	2658	-738	-575	1368	-2694	2381	1259	-200	4262
K14	5,6,8,10	Landesinneres NO 20	7197	14359	2676	4362	1276	1148	2980	-167	4449	3624	2563	6472
K15	5,6,8	Landesinneres N 10	4066	8939	1278	2781	-490	460	2107	-466	3599	1989	2469	5634
K16	10	fKGK Manych 8	7259	16086	-93	597	-3335	-3430	-667	-4455	1459	1606	-929	3777
K17	3,5,7,8	eKGK Pruth- Don 10	1205	5016	-22	215	-1935	-1843	349	-4460	1043	276	-3007	2448
K18	7,8	eKGK Dnjepr- Don 8	331	2916	369	56	-2127	-599	1231	-2232	2243	725	-287	3824
P1	POL 11	Samara (Wolga) 6	-3268	-4848	-2341	-556	-83	2118	-1049	81	-2924	-4436	-1189	-3625
	FPE	Region / N	Ä1	Ä2	J1	J2	J3	J4	J5	J6	J7	J8	J9	J10

Tabelle 13.10: Steppeninterne F_{ST}-Werte (Seite 2 von 4)

Ä1	ÄNL	gesamt 10	48233 ± 480	12385 ± 330	8781 ± 250	29799 ± 460	3227 ± 200	17731 ± 360	37264 ± 390	54925 ± 520	41382 ± 500	12147 ± 310	14434 ± 330	41758 ± 490
Ä2	2,4,5,7	Steppe 6	27869 ± 430	5356 ± 200	3861 ± 200	16266 ± 350	2792 ± 150	17226 ± 370	25067 ± 420	49520 ± 470	31244 ± 470	13573 ± 380	17741 ± 380	19661 ± 380
J1	JAM	gesamt 22	99980 ± 10	86981 ± 340	83526 ± 400	99030 ± 90	13197 ± 300	42689 ± 540	99545 ± 60	95080 ± 220	99653 ± 50	43758 ± 500	41224 ± 460	85289 ± 330
J2	2	Donau 6	58192 ± 450	26294 ± 450	18780 ± 410	39075 ± 450	3386 ± 170	13415 ± 340	87239 ± 320	99990 ± 0	90555 ± 290	15553 ± 360	21612 ± 370	76656 ± 430
J3	2,3,4	Donau-Dnjestr 10	60697 ± 530	16167 ± 380	7227 ± 250	26037 ± 470	792 ± 90	5485 ± 220	73844 ± 450	99990 ± 0	86486 ± 360	11494 ± 270	20503 ± 440	96674 ± 170
J4	2,3,4,5	Donau-Dnjepr 14	52500 ± 510	18929 ± 360	6207 ± 250	20889 ± 400	129 ± 40	891 ± 90	46490 ± 490	99990 ± 0	67538 ± 440	2336 ± 150	5970 ± 260	98436 ± 120
J5	2,3,4,5, 6,7	Donau-Don 19	99564 ± 60	92397 ± 290	72023 ± 470	94307 ± 240	9930 ± 280	24265 ± 470	96664 ± 160	94713 ± 220	99267 ± 90	28057 ± 430	38501 ± 500	90367 ± 290
J6	3,4,5	Pruth-Dnjepr 8	64192 ± 480	31967 ± 480	15117 ± 370	30957 ± 410	356 ± 60	3198 ± 180	31591 ± 450	94040 ± 250	49193 ± 440	7039 ± 280	14187 ± 350	99990 ± 0
J7	3,4,5,6, 7	Pruth-Don 13	99960 ± 20	99950 ± 20	95020 ± 200	98772 ± 100	24998 ± 470	49490 ± 530	91724 ± 240	57361 ± 500	94862 ± 210	58935 ± 440	66785 ± 470	79210 ± 400
J8	3,4,5,6, 7,9,10	Pruth-Wolga 16	99990 ± 0	97772 ± 160	96852 ± 180	99950 ± 20	36095 ± 460	78715 ± 370	98960 ± 90	62618 ± 450	98495 ± 120	79200 ± 440	69330 ± 490	75893 ± 430
J9	4,5	Dnjestr-Dnjepr 7	71250 ± 510	35957 ± 550	18107 ± 400	34818 ± 470	1386 ± 120	4277 ± 200	41590 ± 430	96971 ± 160	56767 ± 500	9247 ± 290	21067 ± 380	94535 ± 240
J10	4,5,6,7	Dnjestr-Don 12	99990 ± 0	99990 ± 0	96258 ± 170	99574 ± 60	30363 ± 450	61588 ± 440	95644 ± 190	58252 ± 500	98099 ± 140	68449 ± 490	76517 ± 430	67330 ± 520
J11	4,5,6,7, 8,10	Dnjestr-Wolga 15		97881 ± 150	97584 ± 140	99990 ± 0	44174 ± 490	84229 ± 350	99416 ± 70	62142 ± 480	99050 ± 80	84952 ± 350	77052 ± 390	62301 ± 440
J12	5,6,7	S.Bug-Don 9	-6253		99990 ± 0	99990 ± 0	65449 ± 440	83002 ± 320	87427 ± 330	17048 ± 420	84279 ± 380	72666 ± 400	65568 ± 400	40372 ± 460
J13	5,6,7,9	S.Bug-Kuban 11	-5649	-9596		99990 ± 0	90417 ± 310	96268 ± 200	93199 ± 250	7871 ± 270	73013 ± 470	89605 ± 360	59638 ± 530	20612 ± 380
J14	5,6,7, 9,10	S.Bug-Wolga 12	-7104	-8618	-8735		82210 ± 360	96634 ± 170	99594 ± 60	26146 ± 460	96584 ± 200	93803 ± 240	74319 ± 430	42996 ± 490
J15	6,7,9	Dnjepr-Kuban 7	-390	-4299	-7209	-5278		99990 ± 0	57717 ± 480	1307 ± 110	30868 ± 430	99990 ± 0	79853 ± 370	772 ± 90
J16	6,7,9, 10	Dnjepr-Wolga 8	-4444	-5409	-7409	-7754	-13259		90793 ± 300	7405 ± 260	68320 ± 460	99990 ± 0	95644 ± 190	5732 ± 240
J17	2,4,7,9	Küste 14	-5273	-3974	-4230	-5504	-1761	-5112		77220 ± 380	99990 ± 0	95466 ± 180	89090 ± 320	18206 ± 360
J18	2,4	Küste W 9	-1580	3889	5422	1815	14660	7735	-3482		90437 ± 250	15177 ± 400	25908 ± 390	86853 ± 360
J19	2,4,7	Küste NW 12	-5115	-4175	-2612	-4439	1206	-3044	-7594	-5725		83427 ± 370	88516 ± 370	31007 ± 480
J20	4,7,9	Küste NO 8	-4707	-4179	-5912	-6234	-10188	-11355	-6559	5592	-5038		99990 ± 0	3188 ± 180
J21	4,7	Küste N 6	-4594	-5109	-3439	-4664	-6566	-9292	-6408	4313	-7498	-13973		5356 ± 240
J22	3,5,6,10	Landesinneres 8	-1767	1265	3752	513	17908	10559	2790	-3714	1702	11933	12713	
J23	3,5,6	Landesinneres NW 7	-213	14	2687	1154	18534	12706	4522	235	3975	13978	15867	-12982
J24	5,6,10	Landesinneres NO 7	-3084	254	2908	-622	17480	9654	1739	-5309	512	11065	11695	-14214
J25	5,6	Landesinneres N 6	-1715	-1774	1187	-336	17826	11751	3386	-1170	2762	13062	14945	-12869
K1	KGK	gesamt 23	4142	9264	9568	6272	22034	15180	7027	3023	7082	15877	18296	-2686
K2	6,8,10	fKGK gesamt 11	3273	10395	10147	6197	24212	15772	5232	779	5205	15863	19049	-3116
K3	3,5,7,8 10	eKGK gesamt 12	2076	6155	6543	3546	17375	10959	4880	1725	4943	11638	13531	-3770
K4	3,5,6,7, 8	Pruth-Don 13	2189	8138	8103	4272	19183	11625	4897	1165	4999	12382	14586	-3731
K5	5,6,7,8	S.Bug-Don 12	2980	9490	9128	5044	19553	11673	5119	1620	5303	12717	15113	-2229
K6	5,6,7,8, 10	S.Bug-Wolga 22	4650	10071	10193	6771	22341	15328	7204	3339	7309	16183	18730	-1805
K7	6,7,8	Dnjepr-Don 11	3107	10061	9504	5250	19178	11133	5073	1762	5311	12138	14466	-1521
K8	6,7,8,10	Dnjepr-Wolga 21	4629	10237	10263	6783	22033	14967	7092	3327	7221	15787	18301	-1547
K9	8,10	Don-Wolge 18	6233	12279	12951	9119	26784	19006	8868	3894	8454	19584	21829	-2061
K10	8	Donez 8	6058	14366	14780	9600	27612	17834	8097	2562	7245	18367	20007	-1714
K11	10	Manych 10	2068	6381	6895	4141	20738	13867	4224	330	4092	13885	16532	-4749
K12	3,5,6,8, 10	Landesinneres 21	4680	9952	10389	7071	24226	17111	7694	3483	7687	17687	20362	-3096
K13	3,5,6,8	Landesinneres NW 11	2716	9017	9130	5195	22208	14087	5594	1549	5572	14639	17213	-4578
K14	5,6,8,10	Landesinneres NO 20	5221	10856	11079	7614	24621	17309	7863	3802	7908	18056	20894	-2136
K15	5,6,8	Landesinneres N 10	3544	10605	10294	6031	22724	14152	5732	1941	5797	15045	17895	-2872
K16	10	fKGK Manych 8	3400	9812	10604	6793	26939	18136	5297	-561	4495	17590	20247	-5090
K17	3,5,7,8	eKGK Pruth- Don 10	1394	6555	7350	3599	19368	11777	4350	-301	3921	12283	13738	-6135
K18	7,8	eKGK Dnjepr- Don 8	1951	8319	8582	4257	18934	10693	4033	-366	3661	11510	12923	-4286
P1	POL 11	Samara (Wolga) 6	-5003	-2051	-2085	-4645	1337	-3878	-3935	-588	-3705	-4848	-5143	1308
	FPE	Region N	J11	J12	J13	J14	J15	J16	J17	J18	J19	J20	J21	J22

Tabelle 13.10: Steppeninterne F_{ST}-Werte (Seite 3 von 4)

Ä1	ÄNL	gesamt 10	13118 ± 340	48490 ± 450	18998 ± 410	7623 ± 250	9059 ± 260	22384 ± 380	24206 ± 360	22641 ± 410	7514 ± 310	23780 ± 380	8811 ± 260	5168 ± 210
Ä2	2,4,5,7	Steppe 6	3524 ± 170	24245 ± 430	16236 ± 360	4069 ± 200	6673 ± 280	13801 ± 330	13553 ± 380	13949 ± 320	4990 ± 200	18553 ± 360	4861 ± 210	2257 ± 150
J1	JAM	gesamt 22	65647 ± 420	88377 ± 370	72498 ± 460	10722 ± 320	35363 ± 460	26106 ± 430	28324 ± 440	23235 ± 400	9068 ± 290	22622 ± 380	9059 ± 290	6643 ± 240
J2	2	Donau 6	39659 ± 450	79170 ± 440	69983 ± 470	16375 ± 360	32423 ± 460	31997 ± 400	29938 ± 490	31304 ± 480	17870 ± 340	29017 ± 400	17434 ± 350	14424 ± 360
J3	2,3,4	Donau-Dnjestr 10	68330 ± 430	97089 ± 150	65162 ± 460	29076 ± 460	62014 ± 480	42491 ± 500	45213 ± 480	41441 ± 490	27591 ± 430	41164 ± 500	28294 ± 470	27077 ± 420
J4	2,3,4,5	Donau-Dnjepr 14	91456 ± 250	98842 ± 100	91862 ± 260	30165 ± 430	56816 ± 470	43956 ± 490	41907 ± 510	29165 ± 450	23602 ± 460	22849 ± 420	21523 ± 440	24908 ± 470
J5	2,3,4,5, 6,7	Donau-Don 19	82804 ± 300	92347 ± 240	85536 ± 360	8346 ± 310	32116 ± 460	22859 ± 380	19097 ± 390	14731 ± 330	6663 ± 230	13325 ± 350	6514 ± 250	7484 ± 260
J6	3,4,5	Pruth-Dnjepr 8	98792 ± 110	99129 ± 100	98426 ± 130	49985 ± 470	54707 ± 510	63677 ± 440	58271 ± 440	41362 ± 500	40145 ± 480	37363 ± 460	38798 ± 510	42471 ± 500
J7	3,4,5,6, 7	Pruth-Don 13	74983 ± 370	79319 ± 400	77289 ± 390	8583 ± 280	21721 ± 420	24215 ± 440	20770 ± 400	13246 ± 360	6168 ± 230	12702 ± 320	5900 ± 210	5198 ± 230
J8	3,4,5,6, 7,9,10	Pruth-Wolga 16	58697 ± 470	76903 ± 410	61499 ± 460	10761 ± 330	25552 ± 430	30551 ± 520	28977 ± 520	23423 ± 450	9474 ± 280	23206 ± 420	9801 ± 280	5485 ± 230
J9	4,5	Dnjestr-Dnjepr 7	99990 ± 0	99990 ± 0	97594 ± 160	32452 ± 470	36630 ± 490	51302 ± 460	44105 ± 450	31433 ± 470	25394 ± 460	28106 ± 420	25948 ± 420	22463 ± 410
J10	4,5,6,7	Dnjestr-Don 12	61548 ± 530	73201 ± 440	71825 ± 440	4208 ± 210	11276 ± 270	19087 ± 370	12078 ± 340	9425 ± 310	3119 ± 170	7970 ± 260	3069 ± 180	1841 ± 130
J11	4,5,6,7, 8,10	Dnjestr-Wolga 15	43243 ± 480	71884 ± 440	57717 ± 480	6098 ± 230	15652 ± 360	22740 ± 420	18176 ± 390	15781 ± 370	4554 ± 210	16187 ± 350	5128 ± 230	2386 ± 160
J12	5,6,7	S.Bug-Don 9	47698 ± 480	46877 ± 520	54945 ± 430	1703 ± 160	2812 ± 160	8534 ± 290	3653 ± 180	2604 ± 160	1406 ± 120	2465 ± 160	1475 ± 130	762 ± 90
J13	5,6,7,9	S.Bug-Kuban 11	29363 ± 490	27819 ± 460	35323 ± 490	624 ± 70	1911 ± 150	4158 ± 190	2257 ± 150	1673 ± 120	515 ± 70	2000 ± 140	535 ± 70	109 ± 30
J14	5,6,7, 9,10	S.Bug-Wolga 12	35937 ± 500	49213 ± 570	45411 ± 530	2742 ± 170	6336 ± 260	18077 ± 350	12415 ± 360	11989 ± 290	2158 ± 150	10850 ± 280	2554 ± 150	1000 ± 100
J15	6,7,9	Dnjepr-Kuban 7	2178 ± 140	2198 ± 160	3683 ± 190	40 ± 20	99 ± 30	188 ± 40	119 ± 30	218 ± 50	59 ± 20	436 ± 70	99 ± 30	20 ± 10
J16	6,7,9, 10	Dnjepr-Wolga 8	3782 ± 190	7286 ± 300	6287 ± 250	455 ± 70	1000 ± 90	2218 ± 150	2178 ± 150	3267 ± 200	525 ± 70	3930 ± 220	644 ± 80	228 ± 50
J17	2,4,7,9	Küste 14	11573 ± 290	27859 ± 410	19097 ± 370	1188 ± 100	8940 ± 270	7178 ± 240	7158 ± 270	7514 ± 270	1386 ± 100	8425 ± 290	1624 ± 120	505 ± 70
J18	2,4	Küste W 9	47342 ± 540	94139 ± 240	57984 ± 470	16790 ± 360	32185 ± 490	31334 ± 460	27690 ± 460	25621 ± 420	16840 ± 420	26077 ± 430	17998 ± 370	14226 ± 330
J19	2,4,7	Küste NW 12	16335 ± 380	37472 ± 430	26582 ± 490	1841 ± 130	12157 ± 300	9445 ± 300	8177 ± 270	8306 ± 290	1802 ± 150	9187 ± 270	2336 ± 150	1247 ± 110
J20	4,7,9	Küste NO 8	2138 ± 130	4277 ± 220	3722 ± 180	257 ± 50	891 ± 90	1297 ± 120	1703 ± 140	1891 ± 140	238 ± 40	2792 ± 160	248 ± 50	50 ± 20
J21	4,7	Küste N 6	1772 ± 120	9643 ± 270	10415 ± 260	792 ± 90	1475 ± 120	2525 ± 160	1693 ± 120	1782 ± 140	812 ± 90	2534 ± 170	931 ± 100	436 ± 60
J22	3,5,6,10	Landesinneres N 6	99990 ± 0	99990 ± 0	99990 ± 0	68746 ± 490	69409 ± 450	74785 ± 460	77220 ± 420	57796 ± 420	56678 ± 530	50975 ± 540	51965 ± 450	61251 ± 500
J23	3,5,6	Landesinneres NW 7	99990 ± 0	99990 ± 0	99990 ± 0	40253 ± 450	40580 ± 500	55529 ± 470	42481 ± 510	25522 ± 450	29324 ± 480	18761 ± 380	27275 ± 380	32581 ± 420
J24	5,6,10	Landesinneres NO 7	-12000		99990 ± 0	53213 ± 490	48154 ± 550	71290 ± 430	60608 ± 430	43966 ± 520	41818 ± 480	37481 ± 470	40501 ± 460	39283 ± 520
J25	5,6	Landesinneres N 6	-16258	-14833		26374 ± 500	21265 ± 380	45659 ± 520	30225 ± 470	17889 ± 410	18909 ± 400	12177 ± 330	16117 ± 320	18226 ± 360
K1	KGK	gesamt 23	-86	-1401	1567		96208 ± 190	98911 ± 90	99901 ± 30	99307 ± 80	99990 ± 0	96406 ± 180	99990 ± 0	99020 ± 100
K2	6,8,10	fKGK gesamt 11	592	-664	3894	-5073		62113 ± 490	95387 ± 220	94149 ± 230	96149 ± 180	95575 ± 210	95961 ± 210	99990 ± 0
K3	3,5,7,8 10	eKGK gesamt 12	-1385	-3299	-677	-5254	-2453		99525 ± 80	99990 ± 0	97198 ± 150	96000 ± 200	97070 ± 160	77270 ± 400
K4	3,5,6,7, 8	Pruth-Don 13	235	-2798	1704	-5561	-5523	-6949		99990 ± 0	99129 ± 90	99990 ± 0	99109 ± 90	92060 ± 270
K5	5,6,7,8	S.Bug-Don 12	2351	-1051	4180	-5489	-5673	-6684	-8267		99426 ± 70	99990 ± 0	99426 ± 70	90347 ± 310
K6	5,6,7,8, 10	S.Bug-Wolga 22	1096	-361	2972	-4513	-5208	-5061	-5434	-5857		97495 ± 160	99990 ± 0	99218 ± 90
K7	6,7,8	Dnjepr-Don 11	3496	-221	5511	-5315	-5553	-6559	-8334	-9262	-5729		98921 ± 110	85714 ± 350
K8	6,7,8,10	Dnjepr-Wolga 21	1536	-30	3520	-4535	-5307	-5063	-5500	-5965	-4805	-6129		99238 ± 80
K9	8,10	Don-Wolge 18	1166	410	4299	-4180	-6247	-3356	-4466	-4687	-4359	-4475	-4411	
K10	8	Donez 8	4930	1465	9434	-3959	-6404	-2882	-6602	-7197	-4201	-7437	-4391	-6639
K11	10	Manych 10	-3658	-2974	-1524	-6157	-7011	-6459	-5028	-4929	-6233	-4536	-6230	-6459
K12	3,5,6,8, 10	Landesinneres 21	-837	-1208	1483	-4339	-5921	-3977	-4954	-4612	-4184	-4116	-4072	-4814
K13	3,5,6,8	Landesinneres NW 11	-795	-2677	1826	-5145	-6897	-4842	-7695	-7131	-4762	-6683	-4596	-5405
K14	5,6,8,10	Landesinneres NO 20	484	-26	3109	-4366	-6239	-3754	-4844	-5067	-4533	-4602	-4437	-5159
K15	5,6,8	Landesinneres N 10	1785	-542	5023	-5272	-7621	-4516	-7780	-8325	-5500	-7944	-5357	-6136
K16	10	fKGK Manych 8	-2455	-2009	1692	-4914	-9942	-1967	-3830	-3485	-4867	-2750	-4714	-7492
K17	3,5,7,8	eKGK Pruth- Don 10	-2573	-5273	-1215	-6642	-5479	-8874	-9054	-8619	-6321	-8288	-6193	-5892
K18	7,8	eKGK Dnjepr- Don 8	730	-2913	2861	-7337	-6367	-9503	-9992	-10856	-7743	-11334	-8029	-7132
P1	POL 11	Samara (Wolga) 6	4743	-184	3333	9135	10078	3680	4816	5808	9871	5251	9519	12116
	FPE	Region N	J23	J24	J25	K1	K2	K3	K4	K5	K6	K7	K8	K9

Tabelle 13.10: Steppeninterne F_{ST}-Werte (Seite 4 von 4)

Ä1	ÄNL	gesamt 10	18394 ± 370	10494 ± 280	4029 ± 200	17384 ± 350	4574 ± 200	21097 ± 390	10593 ± 310	34828 ± 480	36076 ± 530	62350 ± 480
Ä2	2,4,5,7	Steppe 6	13771 ± 350	1515 ± 120	1752 ± 140	12078 ± 300	1921 ± 130	10732 ± 330	4841 ± 210	14553 ± 290	22651 ± 400	69944 ± 470
J1	JAM	gesamt 22	15494 ± 330	46936 ± 460	9999 ± 300	32413 ± 440	8029 ± 270	26770 ± 490	44788 ± 500	43827 ± 490	39768 ± 470	68617 ± 490
J2	2	Donau 6	21889 ± 450	45936 ± 460	14949 ± 400	25552 ± 450	14900 ± 320	28027 ± 480	46104 ± 490	43748 ± 490	39818 ± 500	59905 ± 510
J3	2,3,4	Donau-Dnjestr 10	42055 ± 510	69152 ± 470	30621 ± 460	50183 ± 510	29453 ± 450	52896 ± 500	72735 ± 460	73409 ± 460	64241 ± 460	45738 ± 460
J4	2,3,4,5	Donau-Dnjepr 14	31541 ± 470	77329 ± 420	32165 ± 430	50856 ± 530	25344 ± 450	37412 ± 430	83358 ± 400	68231 ± 440	48461 ± 460	24423 ± 390
J5	2,3,4,5, 6,7	Donau-Don 19	13850 ± 380	50767 ± 510	9841 ± 330	26908 ± 440	8177 ± 250	20384 ± 370	51104 ± 470	38986 ± 460	29987 ± 490	53688 ± 510
J6	3,4,5	Pruth-Dnjepr 8	36957 ± 500	81041 ± 380	58747 ± 590	66597 ± 510	43204 ± 460	43610 ± 500	85091 ± 410	87338 ± 280	70458 ± 480	48203 ± 570
J7	3,4,5,6, 7	Pruth-Don 13	10702 ± 300	34947 ± 460	7346 ± 270	23641 ± 420	5821 ± 230	16345 ± 370	30700 ± 470	34492 ± 510	23691 ± 430	63093 ± 460
J8	3,4,5,6, 7,9,10	Pruth-Wolga 16	12969 ± 310	34848 ± 480	9128 ± 280	29878 ± 430	7484 ± 260	24869 ± 410	28284 ± 440	39244 ± 490	38412 ± 500	83516 ± 360
J9	4,5	Dnjestr-Dnjepr 7	22859 ± 390	59846 ± 460	34422 ± 480	46817 ± 470	22354 ± 380	27027 ± 460	55559 ± 570	68399 ± 470	39026 ± 420	59351 ± 510
J10	4,5,6,7	Dnjestr-Don 12	6821 ± 270	21166 ± 430	3376 ± 180	12573 ± 310	2396 ± 150	8969 ± 250	18770 ± 360	24413 ± 400	17087 ± 380	69280 ± 470
J11	4,5,6,7, 8,10	Dnjestr-Wolga 15	8098 ± 260	23117 ± 420	4010 ± 190	18137 ± 380	3376 ± 200	15840 ± 350	16553 ± 350	29977 ± 460	26681 ± 450	86288 ± 360
J12	5,6,7	S.Bug-Don 9	812 ± 100	11405 ± 260	1297 ± 110	3911 ± 180	891 ± 90	2940 ± 170	9989 ± 320	9326 ± 280	7910 ± 240	49480 ± 480
J13	5,6,7,9	S.Bug-Kuban 11	653 ± 80	5396 ± 230	406 ± 70	2158 ± 130	228 ± 50	1812 ± 120	3495 ± 180	4257 ± 190	3851 ± 210	56044 ± 460
J14	5,6,7, 9,10	S.Bug-Wolga 12	4138 ± 200	11880 ± 370	1841 ± 140	9247 ± 300	1653 ± 130	8801 ± 280	8118 ± 290	17167 ± 370	16127 ± 380	81160 ± 420
J15	6,7,9	Dnjepr-Kuban 7	59 ± 20	188 ± 40	40 ± 20	89 ± 30	30 ± 20	119 ± 30	119 ± 40	119 ± 30	693 ± 70	29205 ± 450
J16	6,7,9, 10	Dnjepr-Wolga 8	2099 ± 150	1099 ± 90	178 ± 40	1168 ± 110	267 ± 50	2109 ± 140	1703 ± 130	2970 ± 160	5099 ± 220	66736 ± 440
J17	2,4,7,9	Küste 14	4435 ± 230	10850 ± 340	762 ± 90	6762 ± 240	842 ± 80	7445 ± 240	10811 ± 280	10524 ± 310	14315 ± 340	78992 ± 410
J18	2,4	Küste W 9	21622 ± 370	35927 ± 440	13118 ± 360	27423 ± 450	11860 ± 340	23948 ± 430	41451 ± 520	45827 ± 440	43154 ± 510	48817 ± 420
J19	2,4,7	Küste NW 12	8177 ± 250	13672 ± 340	931 ± 100	9039 ± 290	1257 ± 110	10138 ± 280	16949 ± 340	14325 ± 340	18117 ± 350	74161 ± 420
J20	4,7,9	Küste NO 8	1386 ± 110	1059 ± 90	168 ± 40	1020 ± 100	149 ± 40	1109 ± 110	1099 ± 110	1584 ± 110	4089 ± 210	73656 ± 500
J21	4,7	Küste N 6	2356 ± 150	1475 ± 120	228 ± 40	1832 ± 130	366 ± 60	1901 ± 140	1653 ± 120	3307 ± 180	5792 ± 220	72745 ± 460
J22	3,5,6,10	Landesinneres 8	50559 ± 510	91219 ± 310	77586 ± 370	83170 ± 420	66142 ± 470	63489 ± 480	88991 ± 280	96495 ± 180	92615 ± 280	36957 ± 510
J23	3,5,6	Landesinneres NW 7	17276 ± 400	89001 ± 330	52430 ± 530	50549 ± 410	38877 ± 550	31442 ± 430	70765 ± 500	68785 ± 450	44481 ± 490	10771 ± 320
J24	5,6,10	Landesinneres NO 7	26037 ± 430	73953 ± 410	59945 ± 510	64380 ± 520	43471 ± 460	41293 ± 510	61053 ± 540	87239 ± 330	58460 ± 470	50807 ± 480
J25	5,6	Landesinneres N 6	5772 ± 250	63687 ± 480	31116 ± 440	31235 ± 410	23087 ± 400	17721 ± 350	30571 ± 450	55638 ± 430	29819 ± 460	27492 ± 370
K1	KGK	gesamt 23	76656 ± 390	99040 ± 100	99990 ± 0	98030 ± 120	99990 ± 0	96792 ± 180	90437 ± 280	99812 ± 40	99990 ± 0	4693 ± 210
K2	6,8,10	fKGK gesamt 11	86566 ± 330	98584 ± 130	99960 ± 20	95515 ± 210	99931 ± 20	89862 ± 290	99990 ± 0	90031 ± 320	89546 ± 290	4396 ± 230
K3	3,5,7,8 10	eKGK gesamt 12	62994 ± 550	95812 ± 210	88853 ± 280	83863 ± 380	85120 ± 360	78240 ± 420	57638 ± 440	99990 ± 0	99990 ± 0	19087 ± 410
K4	3,5,6,7, 8	Pruth-Don 13	92535 ± 280	83328 ± 360	98317 ± 130	99990 ± 0	97040 ± 190	99990 ± 0	69785 ± 450	99990 ± 0	99990 ± 0	16484 ± 420
K5	5,6,7,8	S.Bug-Don 12	93110 ± 250	80903 ± 350	94991 ± 200	99990 ± 0	96307 ± 190	99990 ± 0	64578 ± 410	99990 ± 0	99990 ± 0	14969 ± 380
K6	5,6,7,8, 10	S.Bug-Wolga 22	77359 ± 410	98475 ± 130	99980 ± 10	94822 ± 230	99990 ± 0	95990 ± 180	87348 ± 330	98010 ± 130	99990 ± 0	4514 ± 210
K7	6,7,8	Dnjepr-Don 11	90308 ± 300	78646 ± 380	88001 ± 300	99990 ± 0	89823 ± 360	99990 ± 0	56598 ± 470	99990 ± 0	99990 ± 0	17295 ± 450
K8	6,7,8,10	Dnjepr-Wolga 21	79121 ± 370	97654 ± 150	99990 ± 0	91931 ± 280	99980 ± 10	93337 ± 230	84061 ± 350	97129 ± 170	99990 ± 0	4554 ± 210
K9	8,10	Don-Wolga 18	97753 ± 140	99010 ± 100	99960 ± 20	98525 ± 110	99990 ± 0	99317 ± 80	98475 ± 110	95387 ± 220	99990 ± 0	2257 ± 140
K10	8	Donez 8		54331 ± 480	87209 ± 330	99990 ± 0	89288 ± 300	99990 ± 0	80744 ± 360	93951 ± 240	99990 ± 0	12058 ± 290
K11	10	Manych 10	-2534		99881 ± 30	86665 ± 330	99901 ± 30	86298 ± 310	99990 ± 0	92090 ± 290	80289 ± 390	4000 ± 220
K12	3,5,6,8, 10	Landesinneres 21	-4696	-6266		99990 ± 0	99990 ± 0	99515 ± 70	98772 ± 110	99881 ± 40	98258 ± 130	2940 ± 210
K13	3,5,6,8	Landesinneres NW 11	-8754	-4426	-6055		99020 ± 100	99990 ± 0	90189 ± 300	98891 ± 100	98921 ± 100	9316 ± 300
K14	5,6,8,10	Landesinneres NO 20	-5146	-6452	-4940	-5730		99990 ± 0	98238 ± 130	98673 ± 120	98792 ± 120	2168 ± 140
K15	5,6,8	Landesinneres N 10	-10266	-4507	-6020	-9767	-6440		84952 ± 340	99990 ± 0	98040 ± 140	8623 ± 300
K16	10	fKGK Manych 8	-6004	-8814	-6545	-6064	-6708	-6238		87863 ± 330	88813 ± 330	1990 ± 160
K17	3,5,7,8	eKGK Pruth- Don 10	-7773	-6653	-6223	-8678	-5894	-8298	-5256			23176 ± 430
K18	7,8	eKGK Dnjepr- Don 8	-10162	-7169	-6166	-8306	-6650	-9496	-4969	-11294		25215 ± 430
P1	POL 11	Samara (Wolga) 6	8767	7536	10704	6464	11559	7653	10439	3073	2789	
	FPE	Region N	K10	K11	K12	K13	K14	K15	K16	K17	K18	P1

Tabelle 13.11: F_{ST} -Werte zwischen den Kulturen Westasiens

Die Tabelle befindet sich auf den folgenden drei Seiten. Allen Zahlenwerten (außer den Individuenzahlen) in der Matrix ist der Zusatz „ $\times 10^{-5}$ “ anzuhängen. Die Werte unterhalb der Diagonale der Matrix geben die paarweisen F_{ST} -Werte, basierend auf paarweisen Unterschieden, an. In grauer Schrift sind Vergleiche zwischen Gruppen angegeben, bei denen ein Teil der Individuen in beiden Populationen vorkommen. Fettgedruckt sind signifikante F_{ST} -Werte ($p < 0,05$). In der oberen Diagonale sind die p -Werte, basierend auf 10 000 Permutationen, aufgeführt.

Folgende Populationen wurden verglichen, wobei die Zahl hinter dem Code die Anzahl der Individuen angibt): HGC – Prä-neolithische Jäger/Sammler (Mitteleuropa), LBK – Linearbandkeramik, RSC – Rössener Kultur, SCG – Schöninger Gruppe, NHG/NHG2 – Neolithische Jäger/Sammler/Fischer (mit/ohne Proben aus Ostorf), LNF – Spätneolithische Bauern aus der Blätterhöhle, BAC – Baalberger Kultur, SMC – Salzmünder Kultur, BEC – Bernburger Kultur, FBC – Trichterbecherkultur, PWC – Grübchenkeramik, CWC – Schnurkeramik, BBC – Glockenbecherkultur, UNC – Aunjetitzer Kultur, HGE – osteuropäische Prä-Neolithiker, DDC – Dnepr-Donetz-Kultur, ENL – äneolithische Gruppen des Nordpontikums), YAM – Jamnaja-Kultur, YAW/YAO – Jamnaja-Gruppen westlich / östlich des Dnjepr (Fundplatzeinheiten 2 bis 6 / 7 bis 11; siehe Abbildung 9.9), POL – Poltavka-Kultur, CAT – Katakombengrabkultur, eCAT/dCAT – frühe / entwickelte Katakombengrabkultur, UTC – Ust-Tartas Kultur, ODI – Odinovo-Kultur, KRO – Krotovo-Kultur, ANF – Andronovo- (Fedorovo-) Kultur, AND – Andronovo-Kultur, ANDF – ANF + AND, BAK – Bronzezeit Kasachstan, mALL – Gesamtdatensatz moderner Daten, mOEu – moderne Daten aus Osteuropa (mBul + mUkr + mRus), mBul – heutige Bulgaren, mUkr – heutige Ukrainer, mRus – heutige Westrussen, mCEu – moderne Daten aus Mittel- und Nordeuropa (mGer + mLit + mPol), mGer – heutige Deutsche, mLit – heutige Litauer, mPol – heutige Polen. Eine Liste mit den Individuen aller Datensätze findet sich auf der Daten-DVD.

Tabelle 13.11: F_{ST}-Werte zwischen den Kulturen Westasiens

(Seite 1 von 3) Informationen zur Tabelle siehe auf der vorigen Seite

HGC 16		20 ± 10	59 ± 20	0 ± 0	25304 ± 410	11266 ± 290	2475 ± 170	10 ± 10	0 ± 0	3079 ± 170	723 ± 100	277 ± 50	79 ± 30	168 ± 40
LBK 102	16209		57965 ± 480	24077 ± 440	0 ± 0	10 ± 1	4594 ± 210	38986 ± 480	16355 ± 370	7435 ± 290	16632 ± 330	119 ± 40	307 ± 50	89 ± 30
RSC 11	15666	-1117		38066 ± 450	30 ± 20	0 ± 0	24295 ± 450	91436 ± 280	53222 ± 510	60608 ± 450	16553 ± 390	485 ± 70	48787 ± 480	33304 ± 430
SCG 33	16381	507	224		0 ± 0	0 ± 0	12623 ± 350	69845 ± 470	16711 ± 320	28661 ± 430	17652 ± 410	79 ± 30	12583 ± 340	624 ± 80
NHG 20	1260	15401	17417	15510		58182 ± 510	1564 ± 120	30 ± 20	0 ± 0	782 ± 70	4613 ± 220	30 ± 20	40 ± 20	40 ± 20
NHG2 13	4186	22619	28995	24010	-1194		228 ± 40	0 ± 0	0 ± 0	40 ± 20	59 ± 20	0 ± 0	0 ± 0	0 ± 0
LNF 11	9134	5386	2222	2947	10130	21625		17889 ± 340	15503 ± 370	63964 ± 500	9128 ± 290	1049 ± 90	24146 ± 450	18068 ± 380
BAC 19	12189	84	-3487	-1030	13004	22011	2194		20770 ± 390	50599 ± 520	59756 ± 480	99 ± 30	89575 ± 340	17652 ± 410
SMC 29	17322	980	-667	1284	17079	26068	2793	1212		10831 ± 280	19038 ± 400	10 ± 10	3218 ± 180	624 ± 80
BEC 17	5681	2774	-1147	635	7696	16084	-1663	-415	2403		26195 ± 400	4386 ± 210	66003 ± 500	51223 ± 530
FBC 10	13307	2336	2628	2311	8820	26110	6273	-1066	2079	1630		1594 ± 120	44134 ± 480	12504 ± 300
PWC 19	12711	8774	8919	8416	13933	25206	9622	7663	9389	4638	9618		673 ± 80	1416 ± 120
CWC 44	8369	3253	-320	1162	9689	17355	1223	-1577	2304	-876	-124	5384		52836 ± 560
BBC 35	7888	5000	512	3933	9270	17303	1773	1175	3361	-312	2765	4251	-255	
UNC 94	6386	4532	-876	1949	8204	12287	-260	-16	2758	-338	2931	5335	558	278
HGE 14	7693	11609	4420	10010	13896	19784	6820	4939	11032	3733	10351	5388	4312	4740
DDC 6	16387	7379	572	8306	19903	34202	11512	4108	3641	3760	10730	8078	5202	3341
ENL 10	4472	5298	2492	6853	8320	20566	5593	2280	5452	-1449	-385	9081	1567	150
YAM 22	8493	2807	-1224	1625	10510	19688	1546	-1714	2092	-981	-478	6301	-2142	-1366
YAW 14	8175	3290	-1574	1148	12015	22161	799	22	2848	-2631	4556	4409	-740	-1068
YAO 8	21440	8174	7117	9366	22352	37411	14369	1777	9059	9256	1013	21331	3571	7884
POL 6	10165	3858	-36	5518	15340	27307	8043	-2682	6496	1712	1938	15263	645	2193
CAT 23	12369	8035	5831	8153	15488	26187	7794	6796	5861	5025	7749	1798	3573	1784
eCAT 11	13445	6526	4851	5940	16615	33470	6725	5042	4793	3103	8241	-2303	2073	152
dCAT 12	11063	7009	3684	7249	14865	26360	7710	4677	4445	3965	5158	4086	2606	1488
UTC 17	20531	16388	8043	15642	24313	28885	15268	12415	15769	11523	18114	18525	13483	13558
ODI 10	34563	20735	16305	22957	38868	50288	31920	18824	22889	19944	30501	32009	19931	21435
KRO 26	15720	13860	6506	13452	18972	23249	12787	9750	13306	8533	13435	16425	10203	10174
ANF 20	19356	16930	9305	17019	22696	27046	16750	12830	16615	11670	15752	19512	13945	14388
AND 9	8500	3419	-2259	1291	12746	22468	3490	-2689	3326	-873	757	3108	-3026	-897
ANDF 29	11925	10714	3188	9902	14937	19195	9163	6147	9759	5188	8380	10759	6398	6672
BAK 8	10461	5686	2610	6228	12964	24638	7830	-1916	8234	2739	146	13334	535	3861
mAll 680	9105	3979	-1079	2410	9869	15692	419	282	1374	366	2042	4477	-9	-51
mOEU 246	8980	3926	-997	2413	9816	15830	716	96	1448	143	1621	4456	-146	-199
mBul 30	11134	2809	-1361	786	12772	20702	682	-815	664	-73	1593	5774	-486	429
mUkr 18	10369	1977	-2968	776	12050	21374	755	-2369	698	-1758	279	3364	-2219	-1408
mRus 198	8870	4075	-923	2658	9707	15706	746	245	1536	211	1666	4556	-44	-253
mCEu 434	9323	3968	-1112	2416	10109	15937	303	382	1359	511	2315	4603	86	67
mGer 262	10385	4075	-960	2623	11139	17350	507	766	1389	750	2857	4973	298	251
mLit 97	7764	4516	-1269	3018	9213	14722	14	401	1964	160	2832	3949	138	-401
mPol 75	9302	2644	-1440	1004	10028	16264	396	-1051	641	215	195	5299	-644	245
x 10⁻⁵	HGC 16	LBK 102	RSC 11	SCG 33	NHG 20	NHG2 13	LNF 11	BAC 19	SMC 29	BEC 17	FBC 10	PWC 19	CWC 44	BBC 35

Tabelle 13.11: F_{ST}-Werte zwischen den Kulturen Westeurasiens

(Seite 2 von 3)

HGC 16	149 ± 40	4069 ± 190	1040 ± 100	11623 ± 330	396 ± 70	1208 ± 120	238 ± 40	3920 ± 180	149 ± 40	1040 ± 100	1040 ± 90	0 ± 0	0 ± 0	40 ± 20
LBK 102	0 ± 0	149 ± 40	5514 ± 210	5138 ± 220	4406 ± 200	7336 ± 270	2911 ± 170	13613 ± 360	50 ± 20	3000 ± 180	1891 ± 140	0 ± 0	30 ± 20	0 ± 0
RSC 11	51440 ± 490	9445 ± 320	39283 ± 460	22265 ± 400	66587 ± 520	68686 ± 440	5762 ± 230	44213 ± 460	2713 ± 150	5792 ± 220	11237 ± 330	1406 ± 130	158 ± 40	2643 ± 160
SCG 33	1307 ± 110	50 ± 20	3821 ± 190	2228 ± 150	13088 ± 340	22641 ± 400	1346 ± 130	10514 ± 20	30 ± 20	2881 ± 160	812 ± 90	0 ± 0	0 ± 0	0 ± 0
NHG 20	10 ± 10	99 ± 30	277 ± 50	3851 ± 190	99 ± 30	109 ± 30	347 ± 60	1307 ± 110	0 ± 0	99 ± 30	109 ± 30	0 ± 0	0 ± 0	0 ± 0
NHG2 13	10 ± 10	50 ± 20	0 ± 0	40 ± 20	0 ± 0	0 ± 0	0 ± 0	79 ± 30	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	10 ± 10
LNF 11	36878 ± 450	5693 ± 210	4257 ± 210	10732 ± 310	23374 ± 410	31413 ± 400	1871 ± 140	6287 ± 230	2119 ± 130	10286 ± 330	4485 ± 190	287 ± 50	0 ± 0	297 ± 60
BAC 19	39917 ± 490	2940 ± 170	12821 ± 400	17761 ± 370	85764 ± 340	42629 ± 490	23235 ± 410	70637 ± 360	40 ± 20	3089 ± 180	2891 ± 150	0 ± 0	0 ± 0	40 ± 20
SMC 29	446 ± 70	0 ± 0	13860 ± 320	5336 ± 230	8296 ± 280	8950 ± 250	2732 ± 180	7019 ± 260	188 ± 40	6366 ± 250	4811 ± 230	0 ± 0	0 ± 0	0 ± 0
BEC 17	46589 ± 470	9771 ± 290	17266 ± 390	57519 ± 520	64152 ± 450	87298 ± 330	2317 ± 140	28443 ± 430	1802 ± 130	14276 ± 340	8663 ± 260	89 ± 30	0 ± 0	366 ± 60
FBC 10	12246 ± 350	1465 ± 120	4534 ± 210	44085 ± 450	48688 ± 590	8771 ± 310	35739 ± 520	28195 ± 480	2782 ± 170	5009 ± 200	11336 ± 300	20 ± 10	0 ± 0	287 ± 50
PWC 19	99 ± 30	5821 ± 220	6514 ± 250	2703 ± 180	723 ± 80	5762 ± 240	69 ± 30	426 ± 60	18404 ± 440	67944 ± 440	10247 ± 270	0 ± 0	0 ± 0	0 ± 0
CWC 44	15226 ± 300	2940 ± 160	7841 ± 270	21780 ± 460	99149 ± 90	60271 ± 530	11732 ± 340	35204 ± 480	1792 ± 120	16365 ± 400	10534 ± 280	0 ± 0	0 ± 0	0 ± 0
BBC 35	30116 ± 460	2099 ± 130	16008 ± 350	39590 ± 470	85704 ± 410	69755 ± 460	2336 ± 150	23582 ± 380	9405 ± 270	38372 ± 430	19949 ± 410	0 ± 0	0 ± 0	10 ± 10
UNC 94		10326 ± 320	20998 ± 400	16276 ± 400	52490 ± 480	90823 ± 290	7673 ± 260	22631 ± 370	921 ± 90	18107 ± 380	15226 ± 390	0 ± 0	89 ± 30	0 ± 0
HGE 14	3077		46184 ± 470	9227 ± 290	3406 ± 190	17186 ± 350	970 ± 100	21572 ± 380	2366 ± 140	14593 ± 350	11098 ± 330	4554 ± 240	960 ± 100	6395 ± 220
DDC 6	930	-953		16672 ± 390	9920 ± 300	18553 ± 380	2990 ± 170	29037 ± 440	10484 ± 310	13721 ± 350	22285 ± 460	42847 ± 500	7712 ± 250	56955 ± 480
ENL 10	1849	6494	6006		44708 ± 450	22414 ± 390	16295 ± 360	56856 ± 490	4732 ± 190	5514 ± 230	17444 ± 350	406 ± 60	69 ± 30	1891 ± 140
YAM 22	-312	5142	4754	-279		90674 ± 320	34739 ± 500	58034 ± 480	6227 ± 240	27086 ± 400	19424 ± 410	0 ± 0	0 ± 0	50 ± 20
YAW 14	-1986	2984	2901	2093	-2818		881 ± 90	17048 ± 400	22532 ± 420	48599 ± 440	33858 ± 490	485 ± 70	20 ± 10	950 ± 110
YAO 8	6667	13996	16685	5256	590	12502		62370 ± 500	238 ± 40	525 ± 80	1455 ± 140	20 ± 10	20 ± 10	139 ± 40
POL 6	703	3652	3750	-2531	-1541	3299	-3276		2772 ± 150	1535 ± 110	15127 ± 390	4168 ± 200	366 ± 60	11316 ± 320
CAT 23	3183	6640	5839	6660	2950	1502	17044	10822		96010 ± 190	98753 ± 110	10 ± 10	0 ± 0	10 ± 10
eCAT 11	1416	4622	6392	7450	1276	-495	18247	11825	-4874		57519 ± 500	129 ± 40	10 ± 10	228 ± 50
dCAT 12	2007	4780	3456	3689	1949	1113	12706	4931	-5120	-1830		178 ± 40	40 ± 20	198 ± 40
UTC 17	8641	6216	-470	14116	12995	9940	21150	9426	17269	15664	13751		8900 ± 240	59459 ± 520
ODI 10	12446	14108	12216	25161	20705	18734	34259	17852	26573	31783	21451	5709		5752 ± 220
KRO 26	6980	4376	-1844	8904	9481	7614	16479	5270	14687	13435	11561	-978	5873	
ANF 20	10738	5984	-1499	11526	12992	11522	18222	7416	18147	16674	14201	1055	6541	-2628
AND 9	-1492	-1344	246	1570	-3425	-3001	1512	-2465	2843	595	585	6373	14765	5909
ANDF 29	5358	1281	-4050	5530	5563	4309	10723	2272	9708	7894	7050	444	5221	-2182
BAK 8	1839	6217	8971	-217	-2002	4214	-5751	-5767	13080	12318	8964	13565	25367	9780
mAll 680	1306	5572	2921	2464	-493	-444	6840	3876	2052	578	1481	13911	18929	11714
mOEu 246	953	5165	3148	1982	-618	-643	6775	3556	1966	632	1194	13592	18141	11149
mBul 30	320	6131	2696	2927	-1283	-1907	6952	3109	2566	707	1694	13478	21001	11388
mUkr 18	-759	2160	2614	602	-2660	-2284	4566	445	1835	-18	876	11191	16926	8217
mRus 198	929	5036	3182	1865	-471	-410	6809	3551	1998	795	1094	13400	18109	10961
mCEu 434	1439	5785	2824	2768	-405	-303	6895	4024	2185	626	1683	14000	19505	11925
mGer 262	1635	6420	3142	3447	-274	-192	7689	5059	2373	787	1987	14625	20655	12538
mLit 97	754	4162	1544	2005	-253	-705	7487	3359	2122	348	1553	12094	18652	10276
mPol 75	1002	5505	3524	1755	-1001	-61	3738	1128	2418	1104	1152	13499	18310	11078
x 10⁻⁵	UNC 94	HGE 14	DDC 6	ENL 10	YAM 22	YAW 14	YAO 8	POL 6	CAT 23	eCAT 11	dCAT 12	UTC 17	ODI 10	KRO 26

Tabelle 13.11: F_{ST} -Werte zwischen den Kulturen Westasiens

(Seite 3 von 3)

HGC 16	30 ± 20	3000 ± 190	109 ± 40	3079 ± 190	0 ± 0	0 ± 0	0 ± 0	218 ± 50	20 ± 10	0 ± 0	0 ± 0	20 ± 10	0 ± 0
LBK 102	0 ± 0	11781 ± 330	0 ± 0	7138 ± 250	0 ± 0	0 ± 0	2198 ± 160	11157 ± 310	0 ± 0	0 ± 0	0 ± 0	0 ± 0	79 ± 30
RSC 11	1445 ± 130	79556 ± 390	10078 ± 290	19038 ± 340	74498 ± 420	70191 ± 500	73260 ± 370	92664 ± 270	67478 ± 480	76062 ± 430	71240 ± 490	78279 ± 390	81744 ± 400
SCG 33	0 ± 0	26294 ± 410	0 ± 0	4732 ± 220	129 ± 40	267 ± 50	20018 ± 350	25394 ± 380	149 ± 40	119 ± 40	89 ± 30	109 ± 30	8930 ± 320
NHG 20	0 ± 0	446 ± 70	0 ± 0	1455 ± 120	0 ± 0	0 ± 0	0 ± 0	30 ± 20	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
NHG2 13	10 ± 10	0 ± 0	20 ± 10	69 ± 30	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
LNF 11	129 ± 40	13127 ± 340	832 ± 80	7296 ± 260	31710 ± 450	26047 ± 420	31126 ± 420	32888 ± 460	24572 ± 430	34620 ± 450	30363 ± 520	42867 ± 460	34244 ± 570
BAC 19	30 ± 20	86229 ± 290	545 ± 70	69350 ± 450	28908 ± 390	38452 ± 530	69161 ± 500	94268 ± 210	32264 ± 450	26968 ± 400	16335 ± 390	27849 ± 440	87645 ± 370
SMC 29	0 ± 0	11365 ± 340	0 ± 0	2911 ± 150	2247 ± 140	2624 ± 160	23661 ± 470	29225 ± 440	2495 ± 160	2327 ± 160	2485 ± 150	1287 ± 110	16622 ± 380
BEC 17	198 ± 40	56529 ± 480	1802 ± 140	20186 ± 420	28344 ± 430	36264 ± 540	44966 ± 510	77369 ± 440	33739 ± 460	24057 ± 380	18563 ± 360	37719 ± 490	35729 ± 510
FBC 10	337 ± 50	35155 ± 460	1485 ± 130	38729 ± 500	10573 ± 340	15503 ± 340	20453 ± 420	39392 ± 440	15068 ± 350	8851 ± 310	5861 ± 250	7029 ± 250	39194 ± 490
PWC 19	0 ± 0	15098 ± 360	30 ± 20	871 ± 100	89 ± 30	139 ± 40	109 ± 30	7237 ± 260	109 ± 30	79 ± 30	99 ± 30	307 ± 60	30 ± 20
CWC 44	0 ± 0	90565 ± 290	40 ± 20	35096 ± 450	44164 ± 480	57529 ± 510	63687 ± 490	97376 ± 150	46985 ± 530	33908 ± 490	20186 ± 370	33284 ± 420	89664 ± 290
BBC 35	0 ± 0	55737 ± 510	40 ± 20	9474 ± 290	47698 ± 510	60499 ± 470	27928 ± 410	84902 ± 410	63271 ± 510	37729 ± 490	25047 ± 450	74834 ± 460	29235 ± 410
UNC 94	0 ± 0	58539 ± 490	0 ± 0	16286 ± 360	10 ± 10	277 ± 50	30858 ± 500	63370 ± 530	366 ± 60	20 ± 10	0 ± 0	2257 ± 140	1089 ± 100
HGE 14	5534 ± 220	56717 ± 470	23275 ± 410	9900 ± 280	188 ± 40	337 ± 50	356 ± 70	19117 ± 380	416 ± 70	99 ± 30	30 ± 20	822 ± 90	198 ± 40
DDC 6	50500 ± 480	44015 ± 490	82655 ± 350	10019 ± 300	14385 ± 330	12801 ± 310	18919 ± 390	22513 ± 440	12999 ± 360	14553 ± 350	12712 ± 330	25007 ± 470	11543 ± 340
ENL 10	1475 ± 120	22691 ± 460	5544 ± 220	41996 ± 420	8336 ± 280	12177 ± 310	10910 ± 300	34799 ± 460	13622 ± 300	6336 ± 230	4109 ± 190	11939 ± 300	14929 ± 350
YAM 22	0 ± 0	91555 ± 270	663 ± 70	66112 ± 470	73904 ± 460	78319 ± 360	84912 ± 390	96733 ± 190	69092 ± 460	67746 ± 550	59430 ± 400	55757 ± 450	90149 ± 260
YAW 14	297 ± 60	86457 ± 350	4415 ± 190	11692 ± 330	58559 ± 450	65340 ± 380	91625 ± 280	86665 ± 390	55608 ± 460	54113 ± 470	50134 ± 510	68597 ± 460	46174 ± 500
YAO 8	307 ± 60	27997 ± 500	1198 ± 110	87199 ± 350	1129 ± 100	1406 ± 120	2267 ± 150	10247 ± 290	1515 ± 110	1099 ± 100	703 ± 90	1010 ± 100	6415 ± 290
POL 6	8831 ± 330	67379 ± 450	25097 ± 410	80685 ± 360	9494 ± 330	11266 ± 330	16127 ± 390	41095 ± 490	12306 ± 330	9039 ± 260	5831 ± 260	12474 ± 330	29750 ± 460
CAT 23	0 ± 0	15513 ± 340	20 ± 10	545 ± 70	1277 ± 110	2346 ± 160	3198 ± 160	15157 ± 350	2158 ± 140	931 ± 90	871 ± 80	1733 ± 130	1416 ± 110
eCAT 11	168 ± 40	38115 ± 460	1346 ± 110	1643 ± 120	28938 ± 470	27740 ± 400	30740 ± 450	44243 ± 460	25057 ± 440	27611 ± 430	24582 ± 430	34432 ± 510	21067 ± 430
dCAT 12	218 ± 50	42006 ± 460	1455 ± 120	3653 ± 190	12712 ± 300	16899 ± 350	16563 ± 400	33056 ± 430	19077 ± 360	10385 ± 320	8762 ± 250	13157 ± 310	19622 ± 410
UTC 17	27591 ± 460	5306 ± 220	33413 ± 410	604 ± 70	0 ± 0	0 ± 0	0 ± 0	89 ± 30	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
ODI 10	6207 ± 270	228 ± 50	5613 ± 240	10 ± 10	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
KRO 26	93070 ± 230	5277 ± 210	96357 ± 180	1871 ± 150	0 ± 0	0 ± 0	0 ± 0	218 ± 50	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
ANF 20		2653 ± 160	95099 ± 220	1564 ± 120	0 ± 0	0 ± 0	0 ± 0	119 ± 40	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
AND 9	8645		33215 ± 510	84180 ± 370	46540 ± 510	50767 ± 520	64568 ± 420	91961 ± 230	49658 ± 510	43283 ± 440	36313 ± 510	64944 ± 430	53480 ± 440
ANDF 29	-2633	707		7940 ± 270	0 ± 0	0 ± 0	0 ± 0	2584 ± 150	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
BAK 8	12227	-4530	5188		5000 ± 210	6247 ± 230	11266 ± 320	37353 ± 480	6722 ± 220	4455 ± 210	2554 ± 150	7791 ± 290	12029 ± 310
mAll 680	15816	-191	8746	4078		98663 ± 110	74309 ± 420	96862 ± 170	94426 ± 220	99990 ± 00	98485 ± 120	92426 ± 250	73082 ± 440
mOEu 246	15207	-445	8175	3722	-131		78458 ± 420	99634 ± 60	99990 ± 00	27671 ± 440	12256 ± 320	59727 ± 500	62796 ± 470
mBul 30	14994	-1085	7793	3246	-369	-473		90555 ± 270	47540 ± 530	68241 ± 490	65835 ± 450	59311 ± 470	74418 ± 430
mUkr 18	11345	-3368	4429	411	-1293	-1740	-1657		98505 ± 120	90882 ± 300	82437 ± 340	87902 ± 270	98594 ± 110
mRus 198	15009	-375	8010	3807	-125	-418	-71	-1539		25463 ± 430	10662 ± 350	65756 ± 490	59469 ± 560
mCEu 434	15995	-69	8885	4259	-142	41	-298	-1024	56	99990 ± 00	97406 ± 140	74715 ± 480	
mGer 262	16662	288	9258	5024	-121	130	-282	-818	165	-243		64112 ± 440	33096 ± 470
mLit 97	14221	-963	7296	3449	-206	-80	-243	-1032	-117	-267	-97		32452 ± 490
mPol 75	14600	-493	8025	2527	-149	-120	-476	-1693	-105	-164	66	92	
x 10⁻⁵	ANF 20	AND 9	ANDF 29	BAK 8	mAll 680	mOEu 246	mBul 30	mUkr 18	mRus 198	mCEU 434	mGer 262	mLit 97	mPol 75

13.3 Genotypen

Tabelle 13.12: Ergebnisse der Typisierung der autosomalen loci und der Geschlechtsbestimmung

Zusätzlich ist zu jedem Individuum die jeweilige maternalen Haplogruppenzugehörigkeit angegeben (n. best. = nicht bestimmt). In Spalte 2 zeigt die Allele des SNPs, ancestral > abgeleitet, wobei die putativ selektierte Variante in Kursivschrift ist. Gemäß dem Schema in Abschnitt 8.2.8.1 sicher reproduzierte Genotypen sind fettgedruckt, die vorbehaltlichen in Normalschrift. ¹Genotyp bereits veröffentlicht in (2014).

	Haplogruppe	ABC1 a rs1128503	ABC1 b rs2032582	ABC1 c rs1045642	ABC11 rs17822931	ADH1B a rs3811801	ADH1B b rs1229984	CASP12 rs497116	CYP3A4 rs2740574	CYP3A5 rs776746	EDAR rs3827760	HERC2 rs12913832	LCT a rs4988235	LCT b rs182549	NAT2 rs1801280	SLC12A3 rs1529927	SLC24A5 rs1426654	SLC45A2 rs16891982	TRPV6 b rs4987667	TRPV6 c rs4987682	TYR rs1042602	AMELX/Y
Probe	mtDNA	S01 C > T	S02 G > A/T	S03 C > A/T	S04 C > T	S05 C > T	S06 G > A	S08 C > T	S09 G > A	S10 A > G	S11 T > C	S12 A > G	S13 C > T	S14 G > A	S15 T > C	S16 G > C	S17 G > A	S18 C > G	S20 G > A	S21 C > T	S22 C > A	S23 Indel
<i>Äneolithikum</i>																						
DUR 1	U5a2			C/C									C/C									
KOS 2	n. best.				C/C										T/T						T/T	
MAJ 9	W			T/T		C/C	G/G	T/T				A/A¹	C/C				A/A	C/C		T/T	C/C	X/X
MOB 1	U5a1	C/T	G/T	T/T	C/T			T/T		G/G	T/T	A/A¹	C/C	G/G	T/T	G/G	A/A		A/G	C/T	C/C	
SMY 6	n. best.						G/G							G/G	C/T						T/T	
SMY 9	H7a1		G/G		C/C		G/G				T/T	A/A¹			C/T				A/A	T/T	C/C	
VAP 2	n. best.										T/T		C/C								T/T	
VIN 1	H2a2	C/T	A/T	C/T	T/T	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A	C/G	A/A	T/T	C/C	
<i>Frühe Bronzezeit – Jamnaja-Kultur</i>																						
MAJ 3	U5a1	C/T	G/T	T/T		C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A		A/A	T/T	C/C	X/X
MAJ 4	U5b2a1a		G/T	T/T	T/T							A/A¹	C/C		T/T				A/A		C/C	
MAJ 5	X	C/T	G/T	C/T	C/C	C/C	G/G	T/T	A/A	G/G	T/T	G/G¹	C/C	G/G	T/T	G/G	A/A	C/C	A/G	C/T	C/C	X/Y
OLE 1	T							T/T			T/T	A/A¹	C/C	G/G	T/T				A/A		C/C	X/X
OLE 6	n. best.	T/T		T/T	T/T						T/T		C/C						A/A	T/T		
OVI 2	K1			C/C	C/C			T/T		G/G			C/C	G/G			A/A		A/A	T/T		
OVI 3	n. best.			C/T		C/C		T/T			T/T	A/G¹	C/C	G/G	C/T		A/A				C/C	

	Haplogruppe	ABC81 a rs1128503	ABC81 b rs2032582	ABC81 c rs1045642	ABC811 rs17822931	ADH1B a rs3811801	ADH1B b rs1229984	CASP12 rs497116	CYP3A4 rs2740574	CYP3A5 rs776746	EDAR rs3827760	HERC2 rs12913832	LCT a rs4988235	LCT b rs182549	NAT2 rs1801280	SLC12A3 rs1529927	SLC24A5 rs1426654	SLC45A2 rs16891982	TRPV6 b rs4987667	TRPV6 c rs4987682	TYR rs1042602	AMELX/Y	
Probe	mtDNA	S01 C > T	S02 G > A/T	S03 C > A/T	S04 C > T	S05 C > T	S06 G > A	S08 C > T	S09 G > A	S10 A > G	S11 T > C	S12 A > G	S13 C > T	S14 G > A	S15 T > C	S16 G > C	S17 G > A	S18 C > G	S20 G > A	S21 C > T	S22 C > A	S23 Indel	
PEJ 1	U5a1	T/T	T/T	T/T	C/T			T/T			T/T	A/A¹	C/C	G/G	T/T	G/G	A/A		A/A		C/C¹		
PES 4	H5				T/T			T/T				G/G			C/C								
PES 7	H			T/T				T/T			T/T					A/A	G/G¹	A/A	T/T		C/C¹		
POP 1	T2a1b1a	C/T	G/T	C/T	C/C	C/C	G/G	T/T	A/A		T/T	A/G¹	C/C			G/G	A/A	C/G¹	A/A	T/T	C/C¹	X/Y	
POP 3	U2e1a		G/G		C/C				A/A		T/T	A/A¹	C/C	G/G			G/G¹	A/A			C/C¹		
POP 4	U5a1	C/T			C/T	C/C	G/G	T/T	A/A	G/G	C/T	A/A¹	C/C	G/G	C/T		A/A	C/G	A/A	T/T	C/C¹		
RIL 3	K1			C/C	C/C		G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A	G/G¹	A/A	T/T	A/C¹	X/X	
SUG 2	I	C/C	G/G	C/T	C/T	C/C	G/G	T/T	A/A		T/T	A/A¹	C/C	A/G	C/C	G/G	A/A	C/C¹	A/A	T/T	C/C¹		
SUG 6	H6				C/T		G/G	T/T	A/A	G/G	T/T	A/A¹			C/T		A/A	C/C¹	A/G		C/C¹		
TET 2	U4a1			T/T								A/A		G/G									
VIN 2	T1a	C/C	G/G	C/T	C/T	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	T/T	G/G	A/A	C/G¹	A/A	T/T	C/C¹	X/Y	
VIN 5	T1a											A/A¹											
VIN 12	T1							T/T		G/G		A/A¹	C/C	G/G	C/T					C/T	C/C¹		
ZAG 3	n. best.		G/T		C/C							A/G		G/G									
<i>Frühe bis mittlere Bronzezeit – Katakombengrabkultur</i>																							
KNO 4	U4	C/T		C/T	C/T	C/C		T/T	A/A	G/G	T/T	A/G¹	C/C	A/G	C/T	G/G	A/A		A/A	T/T	C/C¹	X/Y	
LIS 1	U5a1	T/T		C/T	C/C	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C			G/G	A/A		A/A	T/T	C/C¹	X/X	
LIS 2	U4b1a2				C/C		G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T		A/A		A/A		C/C¹	X/X	
LIS 3	H2a1	C/C			C/C	C/C	G/G	T/T	A/A		T/T	A/G¹	C/C		C/T	G/G	A/A	C/C¹	A/A		C/C¹		
NEV 1	U5a1	C/C	G/G	C/T	C/C	C/C	G/G	T/T	A/A		T/T	A/A¹	C/C	G/G	C/T	C/G	A/A	C/C¹	A/A	T/T	C/C¹	X/X	
NEV 3	H6				C/T	C/C		T/T	A/A	G/G		A/A¹	C/C	G/G	C/T		A/A				C/C¹		
NOZ 1	U4	C/T		C/C	C/C	C/C		T/T	A/A	G/G	T/T	G/G¹	C/C		T/T	G/G	A/A	C/G¹	A/A	T/T	C/C¹	X/Y	
NOZ 2	U4										T/T	G/G¹							A/A		C/C¹		
PEJ 3	H6	C/C	G/G				G/G	T/T	A/A		T/T	A/A¹		G/G		G/G	A/A		A/A	T/T	C/C¹	X/X	
PEJ 4	H6		G/G					T/T	A/A			A/A¹	C/C	G/G	C/T	G/G	A/A		A/A	T/T	C/C¹		
PEJ 5	U4	C/C		C/T			G/G	T/T		G/G	T/T	A/A¹			T/T	G/G	A/A	C/C¹	A/A	T/T	C/C¹	X/Y	

	Haplogruppe	ABCB1 a rs1128503	ABCB1 b rs2032582	ABCB1 c rs1045642	ABCC11 rs17822931	ADH1B a rs3811801	ADH1B b rs1229984	CASP12 rs497116	CYP3A4 rs2740574	CYP3A5 rs776746	EDAR rs3827760	HERC2 rs12913832	LCT a rs4988235	LCT b rs182549	NAT2 rs1801280	SLC12A3 rs1529927	SLC24A5 rs1426654	SLC45A2 rs16891982	TRPV6 b rs4987667	TRPV6 c rs4987682	TYR rs1042602	AMELX/Y
Probe	mtDNA	S01 C > T	S02 G > A/T	S03 C > A/T	S04 C > T	S05 C > T	S06 G > A	S08 C > T	S09 G > A	S10 A > G	S11 T > C	S12 A > G	S13 C > T	S14 G > A	S15 T > C	S16 G > C	S17 G > A	S18 C > G	S20 G > A	S21 C > T	S22 C > A	S23 Indel
SAC 2	J2b	C/T	G/T	C/T	C/T	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹		G/G	C/T		A/A		A/A	T/T	C/C¹	X/Y
SUG 5	H6				C/T		G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A				C/C¹	
TEM 1	U4			C/T	C/T		G/G			A/G	T/T	A/A¹			G/G	A/A		A/A	T/T	C/C¹		
TEM 2	H2a2	C/T	G/T	C/T	C/T	C/C	G/G	T/T	A/A		T/T	A/A¹	C/C	G/G	T/T	G/G	A/A	G/G¹	A/A		C/C¹	X/Y
TEM 4	U5a1	T/T	T/T	T/T	C/C	C/C		T/T	A/A		T/T	A/A¹	C/C	G/G	C/T	G/G	A/A	C/G¹	A/A	T/T	C/C¹	X/X
TEM 5	R1		G/G	C/C	C/T	C/C	G/G	T/T			T/T	A/A¹	C/C	A/G	C/T					T/T	C/C¹	X/Y
TEM 6	R1	C/C	G/G		C/C	C/C	G/G	T/T		G/G	T/T	A/A¹			C/C	G/G	A/A		A/A	T/T	C/C¹	X/X
TEM 7	U4		G/G			C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	T/T		A/A		A/A		C/C¹	X/Y
TEM 8	n. best.									G/G		A/G¹									C/C¹	
VIN 3	U5a1	T/T	T/T	T/T	C/C	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A	C/G¹	A/A	T/T	C/C¹	X/X
VIN 8	J1b1a1	C/T	G/T	C/T			G/G				T/T	A/A¹						C/C¹	A/A	T/T	A/C¹	
VIN 13	n. best.												C/C									
<i>Frühe bis mittlere Bronzezeit – Poltavka-Kultur</i>																						
KAL 1	N1a1a	C/T	G/T		C/T	C/C		T/T	A/A	G/G		A/G¹	C/C	G/G	C/T	G/G	A/A	C/C¹	A/A	T/T	A/A¹	X/Y
KAL 2	n. best.			T/T								G/G¹	C/C	G/G								
NIK 1	T1a		T/T	T/T	C/C	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A		A/A	T/T	C/C¹	
NIK 7	H13a1a	T/T	T/T	T/T	C/C		G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	T/T	G/G	A/A	C/G¹	A/A	T/T	C/C¹	X/Y
POD 1	W6	T/T	T/T	T/T	C/C	C/C		T/T	A/A	G/G	T/T	A/G¹	C/C	G/G	C/T	C/G	A/A	G/G¹	A/A	T/T	C/C¹	X/Y
POD 2	T	T/T	T/T	T/T	C/C	C/C	G/G	T/T	A/A	G/G	T/T	A/A¹	C/C	G/G	C/T	G/G	A/A	G/G¹	A/A		C/C¹	
<i>Frühe Bronzezeit - Kugelamphorenkultur</i>																						
CZU 2	K1a1	C/C	G/G	C/C	C/C	C/C				G/G	T/T		C/C	G/G	C/T		A/A		A/A		A/A	X/Y
ZAU 1	H2a2	C/T	G/T	C/T		C/C	G/G	T/T	A/A		T/T	A/G	C/C		C/T	G/G	A/A		A/G	C/T	A/C	X/Y
ZAU 2	J1		G/T	C/T	C/C		G/G	T/T	A/A	G/G	T/T	G/G			C/T	G/G	A/A		A/A	T/T		

