

**Identifizierung T-zellerkannter mutierter  
Neoantigene in einem humanen  
Melanommodell mittels  
Hochdurchsatzsequenzierung  
von Exomen und Transkriptomen**

**Dissertation  
zur Erlangung des Grades  
Doktor der Naturwissenschaften  
  
am Fachbereich Biologie  
der Johannes Gutenberg-Universität Mainz**

Barbara Schrörs, geb. Roider  
geb. am 07.12.1986 in Lich (Hessen)

Mainz, 2015



Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung: 23.05.2016



*Alles ist theoretisch unmöglich, bis es getan wird. (Robert A. Heinlein)*



## Zusammenfassung

Somatische Punktmutationen in Tumoren können zur Generierung von tumorspezifischen Neoantigenen führen, die von patienteneigenen T-Zellen erkannt werden. Nach solchen immunogenen Mutationen wurde im humanen Melanommodell Ma-Mel-86 durch Sequenzieren der Exome und Transkriptome von vier Tumorzelllinien und einer EBV-immortalisierten B-Zelllinie aus dem Blut der gleichen Patientin gesucht. Die Melanomzelllinien waren aus Lymphknotenmetastasen etabliert worden, die über einen Zeitraum von sechs Jahren aufgetreten waren, wobei bei den Zelllinien aus den zuletzt aufgetretenen Metastasen ein partieller bzw. vollständiger Verlust der Oberflächenexpression von HLA-Klasse I-Molekülen vorlag. Durch die Kombination der Daten beider Sequenzierverfahren und die Verwendung von parallel prozessierten Replikaten war es möglich, 181 exprimierte, nicht-synonyme, somatische Punktmutationen sicher zu identifizieren. Mithilfe der Algorithmen von NetMHC und der *Immune Epitope Database* (IEDB) wurden bei 80 ausgewählten Mutationen für die Bindung an HLA-Klasse I-Allele der Patientin insgesamt 174 Okta-, Nona- und Dekamere vorhergesagt und synthetisiert. Die Immunogenität der Peptidkandidaten wurde in IFN $\gamma$ -ELISpot-Assays mit autologen MLTC (engl. *mixed lymphocyte tumor cell culture*)-Responderpopulationen aus Blutlymphozyten überprüft. Insgesamt erwiesen sich mit diesem Verfahren 5% (4/80) der getesteten Punktmutationen als immunogen. Die erkannten Peptide wurden von den punktmutierten Genen HERPUD1<sup>G161S</sup> (engl. *homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1*), INSIG1<sup>S238F</sup> (engl. *insulin induced gene 1*), MMS22L<sup>S437F</sup> (engl. *MMS22-like, DNA repair protein*) und PRDM10<sup>S1050F</sup> (engl. *PR domain containing 10*) kodiert. HERPUD1<sup>G161S</sup> war bereits zuvor durch cDNA-Bank-Screening als ein durch HLA-B\*15:01 restringiertes Neoantigen identifiziert worden. Die drei bisher unbekanntes Neoantigene wurden mit mutationsspezifischen CD8<sup>+</sup> T-Zellklonen weiter charakterisiert. Dabei wurden HLA-A\*24:02 als gemeinsames Restriktionsmolekül bestimmt, die C-Termini der Peptide definiert und deren intrazelluläre Prozessierung bestätigt. Es zeigte sich, dass alle vier mutierten Neoantigene nur auf der Tumorzelllinie erkannt werden konnten, die aus der frühesten Metastase generiert worden war, da diese als einzige noch die restringierenden HLA-Klasse I-Allele exprimierte. Die dargestellten Ergebnisse erweitern das Spektrum bekannter mutierter Tumorantigene und zeigen eine Strategie auf zur raschen Identifizierung von immunologisch relevanten tumorspezifischen Mutationen mithilfe der Hochdurchsatzsequenzierung vorzugsweise im frühen Krankheitsverlauf.

## Summary

Somatic point mutations in tumors can lead to the generation of tumor-specific neoantigens which can be recognized by the patient's own T cells. Such immunogenic mutations were searched for in the human melanoma model Ma-Mel-86 by whole exome and transcriptome sequencing of four tumor cell lines and an autologous EBV-transformed B cell line. The melanoma cell lines had been established from lymph node metastases occurring over six years. Tumor cell lines derived from the later occurring metastases completely or partially lacked cell surface expression of HLA class I molecules. Altogether 181 expressed non-synonymous somatic point mutations could be identified by combining the data from both sequencing procedures and by processing replicates in parallel. Using the NetMHC and the Immune Epitope Database (IEDB) algorithms, for 80 selected mutations 174 octa-, nona- and decamers were predicted to bind to the patient's HLA class I alleles. The immunogenicity of these peptide candidates was tested in IFN $\gamma$  ELISpot assays with MLTC (mixed lymphocyte tumor cell culture) responder populations derived from the patient's blood lymphocytes. With this approach 5% (4/80) of the tested point mutations were found to be immunogenic. The recognized peptides were encoded by the point-mutated genes *HERPUD1*<sup>G161S</sup> (homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1), *INSIG1*<sup>S238F</sup> (insulin-induced gene 1), *MMS22L*<sup>S437F</sup> (MMS22-like, DNA repair protein) and *PRDM10*<sup>S1050F</sup> (PR domain containing 10). *HERPUD1*<sup>G161S</sup> had been identified before as a neoantigen restricted by HLA-B\*15:01 via cDNA expression screening. The three hitherto unknown neoantigens were further characterized with mutation-specific CD8<sup>+</sup> T cell clones. HLA-A\*24:02 was found to serve as a common restriction molecule. The C-termini of the peptides were defined and their intracellular processing was validated. It turned out that all four mutated neoantigens could only be recognized on the tumor cell line established first during the patient's clinical course. It was the only tumor line in this model system that expressed all of the patient's HLA class I alleles. These results broaden the spectrum of known mutated tumor antigens and involved a strategy based on high throughput sequencing which is suitable for the systematic identification of tumor-specific immunogenic mutations in due time preferably in an early disease phase.



# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>vii</b>
<b>Summary</b>	<b>viii</b>
<b>Abbildungsverzeichnis</b>	<b>xv</b>
<b>Tabellenverzeichnis</b>	<b>xvii</b>
<b>Abkürzungsverzeichnis</b>	<b>xix</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Biologischer und medizinischer Hintergrund	1
1.1.1 Maligne Melanome	1
1.1.1.1 Charakteristika	1
1.1.1.2 Behandlung	2
1.1.1.2.1 Kinaseinhibitoren	2
1.1.1.2.2 Immuntherapie	3
1.1.1.3 Das Melanommodell Ma-Mel-86	5
1.1.1.3.1 Krankheitsverlauf und Therapie	5
1.1.1.3.2 Autologe Zelllinien und primäres Patientenmaterial	5
1.1.1.3.3 Vorarbeiten	6
1.1.2 Antigenerkennung durch CD8 <sup>+</sup> T-Lymphozyten	8
1.1.2.1 T-Lymphozyten	8
1.1.2.2 Antigenprozessierung und -präsentation	8
1.1.2.3 Antigenerkennung und T-Zellaktivierung	10
1.1.2.4 Inhibitorische Signale	11
1.1.3 Wechselwirkungen zwischen Immunsystem und Malignomen	12
1.1.3.1 <i>Immunosurveillance</i> und <i>Immunoediting</i>	12
1.1.3.2 <i>Immune Escape</i> -Mechanismen	13
1.2 Methodischer Hintergrund	14
1.2.1 Identifizierung von Tumorantigenen	14
1.2.2 <i>Next-Generation Sequencing</i>	16
1.2.2.1 Prinzip des <i>Next-Generation Sequencing</i>	16
1.2.2.2 <i>Next-Generation Sequencing</i> -Varianten	19
1.2.2.3 NGS als Werkzeug in der Onkologie	20
1.2.3 Vorhersage der Peptid/HLA-Bindung mittels Algorithmen	22
1.3 Zielsetzung der Arbeit	22
<b>2 Material und Methoden</b>	<b>25</b>
2.1 Material	25
2.1.1 Geräte	25
2.1.2 Verbrauchsmaterialien, Glas- und Plastikwaren	26
2.1.3 Chemikalien und Reagenzien	27

---

2.1.4	Antibiotika . . . . .	29
2.1.5	Molekulargewichtsmarker . . . . .	29
2.1.6	Rekombinante humane Zytokine und Mitogene . . . . .	29
2.1.7	Enzyme . . . . .	30
2.1.8	Antikörper . . . . .	30
2.1.9	<i>MicroBeads</i> . . . . .	32
2.1.10	Molekular- und immunbiologische Kits . . . . .	33
2.1.11	Lösungen und Puffer . . . . .	33
2.1.12	Oligonukleotide . . . . .	34
2.1.12.1	HLA-Klasse I-spezifische Primer . . . . .	35
2.1.12.2	Primer zur Amplifizierung von Inserts in <i>pcDNA3</i> - und <i>pDONR</i> -Vektoren . . . . .	35
2.1.12.3	Primer für den Nachweis von $\beta$ -Aktin . . . . .	35
2.1.12.4	<i>INSIG1</i> -spezifische Primer . . . . .	35
2.1.12.5	<i>MMS22L</i> -spezifische Primer . . . . .	35
2.1.12.6	<i>PRDM10</i> -spezifische Primer . . . . .	36
2.1.13	Synthetische Peptide . . . . .	36
2.1.14	Plasmidvektoren . . . . .	38
2.1.15	Bakterien . . . . .	39
2.1.16	Zelllinien . . . . .	40
2.1.16.1	Zelllinien aus dem Melanommodell Ma-Mel-86 . . . . .	40
2.1.16.2	COS-7-Zellen . . . . .	40
2.1.16.3	293T-Zellen . . . . .	40
2.1.16.4	K562-Zellen . . . . .	40
2.1.17	Bakteriennährmedien . . . . .	41
2.1.18	Zellkulturmedien . . . . .	41
2.1.19	Referenzen und Datenbanken . . . . .	42
2.1.20	Software und Tools . . . . .	42
2.2	Methoden . . . . .	43
2.2.1	Zellbiologische Methoden . . . . .	43
2.2.1.1	Zellkultur . . . . .	43
2.2.1.2	Bestimmung von Zellzahl und Viabilität . . . . .	43
2.2.1.3	Kryokonservierung und Auftauen von eukaryontischen Zellen . . . . .	44
2.2.1.4	Behandlung von Zellen mit IFN $\gamma$ . . . . .	44
2.2.1.5	Isolierung peripherer Blutlymphozyten aus Vollblut und Erythrozytenlyse . . . . .	44
2.2.1.6	Anreicherung von CD8 <sup>+</sup> Zellen durch magnetische Zellseparation . . . . .	44
2.2.1.7	Generierung tumorspezifischer T-Lymphozyten in MLTCs . . . . .	45
2.2.1.8	Klonierung von Zellen nach dem Grenzverdünnungsverfahren . . . . .	45
2.2.1.9	Unspezifische Expansion von CD8 <sup>+</sup> T-Zellklonen mit OKT-3 . . . . .	47
2.2.1.10	Antikörperblockade der T-Zellen . . . . .	47
2.2.1.11	Durchflusszytometrie . . . . .	47
2.2.1.12	Transiente Transfektion durch Lipofektion . . . . .	48
2.2.1.13	Transiente und stabile Transfektion mittels Elektroporation . . . . .	49
2.2.1.14	IFN $\gamma$ -ELISpot-Assays . . . . .	50
2.2.1.14.1	Peptiderkennungstest im IFN $\gamma$ -ELISpot-Assay . . . . .	51
2.2.1.14.2	<i>Ex-vivo</i> -Test . . . . .	51
2.2.2	Molekularbiologische Methoden . . . . .	52
2.2.2.1	Kultivierung und Kryokonservierung von Bakterien . . . . .	52

---

2.2.2.2	Isolierung rekombinanter Plasmid-DNA . . . . .	52
2.2.2.3	Photometrische Konzentrationsbestimmungen von Nukleinsäuren . . . . .	52
2.2.2.3.1	Spektralphotometer . . . . .	52
2.2.2.3.2	Fluorometer . . . . .	52
2.2.2.4	Isolation genomischer DNA . . . . .	53
2.2.2.5	Isolation von Gesamt-RNA . . . . .	53
2.2.2.6	cDNA-Synthese . . . . .	53
2.2.2.7	Polymerase-Kettenreaktionen . . . . .	54
2.2.2.8	Gelelektrophoretische Auftrennung von DNA und RNA . .	55
2.2.2.8.1	Agarosegele . . . . .	55
2.2.2.8.2	Denaturierende Formaldehydgele . . . . .	55
2.2.2.9	Restriktionsverdau von doppelsträngiger DNA . . . . .	55
2.2.2.10	Aufreinigung und -konzentrierung von Nukleinsäuren . . .	56
2.2.2.10.1	Aufreinigung von DNA-Fragmenten aus Agarosegelen . . . . .	56
2.2.2.10.2	Aufreinigung von DNA über Säulen . . . . .	56
2.2.2.10.3	Phenol-Chloroform-Extraktion und Ethanol-fällung	56
2.2.2.11	Einbringen von cDNA oder cDNA-Fragmenten in Vektoren	56
2.2.2.11.1	TOPO-TA-Klonierung . . . . .	56
2.2.2.11.2	<i>Gateway</i> -Klonierung . . . . .	57
2.2.2.12	Sequenzierverfahren . . . . .	58
2.2.2.12.1	Sanger-Sequenzierung . . . . .	58
2.2.2.12.2	Hochdurchsatz-Sequenzierung . . . . .	58
2.2.3	Bioinformatische Methoden . . . . .	58
2.2.3.1	Qualitätsbewertung der Illumina-Daten . . . . .	58
2.2.3.2	Nachbearbeitung der Illumina-Rohdaten . . . . .	59
2.2.3.3	<i>Mapping</i> . . . . .	60
2.2.3.3.1	<i>Bowtie2</i> . . . . .	60
2.2.3.3.2	<i>TopHat</i> . . . . .	60
2.2.3.4	Qualitätsbewertung der <i>Mapping</i> -Ergebnisse . . . . .	61
2.2.3.5	Visualisierung der Daten . . . . .	61
2.2.3.6	Detektion und Filtern von Mutationen . . . . .	62
2.2.3.6.1	<i>Variant Calling</i> . . . . .	62
2.2.3.6.2	Annotation der detektierten Mutationen . . . . .	62
2.2.3.6.2.1	rsIDs der <i>dbSNP</i> -Datenbank . . . . .	62
2.2.3.6.2.2	Vorhersage des Effekts von Mutationen . .	63
2.2.3.6.3	Qualitätsfilter zur Identifizierung sicherer Mutationen . . . . .	63
2.2.3.6.4	Filtern somatischer Mutationen . . . . .	63
2.2.3.6.5	Textfilter zur Selektion nicht-synonymer Punktmutationen . . . . .	64
2.2.3.7	Vorhersage der Immunogenität von mutierten Peptiden . .	65
2.2.3.7.1	Generierung mutierter Proteinabschnitte für Vorhersagealgorithmen . . . . .	65
2.2.3.7.2	Vorhersage der Präsentation durch HLA-Klasse I-Moleküle . . . . .	66
2.2.3.7.3	Vorhersage der proteasomalen Spaltung . . . . .	66
2.2.3.8	Expressionsanalyse . . . . .	67
2.2.3.9	Phylogentische Analyse der Ma-Mel-86-Zelllinien . . . . .	68

<b>3</b>	<b>Ergebnisse</b>	<b>71</b>
3.1	Vorbereitung der DNA- und RNA-Proben . . . . .	71
3.1.1	Bestätigung der Identität der Ma-Mel-86-Zelllinien . . . . .	71
3.1.2	Isolation von genomischer DNA und Gesamt-RNA . . . . .	73
3.1.3	Überprüfung der Integrität der DNA- und RNA-Proben . . . . .	74
3.2	Herstellung der Sequenzierbibliotheken und Qualitätskontrollen . . . . .	75
3.3	Statistiken der Sequenzierungen . . . . .	76
3.3.1	Sequenzierrohdaten . . . . .	76
3.3.2	Aufarbeitung der Rohdaten . . . . .	77
3.3.3	Mapping an die Referenz . . . . .	79
3.3.3.1	Exomsequenzierung . . . . .	79
3.3.3.2	Transkriptomsequenzierung . . . . .	82
3.4	Bioinformatische Auswertungen . . . . .	84
3.4.1	Vergleich der Replikate . . . . .	84
3.4.1.1	Übereinstimmung der gefundenen Mutationen . . . . .	84
3.4.1.2	Ähnlichkeit der Expressionsmuster . . . . .	85
3.4.2	Phylogenetische Beziehung der Ma-Mel-86-Zelllinien . . . . .	87
3.4.3	Untersuchung des Defekts im $\beta$ 2-Mikroglobulinen . . . . .	88
3.4.4	Übersicht der somatischen Mutationen . . . . .	90
3.4.5	Detektion von potentiell immunogenen Mutationen . . . . .	91
3.4.5.1	Auswahl von mutierten Proteinen . . . . .	91
3.4.5.2	Vorhersage und Selektion von Peptidkandidaten . . . . .	94
3.5	Identifizierung immunogener Peptide . . . . .	99
3.5.1	Auswahl der T-Zellen für die Analyse der Peptidimmunogenität . . . . .	99
3.5.2	Generierung der MLTC 16C . . . . .	101
3.5.2.1	Generierung der Ma-Mel-86C-HLA-Transfektanten . . . . .	101
3.5.2.2	Kultur der MLTC 16C . . . . .	102
3.5.3	Auswahl antigenpräsentierender Zellen für die Peptidbeladung . . . . .	104
3.5.4	Optimierung der Bedingungen zur Reduktion spontaner IFN $\gamma$ -Sekretion . . . . .	105
3.5.5	Test der Erkennung der mutierten Peptide . . . . .	106
3.5.5.1	<i>Ex-vivo</i> -Test . . . . .	110
3.5.6	Expressionsanalyse der immunogenen Mutationen auf mRNA-Ebene . . . . .	111
3.6	Generierung mutationsspezifischer T-Zellklone . . . . .	113
3.6.1	Klonierung der MLTC 1A.1 . . . . .	113
3.6.2	Klonierung der MLTC 3A . . . . .	114
3.6.3	Klonierung der MLTC 16C . . . . .	114
3.7	Charakterisierung der Peptiderkennung . . . . .	116
3.7.1	PRDM10 <sup>mut</sup> . . . . .	117
3.7.1.1	Bestimmung der HLA-Restriktion und Validierung der Peptidprozessierung . . . . .	118
3.7.1.2	Lokalisation der peptidkodierenden Region . . . . .	119
3.7.2	INSIG1 <sup>mut</sup> . . . . .	120
3.7.2.1	Bestimmung der HLA-Restriktion und Validierung der Peptidprozessierung . . . . .	120
3.7.2.2	Lokalisation der peptidkodierenden Region . . . . .	121
3.7.3	MMS22L <sup>mut</sup> . . . . .	122
3.7.3.1	Bestimmung der HLA-Restriktion und Validierung der Peptidprozessierung . . . . .	122
3.7.3.2	Lokalisation der peptidkodierenden Region . . . . .	123

<b>4</b>	<b>Diskussion</b>	<b>125</b>
4.1	Bewertung der generierten Sequenzierdaten . . . . .	125
4.2	Bioinformatische Analysen . . . . .	127
4.2.1	Stammbaum der Zelllinien . . . . .	127
4.2.2	Analyse des Mutationsprofils von Ma-Mel-86 . . . . .	128
4.3	Identifizierung mutierter Neoantigene . . . . .	130
4.3.1	Strategie der Lokalisierung relevanter Punktmutationen . . . . .	130
4.3.2	Auswahl potentiell HLA-bindender Peptide . . . . .	131
4.3.3	Überprüfung der Peptidimmunogenität . . . . .	132
4.3.4	Charakteristika der identifizierten Neoantigene . . . . .	134
4.3.4.1	INSIG1 <sup>mut</sup> . . . . .	134
4.3.4.2	PRDM10 <sup>mut</sup> . . . . .	135
4.3.4.3	MMS22L <sup>mut</sup> . . . . .	136
4.3.4.4	Eigenschaften der identifizierten Neoantigene . . . . .	138
4.3.5	Bedeutung mutierter Antigene . . . . .	139
4.4	<i>Immune Escape</i> und <i>Immunoediting</i> in Ma-Mel-86 . . . . .	140
<b>5</b>	<b>Fazit</b>	<b>143</b>
<b>A</b>	<b>Anhang</b>	<b>145</b>
A.1	Vektorkarten . . . . .	145
A.2	Sequenzen . . . . .	146
A.2.1	<i>IUPAC Ambiguity Code</i> . . . . .	146
A.2.2	Illumina-Adaptersequenzen . . . . .	146
A.2.3	Sequenzen der immunogen mutierten Gene und Proteine . . . . .	147
A.2.3.1	HERPUD1 <sup>mut</sup> . . . . .	147
A.2.3.2	INSIG1 <sup>mut</sup> . . . . .	148
A.2.3.3	MMS22L <sup>mut</sup> . . . . .	149
A.2.3.4	PRDM10 <sup>mut</sup> . . . . .	151
A.3	Anhang zum bioinformatischen Teil . . . . .	153
A.3.1	Workflow . . . . .	153
A.3.2	Eigene <i>Perl</i> -Skripte . . . . .	154
A.3.2.1	mappedReads_lengthDistribution.pl . . . . .	154
A.3.2.2	compareRefSeqUCSC.pl . . . . .	154
A.3.2.3	peptideCandidates_TAB.pl . . . . .	156
A.3.2.4	peptideCandidates_TAB_to_FASTA.pl . . . . .	158
A.3.2.5	adjust_zygosity_stringent.pl . . . . .	159
A.3.2.6	VCF_tab_to_fasta_alignment_advanced.pl . . . . .	161
A.4	Inhaltsverzeichnis des elektronischen Anhangs . . . . .	164
	<b>Literaturverzeichnis</b>	<b>xxi</b>
	<b>Eidesstattliche Erklärung</b>	<b>xliii</b>
	<b>Danksagung</b>	<b>xlvi</b>
	<b>Lebenslauf</b>	<b>xlvii</b>



# Abbildungsverzeichnis

1.1	Krankheitsverlauf und Behandlung der Patientin INTH. . . . .	5
1.2	Chronologische Einordnung des primären Patientenmaterials des Melanommodells Ma-Mel-86. . . . .	6
1.3	Antigenprozessierung und -präsentation über MHC-Klasse I-Moleküle. . . . .	9
1.4	Aktivierende und hemmende Signale für CD8 <sup>+</sup> T-Zellen. . . . .	10
1.5	Illumina/Solexa-Sequenzierbibliotheken. . . . .	17
1.6	Cluster-Generierung in der <i>Flow Cell</i> und Vorbereitung der Sequenzierung. . . . .	18
1.7	Ablauf der Exomanreicherung in genomischen Sequenzierbibliotheken. . . . .	19
1.8	Generierung der Sequenzierbibliothek für die Transkriptomsequenzierung. . . . .	20
2.1	Filtern somatischer Mutationen. . . . .	64
3.1	Validierung der Identität der Ma-Mel-86-Zelllinien. . . . .	72
3.2	Zeitschema der Probengewinnung für die Exom- bzw. Transkriptomsequenzierungen der Ma-Mel-86-Zelllinien. . . . .	73
3.3	Überprüfung der Probenintegrität. . . . .	75
3.4	Bioanalyzertläufe zur Überprüfung der Sequenzierbibliotheken. . . . .	76
3.5	Ergebnis der Exom- und Transkriptomsequenzierung. . . . .	77
3.6	Qualitätsbereinigung der Rohdaten. . . . .	78
3.7	Erfolg des Exom-Mappings. . . . .	81
3.8	Übereinstimmung der gefundenen Mutationen in den Replikaten. . . . .	85
3.9	Distanzmatrix der Genexpressionsprofile. . . . .	86
3.10	Phylogenetische Beziehungen der Ma-Mel-86-Linien. . . . .	88
3.11	Coverage-Profile der Exomsequenzierungen am B2M-Lokus. . . . .	89
3.12	Einzelnukleotidaustausche in den somatischen Mutationen der Ma-Mel-86-Exome. . . . .	90
3.13	Strategie zur Auswahl mutierter Proteine für die Vorhersage von potentiell immunogenen Peptiden. . . . .	92
3.14	Auswahl mutierter Proteine für die Vorhersage von potentiell HLA-bindenden Peptiden. . . . .	93
3.15	Somatische nicht-synonyme Punktmutationen exprimierter Gene in Ma-Mel-86. . . . .	94
3.16	Schematische Darstellung der relevanten Proteinabschnitte für die Peptidvorhersagen. . . . .	95
3.17	Ergebnisse der Vorhersagealgorithmen. . . . .	97
3.18	Vortest der Ma-Mel-86-MLTCs. . . . .	100
3.19	Validierung der stabilen Expression der in Ma-Mel-86C transfizierten HLA-Klasse I-Moleküle. . . . .	102
3.20	Wachstum und Charakterisierung der MLTC 16C. . . . .	103
3.21	Auswahl der APCs für die Peptidbeladung. . . . .	104
3.22	Vortests der MLTCs zur Minimierung spontaner IFN $\gamma$ -Sekretion. . . . .	105
3.23	Peptiderkennung durch MLTC 3A. . . . .	107

---

3.24	Validierung der Peptidreaktivität der MLTC 3A in An- und Abwesenheit HLA-transfizierter peptidpräsentierender 293T-Zellen. . . . .	108
3.25	Peptiderkennungstests mit weiteren MLTCs. . . . .	109
3.26	Testung der Ex-vivo-Reaktivität gegen mutierte Peptide. . . . .	110
3.27	mRNA-Expressionslevel der Gene mit immunogenen Mutationen. . . . .	111
3.28	Generierung peptidreaktiver T-Zellklone aus den MLTCs 1A.1, 3A und 16C. . . . .	115
3.29	Erkennung der einzelnen Peptide aus den Peptidpools. . . . .	116
3.30	Fragmentierung der antigenkodierenden cDNA. . . . .	117
3.31	Bestimmung des präsentierenden HLA-Klasse I-Moleküls und Validierung der intrazellulären Prozessierung des Antigens PRDM10 <sup>mut</sup> . . . . .	118
3.32	Bestimmung des C-terminalen Peptidendes des Neoantigens PRDM10 <sup>mut</sup> . . . . .	119
3.33	Bestimmung des präsentierenden HLA-Klasse I-Moleküls und Validierung der intrazellulären Prozessierung des Antigens INSIG1 <sup>mut</sup> . . . . .	120
3.34	Bestimmung des C-terminalen Peptidendes des Neoantigens INSIG1 <sup>mut</sup> . . . . .	121
3.35	Bestimmung des präsentierenden HLA-Klasse I-Moleküls und Validierung der intrazellulären Prozessierung des Antigens MMS22L <sup>mut</sup> . . . . .	123
3.36	Bestimmung des C-terminalen Peptidendes des Neoantigens MMS22L <sup>mut</sup> . . . . .	124
4.1	Übersicht der von den immunogen mutierten Proteinen HERPUD1, INSIG1, MMS22L und PRDM10 abgeleiteten, getesteten Peptide. . . . .	133
A.1	Vektorkarte von <i>pcDNA<sup>TM</sup>3.1/V5-His TOPO<sup>®</sup></i> . . . . .	145
A.2	Vektorkarte von <i>pDONR<sup>TM</sup>221</i> . . . . .	145
A.3	Rechnergestützter Workflow zur Lokalisierung von potentiell immunogenen exprimierten, somatischen, nicht-synonymen Punktmutationen. . . . .	153



# Tabellenverzeichnis

1.1	Übersicht und HLA-Klasse I-Expression der stabilen Zelllinien des Melanommodells Ma-Mel-86. . . . .	6
1.2	Übersicht der bereits im Melanommodell Ma-Mel-86 identifizierten Antigene, die von autologen melanomreaktiven T-Zellen erkannt wurden. . . . .	7
2.1	Liste der auf Immunogenität getesteten synthetischen Peptide. . . . .	37
2.2	Verwendete Plattentypen und Zellen für die Kultivierung und Restimulation von T-Lymphozyten. . . . .	46
2.3	Antikörper für Blockadeexperimente. . . . .	47
2.4	Übersicht der verwendeten Elektroporationsprogramme. . . . .	49
2.5	Übersicht der genutzten PCR-Programme. . . . .	54
2.6	Übersicht der verwendeten Algorithmen zur Vorhersage der HLA-Klasse I-Bindewahrscheinlichkeit von mutierten Okta-, Nona- und Dekameren. . . . .	66
3.1	Isolation von genomischer DNA (gDNA) und Gesamt-RNA aus Ma-Mel-86-Zellen. . . . .	74
3.2	Übersicht der Mapping-Statistiken der Exomsequenzierung. . . . .	80
3.3	Übersicht der Mapping-Statistiken der Transkriptomsequenzierungen. . . . .	83
3.4	Abdeckung der B2M-Exons in den Exomsequenzierungen der Ma-Mel-86-Zelllinien. . . . .	89
3.5	Übersicht der verwendeten Algorithmen zur Peptidbindungsvorhersage. . . . .	96
3.6	Übersicht der verfügbaren MLTCs im Modell Ma-Mel-86. . . . .	99
3.7	Klonierung der HLA-Transfektanten von Ma-Mel-86C. . . . .	101
3.8	Übersicht zur Erkennung der Peptide. . . . .	112
3.9	Übersicht der mittels NGS gefundenen Neoantigene im Melanommodell Ma-Mel-86. . . . .	124
A.1	Ambiguity code zur Bezeichnung nicht eindeutig festgelegter Basen in Nukleinsäuresequenzen. . . . .	146



# Abkürzungsverzeichnis

ANN	artifizielle neuronale Netzwerke, engl. <i>artificial neural networks</i>
APC	antigenpräsentierende Zellen, engl. <i>antigen-presenting cells</i>
cDNA	komplementäre DNA, engl. <i>complementary DNA</i>
CDR	engl. <i>complementarity determining region</i>
cfu	engl. <i>colony forming units</i>
CMV	Cytomegalievirus
dsDNA	doppelsträngige DNA
<i>E. coli</i>	<i>Escherichia coli</i>
EBV	Epstein-Barr-Virus
ELISA	engl. <i>enzyme-linked immunosorbent assay</i>
ELISpot	engl. <i>enzyme-linked immunosorbent spot</i>
ER	endoplasmatisches Retikulum
FACS	engl. <i>fluorescence-activated cell sorting</i>
gDNA	genomische DNA
HLA	humanes Leukozytenantigen, engl. <i>human leukocyte antigen</i>
HMT	Histon-Methyltransferase
IFN	Interferon
IL	Interleukin, z.B. IL-2
MACS	magnetische Zellseparation, engl. <i>magnetic activated cell sorting</i>
MAPK	MAP-Kinase
MDA	melanozytäres Differenzierungsantigen
MHC	Haupthistokompatibilitätskomplex, engl. <i>major histocompatibility complex</i>
MLTC	engl. <i>mixed lymphocyte tumor cell culture</i>
NGS	engl. <i>Next-Generation Sequencing</i>
PCR	engl. <i>polymerase chain reaction</i>
PHA	Phytohämagglutinin
RFP	rotfluoreszierendes Protein, engl. <i>red fluorescent protein</i>
rhIL-2	rekombinantes humanes Interleukin-2
SEREX	engl. <i>serological identification of antigens by recombinant expression cloning</i>
SNVs	Einzelnukleotidvarianten, engl. <i>single nucleotide variants</i>

---

SRA	<i>Sequence Read Archive</i>
TAA	tumorassoziertes Antigen
TAP	engl. <i>transporter associated with antigen processing</i>
TCR	T-Zellrezeptor, engl. <i>T cell receptor</i>
TIL	tumorinfiltrierende T-Zellen
TMD	Transmembrandomäne
TSA	tumorspezifisches Antigen

# 1 Einleitung

## 1.1 Biologischer und medizinischer Hintergrund

### 1.1.1 Maligne Melanome

#### 1.1.1.1 Charakteristika

Die zum ersten Mal im Jahr 1812 von René Theophile Laennec beschriebene Erkrankung „Melanosis“ (McLeod & Davis, 1992; Whiteman *et al.*, 2011) verursacht heutzutage unter der Bezeichnung kutanes oder malignes Melanom ca. 90% der Hautkrebs-Mortalität und stellt somit die wohl gefährlichste Form aller Hautkrebsarten dar (Garbe *et al.*, 2010). Primär handelt es sich um eine Neoplasie der pigmentbildenden Zellen der Haut, den sogenannten Melanozyten. Maligne Melanome können jedoch auch in der Aderhaut des Auges (uveales Melanom), im Innenohr, in den Hirnhäuten, auf Schleimhäuten und als sogenannte akrale Melanome in der unbehaarten Haut der Handflächen und Fußsohlen, sowie unter den Nägeln auftreten (Whiteman *et al.*, 2011; Garbe *et al.*, 2010). Insgesamt weisen Melanome ein breites phänotypisches und genotypisches Spektrum auf. Zudem unterliegen die anatomische Lokalisation und das Alter der Patienten bei der ersten Manifestation einer hohen Diversität (Whiteman *et al.*, 2011).

Risikofaktoren für eine Erkrankung sind ausgedehnte Exposition gegenüber UV-Strahlung (Eklund & Malec, 1978; Teppo *et al.*, 1978), genetische Prädisposition und das Vorhandensein vieler bzw. dysplastischer Nävi (Skolnick *et al.*, 1994). Als prognostische Faktoren für den Befund eines primären Melanoms ohne Metastasen (Stadien I und II), werden die vertikale Tumordicke (*Breslow's depth*), das Bestehen von Ulzerationen, die mitotische Aktivität und das Invasionslevel (*Clark's level*) herangezogen (Garbe *et al.*, 2010). Hier liegt die 5-Jahres-Überlebensrate zwischen 53% (Tumordicke > 4 mm) und 97% (Tumordicke ≤ 1 mm). Liegen Metastasen in den umliegenden Lymphknoten vor (Stadium III), sinkt die 5-Jahres-Überlebensrate in Abhängigkeit von der Anzahl der betroffenen Lymphknoten auf 39-70%. Für Patienten mit entfernten Metastasen (Stadium IV) werden 1-Jahres-Überlebensraten von 33-62% (Balch *et al.*, 2009) und 5-Jahres-Überlebensraten von unter 5% (Cummins *et al.*, 2006) angegeben.

Das zuerst entdeckte melanomassoziierte Onkogen ist *RAS* (Sekiya *et al.*, 1983; Albino *et al.*, 1984, 1989), wobei das Familienmitglied *NRAS* in Melanomen am häufigsten mutiert ist. *NRAS*-Mutationen liegen in etwa 15% der malignen Melanome

vor (Whiteman *et al.*, 2011). Davies *et al.* (2002) zeigten das onkogene Transformationspotential von Mutationen in der Kinase-Domäne der Serin-Threonin-Kinase BRAF. BRAF ist ebenso wie NRAS Bestandteil des MAP-Kinase (MAPK)-Signalweges und beteiligt an der zellulären Antwort auf Wachstumssignale. Über 60% der Melanome weisen eine somatische *Missense*-Mutation in *BRAF* auf, die in über 90% zur Substitution V600E führt, also einem Austausch von Valin gegen Glutaminsäure an Aminosäure-Position 600 (Jang & Atkins, 2014). *NRAS*- und *BRAF*-Mutationen liegen in Melanomen nicht gleichzeitig vor und führen unabhängig voneinander zur Aktivierung des MAPK-Signalweges und somit zu unkontrollierter Proliferation (Jang & Atkins, 2014; Davies *et al.*, 2002). Weiterhin sind in malignen Melanomen diverse genetische Alterationen in Tumorsuppressorgenen (beispielsweise *CDKN2A*, *PTEN* oder *TP53*) oder anderen Genen (wie CyclinD1 oder *MITF*) beschrieben (Gray-Schopfer *et al.*, 2007).

Bei etwa 5-10% der Melanomfälle handelt es sich um sogenannte familiäre Melanome. Hier liegt eine Anhäufung von Melanomerkkrankungen innerhalb einer Familie vor, welche nicht durch einen gemeinsamen Lebensstil, ähnliche Umwelteinflüsse oder durch das Grundrisiko für ein Melanom erklärt werden kann (Florell *et al.*, 2005). Charakteristisch für Patienten mit einer ererbten Melanomprädisposition sind ein geringeres Alter bei der Erstdiagnose, ein erhöhte Anzahl an klinisch atypischen Nävi und die Entwicklung von multiplen Primärtumoren. Bisher wurden Keimbahnmutationen in *CDKN2A* oder *CDK4* mit einem stark erhöhten Melanomrisiko in Verbindung gebracht (Punternvoll *et al.*, 2013).

### 1.1.1.2 Behandlung

In frühen Stadien ist die operative Entfernung des Primärtumors kurativ für etwa 80% der Patienten (Gray-Schopfer *et al.*, 2007). Fortgeschrittene und inoperable Melanome sprechen nur sehr schlecht auf Therapien an. Bis 2011 wurden Stadium IV-Patienten mit dem Zytostatikum Dacarbazin (DITC) oder rekombinantem humanen Interleukin (IL-)2 zur Aktivierung tumorreaktiver T-Zellen behandelt (Jang & Atkins, 2014). Durch die Entwicklung und Zulassung weiterer Therapien wurde der Weg für eine personalisierte Melanombehandlung geebnet. Die nächsten Paragraphen geben eine Übersicht hierzu.

#### 1.1.1.2.1 Kinaseinhibitoren

Da die Mehrheit der malignen Melanome eine konstitutive Kinaseaktivität durch mutiertes *BRAF*-V600E aufweist (siehe oben), wurden immer spezifischere BRAF-Inhibitoren entwickelt. Vemurafenib (Zelboraf®, Roche) war der erste oral applizierte BRAF-Inhibitor (Barrett *et al.*, 2008; Joseph *et al.*, 2010), welcher von der FDA (engl. *food and drug administration*) im Jahr 2011 und in Deutschland im Jahr 2012 für die Behandlung von inoperablen oder metastasierten Melanomen mit *BRAF*-V600E zugelassen wurde (Jang & Atkins, 2014; Livingstone *et al.*, 2012). Chapman *et al.* (2011) zeigten in einer Phase III-Studie mit 675 zuvor unbehandelten Patienten mit metastasierten Melanomen

Ansprechraten von 48% nach Behandlung mit Vemurafenib im Vergleich zu Dacarbazin (5%). Jedoch ist der Effekt oft nur transient, und es entwickeln sich im Median nach etwa 5-7 Monaten Resistenzen gegenüber BRAF-Inhibitoren, die beispielsweise auf weitere Veränderungen in *BRAF* oder auf neue Mutationen in *NRAS* oder *MEK* zurückzuführen sind (Jang & Atkins, 2014). Letztere Resistenzen können teilweise durch die Gabe anderer Kinaseinhibitoren gegen MEK (Gilmartin *et al.*, 2011; Flaherty *et al.*, 2012) oder ERK (Morris *et al.*, 2013) überwunden werden.

### 1.1.1.2 Immuntherapie

Melanome gelten aufgrund der hohen Frequenz an tumorreaktiven T-Zellen im Blut der Patienten und im Tumorgewebe als immunogen (Anichini *et al.*, 2004). Zudem sind die meisten bisher identifizierten tumorassoziierten Antigene aus Untersuchungen an Melanomen hervorgegangen. Dies mag ein Resultat des Umstands sein, dass sich aus Melanombiopsien im Vergleich zu anderen Tumorarten leichter stabile Zelllinien herstellen lassen. Das dadurch in nahezu unbegrenzten Mengen zur Verfügung stehende Tumormaterial ermöglichte eine Reihe von *In-vitro*- und *In-vivo*-Studien zur Optimierung immuntherapeutischer Strategien. Generell unterscheidet man hierbei vier Hauptansätze, nämlich die Zytokintherapie, die therapeutische Vakzinierung, den adoptiven T-Zelltransfer und den Einsatz immunmodulatorischer Antikörper (Immun-*Checkpoint*-Inhibitoren).

Wie bereits erwähnt, ist die Gabe von rekombinantem humanem IL-2 eine der zumindest in den USA zugelassenen Behandlungen für Melanompatienten. Die intravenöse Hochdosisadministration von 720.000 IU/kg führte zu partiellen oder vollständigen Regressionen, die jedoch von sepsisähnlichen Nebenwirkungen mit teils tödlichem Ausgang begleitet wurden (Rosenberg *et al.*, 1994; Atkins *et al.*, 1999). In Deutschland ist die Behandlung mit Interferon- $\alpha$  (IFN- $\alpha$ ) als adjuvante Melanomtherapie zugelassen (AWMF *et al.*, 2013; Mocellin *et al.*, 2010). Auch hier kommt es zu Toxizitäten, die ein Einbußen der Lebensqualität der behandelten Patienten zur Folge haben können. Weitere **Zytokintherapien**, die im Rahmen von Studien untersucht wurden, betrafen beispielsweise IL-21 (Thompson *et al.*, 2008), GM-CSF (engl. *granulocyte macrophage colony-stimulating factor*; Spitler *et al.*, 2000), IL-12, IL-15 oder IL-18 (Nicholas & Lesinski, 2011).

Bei der **Vakzinierung** gegen Melanome kann theoretisch auf das zuvor erwähnte breite Spektrum von bereits identifizierten melanomassoziierten Antigenen zurückgegriffen werden. Jedoch zeigen sich hier bisher nur sehr schwache Antworten, was vermutlich mit der immunsuppressiven Wirkung im Tumormikromilieu und der Fähigkeit etablierter Tumoren, der Erkennung durch das Immunsystem zu entgehen, zusammenhängt (Kaufman, 2012). Insgesamt wurden und werden sehr unterschiedliche Vakzinierungsstrategien verfolgt, welche neben peptid- oder proteinbasierten Impfstoffen beispielsweise die Impfung mit Tumorzellen, Gangliosiden, antigenkodierender Plasmid-DNA oder Hitzeschock-Proteinen

einschließen. Außerdem werden Impfungen unter Verwendung von dendritischen Zellen, rekombinanten viralen Vektoren und onkolytischen Viren verfolgt (Weiss *et al.*, 2014; Blanchard *et al.*, 2013; Senzer *et al.*, 2009).

Das Prinzip des **adoptiven T-Zell-Transfers** beruht auf der *Ex-vivo*-Expansion von tumorreaktiven autologen T-Zellen und deren anschließender Reinfusion in den Patienten, der zuvor mit einer lymphodepletierenden Behandlung konditioniert wurde. Die Konditionierung dient hierbei der Eliminierung insbesondere zellulärer immunregulatorischer Effektoren, die die Aktivität der transferierten Zellen hemmen könnten (Shapira-Frommer & Schachter, 2012). Ein erstes Protokoll für dieses Verfahren stammt von Rosenberg *et al.* (1988). Bei den T-Lymphozyten für den adoptiven Transfer handelt es sich entweder um tumorinfiltrierende T-Zellen (TIL), die basierend auf ihrer Reaktivität gegen den Tumor selektiert werden, um unselektierte junge TIL-Massenkulturen oder um genmodifizierte periphere T-Lymphozyten. Letztere werden aus dem Blut des Patienten gewonnen und mittels retroviraler Transduktion mit tumorreaktiven T-Zellrezeptoren oder chimären Antigenrezeptoren (CARs) ausgestattet. Bisherige Studien zeigten vielversprechende Ergebnisse mit Ansprechraten von 40-72% und einem nachweisbaren Langzeiteffekt der Behandlung (Shapira-Frommer & Schachter, 2012).

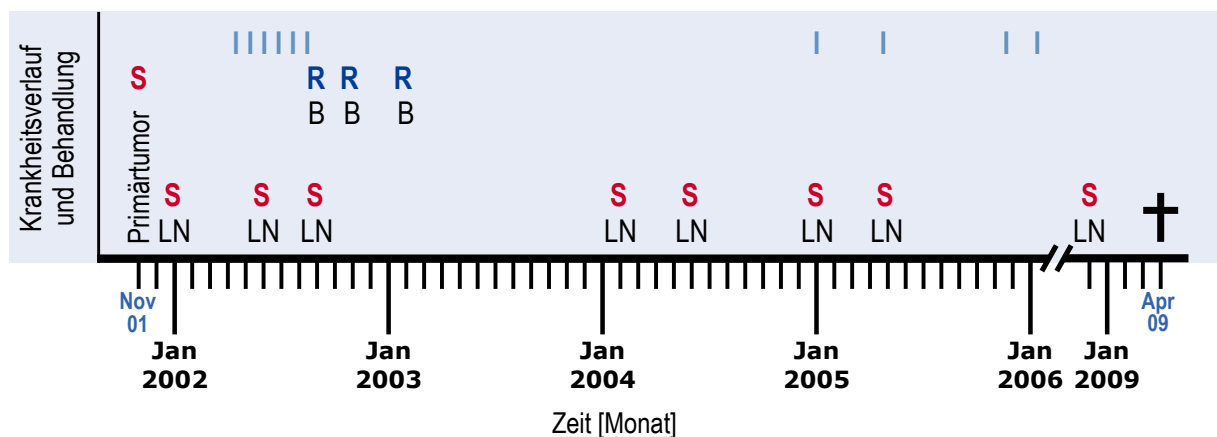
Die wohl größten Erfolge im Bereich der Melanomimmuntherapie konnten bisher durch sogenannte **Immun-Checkpoint-Inhibitoren** erzielt werden. Die immunmodulatorischen Antikörper blockieren die Interaktion von inhibitorischen Korezeptoren auf der T-Zelloberfläche mit ihren Liganden auf der Zelloberfläche von antigenpräsentierenden Zellen (APC, engl. *antigen-presenting cells*). Ein Beispiel ist CTLA-4 (engl. *cytotoxic T-lymphocyte-associated protein 4*), das mit dem aktivierenden Rezeptor CD28 um den gemeinsamen Liganden B7 konkurriert. *Checkpoint*-Inhibitoren wie der monoklonale Antikörper Ipilimumab (Yervoy®, Bristol-Myers Squibb) verhindern die CTLA-4/B7-Interaktion, so dass die T-Zelle aktiviert bleibt. Zudem wurde in Mäusen gezeigt, dass durch  $\alpha$ -CTLA-4-Antikörper auch CTLA-4<sup>+</sup> regulatorische T-Zellen im Tumor eliminiert werden (Simpson *et al.*, 2013) und dass für die maximale therapeutische Effektivität des Antikörpers eine Bindung an Effektor- und an regulatorische T-Zellen nötig ist (Peggs *et al.*, 2009). Ipilimumab ist seit 2011 in den USA und auch in Europa für die Behandlung von inoperablen oder metastasierten Melanomen zugelassen (Sondak *et al.*, 2011; Livingstone *et al.*, 2012). In neueren Studien stehen Immun-*Checkpoint*-Inhibitoren gegen den inhibitorischen Rezeptor PD-1, der im Laufe der T-Zellerschöpfung hochreguliert wird, und gegen seinen Liganden PD-L1 auf Tumorzellen im Fokus (Topalian *et al.*, 2012; Brahmer *et al.*, 2010). Die Kombination von Antikörpern gegen CTLA-4 und PD-1 hat in einer ersten Studie zu einer eindrucksvollen Wirkungssteigerung geführt und wird in aktuellen Studien weiter untersucht (Wolchok *et al.*, 2013). Außerdem wird nach prädiktiven Faktoren für den therapeutischen Erfolg der immunmodulatorischen Antikörper gesucht (Taube *et al.*, 2014). Generell setzt jedoch die Wirkung von Immun-*Checkpoint*-Inhibitoren eine präexistente zelluläre Immunantwort voraus (Cha *et al.*, 2014).



### 1.1.1.3 Das Melanommodell Ma-Mel-86

#### 1.1.1.3.1 Krankheitsverlauf und Therapie

Die Basis der vorliegenden Arbeit bildete das humane Melanommodell Ma-Mel-86 der Patientin INTH. Die Erstdiagnose der Erkrankung erfolgte im November 2001 (vgl. **Abb. 1.1**). Primärtumor und Lymphknotenmetastasen wurden operativ entfernt. Die aufgetretenen Hirnmetastasen wurden radioaktiv bestrahlt. Die Patientin erhielt zudem zwei verschiedene Immuntherapien. Zunächst wurden im Jahr 2002 Peptidvakzinierungen in Kombination mit der Gabe mehrerer Zyklen IFN- $\alpha$  und IL-2 durchgeführt. Ende 2004 bis einschließlich 2006 wurde die Patientin mit Tumorlysat-beladenen dendritischen Zellen vakziniert. Die Patientin wurde zu keinem Zeitpunkt mit mutagenen Chemotherapeutika behandelt. Trotz der hohen Metastasierungsrate überlebte die Patientin fast acht Jahre bis April 2009 (A. Paschen, pers. Mitteilung).



**Abb. 1.1: Krankheitsverlauf und Behandlung der Patientin INTH.** Die Diagnose des Primärtumors der Patientin INTH erfolgte im November 2001. Bis zu ihrem Tod im April 2009 traten mehrere Lymphknoten- (LN) und Hirnmetastasen (B) auf. Der Primärtumor und alle LN wurden operativ entfernt (S). Alle Hirnmetastasen wurden mittels Strahlentherapie (R) behandelt. Zudem erhielt die Patientin zwei verschiedene Immuntherapien (I) im Jahr 2002 und von 2004 bis 2006 (Abb. verändert aus der Dissertation von Silke Lübcke (geb. Lubojanski)).

#### 1.1.1.3.2 Autologe Zelllinien und primäres Patientenmaterial

Im Abstand von ca. ein bis zwei Jahren wurden aus vier der Lymphknotenmetastasen stabile Melanomzelllinien generiert (vgl. **Abb. 1.2**, freundlicherweise von A. Paschen und D. Schadendorf, Klinikum für Dermatologie des Universitätsklinikums Essen, zur Verfügung gestellt), nämlich Ma-Mel-86A (Januar 2002), -86B (Februar 2004), -86C (April 2005) und -86F (Dezember 2008). Neben den vier Tumorzelllinien wurde eine autologe EBV (Epstein-Barr-Virus)-immortalisierte B-Zelllinie (INTH-EBV-B) generiert. Außerdem wurde der Patientin zu verschiedenen Zeitpunkten Blut abgenommen, so dass autologe PBMCs und T-Lymphozyten für die Analyse der T-zellmedierten Immunantwort zur Verfügung standen.

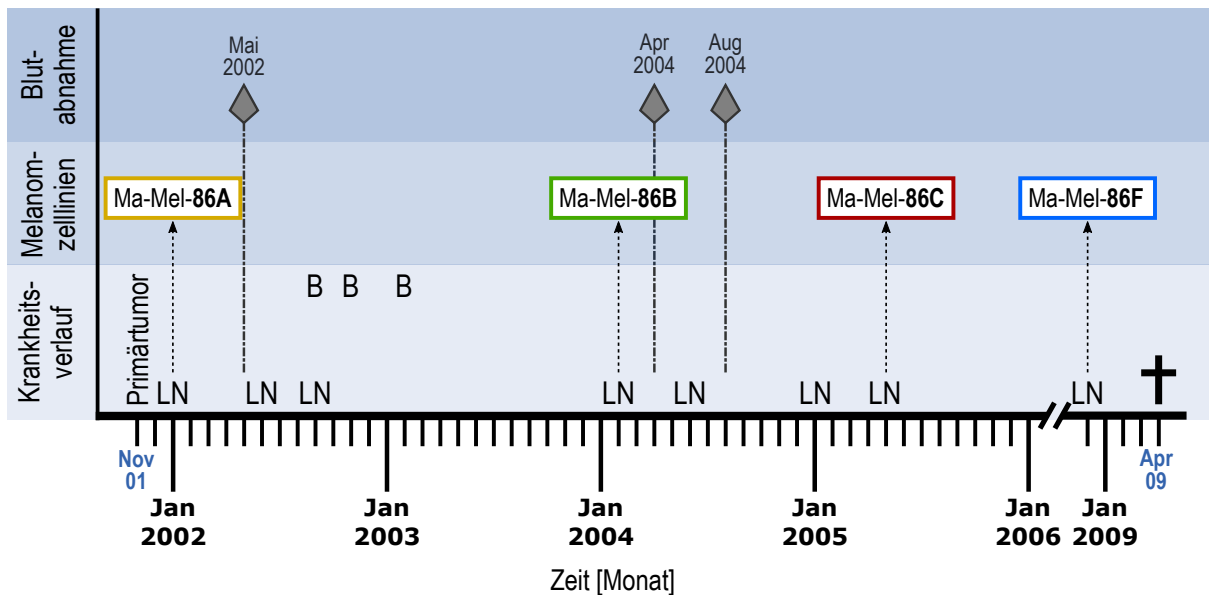
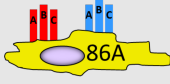

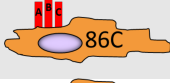


Abb. 1.2: Chronologische Einordnung des primären Patientenmaterials des Melanommodells Ma-Mel-86. Im Krankheitsverlauf der Patientin INTH traten mehrere Lymphknoten- (LN) und Hirnmetastasen (B) auf. Aus vier der Lymphknotenmetastasen wurden die stabilen Melanomzelllinien Ma-Mel-86A, -86B, -86C und -86F generiert. Zudem wurden Lymphozyten aus Blutabnahmen (◆) verschiedener Zeitpunkte isoliert (Abb. verändert aus der Dissertation von Silke Lübcke (geb. Lubojanski)).

### 1.1.1.3.3 Vorarbeiten

Die vier Melanomzelllinien des Modells unterschieden sich charakteristisch in der Expression ihrer HLA-Klasse I-Moleküle (vgl. **Tab. 1.1**). Ma-Mel-86A exprimiert alle sechs HLA-Klasse I-Allele, wohingegen Ma-Mel-86C einen Haplotyp-Verlust und Ma-Mel-86B sowie -86F komplette HLA-Klasse I-Verluste aufwiesen. Zudem war Ma-Mel-86A im Gegensatz zu den anderen Melanomzelllinien negativ für die Expression melanozytärer Differenzierungsantigene (MDA). Alle Zelllinien trugen die zuvor beschriebene Mutation *BRAF-V600E* (V. Lennerz, pers. Mitteilung).

Tab. 1.1: Übersicht und HLA-Klasse I-Expression der stabilen Zelllinien des Melanommodells Ma-Mel-86. Aus vier operativ entfernten Lymphknotenmetastasen der Patientin INTH wurden die vier stabilen Melanomzelllinien Ma-Mel-86A, -86B, -86C und -86F generiert (gelb: negativ für melanozytäre Differenzierungsantigene (MDA), orange: positiv für MDA). Zudem lag eine autologe EBV-immortalisierte B-Zelllinie (INTH-EBV-B) vor. Im Gegensatz zu Ma-Mel-86A stellten die anderen Melanomzelllinien HLA-Klasse I-Verlustvarianten dar (Tabelle verändert aus der Dissertation von Silke Lübcke (geb. Lubojanski)).

Ma-Mel-	HLA-A	HLA-B	HLA-C	Generiert aus Lymphknotenmetastase
	-A*01:01 -A*24:02	-B*08:01 -B*15:01	-C*07:01 -C*03:03	Jan. 2002
	-	-	-	Feb. 2004
	-A*01:01	-B*08:01	-C*07:01	Apr. 2005
	-	-	-	Dez. 2008
	-A*01:01 -A*24:02	-B*08:01 -B*15:01	-C*07:01 -C*03:03	-

Ma-Mel-86A und -86B waren bereits Gegenstand mehrerer Publikationen (Paschen *et al.*, 2006; del Campo *et al.*, 2009; Bernal *et al.*, 2012), die unter anderem die Ursache der fehlenden HLA-Klasse I-Expression in Melanomen und anderen soliden Tumoren untersuchten. Es konnte gezeigt werden, dass Ma-Mel-86B aufgrund fehlender  $\beta$ 2-Mikroglobulin-Expression negativ für HLA-Klasse I-Moleküle auf der Zelloberfläche war. Des Weiteren wurde das Exom der Zelllinie Ma-Mel-86A zusammen mit den Exomen 120 weiterer Melanome sequenziert, um das Repertoire an sogenannten *Driver*-Mutationen (siehe auch Abschnitt 1.2.2.3) zu analysieren (Hodis *et al.*, 2012).

Das Melanommodell Ma-Mel-86 wurde zudem bereits in der Doktorarbeit von Silke Lübcke (geb. Lubojanski, Tag der mündlichen Prüfung: 17.02.2012) bearbeitet. Im Rahmen dieser Arbeit wurde eine Reihe von T-Zellantworten identifiziert und charakterisiert. **Tab. 1.2** zeigt eine Übersicht hierzu. Die gefundenen HLA-abhängigen melanomreaktiven T-Zellen erkannten MDA, überexprimierte Antigene bzw. ein mutiertes Neoantigen. Zudem wurden CSF2RA und TRP-2 als Zielstrukturen HLA-unabhängiger, aber T-Zellrezeptor-vermittelter Immunantworten identifiziert. Für die durchgeführten Analysen wurden bereits alle HLA-Klasse I-Allele der Patientin in Expressionsvektoren kloniert und mehrere autologe MLTCs (engl. *mixed lymphocyte tumor cell cultures*) generiert (Durchführung M. Fatho, S. Lübcke bzw. D. Eberts).

**Tab. 1.2:** Übersicht der bereits im Melanommodell Ma-Mel-86 identifizierten Antigene, die von autologen melanomreaktiven T-Zellen erkannt wurden. Die gelisteten HLA-Klasse I-abhängigen und -unabhängigen Antigene wurden in der Doktorarbeit von S. Lübcke mittels Paneltests mit 42 bekannten tumorassoziierten Antigenen bzw. komplementäre (c-)DNA-Bank-Screenings mit cDNA-Banken der autologen Melanomzelllinien Ma-Mel-86A und -86C identifiziert. Für die HLA-unabhängig erkannten Proteine TRP-2 und CSF2RA sind entsprechend keine Peptidsequenzen angegeben (Tabelle verändert aus der Dissertation von Silke Lübcke (geb. Lubojanski)).

Antigen Peptid	Präsentation über HLA	Antigenkategorie	Identifiziert durch
<b>Tyrosinase</b> <sub>146-156</sub> SSDYVIPIGTY	-A*01:01	Differenzierungsantigen	Paneltest cDNA-Bank-Screening
<b>gp100</b> <sub>472-480</sub> PLDCVLYRY	-A*01:01	Differenzierungsantigen	Paneltest
<b>HERPUD1</b> <sup>mut</sup> <sub>154-162</sub> GLGPGFSSY	-B*15:01	mutiertes Antigen	cDNA-Bank-Screening
<b>CSPG4</b> <sub>554-562</sub> HIIFPHGSL	-C*07:01	überexprimiertes Antigen	cDNA-Bank-Screening
<b>TRP-2</b> -	-	Differenzierungsantigen	Paneltest
<b>CSF2RA</b> -	-	-	cDNA-Bank-Screening

**HERPUD1:** *homocysteine-inducible, ER stress-inducible, ubiquitin-like domain member 1*, **CSPG4:** *chondroitin sulfate proteoglycan 4*, **TRP-2:** *tyrosinase-related protein 2*, **CSF2RA:** *granulocyte-macrophage colony-stimulating factor 2 receptor subunit alpha*

## 1.1.2 Antigenerkennung durch CD8<sup>+</sup> T-Lymphozyten

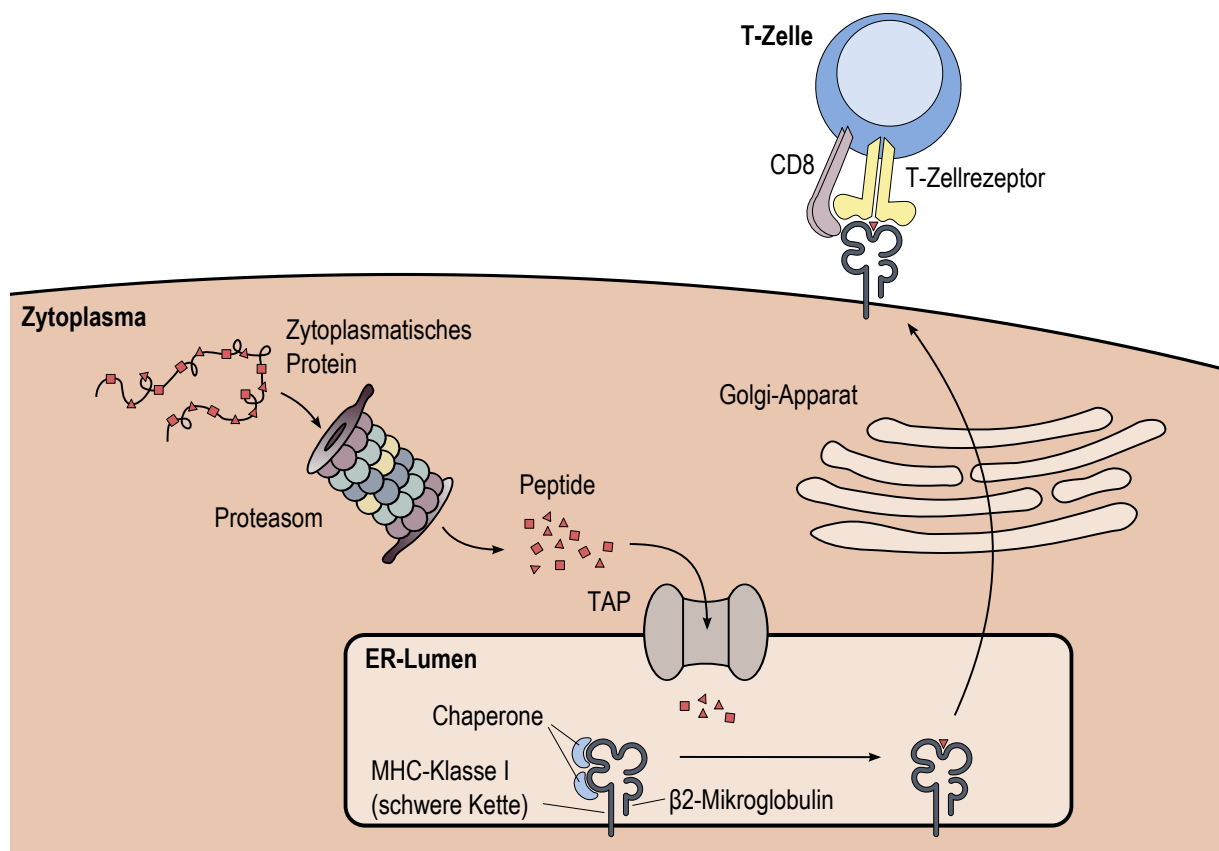
### 1.1.2.1 T-Lymphozyten

T-Lymphozyten sind die Mediatoren der zellvermittelten adaptiven Immunantwort von Wirbeltieren. Sie entwickeln sich aus Vorläuferzellen, die von pluripotenten hämatopoetischen Stammzellen im Knochenmark abstammen. Die Vorläuferzellen wandern vom Knochenmark durch das Blut zum Thymus, wo sie ausreifen. Die Ausreifung im und die Abhängigkeit vom Thymus war namensgebend für T-Lymphozyten oder T-Zellen (Murphy, 2012). Während der Ausreifung im Thymus durchlaufen T-Zellen verschiedene Stadien, die sich unter anderem anhand des Status des T-Zellrezeptors (TCR, engl. *T cell receptor*) und der Expression der Co-Rezeptoren CD4 und CD8 sowie des CD3-Komplexes unterscheiden lassen (Blom & Spits, 2006; Murphy, 2012). Nach Rekombination der TCR-Gene und funktioneller Expression des TCR erfolgt im Thymus die klonale Deletion, die der frühen Eliminierung von autoreaktiven T-Lymphozyten dient. Der Selektionsprozess gliedert sich in zwei Schritte. Zunächst werden T-Lymphozyten, die einen funktionalen TCR exprimieren und mit niedriger Affinität an Selbst-HLA/Peptid-Komplexe binden, durch ein Überlebenssignal selektiert (positive Selektion). Alle anderen T-Lymphozyten sterben aufgrund des fehlenden Überlebenssignals. Anschließend werden durch die negative Selektion alle T-Zellen eliminiert, die mit hoher Affinität gegen körpereigene Antigene präsentiert auf Selbst-HLA reagieren (Jenkins *et al.*, 2010; Murphy, 2012). Nach der Ausreifung besteht das T-Zellrepertoire größtenteils aus naiven CD4<sup>+</sup> bzw. CD8<sup>+</sup> T-Lymphozyten, die den Thymus verlassen und zwischen den peripheren lymphatischen Organen zirkulieren. Solange sie nicht auf ihr Zielantigen treffen, wird das Überleben der naiven T-Zellen durch Interaktion mit Zytokinen (IL-7) und mit Selbst-HLA/Peptid-Komplexen im lymphatischen Gewebe gesichert. Treffen jedoch naive T-Zellen auf ein Antigen, das ihr TCR mit hoher Affinität erkennt, führt dies zu ihrer Aktivierung, klonalen Expansion und Ausdifferenzierung (Burnet, 1959; Murphy, 2012). Während CD4<sup>+</sup> T-Lymphozyten als sogenannte T-Helferzellen unter anderem die Aktivierung von B-Zellen und Makrophagen unterstützen, entwickeln sich CD8<sup>+</sup> T-Lymphozyten zu Effektorzellen mit zytotoxischer Funktion. Nur ein Bruchteil der aktivierten T-Lymphozyten wird zu langlebigen Gedächtnis-T-Zellen, die im Falle eines erneuten Antigenkontaktes mit einer beschleunigten Immunantwort reagieren (Murphy, 2012).

### 1.1.2.2 Antigenprozessierung und -präsentation

Die Präsentation von Antigenen gegenüber T-Zellen erfolgt über MHC (engl. *major histocompatibility complex*)-Moleküle, welche beim Menschen auch als HLA (engl. *human leukocyte antigen*)-Moleküle bezeichnet werden. Man unterscheidet zwei Arten von klassischen MHC, nämlich MHC-Klasse I-Moleküle, die auf allen kernhaltigen Zellen zu finden sind und Antigene für CD8<sup>+</sup> T-Zellen präsentieren, und MHC-Klasse II-Moleküle, die

normalerweise nur von professionellen antigenpräsentierenden Zellen exprimiert werden und Zielantigene von  $CD4^+$  T-Zellen präsentieren (Murphy, 2012). Bei Antigenen, die an MHC-Klasse I-Moleküle binden, handelt es sich um Peptide mit einer Länge von acht bis elf, in der Regel neun oder zehn Aminosäuren (Falk *et al.*, 1991). Die Bindung erfolgt über Wasserstoffbrücken in einer Tasche des HLA-Moleküls, die zwischen zwei  $\alpha$ -Helices liegt und deren Basis von acht  $\beta$ -Faltblättern gebildet wird (Alberts *et al.*, 2008; Murphy, 2012). Die Peptide entstehen durch die Prozessierung von Proteinen im Zytoplasma im Rahmen physiologischer Abraumprozesse. Hierfür ist das ubiquitär exprimierte Proteasom zuständig, welches laut *In-vitro*-Ergebnissen von Kisselev *et al.* (1999) Peptide mit einer Länge von vier bis 20 Aminosäuren produziert (vgl. **Abb. 1.3**). Peptide, die für eine Bindung an HLA-Klasse I-Moleküle noch zu lang sind, werden anschließend durch Aminopeptidasen im Zytoplasma oder im endoplasmatischen Retikulum (ER) gekürzt (Roelse *et al.*, 1994).



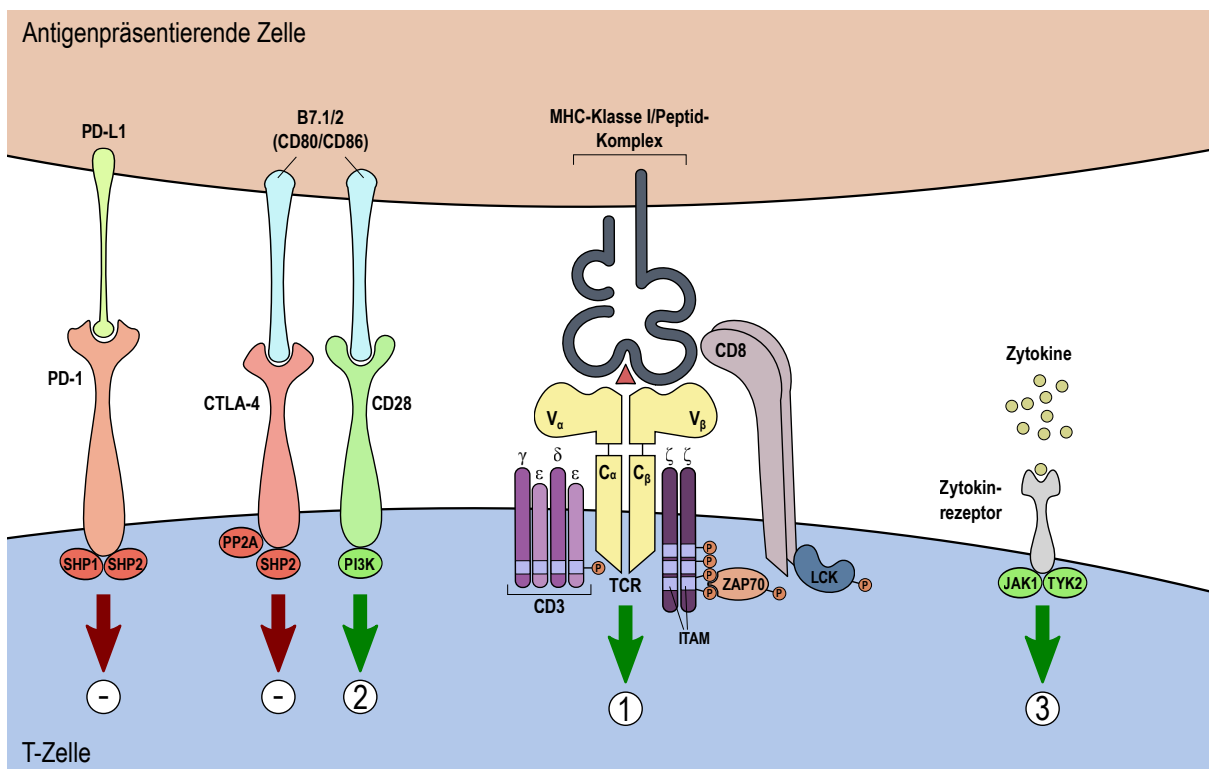
**Abb. 1.3: Antigenprozessierung und -präsentation über MHC-Klasse I-Moleküle.** Für die Präsentation von Peptiden auf MHC-Klasse I-Molekülen werden Proteine im Zytoplasma durch das Proteasom zu Peptiden abgebaut. Die Peptide werden durch TAP-Transporter in das ER-Lumen transloziert. Dort findet die Assemblierung des MHC-Klasse I/Peptid-Komplexes statt. Peptidbeladene MHC-Klasse I-Moleküle werden über den Golgi-Apparat an die Zelloberfläche gebracht, wo die Erkennung des Antigens durch  $CD8^+$  T-Zellen erfolgt. (Abb. verändert nach Yewdell *et al.*, 2003.)

Die Translokation der Peptide vom Zytoplasma ins ER erfolgt mithilfe der ER-ständigen, heterodimeren TAP-Transporter (engl. *transporter associated with antigen processing*). Im ER findet die Assemblierung des MHC-Klasse I/Peptid-Komplexes statt. Dabei wird erst die schwere Kette des MHC-Moleküls gefaltet, woraufhin ein Heterodimer mit  $\beta 2$ -Mikroglobulin („leichte Kette“) gebildet wird. Anschließend erfolgt die Beladung mit dem prozessierten Peptid (Neefjes *et al.*, 1993). Der komplette Assemblierungsvorgang wird durch Chaperone

und TAP unterstützt (Sadasivan *et al.*, 1996). Sobald ein stabiler MHC-Klasse I/Peptid-Komplex vorliegt, kann dieser über den Golgi-Apparat an die Zelloberfläche gebracht werden (Jackson *et al.*, 1994), wo die Erkennung durch CD8<sup>+</sup> T-Zellen stattfindet.

### 1.1.2.3 Antigenerkennung und T-Zellaktivierung

Für die Erkennung von Antigenen auf der Oberfläche von Zielzellen exprimieren T-Lymphozyten einen antigenspezifischen T-Zellrezeptor (TCR), welcher sich aus zwei Ketten zusammensetzt, und zwar entweder einer  $\alpha$ - und einer  $\beta$ - oder einer  $\gamma$ - und einer  $\delta$ -Kette, wobei mehr als 95% der T-Zellen in der Peripherie TCR $\alpha\beta$ <sup>+</sup> sind (Vroom *et al.*, 1991). Die beiden TCR-Ketten (vgl. **Abb. 1.4**, gelb) besitzen jeweils eine membranständige konstante Domäne (C $\alpha/\beta$ ) und eine extrazelluläre variable Domäne (V $\alpha/\beta$ ), die durch somatische Genarrangements verschiedener TCR-Gensegmente entstanden sind. Der variable Teil der TCR- $\beta$ -Kette besteht aus einer V (engl. *variability*)-, einer D (engl. *diversity*)- und einer J (engl. *joining*)-Region. Der variable Abschnitt der TCR- $\alpha$ -Kette setzt sich ebenfalls aus V- und J-Segmenten zusammen, hat jedoch keine D-Region. Die Spezifität des TCR wird durch die drei hypervariablen Bereiche CDR1 (engl. *complementarity determining region*), CDR2 und CDR3 dieser Regionen bestimmt (Murphy, 2012).



**Abb. 1.4: Aktivierende und hemmende Signale für CD8<sup>+</sup> T-Zellen.** Für die Aktivierung einer CD8<sup>+</sup> T-Zelle werden drei Signale benötigt (grüne Pfeile). Der TCR (T-Zellrezeptor) interagiert mit dem MHC-Klasse I/Peptid-Komplex auf der antigenpräsentierenden Zelle. Durch Bildung eines Komplexes mit CD3 kann die Signalweiterleitung ins Zellinnere erfolgen. Das zweite Signal wird durch CD28 nach Bindung an B7 ausgelöst. Durch Zytokine (Signal 3) wird der entstehende Effortyp festgelegt. Kinasen vermitteln diese aktivierenden Signale. Zur Eindämmung der T-Zellantwort (rote Pfeile) werden die inhibitorischen Rezeptoren CTLA-4 (Ligand: B7) und PD-1 (Ligand: PD-L1) exprimiert. Phosphatasen dephosphorylieren hierbei Signalmoleküle des aktivierenden Signalweges. (Abb. zusammengestellt nach Baniyash, 2004; Kershaw *et al.*, 2013; Chen & Flies, 2013; Okazaki *et al.*, 2013.)

Die Signaltransduktion vom TCR ins Zellinnere (Signal 1) basiert auf der Bildung eines Clusters mit den Oberflächenproteinen des CD3-Komplexes, da der TCR keine zytoplasmatischen Signalmodule besitzt (vgl. **Abb. 1.4**). CD3 besteht aus mehreren Ketten ( $\gamma$ ,  $\delta$ ,  $2 \times \epsilon$ , und  $2 \times \zeta$ ) und trägt intrazelluläre ITAM (engl. *immune receptor tyrosine-based activation motif*), die von der Tyrosinkinase LCK (engl. *lymphocyte-specific protein tyrosine kinase*) phosphoryliert werden (Baniyash, 2004). LCK wird vom Korezeptor CD8 rekrutiert (Veillette *et al.*, 1988; Yao *et al.*, 1990), welcher außerhalb der Antigenbindestelle an das MHC-Klasse I-Molekül bindet (Alberts *et al.*, 2008). Durch Wechselwirkung mit weiteren Kinasen, wie ZAP70 (engl. *zeta chain-associated protein kinase 70*), und einer Reihe von Signalmolekülen kann im Zellkern schließlich die Expression von Genen induziert werden, die unter anderem zur Proliferation, Zytokinproduktion und gegebenenfalls Lyse der Zielzellen benötigt werden (Baniyash, 2004; Chen & Flies, 2013).

Die Aktivierung der T-Zelle setzt jedoch neben der Interaktion des TCR mit dem MHC-Klasse I/Peptid-Komplex noch zwei weitere Signale voraus (siehe **Abb. 1.4**, grüne Pfeile). Kostimulation wird durch den Rezeptor CD28 (Signal 2) ausgelöst, der an Membranproteine der B7-Familie (CD80/CD86) bindet, und durch Zytokine (Signal 3), die den entstehenden Effektorotyp bestimmen. An der Weiterleitung der Signale 2 und 3 sind unter anderem die Kinasen PI3K (engl. *phosphoinositide-3-kinase*), JAK1 (engl. *Janus kinase 1*) und TYK2 (engl. *tyrosine kinase 2*) beteiligt (Kershaw *et al.*, 2013; Okazaki *et al.*, 2013). Fehlen die zusätzlichen Signale, führt dies zur Anergie der T-Zelle (Schwartz *et al.*, 1989; Chen & Flies, 2013).

#### 1.1.2.4 Inhibitorische Signale

Zur Regulation einer Immunantwort exprimieren T-Zellen zusätzlich inhibitorische Rezeptoren (vgl. **Abb. 1.4**, rote Pfeile). Dabei handelt es sich unter anderem um CTLA-4 (engl. *cytotoxic T-lymphocyte-associated protein 4*), das mit CD28 um den gemeinsamen Liganden B7 konkurriert, und PD-1 (engl. *programmed cell death 1*), dessen Ligand PD-L1 ebenso auf antigenpräsentierenden Zellen exprimiert wird (Baniyash, 2004; Chen & Flies, 2013; Okazaki *et al.*, 2013). CTLA-4-mRNA wird binnen weniger Stunden nach der T-Zellaktivierung exprimiert (Lindsten *et al.*, 1993), wohingegen PD-1 zu den inhibitorischen Molekülen gehört, die erst während der sogenannten T-Zellerschöpfung aufgrund chronischer Antigenexposition hochreguliert werden (Chen & Flies, 2013). Die Hemmung der T-Zellaktivierung beruht auf der Dephosphorylierung von Molekülen des aktivierenden Signalweges durch Phosphatasen wie SHP1/2 (engl. *SH2 domain-containing tyrosine phosphatase*) und PP2A (engl. *protein phosphatase 2A*). Wie zuvor beschrieben (siehe Abschnitt 1.1.1.2.2), kann die Blockade der Interaktion der inhibitorischen Rezeptoren mit ihren Liganden durch Immun-*Checkpoint*-Inhibitoren zur Behandlung von malignen Melanomen genutzt werden.

### 1.1.3 Wechselwirkungen zwischen Immunsystem und Malignomen

Der sogenannten *Two-Hit*-Hypothese zufolge, die Alfred G. Knudson (Knudson, 1971, 1974) zunächst für Retinoblastome entwickelte, entstehen Tumore aufgrund von zwei Mutationsereignissen in einer somatischen Zelle. Bei vererbten Krebsformen findet die erste Mutation bereits präzygotisch statt. Die betroffenen Gene können in zwei Kategorien unterteilt werden, und zwar Onkogene, die durch eine monoallelische aktivierende Mutation zur Karzinogenese beitragen, und Tumorsuppressorgene, deren biallelische Inaktivierung ein neoplastisches Wachstum ermöglicht. Neben der so erlangten Fähigkeit zum autarken Wachstum mit unbegrenztem replikativem Potential gehören die Insensibilität gegenüber wachstumshemmenden Signalen, Apoptose-Resistenz, Förderung der Angiogenese, die Fähigkeit zur Invasion und Metastasierung, die Deregulierung des zellulären Energiemetabolismus, genomische Instabilität, tumorfördernde Entzündung und die Umgehung der Zerstörung durch das Immunsystem zu den von Hanahan und Weinberg (2011) zusammengestellten Charakteristika von Tumoren. Der Erwerb dieser Merkmale dient der Sicherung des Überlebens, der verbesserten energetischen Versorgung und der Induktion weiterer genetischer Aberrationen zur Steigerung der proliferativen Kapazität und zur Verbreitung und Etablierung in ektopischen Geweben.

#### 1.1.3.1 *Immunosurveillance* und *Immunoediting*

Obwohl Paul Ehrlich im Jahr 1909 bereits annahm, dass das Immunsystem eine ansonsten sehr hohe Frequenz an Karzinomen unterdrücken kann, ist die Rolle des Immunsystems bei der Hemmung der Karzinogenese bis heute umstritten und wurde lange Zeit nur im Zusammenhang mit viral induzierten Tumoren akzeptiert (Dunn *et al.*, 2002). Burnet und Thomas fassten die Fähigkeit des Immunsystems, kleinere Ansammlungen neoplastisch transformierter Zellen zu erkennen und zu eliminieren, unter den Begriff „*Immunosurveillance*“ zusammen (Burnet, 1970; Dunn *et al.*, 2002).

Durch die Etablierung RAG (engl. *recombination activating gene*)-2-defizienter Mäuse, die aufgrund dieser Defizienz keine Lymphozyten bilden (Shinkai *et al.*, 1992), konnte die tumorsuppressive Fähigkeit des Immunsystems und deren Abhängigkeit vom lymphatischen Kompartiment bestätigt werden (Shankaran *et al.*, 2001). Evidenzen für das Vorliegen der *Immunosurveillance* im Menschen ergaben sich zunächst aus Statistiken, denen zufolge Transplantationspatienten ein deutlich erhöhtes Risiko für verschiedene Krebsarten aufwiesen, was auf die immunsuppressive Behandlung zurückgeführt wurde (Buell *et al.*, 2005; Laing *et al.*, 2006). Außerdem konnte für eine Reihe von Tumoren ein Zusammenhang zwischen der Präsenz von tumorinfiltrierenden Lymphozyten und einer guten Prognose hergestellt werden (Zhang *et al.*, 2003; Galon *et al.*, 2006).

Neben der protektiven Funktion des Immunsystems zeigte die Studie von Shankaran und Kollegen (2001) mithilfe der RAG-2<sup>-/-</sup>-Mäuse allerdings auch, dass Tumore, die sich in einer



intakten immunologischen Umgebung entwickeln, weniger immunogen sind als solche, die in immundefizienten Wirten entstehen. Somit trägt das Immunsystem gleichzeitig dazu bei, dass sich Tumore bilden, die in der Lage sind, der immunologischen Erkennung zu entgehen. Um dem Rechnung zu tragen, wurde das Konzept des *Immunoeditings* eingeführt, welches sich in drei Phasen gliedert: *Elimination*, *Equilibrium* und *Escape* (Dunn *et al.*, 2004). Während der Eliminierungsphase sorgen angeborene und erworbene Immunantworten dafür, dass Tumorzellen früh erkannt und zerstört werden. Diese Phase entspricht der *Immunosurveillance*. Während der *Equilibrium*-Phase herrscht ein Gleichgewicht zwischen Tumorwachstum und Zerstörung der Tumorzellen durch das Immunsystem. In dieser Phase entwickeln sich unter dem immunologischen Selektionsdruck Tumorzellvarianten, die nicht mehr durch das Immunsystem erkannt bzw. eliminiert werden können. Sobald das Wachstum dieser Tumorzellen nicht mehr eingeschränkt werden kann, ist die finale Phase (*Escape*) erreicht und der Tumor wird klinisch evident (Mittal *et al.*, 2014).

### 1.1.3.2 Immune Escape-Mechanismen

Aufgrund der bereits angesprochenen genetischen Instabilität von Malignomen entstehen durch die Wechselwirkung mit dem Immunsystem Tumorzellen, die die Auslösung einer Immunantwort verhindern können oder für das Immunsystem „unsichtbar“ werden. Diese sogenannten *Immune Escape*-Mechanismen beruhen entweder auf einer verringerten Erkennbarkeit der Tumorzellen durch das Immunsystem oder auf dem Aufbau eines immunsuppressiven Mikromilieus. Viele Tumore zeigen eine veränderte HLA-Klasse I-Expression, so dass Antigene nicht mehr auf der Zelloberfläche präsentiert werden (Méndez *et al.*, 2008; Paschen *et al.*, 2003; Yeung *et al.*, 2013). Man beobachtet Phänotypen mit kompletten bzw. Haplotyp-, Lokus- oder Allel-Verlusten. Zudem können manche Tumore die HLA-Klasse I-Expression selbst in Anwesenheit von inflammatorischen Zytokinen (z.B. Interferonen) auf einem niedrigen Level halten. Als Schutz gegen natürliche Killerzellen, die HLA-defiziente Zellen erkennen können, werden teilweise nicht-klassische HLA-Klasse I-Moleküle, wie HLA-E, von ansonsten HLA-Klasse I-negativen Tumorzellen exprimiert. Komplette HLA-Klasse I-Verluste werden oftmals durch biallelische Defekte im  $\beta$ 2-Mikroglobulin verursacht. Andere Phänotypen beruhen auf regulatorischen Mechanismen oder sind Folge einer genetischen Aberration im Lokus der schweren Kette des HLA-Moleküls (Garcia-Lora *et al.*, 2003).

Des Weiteren tragen die Herabregulation der Expression von T-zellerkannten Antigenen (Maeurer *et al.*, 1996; Sørensen *et al.*, 2009) oder Defekte in der Antigenprozessierung dazu bei, dass Tumorzellen nicht mehr von CD8<sup>+</sup> T-Zellen erkannt werden können. Betroffen sind dabei zum Beispiel die Peptidtransporter TAP-1 und TAP-2, die Proteasomuntereinheiten LMP (engl. *low molecular mass protein*)-2 und LMP-7 (Restifo *et al.*, 1993; Seliger *et al.*, 1996; Dissemond *et al.*, 2003; Chambers *et al.*, 2007) oder die für die Assemblierung der HLA/Peptid-Komplexe benötigten Chaperone im ER (Dissemond *et al.*, 2004).

Landsberg *et al.* (2012) zeigten außerdem, dass Melanomzellen durch T-zellinduzierte inflammatorische Stimuli reversibel dedifferenzieren und dadurch von MDA-reaktiven T-Zellen nicht mehr erkannt werden können.

Durch die Herabregulation von kostimulatorischen Molekülen, wie B7 (Fujiwara *et al.*, 2004), können T-Zellen aufgrund des Fehlens des zweiten Aktivierungssignals (vgl. Abschnitt 1.1.2.3) in Anergie versetzt werden. Eine erhöhte Expression des inhibitorischen Liganden PD-L1 führt nach Interaktion mit PD-1 auf der T-Zelle (vgl. Abschnitt 1.1.2.4) zu deren funktioneller Erschöpfung, wodurch die Entstehung einer effektiven Anti-Tumorantwort ebenso unterdrückt wird. Wird eine Tumorzelle dennoch erfolgreich von CD8<sup>+</sup> T-Zellen erkannt, kann eine niedrigere Expression bzw. die Expression nicht-funktioneller Varianten des „Todesrezeptors“ Fas noch vor einer T-zellinduzierten Apoptose schützen (Elnemr *et al.*, 2001). Zudem führen höhere Level des Fas-Liganden (FASLG) auf der Tumorzelle zur Induktion des programmierten Zelltods von Fas<sup>+</sup> T-Zellen (Müschen *et al.*, 2000).

Zur Ausbildung eines immunsuppressiven Tumormikromilieus werden beispielsweise regulatorische Immunzellen rekrutiert, wie regulatorische T-Zellen (Grauer *et al.*, 2007) oder MDSCs (engl. *myeloid derived suppressor cells*; Almand *et al.*, 2001; Diaz-Montero *et al.*, 2009). Außerdem können Tumorzellen selbst immunsuppressive Faktoren, wie z. B. IL-10 und TGF (engl. *transforming growth factor*)- $\beta$ , ausschütten (Conrad *et al.*, 1999) und Enzyme, wie Indoleamin-2,3-Dioxygenase (IDO) exprimieren, die zur Deprivation der Aminosäure Tryptophan führen (Weinlich *et al.*, 2007; Karanikas *et al.*, 2007). Dadurch wird die T-Zellproliferation und -differenzierung gehemmt.

## 1.2 Methodischer Hintergrund

### 1.2.1 Identifizierung von Tumorantigenen

Das Vorkommen von tumorspezifischen Antigenen (TSA) wurde bereits vor 50 Jahren diskutiert (Old & Boyse, 1964). Ihre Charakterisierung zeigte jedoch, dass die Antigene tumorreaktiver T-Zellen nicht grundsätzlich nur von Tumorzellen exprimiert werden. Daher werden diese Antigene als tumorassoziierte Antigene (TAA) bezeichnet. Man unterscheidet vier verschiedene Kategorien (Novellino *et al.*, 2005):

- 1) Differenzierungsantigene, die auch im Normalgewebe, von dem der Tumor abgeleitet ist, exprimiert werden,
- 2) überexprimierte Antigene, die niedriger exprimiert auch in Normalgeweben nachzuweisen sind,
- 3) *Cancer/Germine*-Antigene, die außer in Keimzellen und in Trophoblastzellen der Plazenta in normalen Zellen nicht exprimiert werden, und

- 4) mutierte Antigene, die tatsächlich tumorspezifisch sind.

Weiterhin findet man T-Zellen gegen virale Antigene, die von Malignomen exprimiert werden, die auf Virusinfektionen zurückzuführen sind (Murray *et al.*, 1992).

Eine Methode zur Identifikation von TAA bzw. TSA stellt das **T-zellbasierte cDNA-Bank-Screening** dar (Brichard *et al.*, 1993; Lennerz *et al.*, 2005; Tschiedel *et al.*, 2008). Hierfür werden Tumor-cDNA-Banken in eukaryontischen Expressionsvektoren konstruiert. Durch Kotransfektion der Tumor-cDNA mit den patienteneigenen HLA-cDNAs in geeignete Rezipienten, wie COS-7- oder 293T-Zellen, kann die Erkennung durch tumorreaktive T-Lymphozyten beispielsweise in IFN $\gamma$ -ELISpot-Assays getestet werden. Zunächst werden dabei Pools aus 100 cDNA-Bank-Klonen transfiziert. Erkannte Pools können schrittweise subkloniert werden.

Des Weiteren können MHC-gebundene Peptide durch **biochemische Verfahren** identifiziert werden. MHC-/Peptid-Komplexe werden dafür von den Tumorzellen isoliert und die Peptide bei niedrigem pH-Wert eluiert. Danach werden sie mittels HPLC (engl. *high performance liquid chromatography*) fraktioniert und massenspektrometrisch identifiziert (Röttschke *et al.*, 1990; Schirle *et al.*, 2000; Castelli *et al.*, 1995; Skipper *et al.*, 1999). Die so gewonnen Peptidsequenzdaten können auch für die Bestimmung von spezifischen HLA-Bindemotiven herangezogen werden (Falk *et al.*, 1991). Die Plattform XPRESIDENT<sup>TM</sup> der Firma immatics Biotechnologies GmbH (Tübingen) kombiniert die beschriebene biochemische Identifizierung von Tumor-MHC-Liganden mit Genexpressionsanalysen, auf deren Grundlage Peptidkandidaten für eine individualisierte Immuntherapie ausgewählt werden (Weinschenk *et al.*, 2002; Walter *et al.*, 2012).

Das **SEREX**-Verfahren (engl. *serological identification of antigens by recombinant expression cloning*) dient der direkten Identifizierung von Tumorantigenen, die eine IgG-vermittelte Immunantwort in Patienten induzieren. Hierbei werden Tumor-cDNA-Banken in  $\lambda$ -Phagen-Vektoren zur Expression in *Escherichia coli* generiert und die an die rekombinant exprimierten Proteine gebundenen Serumantikörper detektiert. Obwohl durch SEREX primär B-Zellantworten untersucht werden, ist davon auszugehen, dass Peptide aus den identifizierten Proteinen zumindest auch von CD4<sup>+</sup> T-Helferzellen erkannt werden, da diese an der Aktivierung der B-Zellen beteiligt sind (Sahin *et al.*, 1995; Türeci *et al.*, 1996). Für das durch SEREX identifizierte Antigen NY-ESO-1 wurden neben humoralen (Chen *et al.*, 1997) auch durch CD4<sup>+</sup> und CD8<sup>+</sup> T-Zellen vermittelte Immunantworten (Jäger *et al.*, 2000, 1998) detektiert.

Die **reverse Immunologie** beruht auf der theoretischen Vorhersage von Peptidkandidaten basierend auf bekannten HLA-Bindemotiven oder auch Proteasomschnittstellen (Noppen *et al.*, 2000; Kessler *et al.*, 2001; Asemissen *et al.*, 2009). Hierbei werden Proteine ausgewählt, die aufgrund einer Mutation oder ihres Expressionsprofils möglichst tumorspezifisch sind, und die dazugehörige Proteinsequenz mittels Vorhersagealgorithmen analysiert. Für die Vorauswahl von Proteinen werden Methoden zur Erstellung

von Genexpressionsprofilen (z. B. serielle Analyse der Genexpression (SAGE), *Microarrays* oder *Next-Generation Sequencing*), sowie molekularbiologische (z. B. repräsentative Differenzanalyse) oder biochemische Verfahren (z. B. *Western Blot*) eingesetzt. Die vorhergesagten Peptide werden dann synthetisiert und auf Erkennung durch T-Zellen getestet (Viatte *et al.*, 2006).

Eine Übersicht zu bisher identifizierten immunogenen Peptiden findet sich auf der *Cancer Immunity*-Webseite<sup>1</sup>. Für die dortige Listung als validiertes Peptid wird die Erfüllung folgender Kriterien vorausgesetzt (Vigneron *et al.*, 2013; van der Bruggen *et al.*, 2013):

- Isolation von stabilen humanen T-Zellklonen oder -linien, die das Peptid erkennen,
- Identifikation des T-zellerkannten Peptids,
- Bestimmung des präsentierenden HLA-Moleküls,
- Evidenz der Prozessierung und Präsentation des Peptids durch Tumorzellen,
- Bestimmung des kürzesten erkannten Peptids, das in einer Konzentration von  $\leq 1 \mu\text{M}$  erkannt werden soll (nicht verpflichtend für Peptide, die durch reverse Immunologie identifiziert wurden) und
- Tumor- oder Gewebespezifität in nicht näher spezifiziertem Ausmaß.

Etwa die Hälfte (25) der 46 Publikationen, die 2013 für die Aktualisierung der Datenbank herangezogen wurden, erfüllten die beschriebenen Kriterien nicht. Diese Peptide wurden als „potentielle Antigene“ eingestuft (Vigneron *et al.*, 2013).

### 1.2.2 *Next-Generation Sequencing*

Seit der Entwicklung des Sanger-Sequenzierverfahrens (Sanger *et al.*, 1977) haben einige technische Verbesserungen dazu beigetragen, dass im April 2003 eine qualitativ hochwertige Referenzsequenz des menschlichen Genoms veröffentlicht werden konnte (Collins *et al.*, 2003; Metzker, 2005). Dennoch ist dieses Standardverfahren mit einigen Limitationen verbunden, die letztlich zur Entwicklung von Sequenzierverfahren der nächsten Generation (engl. *Next-Generation Sequencing*, NGS) führten (Metzker, 2010).

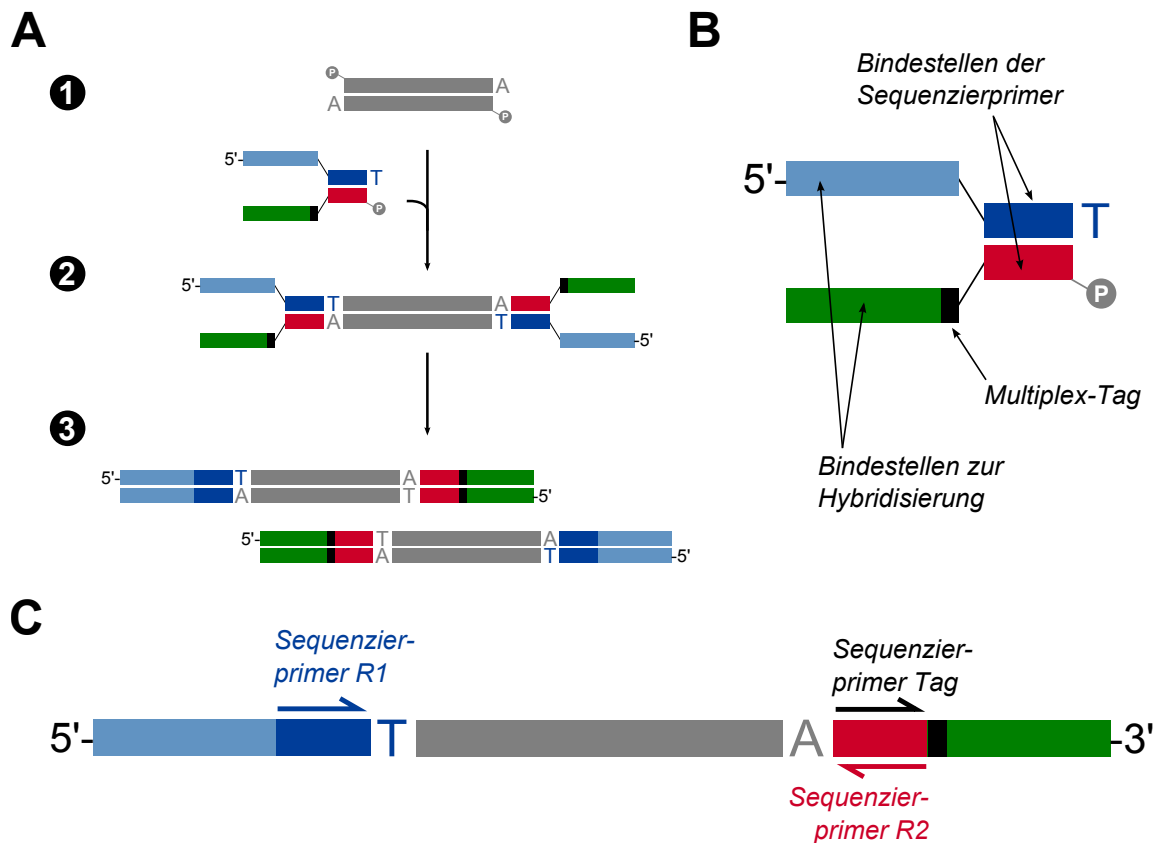
#### 1.2.2.1 Prinzip des *Next-Generation Sequencing*

Die aktuell gängigsten NGS-Verfahren (Illumina/Solexa, Roche/454, Applied Biosystems/SOLiD) beruhen auf drei Grundsritten, nämlich die Generierung einer Sequenzierbibliothek, die klonale Amplifikation der DNA-*Templates* und die Sequenzierung (Mutz *et al.*, 2013). Diese Schritte sollen im Folgenden für das System von Illumina/Solexa näher erläutert werden.

---

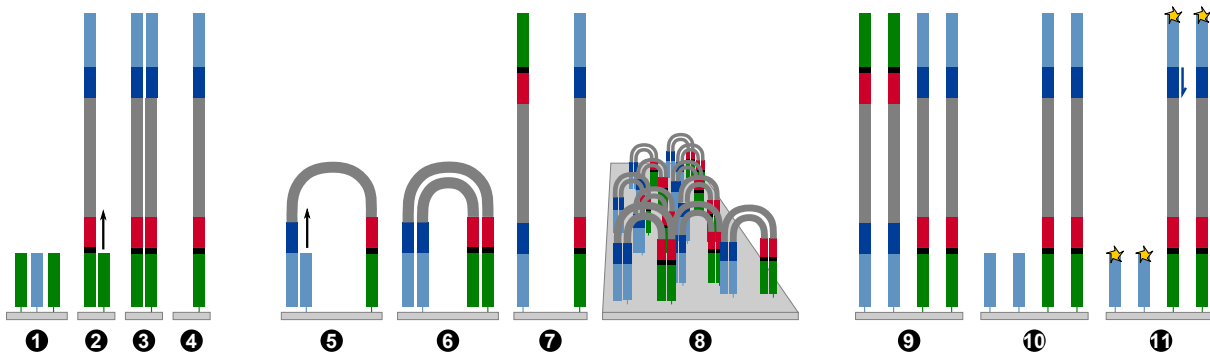
<sup>1</sup><http://cancerimmunity.org/peptide/> (zuletzt aufgerufen am 01.07.2014)

Eine Sequenzierbibliothek besteht aus etwa 300-500 bp langen DNA-Fragmenten, die von Adaptersequenzen flankiert werden, welche es ermöglichen, die flankierten DNA-Stücke (*Inserts*) mittels Adapter-spezifischer Primer zu sequenzieren. Für die Herstellung einer *TruSeq*-Sequenzierbibliothek (Illumina) wird die zu sequenzierende DNA zunächst mechanisch fragmentiert (*TruSeq<sup>®</sup> DNA Sample Prep Kits Data Sheet*, Illumina, 2012). Nachdem die Fragmentenden repariert und phosphoryliert worden sind, werden A-Überhänge geschaffen (**Abb. 1.5A**, 1). Diese A-Überhänge dienen der gerichteten Ligation von sogenannten Y-Adaptoren (**Abb. 1.5A**, 2), die neben komplementären T-Überhängen die Hybridisierungssequenzen für die Immobilisierung und die Bindestellen für die Sequenzierprimer enthalten (vgl. **Abb. 1.5B**). Außerdem enthält der Y-Adapter einen *Multiplex-Tag*, der die Identifizierung der Sequenzierbibliothek ermöglicht. Nach Fertigstellen der Bibliothek liegen die Inserts flankiert von den beiden Adaptersequenzen als doppelsträngige DNA vor (**Abb. 1.5A**, 3). **Abb. 1.5C** zeigt schematisch, wie die drei Sequenzierprimer „R1“, „R2“ und „Tag“ an die Adapter-Sequenzen binden.



**Abb. 1.5:** Illumina/Solexa-Sequenzierbibliotheken. (A) **Generierung der Sequenzierbibliothek.** Nachdem die Enden der zuvor fragmentierten DNA (grau) repariert und phosphoryliert (P) wurden, werden A-Überhänge generiert (1). Diese werden für die Ligation des Y-Adapters benötigt (2). Die DNA-Fragmente werden nach der Ligation denaturiert und amplifiziert, so dass doppelsträngige DNA-Stücke entstehen, die von zwei verschiedenen Adaptersequenzen flankiert werden (3). (B) **Schematische Darstellung des Y-Adapters.** Neben dem für die Ligation benötigten T-Überhang beinhaltet der Y-Adapter zwei verschiedene Sequenzen zur Hybridisierung (hellblau bzw. grün) und Bindestellen für die Sequenzierprimer (dunkelblau bzw. rot). Außerdem trägt der Adapter einen *Multiplex-Tag* (schwarz) zur Identifizierung der Sequenzierbibliothek. (C) **Bindestellen der Sequenzierprimer.** Für die Sequenzierung der Inserts (grau) werden die Sequenzierprimer „R1“ und „R2“ verwendet. Zur Detektion des *Multiplex-Tags* wird der Primer „Tag“ verwendet. (Abb. verändert nach dem Datenblatt zum *TruSeq<sup>®</sup> DNA Sample Prep Kit* (Illumina, 2012).)

Vor der Sequenzierung wird die Sequenzierbibliothek auf einer Glasoberfläche, der *Flow Cell*, immobilisiert. Hierfür werden die Hybridisierungssequenzen der Y-Adapter (**Abb. 1.5B**, hellblau bzw. grün) benötigt. Die *Flow Cell* ist mit Sonden beschichtet, die komplementär zu diesen Adaptersequenzen sind (vgl. **Abb. 1.6**, Schritt 1). Nachdem die denaturierte Sequenzierbibliothek an die *Flow Cell* gebunden ist, erfolgt die 3'-Verlängerung der *Flow Cell*-Sonden komplementär zur hybridisierten DNA (**Abb. 1.6**, Schritt 2). Die nun doppelsträngige DNA (**Abb. 1.6**, Schritt 3) wird denaturiert und die nicht mit der *Flow Cell* verbundene DNA abgewaschen (**Abb. 1.6**, Schritt 4). Das freie Ende der gebundenen DNA kann im nächsten Schritt an eine noch freie komplementäre Sonde binden (**Abb. 1.6**, Schritt 5), so dass wiederum eine 3'-Verlängerung ausgehend von der neu gebundenen Sonde erfolgen kann (**Abb. 1.6**, Schritt 6). Nach einem weiteren Denaturierungsschritt liegen die DNA-Moleküle der Sequenzierbibliothek in beiden Orientierungen und gebunden an die *Flow Cell* vor (**Abb. 1.6**, Schritt 7). Durch mehrfaches Wiederholen dieser Brückenamplifikation entstehen Cluster von DNA-Molekülen (**Abb. 1.6**, Schritt 8 und 9). In der *Flow Cell* können laut Hersteller über 200 Millionen Cluster mit etwa 1.000 DNA-Molekülen pro Cluster entstehen (*cBot Specification Sheet*, Illumina, 2011).



**Abb. 1.6: Cluster-Generierung in der *Flow Cell* und Vorbereitung der Sequenzierung.** Die *Flow Cell* ist beschichtet mit Sonden, die komplementär zu den Hybridisierungssequenzen der Adapter (grün bzw. hellblau) sind (1). Die denaturierte Sequenzierbibliothek bindet an die Sonden (2). Nach der 3'-Verlängerung der Sonde (3) werden die DNA-Stränge erneut denaturiert und ungebundene DNA wird abgewaschen (4). Das freie Ende bindet an eine noch freie Sonde in der *Flow Cell* (5) und der komplementäre Strang wird synthetisiert (6). Die DNA wird wiederum denaturiert, woraufhin beide Stränge eines DNA-Fragments an die *Flow Cell* gebunden sind (7). Durch Wiederholen der Brückenamplifikation (8) entstehen DNA-Cluster (9). Für die Sequenzierung werden alle DNA-Einzelstränge enzymatisch innerhalb der Adaptersequenzen geschnitten (10), die freien 3'-Enden werden blockiert (11, gelbe Sterne), und die Sequenzierung mit dem ersten Sequenzierprimer (dunkelblau) wird gestartet. (Abb. verändert nach dem *cBot Specification Sheet* (Illumina, 2011).)

Die anschließende Sequenzierung erfolgt parallel für alle Cluster in der *Flow Cell*, jedoch getrennt in *Forward*- bzw. *Reverse*-Richtung. Hierfür werden zunächst die DNA-Fragmente einer Orientierung – beispielsweise *Reverse* – durch enzymatischen Verdau innerhalb der *Reverse*-Adaptersequenz entfernt und alle 3'-Enden blockiert. Der erste Sequenzierprimer bindet dann an die *Forward*-Adaptersequenz (**Abb. 1.6**, Schritt 10 und 11). Die Sequenzierung erfolgt während der Synthese des komplementären DNA-Strangs. Im Gegensatz zur Sanger-Sequenzierung werden nur reversible, fluoreszenzmarkierte Terminatoren (3'-O-Azidomethyl-2'-Deoxynukleosid-Triphosphate, 3'-O-N<sub>3</sub>-dNTPs) verwendet, so dass immer nur ein neues Nukleotid angefügt wird (Bentley *et al.*, 2008). Das Fluoreszenzsignal wird direkt zusammen mit den Koordinaten innerhalb der *Flow Cell* detektiert. Durch die

Zugabe von TCEP (Tris-(2-carboxyethyl)-phosphin) wird der Farbstoff entfernt und die 3'-OH-Gruppe des entstehenden DNA-Strangs wiederhergestellt, sodass wiederum die vier farbstoffgekoppelten reversiblen Terminatoren in die *Flow Cell* gegeben und das nächste Nukleotid bestimmt werden kann. Dies wird so oft wiederholt bis die gewünschte Sequenzlänge erreicht ist (beispielsweise 100 bp).

Zur Sequenzierung in *Reverse*-Richtung müssen zunächst die komplementären Stränge der bereits sequenzierten DNA-Moleküle durch die oben beschriebene Brückenamplifizierung wiederhergestellt werden. Dann werden die zuvor sequenzierten DNA-Stränge gleichermaßen entfernt und die verbliebene DNA in reverser Orientierung sequenziert. Mit einem dritten Sequenzierprimer wird außerdem der *Multiplex-Tag* ausgelesen, was das Demultiplexen gepoolter Sequenzierbibliotheken ermöglicht.

### 1.2.2.2 Next-Generation Sequencing-Varianten

Das Ausgangsmaterial für die Generierung der Sequenzierbibliotheken ist DNA. Im Falle der **Exomsequenzierung** werden in einer genomischen Sequenzierbibliothek vor der Brückenamplifikation und Sequenzierung die Fragmente angereichert, die Exomsequenzen enthalten. Hierfür werden biotinylierte Sonden eingesetzt, die an Exonabschnitte hybridisieren (vgl. Abb. 7). Mit Hilfe magnetischer Streptavidin-*Beads* werden dann die „markierten“ Fragmente aus der Gesamtbibliothek isoliert (*TruSeq Exome Enrichment Kit Data Sheet*, Illumina, 2011). Da die Fragmentierung der genomischen DNA zufällig erfolgt, sind unter den angereicherten Fragmenten zum Teil auch Intron-Sequenzen, so dass Intron-Exon- bzw. Exon-Intron-Übergänge mitsequenziert werden. Durch eine Selektion bestimmter Gene erlaubt dieses Verfahren außerdem die Herstellung und NGS-Analyse benutzerdefinierter Bibliotheken.

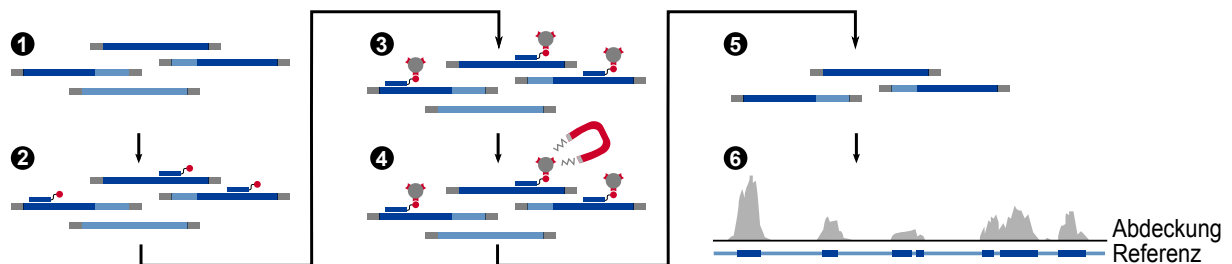
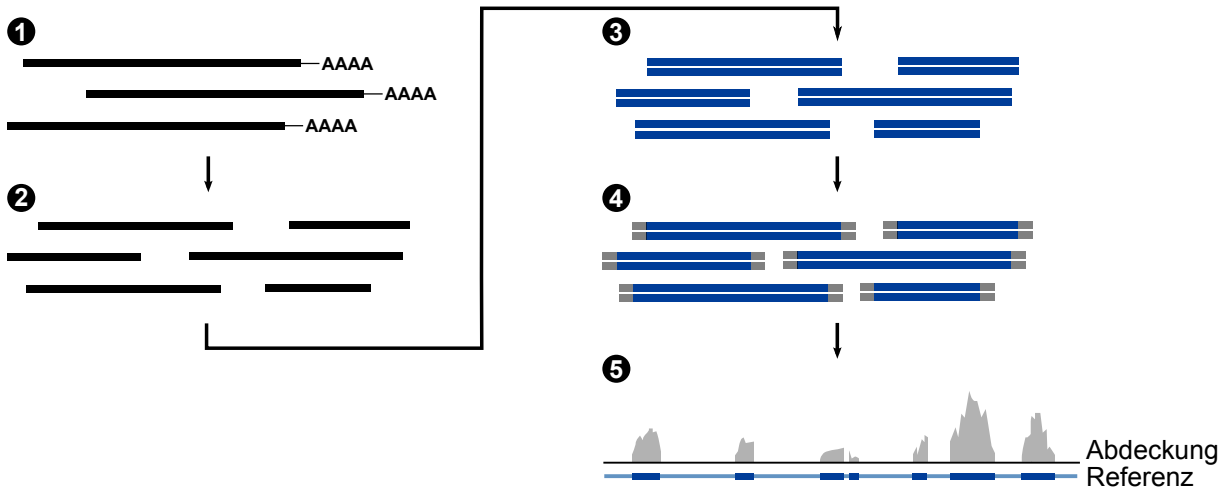


Abb. 1.7: Ablauf der Exomanreicherung in genomischen Sequenzierbibliotheken. In einer genomischen Sequenzierbibliothek (1, Adapter: grau, Exons: dunkelblau, Introns: hellblau) werden zur Anreicherung von DNA-Fragmenten mit Exomsequenzen biotinylierte Sonden eingesetzt, die spezifisch mit Exomsequenzen hybridisieren (2). Durch die Zugabe magnetischer Streptavidin-*Beads* (3) können die biotinmarkierten DNA-Fragmente nach Anlegen eines Magnetfelds isoliert werden (4). Anschließend werden die DNA-Fragmente eluiert (5). Aufgrund der zufälligen Fragmentierung der genomischen DNA werden in Exombibliotheken auch Intron-Exon- bzw. Exon-Intron-Übergänge mitsequenziert (6). (Abb. verändert nach dem Datenblatt zum *TruSeq Exome Enrichment Kit* (Illumina, 2011).)

Für die **Transkriptomsequenzierung** („*RNA-Seq*“) wird zunächst Gesamt-RNA isoliert, aus der mRNA mittels Poly-A-Sonden angereichert wird (vgl. **Abb. 1.8**). Die mRNA wird fragmentiert und die Fragmente mithilfe von *Random-Hexamer*-Primern in cDNA (komplementäre DNA, engl. *complementary DNA*) revers transkribiert. An die cDNA-Fragmente werden dann die Y-Adapter für die Brückenamplifikation und Sequenzierung

ligiert (*TruSeq™ RNA and DNA Sample Preparation Kits v2 Data Sheet*, Illumina, 2011). Beim *RNA-Seq* werden nur die tatsächlich exprimierten Abschnitte sequenziert, da nicht-kodierende (Intron-)Sequenzen beim Spleißen der mRNA entfernt wurden.



**Abb. 1.8: Generierung der Sequenzierbibliothek für die Transkriptomsequenzierung.** Zur Anreicherung von mRNA (schwarz) aus Gesamt-RNA werden Poly-A-Sonden eingesetzt (1). Nach Fragmentierung der mRNA (2) wird mittels *Random-Hexamer*-Primern die cDNA (blau) synthetisiert (3), an die anschließend die Sequenzieradapter (grau) ligiert werden (4). Bei der Transkriptomsequenzierung werden ausschließlich die exprimierten Abschnitte sequenziert (5, Exon: dunkelblau, Intron: hellblau). (Abb. verändert nach dem Datenblatt zu den *TruSeq™ RNA and DNA Sample Preparation v2 Kits* (Illumina, 2011).)

Neben diesen drei Grundformen der Hochdurchsatz-Sequenzierung (Genom-, Exom- und Transkriptomsequenzierung) gibt es weitere Varianten, die beispielsweise die genomweite Identifizierung von Transkriptionsfaktorbindestellen (*ChIP-Seq*, Robertson *et al.*, 2007), die Bestimmung von DNA-Methylierungen (Bisulfit-Sequenzierung, Taylor *et al.*, 2007) oder die Beurteilung von Synthese und Stabilität von RNA (*Bru-Seq*, Paulsen *et al.*, 2014) ermöglichen. In der vorliegenden Arbeit wurden ausschließlich Exom- und Transkriptomsequenzierungen durchgeführt, weswegen auf die weiteren Spezialanwendungen nicht weiter eingegangen wird.

### 1.2.2.3 NGS als Werkzeug in der Onkologie

Die Publikationen der vergangenen Jahre zeigen, dass NGS sich bereits als ein mächtiges Werkzeug in der Onkologie etabliert hat. So sollen beispielsweise zum besseren Verständnis der Onkogenese die Mutationen bestimmt werden, die zur Entstehung eines Tumors beitragen. Man unterscheidet *Driver*-Mutationen, die der Tumorzelle einen Wachstumsvorteil verschaffen, und *Passenger*-Mutationen, die der Tumorzelle keinen Selektionsvorteil verleihen (Vogelstein *et al.*, 2013). Dementsprechend sind Mutationen, die ein Onkogen aktivieren bzw. ein Tumorsuppressorgen ausschalten (vgl. Abschnitt 1.1.3), *Driver*-Mutationen.

Die Gene mit bisher identifizierten *Driver*-Mutationen lassen sich zwölf verschiedenen Signalwegen zuordnen, die das Schicksal und Überleben der Zelle und die Aufrechterhaltung der genomischen Integrität regulieren (Vogelstein *et al.*, 2013). Mittlerweile sind



einige Programme zur Vorhersage des onkogenen Potentials einer Mutation verfügbar (Krishnan & Ng, 2012). Die Identifizierung neuer *Driver*-Mutationen ist insbesondere für Patientengruppen bedeutsam, deren Tumore keine der bekannten Mutationen aufweisen. In einer aktuellen Metaanalyse von Xia *et al.* (2014) wurde in NGS-Daten von 241 Melanomen nach klinisch relevanten Mutationen gesucht. Ein Drittel der Melanome war negativ für wiederkehrende *Driver*-Mutationen in fünf verschiedenen Genen, die im Fokus aktueller Melanomtherapien stehen. Bei Mutationen, die in diesen „negativen“ Tumoren häufiger vorkommen, könnte es sich um weitere, bislang unbekannte *Driver*-Mutationen handeln.

Durch *RNA-Seq* können neben Mutationen auch andere Faktoren untersucht werden, die in direktem oder indirektem Zusammenhang mit dem Tumorwachstum stehen (Berger *et al.*, 2010). Dazu gehören die Analyse differentiell exprimierter Gene (Cheng *et al.*, 2012; van Delft *et al.*, 2012) und die Detektion neuer Transkripte, Spleißvarianten (Chen *et al.*, 2013; Jakhesara *et al.*, 2013) oder Gen-Fusionen (Maher *et al.*, 2009). Die Expressionsprofile können Aufschluss über deregulierte Signalwege geben (Ma *et al.*, 2012; Wu *et al.*, 2014).

Weiterhin war es möglich, bei einzelnen Patienten die Heterogenität ihrer Tumore mithilfe von NGS zu analysieren. In Abhängigkeit von der Tiefe der Sequenzierung können innerhalb einer gemischten Population auch Zellklone mit einer Frequenz von 1 in 5.000 erfasst werden (Campbell *et al.*, 2008). Dies ist beispielsweise von Bedeutung, wenn ein Tumor therapieresistente Subpopulationen aufweist. Gerlinger und Kollegen (2012; 2014) untersuchten Biopsien von räumlich getrennten Tumorarealen in Nierenkarzinomen und abgeleiteten Metastasen und konnten nur maximal ein Drittel der somatischen Mutationen in allen Regionen eines Tumors detektieren. Zudem wurden phylogenetische Analysen durchgeführt, die auf einen verzweigten und nicht auf einen linearen Verlauf der Entwicklung der Tumorsubpopulationen hindeuteten. Ein solches „Evolutionsschema“ ergab sich auch für die untersuchten Subklone zweier chronisch lymphatischer B-Zell-Leukämien (Campbell *et al.*, 2008).

Darüber hinaus lässt sich aus der Tumorerheterogenität die Notwendigkeit ableiten, weitere TAA zu identifizieren, wenn einige TAA nur in Subpopulationen exprimiert werden. Durch die Hochdurchsatzsequenzierung von Tumorzellen können überexprimierte und aberrant exprimierte sowie bisher nicht annotierte Proteine identifiziert werden, welche eine Quelle für potentiell immunogene Peptide sein können (reverse Immunologie, vgl. Abschnitt 1.2.1). Mittels NGS wurden bisher insbesondere Neoantigene aus Proteinen mit *Missense*-Mutationen untersucht (Castle *et al.*, 2012; Tran *et al.*, 2014; Rajasagi *et al.*, 2014). NGS eignet sich allerdings auch zur Detektion anderer Mutationstypen, wie beispielsweise interne Tandemduplikationen (ITDs) in FLT3 (Spencer *et al.*, 2013), die für immunogene Peptide kodieren können (Graf *et al.*, 2007).

### 1.2.3 Vorhersage der Peptid/HLA-Bindung mittels Algorithmen

Potentiell HLA-bindende Peptide in einem zu untersuchenden Protein können mithilfe von Algorithmen vorhergesagt werden, die die Wahrscheinlichkeit für die Bindung eines Peptids an ein MHC-Molekül wiedergeben. Die am häufigsten verwendeten Algorithmen basieren auf allelspezifischen Matrizen, wie z. B. SYFPEITHI (Rammensee *et al.*, 1999) und BIMAS (Parker *et al.*, 1994), bzw. auf neuronalen Netzwerken, wie NetMHC (Zhang *et al.*, 2008). Bei den matrixbasierten Verfahren werden die Bindeaffinitäten der einzelnen Aminosäuren der Peptidsequenz unabhängig voneinander betrachtet. Die Vorhersage der Bindeaffinität mithilfe von neuronalen Netzwerken (ANN, engl. *artificial neural networks*) bzw. Maschinelernen ermöglicht es, auch den Einfluss der anderen Aminosäuren miteinzubeziehen. Im Vergleich wurde durch die neueren ANN-Verfahren eine höhere Vorhersagegenauigkeit erzielt (Lundegaard *et al.*, 2010). Des Weiteren können aufgrund der stetig wachsenden Zahl identifizierter immunogener Peptide Bindewahrscheinlichkeiten für immer mehr MHC-Allele mit steigender Sicherheit vorhergesagt werden (Lafuente & Reche, 2009).

## 1.3 Zielsetzung der Arbeit

Das Ziel der vorliegenden Arbeit war die Identifizierung mutierter Neoantigene im humanen Melanommodell Ma-Mel-86 basierend auf den Exom- und Transkriptomsequenzierungen der vier vorhandenen Tumorzelllinien und der autologen EBV-immortalisierten B-Zelllinie. Der Fokus sollte dabei auf der systematischen Detektion von somatischen, nicht-synonymen Punktmutationen in exprimierten Genen liegen, die zur Generierung von potentiell immunogenen Peptiden führen, welche von den HLA-Klasse I-Molekülen der Patientin präsentiert werden können. Mit den zur Verfügung stehenden autologen T-Zellen sollte die Immunogenität der Peptidkandidaten *in vitro* überprüft und erkannte Antigene abschließend näher charakterisiert werden.

Hierfür waren die folgenden Hauptarbeitsschritte geplant:

- Gewinnung des Probenmaterials für die Sequenzierung,
- Sequenzierung der Exome und Transkriptome der Zelllinien des Modells (in Kooperation mit Prof. Dr. T. Hankeln, Institut für Molekulargenetik, Johannes Gutenberg-Universität, Mainz),
- Aufbereitung und *Alignment* der Sequenzierdaten,
- Detektion und Filterung der relevanten Mutationen,
- Vorhersage von Peptidkandidaten,
- Auswahl autologer MLTCs und Optimierung der Testbedingungen für die Bestimmung der Immunogenität der Peptidkandidaten,
- Überprüfung der Erkennung der Peptide durch autologe MLTCs,

- Generierung von peptidreaktiven T-Zellklonen,
- Identifizierung der restringierenden HLA-Moleküle,
- Ermittlung der C-terminalen Peptidenden und
- Sicherstellung der intrazellulären Prozessierung der Peptide.

Die Arbeit sollte zeigen, ob sich immunogene mutierte Peptide anhand der Sequenzierung der Exome und Transkriptome vorhersagen lassen, wie die relevanten Mutationen auf die vier Melanomzelllinien verteilt waren und zu welchem Anteil die Mutationen tatsächlich immunogen waren.



## 2 Material und Methoden

### 2.1 Material

#### 2.1.1 Geräte

Gerät	Bezeichnung/ Spezifikation	Hersteller
Autoklav	Durchreichautoklav	Holzner GmbH, Nußloch
Autoklav	VX-150	Systec GmbH, Wettenberg
Bakterienbrutschrank		Heraeus Instruments, Hanau
Brucheisautomat		Ziegra, Isernhagen
Chemikalienwaage		Sartorius, Göttingen
<sup>137</sup> Cäsium-Quelle	Gammacell 2000	Nuklear Data, Frankfurt a. M.
CO <sub>2</sub> -Inkubator	BBD 6620	Heraeus Instruments, Hanau
Durchflusszytometer	FACS Canto II	Becton Dickinson, Heidelberg
Elektroporationsgerät	Gene Pulser <sup>®</sup> Xcell <sup>TM</sup>	Bio-Rad GmbH, München
ELISpot-Analyse-System	ImmunoSpot <sup>®</sup> Analyzer	CTL Europe, Bonn
Feinwaage	Precisia 40SM-200A	PAG Oerlikon AG, Zürich (Schweiz)
Fluorometer	Qubit <sup>®</sup> 2.0	Invitrogen, Darmstadt
Gefrierschrank (-80 °C)	HERAfreeze	Thermo Scientific, Schwerte
Geldokumentationsgerät	BioDocAnalyze	Biometra, Göttingen
Gelelektrophoresekammern		PEQLAB Biotechnologie GmbH, Erlangen
Heizblöcke	Thermomixer 5436 Thermomixer comfort	Eppendorf, Hamburg
Kolbenhubpipetten	Maximales Pipettiervolumen: 10 µl, 100 µl, 200 µl, 1.000 µl	Eppendorf, Hamburg Gilson, Middleton (USA) STARLAB, Hamburg
Kühl- und Gefrierschränke		Liebherr, Biberach a. d. Riß
Magnetrührer mit Heizplatte	IKAMAG RET G	Janke & Kunkel, Staufen
Magnetständer / Magnet (für Zellseparation)	MidiMACS, OctoMACS	Miltenyi Biotec, Bergisch Gladbach
Mikrowelle	MWS 2819	Bauknecht, Neunkirchen
Multikanalpipette	m300	Biohit/Sartorius, Göttingen
Multikanalpipette	10-100 µl bzw. 30-300 µl ErgoOne <sup>®</sup> 12-Channel Pipette	STARLAB, Hamburg
Multisteppipette	Multipipet plus	Eppendorf, Hamburg
Peristaltikpumpe	Econo Pump EP-1	Bio-Rad GmbH, München
Phasenkontrastmikroskop	ECLIPSE TS100	Nikon, Japan
pH-Meter	FiveEasy	Mettler Toledo GmbH, Gießen

wird fortgesetzt...

## 2.1 Material

Gerät	Bezeichnung/ Spezifikation	Hersteller
Pipettierhilfen	accu-jet <sup>®</sup> pro PIPETBOY acu	Brand, Wertheim INTEGRA Biosciences GmbH, Fernwald
Schüttelinkubator	Certomat SII mit Wärmehaube Certomat H	Sartorius, Göttingen
Spannungsgerät	GPS 200/400	Pharmacia Biotech, Freiburg
Spektralphotometer	Ultrospec 3000	Pharmacia Biotech, Freiburg
Sterilwerkbank	HERAsafe	Haereus Instruments, Hanau
Stickstoffbank	1400 Series mit MVE TEC 3000 Control System	Chart Biomed, Ball Ground (USA)
Stickstofftank	XL-180	Taylor-Wharton, Minnetonka (USA)
Thermocycler	T1, TGradient	Biometra, Göttingen
Tischzentrifuge	Centrifuge 5424	Eppendorf, Hamburg
Tischzentrifuge, gekühlt	Centrifuge 5415	Eppendorf, Hamburg
Trockenschrank	OV 5	Biometra, Göttingen
Vakuumpumpe	ME 2C	Vacuubrand GmbH + Co. KG, Weinheim
Vortexer	VF2	Janke & Kunkel, Staufen
Wasserbad, beheizbar		GFL, Großburgwedel
Wasserdeionisierungsanlage	PURELAB Classic DI	Elga Labwater, Celle
Zentrifuge	Rotina 420	Hettich, Bäch (Schweiz)

### 2.1.2 Verbrauchsmaterialien, Glas- und Plastikwaren

Material	Bezeichnung/ Spezifikation	Hersteller/Bezugsquelle
Bechergläser, Glas	500 ml, 1.000 ml, 2.000 ml	Jenaer Glas bzw. Schott, Mainz
Bechergläser, Kunststoff	500 ml, 1.000 ml, 2.000 ml	Vitlab, Seeheim-Jugenheim
Duran-Glasflaschen	0,5l, 1l, 2l und 5l	Schott, Mainz
Einfriercontainer	NALGENE <sup>™</sup> Cryo 1 °C Freezing Container	Thermo Scientific, Schwerte
Einfrierröhrchen	2 ml Cryotube PP	greiner bio-one, Frickenhausen
Elektroporationsküvetten	GenePulser <sup>®</sup> Cuvette, 0,2 cm	Bio-Rad GmbH, München
ELISpot-Platten	MultiScreen-IP, 0,45 µm, weiß, steril MultiScreen <sub>HTS</sub> -IP, 0,45 µm, transparent, steril	Millipore, Eschborn
Eppendorf-Röhrchen	0,5 ml, 1,5 ml und 2 ml	Eppendorf, Hamburg
Erlenmeyerkolben	500 ml, 1.000 ml und 2.000 ml	Jenaer Glas bzw. Schott, Mainz
FACS-Röhrchen	5 ml	Becton Dickinson GmbH, Heidelberg
Falcon-Röhrchen	Polypropylen, 15 ml und 50 ml	greiner bio-one, Frickenhausen
Filtereinheiten	0,2 µm 0,45 µm	Whatman GmbH, Dassel
Gewebekulturflaschen	25 cm <sup>2</sup> , 75 cm <sup>2</sup> , 175 cm <sup>2</sup>	greiner bio-one, Frickenhausen
Gewebekulturplatten	24-Well-, 48-Well- und 96-Well-Gewebekulturplatten (Flachboden), 96-Well-Platten (Rundboden), 6-Well-Gewebekulturplatten	greiner bio-one, Frickenhausen
Handschuhe	sempercare <sup>®</sup> Latex Powder-Free	sempermed/semperit Technische Produkte GmbH, Wien (Österreich)

wird fortgesetzt...

Material	Bezeichnung/ Spezifikation	Hersteller/Bezugsquelle
Impfösen	10 µl	Nunc/Thermo Scientific, Roskilde, Denmark
MACS-Säulen (für Zellseparation)	LS	Miltenyi Biotec, Bergisch Gladbach
Parafilm®		Pechiney Plastic Packaging, Menasha, USA
PCR-Reaktionsgefäße	0,2 ml	greiner bio-one, Frickenhausen
Petrischalen	ø35 mm, 60 mm, 100 mm	greiner bio-one, Frickenhausen
Pipettenspitzen	gestopft (10 µl, 20 µl, 100 µl, 200 µl, 1.000 µl) ungestopft (10 µl, 200 µl, 1.000 µl)	STARLAB, Bonn
Prä-Separationsfilter	30 µm	Miltenyi Biotec, Bergisch Gladbach
Quartzküvetten	10 mm	Hellma Analytics, Müllheim
Säulen zur Antikörperaufreinigung	Affinitätssäule HiTrap™ Protein G HP Säulen, 1 ml	VWR, Darmstadt
Serologische Pipetten	2 ml, 5 ml, 10 ml, 25 ml und 50 ml	greiner bio-one, Frickenhausen
Skalpelle	Feather disposable Scalpel	pfm Medical, Köln
Sterilfilter	Steritop™, 0,22 µm/0,45 µm	Millipore, Eschborn
Zählkammer	Fuchs-Rosenthal	Marienfeld Superior, Lauda Königshofen

### 2.1.3 Chemikalien und Reagenzien

Chemikalie/Reagenz	Bezeichnung/ Spezifikation	Hersteller/Bezugsquelle
3-Amino-9-Ethyl-Carbazol (AEC)	Tabletten	Sigma, Steinheim
Agarose	Star Pure AGAROSE	STARLAB, Hamburg
AIM-V® Medium	<i>serum free medium</i> , 1×	Gibco, Berlin
Ammoniumacetat	<i>pro analysi</i>	Merck, Darmstadt
Aqua	Ecotainer® 1.000 ml	B. Braun Melsungen AG, Melsungen
Aqua	<i>ad iniectabilia</i> , 100 ml	B. Braun Melsungen AG, Melsungen
Bacto Agar		Becton Dickinson GmbH, Heidelberg
Bacto Tryptone	<i>Pancreatic Digest of Casein</i>	Becton Dickinson GmbH, Heidelberg
Bacto Yeast Extract	<i>Extract of Autolysed Yeast Cells</i>	Becton Dickinson GmbH, Heidelberg
Bovines Serumalbumin (BSA)	<i>lyophilized powder</i> , ≥96%	Sigma, Steinheim
Diethyl-Pyrocbonat (DEPC)	<i>approx. 97% NMR</i>	Sigma, Steinheim
Dimethylsulfoxid (DMSO)	<i>for synthesis</i>	Merck, Darmstadt
Dimethylformamid (DMF)	>99,8%, zur Synthese	Carl Roth AG, Karlsruhe
DNA-Ladepuffer	6×	Thermo Scientific, Schwerte
dNTP-Mix	10 mM <i>each</i>	NEB, Frankfurt a. M.
Essigsäure	100%, <i>per analysi</i>	Carl Roth AG, Karlsruhe
Ethanol	absolut, zur Synthese	AppliChem, Darmstadt
Ethylendiamin-Tetraessigsäure (EDTA)	Dihydrat, 99+%	Sigma, Steinheim
FACSClean		Becton Dickinson GmbH, Heidelberg

wird fortgesetzt...

## 2.1 Material

Chemikalie/Reagenz	Bezeichnung/ Spezifikation	Hersteller/Bezugsquelle
FACSFlow™		Becton Dickinson GmbH, Heidelberg
FACSRinse		Becton Dickinson GmbH, Heidelberg
Fötales Kälberserum (FCS)		Sigma, Steinheim
Formaldehyd	37%, zur Analyse	AppliChem, Darmstadt
GelRed	10.000× <i>in water</i>	Biotrend, Köln
Glycerol	BioXtra, ≥99% (GC)	Sigma, Steinheim
Histopaque 1077		Sigma, Steinheim
Lipofectamine™	2000, 3000	Invitrogen, Leek (Niederlande)
Methanol	zur Analyse	J. T. Baker, Deventer (Niederlande)
β-Mercaptoethanol	14,3 M	Sigma, Steinheim
(N-Morpholino)- Propansulfonsäure (MOPS)	<i>molecular biology grade</i>	VWR, Haasrode (Belgien)
Natriumacetat	wasserfrei, <i>per analysis</i>	Merck, Darmstadt
Natriumacetat	Lösung (3 M, pH 5,2) für die Molekularbiologie	AppliChem GmbH, Darmstadt
Natriumchlorid	≥99,5%, <i>per analysis</i>	Carl Roth AG, Karlsruhe
Natriumhydrogencarbonat	zur Analyse	Merck, Darmstadt
Natriumhydrogenphosphat	Monohydrat, zur Analyse	Merck, Darmstadt
OptiMEM®	<i>Reduced Serum Medium, no phenol red</i>	Gibco, Berlin
<i>Phosphate-buffered saline</i> (PBS), Pulver	w/o Ca <sup>2+</sup> , w/o Mg <sup>2+</sup>	Biochrom AG, Berlin
<i>Phosphate-buffered saline</i> (PBS), steril	1×	Sigma, Steinheim
Phenol-Chloroform	<i>Roti Phenol Chloroform for separation of DNA/RNA</i>	Carl Roth AG, Karlsruhe
Propan-2-ol (Isopropanol)	<i>Analytical Reagent Grade</i>	Fisher-Scientific, Schwerte
RNA-Ladepuffer	2×	Fermentas/Thermo Scientific, Schwerte
RPMI-1640	<i>with L-glutamine and NaHCO<sub>3</sub></i>	Sigma, Steinheim
S.O.C. Medium		Invitrogen, Leek (Niederlande)
TBE-Puffer	10×	Serva Electrophoresis GmbH, Heidelberg
Trypanblau	für die Mikroskopie	Merck, Darmstadt
Trypsin-EDTA	1×, <i>sterile-filtered</i>	Sigma, Steinheim
Tween®20	Polyoxyethylen-Sorbitan- Monolaurat	Sigma, Steinheim
Wasserstoffperoxid	30%, medizinisch reinst, stabilisiert	Carl Roth AG, Karlsruhe



## 2.1.4 Antibiotika

Antibiotikum	Spezifikation	Bezugsquelle	Verwendete Konzentration
Ampicillin	5 g	Sigma, Steinheim	100 µg/ml
G418	flüssig (gebrauchsfertig), steril, 50 mg/ml	Biochrom AG, Berlin	0,8-1 mg/ml
Kanamycin	<i>Kanamycin sulphate</i> , 5 g	Roth, Karlsruhe	50 µg/ml
Penicillin-Streptomycin	100.000 U Penicillin und 10 mg Streptomycin pro ml	Sigma, Steinheim	1%

## 2.1.5 Molekulargewichtsmarker

Die DNA-Molekulargewichtsmarker (500 ng) wurden *ad* 5 bzw. 10 µl dH<sub>2</sub>O (*ad inieci-tabilia*) verdünnt und DNA-Ladepuffer wurde vor Beladen des Gels hinzugefügt. Der RNA-Molekulargewichtsmarker wurde mit RNA-Ladepuffer versetzt und für 5 min bei 72 °C inkubiert (siehe auch Abschnitt 2.2.2.8.2).

Bezeichnung	Spezifikation	Bezugsquelle
1 kb DNA Ladder	500 µg/ml, 0,5-10,0 kb	NEB, Frankfurt a. M.
100 bp DNA Ladder	500 µg/ml, 100-1.517 bp	NEB, Frankfurt a. M.
<i>GeneRuler</i> <sup>TM</sup> 100 bp DNA Ladder	0,5 µg/µl, 100-1.000 bp	Fermentas/Thermo Scientific, Schwerte
<i>RiboRuler</i> <sup>TM</sup> High Range RNA Ladder	200-6.000 bp	Fermentas/Thermo Scientific, Schwerte

## 2.1.6 Rekombinante humane Zytokine und Mitogene

Zytokin/Mitogen	Spezifikation	Bezugsquelle	Lagerung
<b>Interleukin-2</b> (rhIL-2)	Proleukin <sup>®</sup> S 18 Mio. I.E. Pulver	Novartis Pharma, Nürnberg	gelöst in AIM-V (1×10 <sup>6</sup> U/ml bzw. 2,5×10 <sup>4</sup> U/ml), gelagert bei -20 °C
<b>Interleukin-7</b> (rhIL-7)	<i>research grade</i>	Miltenyi Biotech, Bergisch Gladbach	gelöst in AIM-V (10 µg/ml), gelagert bei -20 °C
<b>Interleukin-15</b> (rhIL-15)	<i>premium grade</i>	Miltenyi Biotech, Bergisch Gladbach	gelöst in AIM-V (10 µg/ml), gelagert bei -20 °C
<b>Rekombinantes humanes IFN<math>\gamma</math></b>	>97%	R&D Systems, Wiesbaden-Nordenstadt	gelöst in PBS (10 <sup>5</sup> U/ml), gelagert bei -20 °C
<b>Phytohämagglutinin E</b> (PHA)	lyophilisiert	Biochrom AG, Berlin	gelöst in PBS (500 µg/ml), gelagert bei -20 °C

### 2.1.7 Enzyme

Die gelisteten Enzyme wurden entsprechend den Herstellerangaben mit den mitgelieferten Puffern und Reagenzien verwendet.

Enzym	Bezeichnung/ Spezifikation	Hersteller/Bezugsquelle
Advantage2-Polymerase	Advantage <sup>®</sup> 2 PCR Kit	Clontech, Saint-Germain-en-Laye (Frankreich)
Alkaline Phosphatase, Calf Intestinal (CIP)	10.000 U/ml	NEB, Frankfurt a. M.
BP Clonase	Gateway <sup>®</sup> BP Clonase <sup>™</sup> II Enzyme mix	Invitrogen, Leek (Niederlande)
DNase I	grade II, from bovine pancreas	Roche Diagnostics, Mannheim
Easy-A-Polymerase	Easy-A High-Fidelity PCR Cloning Enzyme, 5 U/μl	Agilent Technologies, Waldbronn
LR Clonase	LR Clonase <sup>®</sup> II Plus enzyme	Invitrogen, Leek (Niederlande)
Q5-Polymerase	Q5 <sup>®</sup> High-Fidelity DNA Polymerase, 2.000 U/ml	NEB, Frankfurt a. M.
Restriktionsendonukleasen	8-20 U/μl	Promega, Mannheim NEB, Frankfurt a. M.
T4 DNA Ligase	400.000 U/ml	NEB, Frankfurt a. M.
Taq Polymerasen	VWR <sup>®</sup> Taq DNA-Polymerase, 5 U/μl 5PRIME MasterMix, 2,5×	VWR, Darmstadt 5prime GmbH, Hilden

### 2.1.8 Antikörper

Antikörper	Isotyp	Konjugat <sup>a</sup>	Spezifität <sup>b</sup>	Klon	Hersteller/ Bezugsquelle	Verwendung <sup>c</sup>
Anti-Human Interferon-γ mAb	Maus IgG1	-	IFNγ	1-D1K	Mabtech AB, Nacka Strand (Schweden)	E
Anti-Human Interferon-γ mAb	Maus IgG1	Bt	IFNγ	7-B6-1	Mabtech AB, Nacka Strand (Schweden)	E
B1.23.2*	Maus IgG2a	-	HLA-B, -C, -Aw19	-	Hybridomüberstand <sup>d</sup> (Rebai & Malissen, 1983)	D, B
C7709.A2.6*	Maus IgG2a	-	HLA-A02, -A24	-	Hybridomüberstand (Ikeda <i>et al.</i> , 1997)	D, B
CD16	Maus IgG1	F	CD16	3G8	Beckman Coulter GmbH, Krefeld	D
CD3	Maus IgG1	F	CD3	UCHT1	Beckman Coulter GmbH, Krefeld	D
CD3	Maus IgG1	P	CD3	UCHT1	Beckman Coulter GmbH, Krefeld	D

wird fortgesetzt. . .

Antikörper	Isotyp	Konjugat <sup>a</sup>	Spezifität <sup>b</sup>	Klon	Hersteller/ Bezugsquelle	Verwendung <sup>c</sup>
CD3 / CD8	Maus IgG1	F/P	CD3, CD8	UCHT1 / SFC121Thy2D3	Beckman Coulter GmbH, Krefeld	D
CD3 / CD4	Maus IgG1	F/P	CD3, CD4	UCHT1 / 13B8.2	Beckman Coulter GmbH, Krefeld	D
CD4	Maus IgG1	F	CD4	13B8.2	Beckman Coulter GmbH, Krefeld	D
CD56	Maus IgG1	P	CD56	N901	Beckman Coulter GmbH, Krefeld	D
CD8	Maus IgG1	F	CD8	B9.11	Beckman Coulter GmbH, Krefeld	D
<i>Goat</i> F(ab') <sub>2</sub> Fragment Anti- <i>Mouse</i> IgG + IgM (H+L)-FITC	Ziege F(ab') <sub>2</sub> Fragment	F	murines IgG + IgM (H+L)	polyklonal	Beckman Coulter GmbH, Krefeld	D
HLA-A01/A11/A26+	Maus IgM	-	HLA-A01, -A11, -A26	-	OneLambda, Canoga Park (USA)	D
HLA-A2/A24	Maus IgM	-	HLA-A02, -A24	-	OneLambda, Canoga Park (USA)	D
HLA-B13/B62/B15	Maus IgG	-	HLA-B13, -B62, -B15	-	OneLambda, Canoga Park (USA)	D
<i>Human beta-2 microglobulin</i>	Maus IgG1	-	β <sub>2</sub> -Mikroglobulin	B2M-02	Thermo Scientific, Schwerte	D
Isotyp-Kontrolle IgG1	Maus IgG1	-	unspezifisch	679.1Mc7	Beckman Coulter GmbH, Krefeld	D
Isotyp-Kontrolle IgG1	Maus IgG1	F/P	unspezifisch	679.1Mc7	Beckman Coulter GmbH, Krefeld	D
OKT-3	Maus IgG2a	-	CD3	-	Hybridomüberstand (Kung <i>et al.</i> , 1979)	D, B, T
TCR Vbeta 1	Ratte IgG1	P	Vβ1 (TCRBV1S1, S2)	BL37.2	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 2	Maus IgG1	P	Vβ2 (TCRBV2S1)	MPB2D5	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 3	Maus IgM	F	Vβ3 (TCRBV3S1)	CH92	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 4	Ratte IgM	P	Vβ4 (TCRBV4)	WJF24	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 5.1	Maus IgG2a	P	Vβ5.1 (TCRBV5S1)	IMMU 157	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 5.2	Maus IgG1	P	Vβ5.2 (TCRBV5S2)	36213	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 5.3	Maus IgG1	P	Vβ5.3 (TCRBV5S3)	3D11	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 7.1	Maus IgG2a	P	Vβ7.1 (TCRBV7S1, S2)	ZOE	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 7.2	Maus IgG2a	P	Vβ7.2 (TCRBV7S1)	ZIZOU4	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 8	Maus IgG2a	P	Vβ8 (TCRBV8S1, S2)	56C5.2	Beckman Coulter GmbH, Krefeld	D

wird fortgesetzt...

## 2.1 Material

Antikörper	Isotyp	Konjugat <sup>a</sup>	Spezifität <sup>b</sup>	Klon	Hersteller/ Bezugsquelle	Verwendung <sup>c</sup>
TCR Vbeta 9	Maus IgG2a	P	Vβ9 (TCRBV9S1)	FIN9	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 11	Maus IgG2a	P	Vβ11 (TCRBV11S1)	C21	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 12	Maus IgG2a	P	Vβ12 (TCRBV12S1)	VER2.32.1	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 13.1	Maus IgG2b	P	Vβ13.1 (TCRBV13S1)	IMMU 222	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 13.2	Maus IgG1	P	Vβ13.2 (TCRBV13S2)	H132	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 13.6	Maus IgG1	F	Vβ13.6 (TCRBV13S6)	JU74.3	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 14	Maus IgG1	P	Vβ14 (TCRBV14S1)	CAS1.1.3	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 16	Maus IgG1	F	Vβ16 (TCRBV16S1)	TAMAYA1.2	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 17	Maus IgG1	F	Vβ17 (TCRBV17S1)	E17.5F3.15.13	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 18	Maus IgG1	P	Vβ18 (TCRBV18S1)	BA62.6	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 20	Maus IgG2a	P	Vβ20 (TCRBV20S1)	ELL1.4	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 21.3	Maus IgG2a	F	Vβ21.3 (TCRBV21S3)	IG125	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 22	Maus IgG1	P	Vβ22 (TCRBV22S1)	IMMU 546	Beckman Coulter GmbH, Krefeld	D
TCR Vbeta 23	Maus IgG1	P	Vβ23 (TCRBV23S1)	AF23	Beckman Coulter GmbH, Krefeld	D
W6/32	Maus IgG2a	-	HLA-A/B/C	-	Hybridomüberstand (Barnstable <i>et al.</i> , 1978)	D, B

<sup>a</sup> -: unkonjugiert, Bt: Biotin, F: Fluoresceinisothiocyanat (FITC), P: Phycoerythrin (PE)

<sup>b</sup> Spezifität gegen humane Proteine (sofern nicht anders angegeben)

<sup>c</sup> D: Durchflusszytometrie, B: Blockadeexperimente, T: T-Zellstimulation

<sup>d</sup> Für Blockadeexperimente wurden die Hybridomüberstände mithilfe des MAbTrap™ Kits (GE Healthcare) nach dem Herstellerprotokoll über Affinitätssäulen (HiTrap™ Protein G HP, GE Healthcare) aufgereinigt.

\* Freundlicherweise von P. G. Coulie (LICR, Brüssel, Belgien) zur Verfügung gestellt

### 2.1.9 MicroBeads

Die Isolation CD8<sup>+</sup> Zellen aus PBMCs oder Pan-T-Zellpopulationen mittels magnetischer Zellseparation (siehe Abschnitt 2.2.1.6) erfolgte mithilfe der CD8-*MicroBeads* der Firma Miltenyi Biotec.

## 2.1.10 Molekular- und immunbiologische Kits

Die Verwendung der folgenden Kits erfolgte nach den Angaben des Herstellers, sofern nicht anders im Methodenteil angegeben.

Kit	Hersteller
<i>RNase-free DNase Set</i>	QIAGEN, Hilden
<i>High Pure<sup>®</sup> PCR Product Purification Kit</i>	Roche, Mannheim
<i>HiSpeed<sup>®</sup> Plasmid Maxi Kit</i>	QIAGEN, Hilden
<i>MABTrap<sup>™</sup> Kit</i>	GE Healthcare, Freiburg
<i>pcDNA3.1/V5-His TOPO TA Expression Kit<sup>®</sup></i>	Invitrogen, Leek (Niederlande)
<i>QIAamp<sup>®</sup> DNA Mini Kit</i>	QIAGEN, Hilden
<i>QIAprep<sup>®</sup> Spin Miniprep Kit</i>	QIAGEN, Hilden
<i>QIAquick<sup>®</sup> Gel Extraction Kit</i>	QIAGEN, Hilden
<i>QIAquick<sup>®</sup> PCR Purification</i>	QIAGEN, Hilden
<i>Qubit<sup>®</sup> dsDNA BR Assay Kit</i>	Invitrogen, Leek (Niederlande)
<i>Qubit<sup>®</sup> dsDNA HS Assay Kit</i>	Invitrogen, Leek (Niederlande)
<i>Qubit<sup>®</sup> RNA HS Assay Kit</i>	Invitrogen, Leek (Niederlande)
<i>RNeasy Mini Kit</i>	QIAGEN, Hilden
<i>SuperScript<sup>®</sup> III First Strand Synthesis System</i>	Invitrogen, Leek (Niederlande)
<i>Taq-DNA-Polymerase Kit</i>	VWR, Darmstadt
<i>Vectastain<sup>®</sup> Elite Kit</i>	Vector Laboratories, Peterborough (Vereinigtes Königreich)
<i>TruSeq<sup>™</sup> DNA Sample Preparation Kit<sup>*</sup></i>	Illumina, Eindhoven (Niederlande)
<i>TruSeq<sup>™</sup> RNA Sample Preparation v2 Kit<sup>*</sup></i>	Illumina, Eindhoven (Niederlande)
<i>TruSeq<sup>™</sup> Exome Enrichment Kit<sup>*</sup></i>	Illumina, Eindhoven (Niederlande)

\* Die Kits zur Generierung der Sequenzierbibliotheken wurden von der Firma Genterprise GENomics (Mainz) verwendet.

## 2.1.11 Lösungen und Puffer

Folgende Puffer und Lösungen wurden angesetzt, wie beschrieben, und bei Raumtemperatur gelagert, sofern nichts anderes angegeben ist.

Puffer/Lösung	Zusätze	Bemerkungen
<b>Acetatpuffer für IFN<math>\gamma</math>-ELISpot-Assays (1l)</b>	73,6 ml 0,2 N Essigsäure 176,0 ml 0,2 N Natriumacetat 750,4 ml dH <sub>2</sub> O	
<b>AEC-Substrat für IFN<math>\gamma</math>-ELISpot-Assays (50 ml)</b>	1 Tablette 3-Amino-9-Ethyl-Carbazol 2,5 ml Dimethylformamid 47,5 ml Acetatpuffer	filtriert (0,45 $\mu$ g) und maximal eine Woche bei 4 °C gelagert
<b>Ampicillin (50 mg/ml)</b>	5,0 g Ampicillin ad 100 ml dH <sub>2</sub> O (steril)	in Aliquots bei -20 °C gelagert
<b>Avidin-Peroxidase-Komplex für IFN<math>\gamma</math>-ELISpot-Assays (10 ml)</b>	1 Tropfen <i>Vectastain<sup>®</sup></i> -Reagenz A 1 Tropfen <i>Vectastain<sup>®</sup></i> -Reagenz B 0,1 ml 10% Tween <sup>®</sup> 20 10,0 ml PBS	30 min lichtgeschützt bei Raumtemperatur inkubiert
<b>DEPC-behandeltes Wasser (2l)</b>	2,0 ml DEPC ad 2,0l dH <sub>2</sub> O	20 min gemischt und autoklaviert wird fortgesetzt...

## 2.1 Material

Puffer/Lösung	Zusätze	Bemerkungen
<b>Erythrozytenlysepuffer</b> (11)	8,29 g NH <sub>4</sub> Cl 1,0 g KHCO <sub>3</sub> 0,0372 g EDTA ad 1,0l dH <sub>2</sub> O	eingestellt auf pH 7,29 und sterilfiltriert
<b>Essigsäure 0,2 N</b> (11)	11,3 g Essigsäure 96% ad 1,0l dH <sub>2</sub> O	
<b>FACS-Puffer</b> (500 ml)	500 ml PBS 2,5 g BSA	sterilfiltriert und bei 4 °C gelagert
<b>FACS-Fixans</b> (1% Formaldehyd, 50 ml)	1,35 ml 37% Formaldehyd ad 50 ml PBS	bei 4 °C gelagert
<b>Kanamycin</b> (50 mg/ml)	5,0 g Kanamycin ad 100 ml dH <sub>2</sub> O (steril)	in Aliquots bei -20 °C gelagert
<b>MACS-Puffer</b> (500 ml)	500 ml PBS 2,5 g BSA 0,29 g EDTA	sterilfiltriert und bei 4 °C gelagert
<b>MOPS-Puffer, 10× (pH 7)</b> (11)	41,9 g (N-Morpholino)- Propansulfonsäure (MOPS) 8,2 g Natriumacetat 3,72 g EDTA ad 1,0l DEPC-behandeltes Wasser	sterilfiltriert und bei 4 °C lichtgeschützt gelagert
<b>Natriumacetat 0,2 N</b> (11)	16,4 g Natriumacetat ad 1,0l dH <sub>2</sub> O	
<b>Natriumchlorid, 4,5%</b> (500 ml)	22,5 g NaCl ad 500 ml dH <sub>2</sub> O (steril)	
<b>PBS, 1×</b> (10l)	95,5 g PBS-Pulver ad 10l dH <sub>2</sub> O	
<b>TBE, 0,5×</b> (2l)	100 ml TBE (10×) ad 2,0l dH <sub>2</sub> O	
<b>Trypanblau, 0,2%</b> (Stammlösung, 500 ml)	1,0 g Trypanblau ad 500 ml dH <sub>2</sub> O (steril)	
<b>Trypanblau</b> (Gebrauchslösung, 500 ml)	375 ml 0,2% Trypanblau 125 ml 4,5% NaCl	sterilfiltriert
<b>Tween<sup>®</sup>20, 10%</b> (500 ml)	50 ml Tween <sup>®</sup> 20 ad 500 ml PBS	

### 2.1.12 Oligonukleotide

Für die Amplifikation spezifischer DNA-Abschnitte in PCRs wurden entsprechende Primer von der Firma Eurofins MWG Operon (Ebersberg) bezogen. Die Primer für die Generierung der *attB*-PCR-Produkte für die *Gateway*-Klonierung wurden von der Firma IDT (Leuven, Belgien) synthetisiert. Die lyophilisierten Oligonukleotide wurden in dH<sub>2</sub>O (*ad iniectionem*) für 10 min bei 50 °C und 500 rpm gelöst (Stockkonzentration: 100 µM) und bei -20 °C gelagert. Die Verbrauchslösung wurde auf 5 µM eingestellt und ebenso gelagert. Je PCR-Ansatz wurden, sofern nicht anders angegeben, 5-10 pmol Primer eingesetzt.

Das in den Primerbezeichnungen verwendete Symbol \$ zeigt an, dass mithilfe des Primers ein Stopcodon an die zu amplifizierende Sequenz angefügt wurde. In den Primersequenzen

sind diese Stopcodons fett gedruckt und unterstrichen. Weiterhin sind mutierte Basen in Rot dargestellt.

### 2.1.12.1 HLA-Klasse I-spezifische Primer

Bezeichnung	Orientierung	Sequenz (5'→3')
HLA-5p2	<i>sense</i>	GGGCGAATTCGGACTCAGAATCTCCCCAGACGCCGAG
HLA-3pA	<i>antisense</i>	CCGCGTCGACTTGGGAGGGAGCACAGGTCAGCGTGGGAAG
HLA-3pB	<i>antisense</i>	GGGCGTCGACTGGGAGGAAACACAGGTCAGCATGGGAAC
HLA-3pC	<i>antisense</i>	GGGCGTCGACCTGCATCTCAGTCCACACAGGC

### 2.1.12.2 Primer zur Amplifizierung von Inserts in *pcDNA3*- und *pDONR*-Vektoren

Bezeichnung	Orientierung	Sequenz (5'→3')
T7	<i>sense</i>	TAATACGACTCACTATAGGG
BGH	<i>antisense</i>	TAGAAGGCACAGTCGAGG
CMV	<i>sense</i>	CGCAAATGGGCGGTAGGCGTG
M13_for	<i>sense</i>	GTAAAACGACGGCCAG
M13_rev	<i>antisense</i>	CAGGAAACAGCTATGAC

### 2.1.12.3 Primer für den Nachweis von $\beta$ -Aktin

Bezeichnung	Orientierung	Sequenz (5'→3')
bAct_for	<i>sense</i>	GGATTCCTATGTGGGCG
bAct_rev	<i>antisense</i>	GGCGTACAGGGATAGC

### 2.1.12.4 *INSIG1*-spezifische Primer

Bezeichnung	Orientierung	Sequenz (5'→3')
INSIG1_for	<i>sense</i>	ATGCCAGATTGCACGACCAC
INSIG1_fl_rev	<i>antisense</i>	TCAATCACTATGGGGCTTTTCAGG
INSIG1_DFLY_\$_r	<i>antisense</i>	CTAATAGAGGAAATCTGGGAATGTATAC
INSIG1_DFL_\$_r	<i>antisense</i>	CTAGAGGAAATCTGGGAATGTATACTG
INSIG1_DF_\$_r	<i>antisense</i>	CTAGAAATCTGGGAATGTATACTGATAG
INSIG1_D_\$_r	<i>antisense</i>	CTAATCTGGGAATGTATACTGATAGAC

### 2.1.12.5 *MMS22L*-spezifische Primer

#### Allgemeine Primer

Bezeichnung	Orientierung	Sequenz (5'→3')
MMS22L_for	<i>sense</i>	ATGGAGAAGTGTCTGCTGC
MMS22L_rev	<i>antisense</i>	TTAAGTATTATCATTTCCAGTCTCTGC
MMS22Lmut_SFFS_\$_r	<i>antisense</i>	CTAACTGAAGAACTATTCAGGTTCTTAC
MMS22Lwt_SFFS_\$_r	<i>antisense</i>	CTAACTGAAGAACTATTCAGGTTCTTAC

wird fortgesetzt...

## 2.1 Material

Bezeichnung	Orientierung	Sequenz (5'→3')
MMS22Lmut_SF\$\$_r	<i>antisense</i>	CTAGAAGAACTATTCAGGTTCTTACT
MMS22Lwt_SSF\$\$_r	<i>antisense</i>	CTAGAAGGAACATTCAGGTTCTTACT
MMS22Lmut_SF\$\$_r	<i>antisense</i>	CTAGAACTATTCAGGTTCTTACTATAATAT
MMS22L_c.1731_for	<i>sense</i>	CCTCTTGATGTATGCCCAGA
MMS22L_c.1750_rev	<i>antisense</i>	TCTGGGCATACATCAAGAGG
MMS22L_c.2507_rev	<i>antisense</i>	AGCAAATCATCAGGCCAGAGAG
MMS22L-screen_rev	<i>antisense</i>	GTTTGGCTCCCAGAAATCACAAAGTG
MMS22L_c.3145_rev	<i>antisense</i>	CAGGATGTTGTCCAAGATCTG

### Gateway-Primer

Bezeichnung (Orientierung)	Sequenz (5'→3')
MMS22L_for_attB1 ( <i>sense</i> )	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGAGAAGCTTCTGCTGCA TC
MMS22L_fl_rev_attB2 ( <i>antisense</i> )	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAGTATTATCATTTTCCAGTCTC TGC

### 2.1.12.6 PRDM10-spezifische Primer

#### Gateway-Primer

Bezeichnung (Orientierung)	Sequenz (5'→3')
PRDM10_for_opt_attB1 ( <i>sense</i> )	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGATTGAAAGATGAAAGC TC
PRDM10_fl_rev_attB2 ( <i>antisense</i> )	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATGGTTTGGTGATATGCACTTC
PRDM10_M845Start_attB1* ( <i>sense</i> )	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGTCCAGCACATTCGAAAG
PRDM10_NFF\$\$_attB2_rev ( <i>antisense</i> )	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGAAGAAATTCCAAGCACTGG
PRDM10_NF\$\$_attB2_rev ( <i>antisense</i> )	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGAAATTCCAAGCACTGGGCAG
PRDM10_N\$\$_attB2_rev ( <i>antisense</i> )	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAATTCCAAGCACTGGGCAG

\* genutzt für die Amplifizierung eines internen Leserahmens beginnend mit dem für Methionin (M) kodierenden Codon 845 der PRDM10-Isoform 2

### 2.1.13 Synthetische Peptide

Zur Überprüfung der Immunogenität der Peptidkandidaten wurden die vorhergesagten Okta-, Nona- und Dekamere von der Firma Peptide 2.0 (Chantilly, Virginia, USA) mit einer Reinheit von 20-80% synthetisiert (3-4,5 mg je Peptid). Die lyophilisierten Peptide wurden in DMSO gelöst (Stockkonzentration: 40 mg/ml) und bei -20 °C gelagert. Die Gebrauchslösungen wurden auf 200 µg/ml in PBS eingestellt und bei -80 °C gelagert. Schwer-



lösliche Peptide wurden bei 37 °C und 500 rpm in Lösung gebracht. In den durchgeführten Peptiderkennungstests wurden die Peptide in einer Endkonzentration zwischen 1 µg/ml und 10 µg/ml eingesetzt (siehe Abschnitt 2.2.1.14.1 und Abschnitt 2.2.1.14.2).

**Tab. 2.1:** Liste der 174 auf Immunogenität getesteten synthetischen Peptide. Die Übersicht zeigt die nach Peptidnummer (#) sortierten Peptidsequenzen unter Angabe des jeweils mutierten Gens und dem zugrundeliegenden Aminosäure (AS)-Austausch im genannten Transkript.

#	Peptid	Gen-symbol	Transkript ID (RefSeq)	AS-Aus-tausch	#	Peptid	Gen-symbol	Transkript ID (RefSeq)	AS-Aus-tausch
1	STAPISSVPY	RAPGEF1	NM_005312	P500S	88	LGYSYVKL	RFC1	NM_002913	E673K
2	DYHAFLILI	PTPLB	NM_198402	Y215H	89	LSLNYLVL	LRRC58	NM_001099678	P189L
3	EYYSKNLNSF	MMS22L	NM_198468	S437F	90	MAAGSSCGHI	WDR36	NM_139281	P310S
4	FFSISWLPF	MMS22L	NM_198468	S437F	91	MAVPPSAPQQ	ETNK2	NM_018208	P10Q
5	HYLRENKEKF	OTUD4	NM_001102653	R11K	92	MLLQALPL	C19orf6	NM_001033026	P468L
6	IHSQISSIF	INSIG1	NM_198337	P141S	93	MVSHGFSPSL	TRIM56	NM_030961	S6F
7	IYQRANDDL	RRAGC	NM_022157	H196Y	94	RAAGNSGSL	MDC1	NM_014641	P1730S
8	KFVEKCSLI	MCM5	NM_006739	P399L	95	RACLSLHEL	STON1-GTF2A1L	NM_001198594	Q607H
9	LWNMSDDKLF	ATF2	NM_001256094	P24L	96	RAKPSAPVVL	SIRPB1	NM_006065	S153L
10	LWVGTWVWGF	SLC04A1	NM_016354	A309T	97	TAAGIHSQI	INSIG1	NM_198337	P141S
11	NFSFDYHAF	PTPLB	NM_198402	Y215H	98	TVLDQQQTL	LRPPRC	NM_133259	P1137L
12	QYTFPDFLYI	INSIG1	NM_005542	S238F	99	VAGVPSQNNL	ZHX1	NM_001017926	I426L
13	RFHLGFETL	ZNF804A	NM_194250	P821L	100	WAAQHGHV	DAPK1	NM_004938	R421Q
14	RYPQSSLSRL	C11orf41	NM_012194	P1776L	101	FHLGFETL	ZNF804A	NM_194250	P821L
15	SFDYHAFLI	PTPLB	NM_198402	Y215H	102	ARKGRISV	PKP2	NM_001005242	S70I
16	SFFSISWLPF	MMS22L	NM_198468	S437F	103	FHLGFETLEL	ZNF804A	NM_194250	P821L
17	SFLEHVEKF	IQCH	NM_001031715	L301F	104	FQSPNPSTHV	TET2	NM_017628	S460F
18	SYLSHSSADI	TCF4	NM_001243235	P106L	105	GRFHLGFETL	ZNF804A	NM_194250	P821L
19	SYTPYGWLQL	HERPUD1	NM_001010990	G136S	106	LHPHERLSYL	TCF4	NM_001243235	P106L
20	TRLHWFLLF	CCNL1	NM_020307	P265L	107	PPERLSYL	TCF4	NM_001243235	P106L
21	TYLPSAWNF	PRDM10	NM_199437	S1050F	108	PRSRYPQSSL	C11orf41	NM_012194	P1776L
22	TYLPSAWNFF	PRDM10	NM_199437	S1050F	109	QHGHVDTL	DAPK1	NM_004938	R421Q
23	VYNGVYQYTF	INSIG1	NM_005542	S238F	110	SRYPQSSL	C11orf41	NM_012194	P1776L
24	VYQYTFPDF	INSIG1	NM_005542	S238F	111	SRYPQSSLS	C11orf41	NM_012194	P1776L
25	VYQYTFPDFL	INSIG1	NM_005542	S238F	112	SRYPQSSLSR	C11orf41	NM_012194	P1776L
26	YYSKNLNSF	MMS22L	NM_198468	S437F	113	FIDGKVLVY	PABPC5	NM_080832	S263F
27	YYSKNLNSFF	MMS22L	NM_198468	S437F	114	PSSYYSIFNY	TMEM171	NM_001161342	P252S
28	YLPSAWNFF	PRDM10	NM_199437	S1050F	115	SIEVGHTFY	PARS2	NM_152268	G304S
29	HPHERLSYL	TCF4	NM_001243235	P106L	116	WSAKKYPGDY	FAM213B	NM_001195737	S192Y
30	LARKGRISV	PKP2	NM_001005242	S70I	117	LTQADCLALY	SNRNP48	NM_152551	R161C
31	VAKKYHEVL	PDGFC	NM_016205	T302A	118	ELRKHLSL	LIMK2	NM_001031801	N609S
32	RRRKRGRFHL	ZNF804A	NM_194250	P821L	119	EMLLQALPL	C19orf6	NM_001033026	P468L
33	AQKEGTSQAL	CEP85	NM_022778	E607K	120	ENKEKFEAF	OTUD4	NM_001102653	R11K
34	DLQEDVDACY	CNBP	NM_001127194	E67D	121	FIRKLVNPL	MRPL36	NM_032479	M9L
35	FQFSSVLNSH	ELK4	NM_001973	P401S	122	GPLTKTKSI	PARS2	NM_152268	G304S
36	GISQEFSY	AHR	NM_001621	P686S	123	LLKFVEKCSL	MCM5	NM_006739	P399L
37	GLGPGFSSY	HERPUD1	NM_014685	G161S	124	LPLKFSKTL	MYO1E	NM_004998	N875K
38	ILFIPSGWIY	PHF2	NM_005392	T312I	125	MPRANVCVV	DIP2A	NM_001146114	G661R
39	ITKADETSTY	TRAF7	NM_032271	G49E	126	NLEAKSTSA	C1orf9	NM_016227	P287L
40	KLVNPLLY	MRPL36	NM_032479	M9L	127	RERDRHSPV	ZNF608	NM_020747	F1455V
41	KSIEVGHTF	PARS2	NM_152268	G304S	128	SLKCRAPKV	AP1B1	NM_001166019	E904K
42	KSIEVGHTFY	PARS2	NM_152268	G304S	129	SRSARVKPL	RAD51AP1	NM_006479	L322S
43	KVSQHVIYQAY	AP1B1	NM_001127	E934K	130	TLKIQLQSCL	SATB2	NM_001172509	S160L
44	LAWGGIFS	IQCH	NM_001031715	L301F	131	FLNMRGGLSL	UQCRC2	NM_003366	G245S
45	LLFFGGQF	TRMT11	NM_001031712	L32F	132	MTFNNETF	RPA1	NM_002945	S285F
46	LLFFGGQF	TRMT11	NM_001031712	L32F	133	TSFQIQEF	SYNE1	NM_182961	L3104F
47	LLPNENPSSY	TMEM171	NM_001161342	P252S	134	SFSDYHAFLI	PTPLB	NM_198402	Y215H
48	LLPTSSTSSL	EP400	NM_015409	P661L	135	AHDPHSGHFV	CDK4	NM_000075	R24H
49	LQEDDACY	CNBP	NM_001127195	E66D	136	ARRSGGALYI	KANK1	NM_153186	E186A
50	LQEDVDACY	CNBP	NM_001127194	E67D	137	CRAPKVSQHV	AP1B1	NM_001127	E934K

wird fortgesetzt...

#	Peptid	Gen-symbol	Transkript ID (RefSeq)	AS-Aus-tausch	#	Peptid	Gen-symbol	Transkript ID (RefSeq)	AS-Aus-tausch
51	LQKSQVPA	ZHX1	NM_001017926	I426L	138	FRDTNGHCWV	PTP4A2	NM_001195101	C134W
52	LQKSQVPA	ZHX1	NM_001017926	I426L	139	FSFDYHAF	PTPLB	NM_198402	Y215H
53	LSLSGAKANY	UQCRC2	NM_003366	G245S	140	GHEDEEEPQL	HERC1	NM_003922	S2806L
54	LVWSASDY	RANBP2	NM_006267	T3006S	141	HRDVKSSNIL	MAP2K1	NM_002755	P193S
55	NISEVSTSF	SYNE1	NM_182961	L3104F	142	IHSQISSI	INSIG1	NM_198337	P141S
56	NMRGGLSL	UQCRC2	NM_003366	G245S	143	IHSQISSIFV	INSIG1	NM_198337	P141S
57	QQRASFHL	ETNK2	NM_018208	P10Q	144	LRDQELVQIV	BRWD1	NM_033656	K1006Q
58	RIYTGEKPY	ZNF470	NM_001001668	H643Y	145	MHRDVKSSNI	MAP2K1	NM_002755	P193S
59	RLHWFLLF	CCNL1	NM_020307	P265L	146	MTFNNETFVM	RPA1	NM_002945	S285F
60	RQRAAGNSG	MDC1	NM_014641	P1730S	147	QRAAGNSGSL	MDC1	NM_014641	P1730S
61	SLAWGGIFS	IQCH	NM_001031715	L301F	148	RHSPVQQRHL	ZNF608	NM_020747	F1455V
62	SLSGAKANY	UQCRC2	NM_003366	G245S	149	RRSGGALYI	KANK1	NM_153186	E186A
63	SQISSIFVL	INSIG1	NM_198337	P141S	150	SETPANSLLL	SEZ6L2	NM_001114100	P335S
64	SSVPYAPF	RAPGEF1	NM_005312	P500S	151	SHGFSPSL	TRIM56	NM_030961	S6F
65	SSYYSIFNY	TMEM171	NM_001161342	P252S	152	SRARRSGGAL	KANK1	NM_153186	E186A
66	TLFQFSSVL	ELK4	NM_001973	P401S	153	SRVLPSRL	ARHGAP23	NM_001199417	P278L
67	TQADCLAL	SNRNP48	NM_152551	R161C	154	TRLHWFL	CCNL1	NM_020307	P265L
68	TQADCLALY	SNRNP48	NM_152551	R161C	155	TSFQKIQEFL	SYNE1	NM_182961	L3104F
69	VQIVGIRY	BRWD1	NM_033656	K1006Q	156	YHAFLILI	PTPLB	NM_198402	Y215H
70	VVRRPRAAF	METTL22	NM_024109	S132F	157	YHAFLILIMI	PTPLB	NM_198402	Y215H
71	YQYTFPDF	INSIG1	NM_005542	S238F	158	ETLKRPTSL	HMHA1	NM_012292	G60E
72	YQYTFPDFLY	INSIG1	NM_005542	S238F	159	KAFMRFSSL	ADAMTS9	NM_182920	S1452L
73	YSKNLNSFF	MMS22L	NM_198468	S437F	160	NIRINRCAL	PC	NM_000920	G371R
74	YTFPDFLY	INSIG1	NM_005542	S238F	161	HHASNPNNPL	SOBP	NM_018013	S424N
75	FSFDYHAF	PTPLB	NM_198402	Y215H	162	IRINRCAL	PC	NM_000920	G371R
76	ISSVPYAPF	RAPGEF1	NM_005312	P500S	163	MRFSSLTRHM	ZNF791	NM_153358	L112F
77	AAGNSGSL	MDC1	NM_014641	P1730S	164	AAWSTGPWSL	ADAMTS9	NM_182920	S1452L
78	AAQHGHVDTL	DAPK1	NM_004938	R421Q	165	DAVFPEPSL	HMHA1	NM_012292	G60E
79	EASPGTADM	BEND5	NM_024603	E168A	166	FSSLTRHM	ZNF791	NM_153358	L112F
80	FDYHAFLI	PTPLB	NM_198402	Y215H	167	FSVTDVAVSF	HSHP1	NM_006644	P399S
81	FPEPGRSVL	ARHGAP23	NM_001199417	P278L	168	GVYDTPPM	BCAR1	NM_001170720	E117G
82	FSSYTPYGWL	HERPUD1	NM_001010990	G136S	169	ISANKEDL	C5	NM_001735	E1436K
83	FSYKSEMDSM	AHR	NM_001621	P686S	170	KAFSYFKSF	IRS2	NM_003749	G1057D
84	GAVKVKEEPM	HDAC9	NM_001204147	V475M	171	LSLDIQCEEL	RNH1	NM_002939	Q6L
85	ISVGNGNL	PKP2	NM_001005242	S70I	172	MSLDILSL	RNH1	NM_002939	Q6L
86	LAWGGIFSFL	IQCH	NM_001031715	L301F	173	SQYQGQVY	BCAR1	NM_001170720	E117G
87	LEAKSTSAIL	C1orf9	NM_016227	P287L	174	WSTGPWSL	ADAMTS9	NM_182920	S1452L

### 2.1.14 Plasmidvektoren

Der im Zusammenhang mit dem *pcDNA3.1/V5-His TOPO TA Cloning Kit*<sup>®</sup> (Invitrogen, Leek, Niederlande) verwendete Expressionsvektor *pcDNA<sup>TM</sup>3.1/V5-His TOPO*<sup>®</sup> ist eine Weiterentwicklung des Vektors *pcDNA3* (Invitrogen, Leek, Niederlande), der für die konstitutive Expression heterologer DNA in eukaryontischen Zellen entwickelt wurde. Aufgrund der kovalent an den Vektor *pcDNA<sup>TM</sup>3.1/V5-His TOPO*<sup>®</sup> gebundene Topoisomerase I aus dem *Vaccinia*-Virus wird die Ligation des Inserts in den Vektor ohne die Zugabe einer Ligase vermittelt. Weiterhin beinhaltet der Vektor Eigenschaften, die für die Amplifikation in Bakterien bzw. für die Expression in eukaryontischen Zellen benötigt werden. So induziert der pUC-Replikationsursprung eine hohe Plasmidkopienzahl in Bakterien,

welche aufgrund des im Vektor kodierten  $\beta$ -Lactamasegens mit Ampicillin selektioniert werden können. Die Expression des Transgens in Eukaryonten wird durch den CMV (Cytomegalievirus)-Promotor kontrolliert. Durch die Polyadenylierungssequenz aus dem bovinen Wachstumshormon (BGH, engl. *bovine growth hormone*) wird die Expression von Proteinen in eukaryontischen Zellen gewährleistet. Der SV40-Replikationsursprung ermöglicht die episomale Replikation des Plasmids in Zellen, die das große T-Antigen des Virus SV40 (engl. *simian vacuolating virus 40*) stabil exprimieren. Hieraus resultiert eine besonders starke Expression, die nach ca. 72 Stunden zum Absterben der transfizierten Zellen führt. Das Neomycinresistenzgen erlaubt die Selektion stabiler Transfektanten mit G418.

Für die Klonierung nach dem *Gateway*-Verfahren wurden die Vektoren *pDONR<sup>TM</sup>221* (Invitrogen, Leek, Niederlande) und *pcDNA3.1 DEST #6* (Invitrogen, modifiziert durch C. Wölfel) verwendet. Für den ersten Rekombinationsschritt der *Gateway*-Klonierung wird *pDONR<sup>TM</sup>221* benötigt. Dieser Vektor verfügt ebenso über einen pUC-Replikationsursprung für die Amplifikation in *Escherichia coli* (*E. coli*). Außerdem enthält der Vektor die zwei Rekombinationsstellen, *attP1* und *attP2*, die das *ccdB*-Gen flankieren, welches bei der Rekombination gegen das zu klonierende Insert ausgetauscht wird. Das Kanamycinresistenzgen erlaubt die Selektion rekombinanter Bakterien. Beim zweiten Rekombinationsschritt wird das Insert aus *pDONR<sup>TM</sup>221* in einen Destination-Vektor, beispielsweise *pcDNA3.1 DEST #6*, transferiert. Der Vektor *pcDNA3.1 DEST #6* verfügt über die zuvor beschriebenen Eigenschaften des *pcDNA3* und wurde von C. Wölfel durch Einfügen der *RfA*-Kassette mit den *attR1/2*-Rekombinationsstellen kompatibel für die *Gateway*-Klonierung gemacht.

### 2.1.15 Bakterien

Für die Transformation rekombinater Plasmid-DNA wurden chemisch kompetente *E. coli* eingesetzt. Bei der TOPO-TA-Klonierung wurden die im *cDNA3.1/V5-His TOPO TA Cloning Kit<sup>®</sup>* mitgelieferten One Shot<sup>®</sup> TOP10 (Invitrogen, Leek, Niederlande) verwendet, welche besonders geeignet sind für eine effiziente Klonierung und Plasmidamplifikation mit hoher Kopienzahl. Die Transformationseffizienz liegt bei  $1 \times 10^9$  cfu/ $\mu$ g *supercoiled* DNA (cfu, engl. *colony forming units*). Der Stamm verfügt über folgenden Genotyp:

$$F^- \text{ mcrA } \Delta(\text{mrr-hsdRMS-mcrBC}) \phi 80\text{lacZ}\Delta\text{M15 } \Delta\text{lacX74 } \text{recA1 } \text{araD139} \\ \Delta(\text{ara-leu})7697 \text{ galU } \text{galK } \text{rpsL (Str}^R\text{)} \text{ endA1 } \text{nupG}$$

Für die durchgeführten *Gateway*-Klonierungen wurden NEB 10-beta Competent *E. coli* (NEB, Frankfurt a. M.) verwendet, welche laut den Angaben des Herstellers eine Alternative zu den One Shot<sup>®</sup> TOP10 darstellen und eine Transformationseffizienz von 1-3 cfu/ $\mu$ g *pUC19*-DNA aufweisen. Der Genotyp dieser Bakterien lautet:

$$\Delta(\text{ara-leu})7697 \text{ araD139 } \text{fhuA } \Delta\text{lacX74 } \text{galK16 } \text{galE15 } \text{e14- } \phi 80\text{dlacZ}\Delta\text{M15} \\ \text{recA1 } \text{relA1 } \text{endA1 } \text{nupG } \text{rpsL (Str}^R\text{)} \text{ rph } \text{spoT1 } \Delta(\text{mrr-hsdRMS-mcrBC})$$

## 2.1.16 Zelllinien

### 2.1.16.1 Zelllinien aus dem Melanommodell Ma-Mel-86

Die Melanomzelllinien Ma-Mel-86A, -86B, -86C und -86F wurden aus vier verschiedenen Lymphknotenmetastasen der Patientin INTH generiert und freundlicherweise von D. Schadendorf und A. Paschen (Klinik für Dermatologie, Universitätsklinikum Essen) zur Verfügung gestellt. Die Charakteristika der vier Tumorzelllinien sind in Abschnitt 1.1.1.3.2 erläutert. Die autologen INTH-EBV-B-Zellen entstanden durch die Infektion von B-Zellen der Patientin mit dem Epstein-Barr-Virus (EBV).

### 2.1.16.2 COS-7-Zellen

COS-7 (CV-1 origin SV40)-7 ist eine adhärente Zelllinie mit fibroblastenähnlicher Morphologie, die aus CV-1-Nierenzellen eines adulten Männchens der Meerkatzenart *Cercopithecus aethiops* (Grüne Meerkatze) generiert wurde (Gluzman, 1981). Aufgrund der Transformation mit dem defekten Genom des SV40-Virus, in dem der DNA-Replikationsursprung deletiert war, exprimieren COS-7-Zellen das große SV40-T-Antigen konstitutiv und ermöglichen dadurch die episomale Replikation von zirkulären Plasmiden, die einen SV40-Replikationsursprung tragen. Die hier verwendeten Expressionsvektoren *pcDNA<sup>TM</sup>3.1/V5-His TOPO<sup>®</sup>* und *pcDNA3.1 DEST #6* verfügen über einen SV40 *ori*.

### 2.1.16.3 293T-Zellen

Bei der 293-Zelllinie handelt es sich um eine semi-adhärente Zelllinie aus humanen embryonalen Nierenzellen. Die Immortalisierung der Zellen erfolgte durch die Transformation mit gescherter DNA des Adenovirus Typ5 (Graham *et al.*, 1977). Die Zelllinie 293T wurde zusätzlich stabil mit dem großen T-Antigen des SV40-Virus transfiziert und wurde freundlicherweise von N. Shastri (The University of California, Berkeley, USA) zur Verfügung gestellt. Diese Zellen eignen sich alternativ zu COS-7-Zellen für die transiente Transfektion mit Plasmiden, die über einen den SV40-Replikationsursprung verfügen.

### 2.1.16.4 K562-Zellen

Die Suspensionszelllinie K562 wurde abgeleitet von einer Patientin mit chronischer myeloischer Leukämie (CML) in der terminalen Blastenkrise. Die Zellen exprimieren keine HLA-Moleküle (Drew *et al.*, 1977) und dienen daher in der Regel als Zielzellen für die Untersuchung natürlicher Killerzellen. In der vorliegenden Arbeit wurde die Möglichkeit HLA-transfizierte K562-Zellen als peptidpräsentierende Zellen zu verwenden überprüft.

### 2.1.17 Bakteriennährmedien

Die Kultivierung von *E. coli* erfolgte in Suspension mit LB-Medium bzw. adhärent auf LB-Agarplatten. Für die Kryokonservierung der Bakterien wurde ein glycerolhaltiges Einfriermedium verwendet.

Medium	Zusätze
<b>LB-Medium</b> (11)	10,0 g Trypton 5,0 g Hefeextrakt 10,0 g NaCl ad 1,01 dH <sub>2</sub> O
<b>LB-Agar</b> (500 ml)	10,0 g Trypton 5,0 g Hefeextrakt 10,0 g NaCl 7,5 g Bacto-Agar ad 1,01 dH <sub>2</sub> O
<b>Einfriermedium</b> (60% Glycerol, 50 ml)	30,0 ml Glycerol ad 50 ml LB-Medium

Das LB-Medium und der LB-Agar wurden in Duranglasflaschen angesetzt und für 30 min bei 121 °C autoklaviert. Die Zugabe von Antibiotika erfolgte nach Abkühlen des Mediums auf ca. 40 °C. Für die Herstellung von LB-Agarplatten wurde der LB-Agar nach der Zugabe des Selektionsantibiotikums in Petrischalen (Ø 100 mm) gegossen. Die Polymerisierung erfolgte in der Sterilbank bei geöffnetem Deckel. Die Flüssigmedien und die Agarplatten wurden bei 4 °C gelagert.

### 2.1.18 Zellkulturmedien

Für die Kultivierung von Melanomzellen und weiteren eukaryontischen Zelllinien sowie für die T-Zellkultur wurden die gelisteten Kultur- und Einfriermedien verwendet. Bei der Transfektion der Zelllinien wurden die Medien ohne Zusätze eingesetzt. Alle Medien wurden bei 4 °C gelagert.

Medium	Zusätze	Verwendung
<b>RPMI<sub>komplett</sub></b> (550 ml)	500 ml RPMI 1640 + L-Glutamin 50 ml FCS* 500 U/ml Penicillin 500 µg/ml Streptomycin	Kultivierung von COS-7-, 293T-, K562-, Melanom- und EBV-B-Zelllinien
<b>RPMI<sub>FCS</sub></b> (550 ml)	500 ml RPMI 1640 + L-Glutamin 50 ml FCS*	Lipofektion
<b>AIM-V<sub>komplett</sub></b> (1, 11)	11 AIM-V + L-Glutamin 100 ml Humanserum **	Kultivierung von T-Zellen
<b>OptiMEM</b> (500 ml)	500 ml OptiMEM ohne Zusätze	Elektroporation
<b>Einfriermedium</b> (10% DMSO, 50 ml)	45 ml FCS* 5 ml DMSO	Kryokonservierung eukaryontischer Zellen

\* Hitzeinaktiviert für 45 min bei 56 °C

\*\* Das Humanserum stammte aus Vollblutspenden von 10-20 Freiwilligen. Die Blutspende, die Aufarbeitung des Serums sowie der Auftrag für die jeweilige Infektionsserologie des Blutes bzw. des Serums erfolgten durch die Transfusionszentrale der Universitätsmedizin Mainz. Die zusammengeführten Seren wurden für 45 min bei 56 °C hitzeinaktiviert, anschließend unter sterilen Bedingungen in 200 ml Portionen aliquotiert und bei -80 °C gelagert.

## 2.1.19 Referenzen und Datenbanken

Datenbank/Referenz	Spezifikation	Quelle/URL
Database for Short Genetic Variations (dbSNP)	<i>Build 137</i>	ftp: //ftp.ncbi.nih.gov/snp/organisms/ human_9606/VCF/00-All.vcf.gz (heruntergeladen am 31.01.13)
NCBI Reference Sequence Database (RefSeq)	<i>Release 58</i>	http: //www.ncbi.nlm.nih.gov/refseq/
Humanes Referenzgenom hg19	Februar 2009	ftp://igenome: G3nom3s4u@ussd-ftp.illumina.com/ Homo_sapiens/UCSC/hg19/Homo_ sapiens_UCSC_hg19.tar.gz (heruntergeladen am 04.01.2013)
TruSeq Exome Enrichment Kit Targeted Regions	18.03.2011	http://support.illumina.com/ downloads/truseq_exome_targeted_ regions_bed_file.ilmm (heruntergeladen am 04.05.2013)
Sequence Read Archive (SRA)	-	http://www.ncbi.nlm.nih.gov/sra
UniProtKB	-	http://www.uniprot.org/
GeneCards	-	http://www.genecards.org/
UCSC Table Browser	-	https://genome.ucsc.edu/cgi-bin/ hgTables

## 2.1.20 Software und Tools

Folgende Programme und Tools wurden unter Linux (Ubuntu 12.10) verwendet:

Name	Version	Entwickler	Verwendung
FastQC	0.10.1	S. Andrews, Babraham Institute, Babraham, Vereinigtes Königreich <sup>a</sup>	Qualitätsanalyse von FASTQ-Dateien
FASTX-toolkit	0.0.13.2	A. Gordon, Cold Spring Harbor Laboratory, Cold Spring Harbor, USA <sup>b</sup>	Qualitätsanalyse von FASTQ-Dateien
Trimmomatic	v.0.22	Bolger <i>et al.</i> (2014)	Nachbearbeitung der FASTQ-Rohdaten
TopHat	v2.0.3	Trapnell <i>et al.</i> (2012)	<i>Mapping</i> an Referenz (Transkriptom)
Bowtie2	2.0.0-beta6	Langmead & Salzberg (2012)	<i>Mapping</i> an Referenz (Exome)
IGV - Integrative Genomics Viewer	v2.3	Thorvaldsdóttir <i>et al.</i> (2013)	Visualisierung der <i>Alignments</i> und der detektierten Mutationen
SAMtools	0.1.18 (r982:295)	Li <i>et al.</i> (2009)	Prozessierung von SAM/BAM-Dateien
bedtools	v2.16.1	Quinlan & Hall (2010)	<i>Coverage</i> -Analyse der BAM-Dateien
BCFtools	0.1.17-dev	Li <i>et al.</i> (2009)	Prozessierung von VCF/bcf-Dateien
VCFtools	v0.1.9.0	Danecek <i>et al.</i> (2011)	Prozessierung von VCF-Dateien
snpEff	3.1m	Cingolani <i>et al.</i> (2012)	Annotation des Effekts von Mutationen
R	2.15.1	<i>R Core Team, R Foundation for Statistical Computing, Wien, Österreich</i> <sup>c</sup>	Anwendung von Textfiltern, Expressionsanalysen
Perl	v5.14.2		Programmiersprache eigener Programme
Circos	0.67-pre5	Krzywinski <i>et al.</i> (2009)	Erstellen von <i>Circos</i> -Plots

<sup>a</sup> <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

<sup>b</sup> [http://hannonlab.cshl.edu/fastx\\_toolkit/download.html](http://hannonlab.cshl.edu/fastx_toolkit/download.html)

<sup>c</sup> <http://www.R-project.org>

Außerdem wurden die folgenden Programme unter Windows 7 eingesetzt:

Name	Version	Entwickler	Verwendung
FlowJo	v7.6.5	FlowJo, LLC, Ashland (USA)	Auswertung von Durchflusszytometriedaten
GraphPad Prism	6	GraphPad Software, Inc., La Jolla (USA)	Erstellen von Graphen
IGV - Integrative Genomics Viewer	Java Web Start (v2.3)	Thorvaldsdóttir <i>et al.</i> (2013)	Visualisierung der <i>Alignments</i> und der detektierten Mutationen
MEGA	5.1	Tamura <i>et al.</i> (2011)	Phylogenetische Analysen
MS Excel	2010	Microsoft, Redmond (USA)	Tabellenkalkulation

## 2.2 Methoden

### 2.2.1 Zellbiologische Methoden

#### 2.2.1.1 Zellkultur

Die Kultivierung von eukaryontischen Zellen wurde unter sterilen Bedingungen an der Sterilwerkbank nach Desinfizieren der Arbeitsfläche mit 70% Ethanol durchgeführt. Die Zellen wurden in Kulturflaschen bzw. 96- bis 24-Well-Platten in den jeweiligen Kulturmedien im Brutschrank bei 37 °C, einem CO<sub>2</sub>-Partialdruck von 5% und 95% Luftfeuchtigkeit kultiviert. Die verwendeten Medien sind in Abschnitt 2.1.18 gelistet. Die Subkultivierung der Suspensionskulturen wurde zweimal wöchentlich durchgeführt. Adhärenz Zelllinien wurden bei einer Konfluenz von ca. 70-90% passagiert. Das Ablösen adhärenter Zellen erfolgte enzymatisch mit 5% Trypsin-EDTA (Sigma) oder mechanisch durch Abklopfen.

#### 2.2.1.2 Bestimmung von Zellzahl und Viabilität

Die Bestimmung der Anzahl vitaler Zellen erfolgte mithilfe einer Fuchs-Rosenthal-Zählkammer unter dem Phasenkontrastmikroskop. Zur Unterscheidung zwischen lebenden und toten Zellen wurde der Farbstoff Trypanblau verwendet, welcher tote bzw. geschädigte Zellen blau färbt, da der Farbstoff die durchlässig gewordene Membran dieser Zellen passieren kann. Lebende Zellen bleiben ungefärbt. Je nach Zelldichte wurde die Zellsuspension 1:2 bis 1:20 mit der Trypanblau-Gebrauchslösung (siehe Abschnitt 2.1.11) gemischt, so dass 50 bis 150 lebende Zellen pro Großquadrat zu zählen waren. Zur Berechnung der Zelldichte (Zellen/ml) wurde folgende Formel verwendet:

$$\text{Zelldichte [Zellen/ml]} = \frac{\text{Anzahl der gezählten lebenden Zellen pro Großquadrat}}{\text{ausgezählte Fläche [1 mm}^2\text{]} * \text{Kammertiefe [0,2 mm]}} \times \text{Verdünnungsfaktor} \times 10^4$$

### 2.2.1.3 Kryokonservierung und Auftauen von eukaryontischen Zellen

Die Kryokonservierung eukaryontischer Zellen erfolgte in je 1 ml Einfriermedium (siehe Abschnitt 2.1.18) in Kryoröhrchen der Firma Greiner. Zum schonenden Einfrieren ( $-1^{\circ}\text{C}/\text{min}$ ) wurden die Zellen zunächst für mindesten 24 h in Isopropanol-gefüllten Kryocontainern bei  $-80^{\circ}\text{C}$  gelagert und anschließend in Flüssigstickstoff überführt.

Zum Auftauen wurden die Zellen bei  $37^{\circ}\text{C}$  im Wasserbad angetaut, zügig in Medium überführt und nach zweimaligem Waschen mit PBS oder Zellkulturmedium in Kultur genommen.

### 2.2.1.4 Behandlung von Zellen mit $\text{IFN}\gamma$

Die Behandlung adhärenter Zellen mit  $\text{IFN}\gamma$  erfolgte bei einer Konfluenz von ca. 70-80%. Die Zellen wurden mit  $100 \text{ U}/\text{ml}$   $\text{IFN}\gamma$  für vier Tage bei  $37^{\circ}\text{C}$  im Brutschrank (5%  $\text{CO}_2$ , 95% Luftfeuchtigkeit) inkubiert.

### 2.2.1.5 Isolierung peripherer Blutlymphozyten aus Vollblut und Erythrozytenlyse

Für die Isolierung von Lymphozyten wurde eine Dichtegradientenzentrifugation mit Hilfe von Ficoll (Histopaque 1077, Sigma) in 50 ml-Falcon-Röhrchen durchgeführt. Nach Verdünnen des Blutes mit sterilem PBS (1:2) wurden jeweils 15 ml Ficoll mit 35 ml des verdünnten Blutes überschichtet. Die Abtrennung der mononukleären Zellen (T- und B-Lymphozyten sowie Monozyten) als trübe Interphase erfolgte während der Zentrifugation bei 2.500 rpm für 10 min (ohne Bremse). Der sogenannte Lymphozyten-Ring wurde abgenommen und die Zellen wurden mindestens dreimal mit PBS gewaschen.

Im Falle einer starken Verschleppung von Erythrozyten wurde eine Erythrozytenlyse durchgeführt. Dazu wurden die Zellen pelletiert, mit 3-5 ml Erythrozytenlysepuffer (siehe Abschnitt 2.1.11) resuspendiert und für 3 min bei Raumtemperatur inkubiert. Die Lyse wurde durch Zugabe von 20 ml PBS gestoppt und die Zellen wurden gewaschen.

Die isolierten Lymphozyten wurden, gegebenenfalls nach CD8-Isolation (siehe Abschnitt 2.2.1.6), entweder in Kultur genommen oder kryokonserviert (siehe Abschnitt 2.2.1.3).

### 2.2.1.6 Anreicherung von $\text{CD8}^+$ Zellen durch magnetische Zellseparation

$\text{CD8}^+$  T-Zellen wurden mittels magnetischer Zellseparation (MACS, engl. *magnetic activated cell sorting*) mithilfe der magnetischen anti- $\text{CD8-MicroBeads}$  der Firma Miltenyi Biotec über LS-Säulen (Miltenyi Biotec) aus PBMCs bzw. Pan-T-Zellpopulationen isoliert. Abweichend vom Herstellerprotokoll wurden bis zu  $10^9$  Zellen je Ansatz zu  $20 \times 10^7$  Zellen/ml in kaltem MACS-Puffer (siehe Abschnitt 2.1.11) eingestellt und 20  $\mu\text{l}$  anti- $\text{CD8-MicroBeads}$  pro



$10^7$  Zellen wurden hinzugefügt. Die LS-Säule wurde in das Magnetfeld des MACS-Separators eingebracht und mit 3 ml MACS-Puffer vorbereitet. Nach 15-minütiger Inkubation mit den *MicroBeads* bei Raumtemperatur wurden die Zellen mit MACS-Puffer *ad* 3 ml verdünnt und in Aliquots von je 500  $\mu$ l auf die vorbereitete Säule gegeben. Der Durchfluss wurde gesammelt. Nach dreimaligem Spülen der Säule mit je 3 ml MACS-Puffer wurde die Säule aus dem MACS-Separator entfernt und die magnetisch markierten CD8<sup>+</sup> Zellen wurden mit weiteren 3 ml MACS-Puffer durch die Säule gedrückt und in einem zweiten Röhrchen gesammelt. Der Erfolg der Separation wurde anschließend durchflusszytometrisch (vgl. Abschnitt 2.2.1.11) überprüft.

### 2.2.1.7 Generierung tumorspezifischer T-Lymphozyten in MLTCs

Kryokonservierte PBMCs bzw. aus PBMCs isolierte CD8<sup>+</sup> T-Zellen wurden wie in Abschnitt 2.2.1.3 beschrieben aufgetaut. Gegebenenfalls wurden die Zellen 5 min bei Raumtemperatur mit DNase I (1 mg/ml, Roche) inkubiert und anschließend gewaschen, um eine Verklumpung der Zellen aufzulösen. Für die Anreicherung tumorreaktiver CD8<sup>+</sup> T-Zellen in sogenannten gemischten Lymphozyten-Tumorzellkulturen (MLTCs, engl. *mixed lymphocyte tumor cell cultures*) wurden  $1-2 \times 10^6$  PBMCs bzw. CD8<sup>+</sup> T-Zellen pro Kultureinheit in einer 24-Well-Platte in 1 ml AIM-V<sub>komplett</sub> (siehe Abschnitt 2.1.18) ausgesät. Als Stimulatorzellen wurden  $1 \times 10^5$  bestrahlte (10.000 rad bzw. 100 Gy) Tumorzellen in 1 ml AIM-V<sub>komplett</sub> hinzugefügt. An Tag 3 wurde die Hälfte des Mediums durch frisches Medium mit 500 U/ml rhIL-2 (rekombinantes humanes IL-2, Endkonzentration: 250 U/ml) ersetzt. Die MLTCs wurden wöchentlich mit bestrahlten Tumorzellen restimuliert. In MLTC-Responderpopulationen mit einem zu hohen Anteil (>10%) an CD4<sup>+</sup> T-Zellen (durchflusszytometrisch bestimmt) wurden, sobald ausreichend Zellen vorhanden waren, CD8<sup>+</sup> T-Zellen durch CD8-*MicroBeads* isoliert (vgl. Abschnitt 2.2.1.6). Bei MLTCs, die aus vorisolierten CD8<sup>+</sup> T-Zellen generiert wurden, wurde die CD8-negative Zellpopulation ( $2 \times 10^5$  Zellen/Well) teilweise ebenfalls bestrahlt und zusammen mit den bestrahlten Tumorzellen zur Restimulation als *Feeder*-Zellen verwendet.

### 2.2.1.8 Klonierung von Zellen nach dem Grenzverdünnungsverfahren

Die Klonierung von CD8<sup>+</sup> MLTC-Responderpopulationen erfolgte mittels des Grenzverdünnungsverfahrens. T-Zellklone wurden durch wöchentliche Restimulation mit den autologen Melanomzelllinien expandiert. Die Klonierung erfolgte in 96-Well-Platten mit Rundboden. Um klonale T-Zellpopulationen zu erhalten, wurden verschiedene Ausgangsverdünnungen angesetzt und Platten mit je 1 Lymphozyte pro Kultureinheit, 3 Lymphozyten pro Kultureinheit und 0,3 Lymphozyten pro Kultureinheit angelegt. Die CD8<sup>+</sup> Responderpopulation wurde auf  $1 \times 10^6$  Zellen/ml in AIM-V<sub>komplett</sub> eingestellt und folgendermaßen weiterverdünnt:

V1:	100 µl Stammsuspension	+ 900 µl AIM-V <sub>komplett</sub>	(≅10 <sup>4</sup> Zellen/100 µl)
V2:	500 µl V1	+ 4,5 ml AIM-V <sub>komplett</sub>	(≅10 <sup>3</sup> Zellen/100 µl)
V3:	500 µl V2	+ 4,5 ml AIM-V <sub>komplett</sub>	(≅10 <sup>2</sup> Zellen/100 µl)
V4:	3 ml V3	+ 27 ml AIM-V <sub>komplett</sub>	(≅10 Zellen/100 µl)
V5:	11 ml V4	+ 99 ml AIM-V <sub>komplett</sub>	(≅1 Zelle/100 µl)
V6:	15 ml V4	+ 30 ml AIM-V <sub>komplett</sub>	(≅3 Zellen/100 µl)
V7:	3,5 ml V6	+ 31,5 ml AIM-V <sub>komplett</sub>	(≅0,3 Zellen/100 µl)

Fünf bis zehn 96-Well-Rundbodenplatten mit 1 Zelle/Well (100 µl/Well von V5), und bis zu drei Platten mit 3 bzw. 0,3 Zellen/Well (100 µl/Well von V6 bzw. V7) wurden angelegt. Für die Stimulation wurden pro Kultureinheit 3×10<sup>3</sup> Tumorzellen und 5×10<sup>4</sup> allogene EBV-transformierte B-Zellen als *Feeder*-Zellen bestrahlt (10.000 rad bzw. 100 Gy), in 100 µl/Well AIM-V<sub>komplett</sub> mit 500 U/ml rhIL-2 (Endkonzentration: 250 U/ml) aufgenommen und zu den T-Zellen gegeben.

Die Klonierungseffizienz  $z$  wird aus der Anzahl der gesäten ( $N_{Aussaat}$ ) und der Anzahl der tatsächlich gewachsenen Ansätze ( $N_{pos}$ ) berechnet:

$$z = \frac{N_{pos}}{N_{Aussaat}}.$$

Zur Bestimmung der Wahrscheinlichkeit  $P_{clone}$ , dass proliferierende Zellpopulationen aus einer Zelle hervorgingen, wird die Poissonverteilung eingesetzt:

$$P_{clone}(f, x) = \frac{fx \times e^{-fx}}{1 - e^{-fx}},$$

wobei  $f$  die Frequenz der Vorläuferzellen in der Massenkultur und  $x$  die Anzahl ausgesäter Zellen je Kultureinheit ist (Taswell *et al.*, 1980). Die Vorläuferfrequenz  $f$  kann experimentell aus der Anzahl der gesäten Zellen je Kultureinheit  $x$  und dem Anteil negativer Ansätze  $y$  ( $y = 1 - z$ ) geschätzt werden:

$$\ln y = -fx + \ln a,$$

wobei  $y < a \leq 1$  gilt. Da  $f = -\frac{\ln \frac{y}{a}}{x}$ , ist  $f_{max} = -\frac{\ln \frac{1}{a}}{x} = -\frac{\ln a}{x}$ . Mit  $f_{max}$  kann die Mindestwahrscheinlichkeit für Klonalität ermittelt werden, wenn die tatsächliche Vorläuferfrequenz  $f$  unbekannt ist.

Zur Weiterkultivierung und Expansion der T-Zellklone wurde die Anzahl der zur Restimulation eingesetzten Tumor- und *Feeder*-Zellen an die Fläche der Kultureinheiten der verwendeten Platten angepasst (siehe Tab. 2.2).

Tab. 2.2: Verwendete Plattentypen und Zellen für die Kultivierung und Restimulation von T-Lymphozyten.

Plattentyp	Verwendete Stimulator- und <i>Feeder</i> -Zellen je Well
96-Well-Platten, Rund- bzw. Flachboden	3.000 Tumorzellen 30.000-50.000 allogene EBV-transformierte B-Zellen
48-Well-Platten	50.000 Tumorzellen 100.000 allogene EBV-transformierte B-Zellen
24-Well-Platten	100.000 Tumorzellen 200.000 allogene EBV-transformierte B-Zellen

### 2.2.1.9 Unspezifische Expansion von CD8<sup>+</sup> T-Zellklonen mit OKT-3

Der monoklonale Antikörper OKT-3, der an den membranständigen Rezeptor CD3 bindet und dadurch die Aktivierung von T-Zellen induziert, kann zur raschen Expansion von CD8<sup>+</sup> T-Zellklonen eingesetzt werden. Als *Feeder*-Zellen für die Stimulation wurden  $2,5 \times 10^7$  PBMCs aus Leukozytenkonzentraten dreier gesunder Spender (Isolation siehe Abschnitt 2.2.1.5) und  $5 \times 10^6$  allogene EBV-transformierte B-Zellen verwendet. Diese wurden nach Bestrahlung (10.000 rad bzw. 100 Gy) mit  $1-2,5 \times 10^5$  T-Zellen in 25 ml AIM-V<sub>komplett</sub> mit 30 ng/ml OKT-3, 250 U/ml rhIL-2 und 2,5 ng/ml rhIL-15 in 25 cm<sup>2</sup> Kulturf flaschen stehend für 14 Tage im Brutschrank inkubiert. Gegebenenfalls wurde nach etwa einer Woche die Hälfte des Mediums abgenommen durch frisches Medium mit 250 U/ml rhIL-2 ersetzt. Nach 14 Tagen wurde die T-Zellen geerntet und ihre Reaktivität im IFN $\gamma$ -ELISpot-Assay (siehe Abschnitt 2.2.1.14) getestet. Bei zu schwacher Reaktivität wurden die T-Zellen nochmals spezifisch mit dem autologen Melanomzelllinien restimuliert. Andernfalls wurden die T-Zellen direkt kryokonserviert (siehe Abschnitt 2.2.1.3).

### 2.2.1.10 Antikörperblockade der T-Zellen

Zur Charakterisierung von T-Zellantworten im IFN $\gamma$ -ELISpot-Assay (siehe Abschnitt 2.2.1.14) wurden Antikörper eingesetzt, die gegen einzelne HLA-Moleküle, Gruppen von HLA-Molekülen oder gegen das Oberflächenmolekül CD3 gerichtet waren. Dadurch konnte das Restriktionsmolekül bestimmt bzw. die Abhängigkeit der IFN $\gamma$ -Sekretion vom TCR bestätigt werden. Die in dieser Arbeit verwendeten Antikörper wurden aus Hybridomüberständen aufgereinigt und entsprechend der Resultate einer vorherigen Titration verwendet (siehe auch Abschnitt 2.1.8).

Tab. 2.3: Antikörper für Blockadeexperimente.

Antikörper	Spezifität	Verwendete Konzentration
W6/32	HLA-Klasse I-Moleküle	25-50 $\mu$ g/ml
B1.23.2	HLA-B, HLA-C-Allele und HLA-Aw19	100-200 $\mu$ g/ml
C7709.A2.6	HLA-A02 und -A24	100-200 $\mu$ g/ml
OKT-3	CD3	25-50 $\mu$ g/ml

### 2.2.1.11 Durchflusszytometrie

Die Durchflusszytometrie ist ein Verfahren zur Charakterisierung eines Gemisches aus verschiedenen Zellpopulationen. Hierfür werden Antikörper, die gegen unterschiedliche Zielstrukturen gerichtet sind, eingesetzt. Die Antikörper sind entweder direkt mit einem Fluoreszenzfarbstoff gekoppelt oder können durch einen zweiten gegen den konstanten Teil des Erstantikörpers gerichteten Zweitantikörper mit einem Fluoreszenzfarbstoff markiert

werden. Im Durchflusszytometer wird das Passieren des Laserstrahls für jede einzelne Zelle detektiert. Dabei werden zunächst zwei Arten der Streuung des Laserstrahls unterschieden: der *forward scatter*, der Aufschluss über die Größe der Zelle gibt, und der *sideward scatter*, der ein Maß für die Granularität der Zelle ist. Außerdem wird die durch verschiedene Laser angeregte Emission der an die gebundenen Antikörper gekoppelten Fluoreszenzfarbstoffe detektiert.

Für die Untersuchung verschiedener Zellpopulationen wurden pro Ansatz  $0,5-2 \times 10^5$  Zellen in 200  $\mu\text{l}$  RPMI<sub>FCS</sub> (siehe Abschnitt 2.1.18) in FACS (engl. *fluorescence-activated cell sorting*)-Röhrchen vorgelegt. Eine Liste der verwendeten Antikörper ist in Abschnitt 2.1.8 zu finden. Direkt farbstoffgekoppelte Antikörper wurden entsprechend der Herstellerangabe 1:200 bis 1:10 eingesetzt und für 30 min bei 4 °C mit den Zellen inkubiert. Für eine indirekte Färbung durch gekoppelte Zweitantikörper wurden die Zellen zunächst mit dem Primärantikörper für 30 min bei 4 °C inkubiert. Der Primärantikörper wurde ebenfalls nach den Anweisungen des Herstellers eingesetzt. Im Falle von Hybridomüberständen, die den Primärantikörper enthielten, wurden jeweils 100  $\mu\text{l}$  Überstand pro Ansatz verwendet. Nach der Inkubation mit dem Primärantikörper wurden die Zellen mit FACS-Puffer (siehe Abschnitt 2.1.11) gewaschen, um dann für 30 min bei 4 °C mit dem farbstoffgekoppelten Sekundärantikörper (*goat F(ab')<sub>2</sub> Fragment Anti-mouse IgG + IgM (H+L)-FITC*) inkubiert zu werden. Abschließend wurden die Zellen nochmals mit FACS-Puffer gewaschen und in 200-300  $\mu\text{l}$  FACS-Fixans (siehe Abschnitt 2.1.11) aufgenommen. Als Kontrollen und zum Einstellen des Durchflusszytometers wurden jeweils ungefärbte, einzel gefärbte und Isotyp-gefärbte Ansätze verwendet, die gleichermaßen behandelt wurden.

Die Proben wurden mit Hilfe des Durchflusszytometers FACS Canto II der Firma BD mit der Software *FACS Diva* aufgenommen. Für die Auswertung wurde das Programm *FlowJo* (Version v7.6.5, FlowJo, LLC, Ashland, USA) verwendet.

### 2.2.1.12 Transiente Transfektion durch Lipofektion

Bei einer transienten Transfektion wird Fremd-DNA meist in Form eines Plasmids in eukaryontische Zellen eingebracht, ohne dass die eingebrachte DNA in das Genom der transfizierten Zelle integriert, wodurch die Expression des Transgens nur über einen beschränkten Zeitraum erfolgt. Im Falle der Lipofektion ermöglichen Lipidtröpfchen, die mit der Plasmid-DNA komplexieren, die Aufnahme der Fremd-DNA durch Fusion mit der Zellmembran.

Die transiente Transfektion von 293T- bzw. COS-7-Zellen durch Lipofectamin<sup>®</sup>2000 (Invitrogen) mit Plasmiden, die für HLA- bzw. Antigen-cDNA kodierten, erfolgte direkt in den Testeinheiten der ELISpot-Platten (vgl. Abschnitt 2.2.1.14). Hierfür wurden je Testeinheit  $2 \times 10^4$  293T- bzw. COS-7-Zellen in 120  $\mu\text{l}$ /Well RPMI<sub>FCS</sub> (siehe Abschnitt 2.1.18) ausplattiert und für 30 min bei 37 °C inkubiert. Für den Transfektionsansatz wurde das Lipofectamin<sup>®</sup>2000 zunächst vorverdünnt. Je nach zuvor ausgetesteter Effizienz des ver-

wendeten Transfektionsreagenz wurden pro Testeinheit 0,2-1  $\mu\text{l}$  Lipofectamin<sup>®</sup>2000 *ad* 21  $\mu\text{l}$  RPMI (ohne Zusätze) ergänzt und für 5 min bei Raumtemperatur inkubiert. Die zu transfizierende antigenkodierende Plasmid-DNA wurde auf 300  $\text{ng}/5 \mu\text{l}$  RPMI je Ansatz eingestellt und mit der HLA-kodierenden Plasmid-DNA (100  $\text{ng}/7 \mu\text{l}$  RPMI je Ansatz) gemischt. Das vorverdünnte Lipofectamin<sup>®</sup>2000 (21  $\mu\text{l}/\text{Ansatz}$ ) wurde hinzugemischt und zur Bildung des Lipid/DNA-Komplexes für 20 min bei Raumtemperatur inkubiert. Für die Lipofektion wurden 30  $\mu\text{l}$  des Transfektionsansatzes pro Testeinheit zu den ausgesäten 293T- bzw. COS-7-Zellen gegeben und durch Auf- und Abpipettieren gemischt. Nach Inkubation bei 37 °C über Nacht wurden die eingebrachten Transgene exprimiert und der IFN $\gamma$ -ELISpot-Assay durch Zugabe von T-Lymphozyten fortgeführt (siehe Abschnitt 2.2.1.14).

Für die Transfektion mit Lipofectamin<sup>®</sup>3000 (Invitrogen) wurde entsprechend verfahren. Für jeden Transfektionsansatz wurden 0,2  $\mu\text{l}$  Lipofectamine<sup>®</sup>3000 verwendet. Abweichend vom Protokoll für Lipofectamine<sup>®</sup>2000 wurde die Plasmid-DNA vor Zugabe des verdünnten Lipofectamin<sup>®</sup>3000 mit 0,6  $\mu\text{l}$  je Ansatz P3000<sup>TM</sup>-Reagenz versetzt.

### 2.2.1.13 Transiente und stabile Transfektion mittels Elektroporation

Bei der Transfektion durch Elektroporation wird Fremd-DNA dadurch in eukaryontische Zellen eingebracht, dass die Zellmembran der zu transfizierenden Zellen durch einen elektrischen Impuls kurzzeitig permeabel wird. Im Vergleich zur Lipofektion können mehr Zellen pro Ansatz transfiziert werden. Außerdem ist es möglich, durch entsprechende Selektionsmedien stabile Transfektanten zu generieren, bei denen die eingebrachte DNA in das Genom integriert ist. Für die Elektroporation wurden  $3\text{-}5 \times 10^6$  Zellen pro Transfektionsansatz zweimal mit OptiMEM (Gibco) gewaschen und anschließend in je 100  $\mu\text{l}$  OptiMEM je Ansatz aufgenommen. Die Zellen wurden zu 10-15  $\mu\text{g}$  Plasmid-DNA in 0,2 cm-Elektroporationsküvetten gegeben, direkt elektroporiert und anschließend für 5 min bei Raumtemperatur inkubiert. **Tab. 2.4** zeigt die auf die verschiedenen Zelltypen angepassten Elektroporationsprogramme.

**Tab. 2.4:** Übersicht der verwendeten Elektroporationsprogramme. Für die Elektroporation der gelisteten Zelllinien wurden unterschiedliche Programme verwendet, die in der genutzten Spannung und der Dauer des Pulses variierten.

Zelllinie	Spannung	Dauer	Anzahl Pulse
293T	110 V	25 ms	1
COS-7	110 V	20 ms	1
K562	160 V	15 ms	1
Ma-Mel-86C	140 V	25 ms	1

Die Zellen wurden in je 5 ml RPMI<sub>komplett</sub> in 25 cm<sup>2</sup>-Kulturflaschen überführt und über Nacht bei 37 °C, 5% CO<sub>2</sub> und 95% Luftfeuchtigkeit inkubiert, damit die eingebrachte DNA exprimiert wurde. Am nächsten Tag wurden die Zellen geerntet und mit 50.000 Zellen/Well im IFN $\gamma$ -ELISpot-Assay eingesetzt und/oder durchflusszytometrisch mit entsprechenden Antikörpern analysiert, um die Transfektionseffizienz zu bestimmen. Sollten stabile

Transfektanten generiert werden, wurden dem Kulturmedium Selektionsantibiotika (siehe Abschnitt 2.1.4) hinzugefügt und die Zellen wurden weiter in Kultur gehalten. Gegebenenfalls wurde die Massenkultur kloniert, um eine stabil transfizierte klonale Zellpopulation zu erhalten.

### 2.2.1.14 IFN $\gamma$ -ELISpot-Assays

IFN $\gamma$ -ELISpot-Assays (engl. *enzyme-linked immunosorbent spot*, ELISpot) weisen in einer Farbreaktion das von T-Lymphozyten nach Aktivierung freigesetzte Zytokin Interferon- $\gamma$  (IFN $\gamma$ ) durch Bindung an enzymgekoppelte IFN $\gamma$ -spezifische Antikörper nach.

Vor der Primärantikörperbeschichtung wurden die PVDF-Membranen der ELISpot-Platten mit 35% Ethanol aktiviert und durch dreimaliges Waschen mit sterilem PBS vorbereitet. Die Beschichtung der Membranen erfolgte durch Inkubation über Nacht bei 4 °C mit 10  $\mu\text{g}/\text{ml}$  anti-human-IFN $\gamma$ -Erstantikörper (mAb 1-D1K, Mabtech) in PBS, damit der Primärantikörper mit seinem konstanten Teil an die Membran band. Ungebundene Antikörpermoleküle wurden anschließend durch Waschen mit PBS entfernt.

Sollte die T-Zellerkennung von durch Lipofektion transient transfizierten Zielzellen untersucht werden, wurden  $2 \times 10^4$  293T- bzw. COS-7-Zellen pro Testeinheit in RPMI<sub>FCS</sub> (siehe Abschnitt 2.1.18) ausgesät, wodurch auch freie Antikörperbindestellen auf den Membranen belegt wurden. Im Falle anderer Zielzellen wurden die Platten zunächst mit 120  $\mu\text{l}/\text{well}$  RPMI<sub>FCS</sub> für 30 min bzw. über Nacht vorinkubiert.

Die transiente Transfektion von 293T- bzw. COS-7- durch Lipofectamin<sup>®</sup>2000 bzw. Lipofectamin<sup>®</sup>3000 mit Plasmiden, die für HLA- bzw. Antigen-cDNA kodierte, erfolgte direkt in den Testeinheiten der ELISpot-Platten (siehe Abschnitt 2.2.1.12). Nach Inkubation über Nacht bei 37 °C wurden 120  $\mu\text{l}$  des Überstandes je Testeinheit abgenommen und T-Lymphozyten ( $0,5\text{-}5 \times 10^4$  Zellen/well) bzw. MLTCs ( $1\text{-}3 \times 10^4$  Zellen/well) in RPMI<sub>FCS</sub> (gegebenenfalls mit 25 U/well rhIL-2) ausgesät. Weitere Zielzellen (Malignom-, EBV-transformierte B-Zellen bzw. elektroporierte Transfektanten) wurden mit  $3\text{-}5 \times 10^4$  Zellen/well hinzugefügt. Außerdem wurden gegebenenfalls Peptide (1-10  $\mu\text{g}/\text{ml}$ ) bzw. Blockadeantikörper (siehe Tab. 2.3) hinzugegeben.

Der Test wurde nach Inkubation über Nacht bei 37 °C entwickelt. Hierfür wurden die Zellen in den Testeinheiten durch sechsmaliges Waschen mit PBS mit 0,05% Tween<sup>®</sup>20 abgewaschen. Zur Detektion des sezernierten und an die Erstantikörper gebundenen IFN $\gamma$  wurden die Platten für zwei Stunden bei 37 °C mit 2  $\mu\text{g}/\text{ml}$  des biotinylierten anti-human-IFN $\gamma$ -Sekundärantikörpers (mAb7-B6-1, Mabtech) in PBS mit 0,5% BSA (60  $\mu\text{l}/\text{well}$ ) inkubiert. Anschließend wurden die Platten erneut mit PBS mit 0,05% Tween<sup>®</sup>20 gewaschen und dann für eine Stunde im Dunkeln mit dem zuvor vorbereiteten Avidin-Peroxidase-Komplex (in PBS mit 0,1% Tween<sup>®</sup>20, 100  $\mu\text{l}/\text{well}$ , siehe Abschnitt 2.1.11) bei Raumtemperatur inkubiert. Nach abschließendem Waschen mit PBS mit 0,05% und ohne Tween<sup>®</sup>20 wurde die Farbreaktion durch Zugabe des AEC-Substrats (100  $\mu\text{l}/\text{well}$ , siehe Abschnitt 2.1.11)

gestartet. Hierbei wurde das gebundene IFN $\gamma$  als rötliche Spots auf der Plattenmembran visualisiert. Nach 10-minütiger Inkubation bei Raumtemperatur wurde das AEC-Substrat unter fließendem Leitungswasser abgewaschen und die Platten für 20 min bei 56 °C getrocknet.

Zur Auswertung wurden die Platten mit dem ELISpot-Reader der Firma CTL mithilfe der Software *Immuno Capture* (Version 6.4.35.0, Series 5) eingescannt und die Spots pro Testeinheit mit dem Programm *Immuno Spot* (Version 5.0.40 Professional DC) automatisch gezählt. Um ein Mitzählen von dunklen Verfärbungen an den Rändern der Testeinheiten zu umgehen, wurde beim Zählen der Spots standardmäßig nur 90% der Membranfläche berücksichtigt. Es wurde keine Normalisierung auf die Gesamtfläche durchgeführt.

#### 2.2.1.14.1 Peptiderkennungstest im IFN $\gamma$ -ELISpot-Assay

Auf der Grundlage der bioinformatischen Auswertung der Sequenzierdaten der vier Melanomlinien des Modells Ma-Mel-86 und der autologen EBV-B-Zelllinie INTH-EBV-B wurden 174 Peptide (Okta-, Nona- und Dekamere, siehe Tab. 2.1) ausgewählt, die laut der Algorithmen von NetMHC und IEDB potentiell von den HLA-Klasse I-Molekülen der Patientin präsentiert werden konnten. Diese 174 Peptide wurden von der Firma Peptide 2.0 (Chantilly, Virginia, USA) mit einer Reinheit von 20-80% synthetisiert. Die lyophilisierten Peptide wurden in DMSO rekonstituiert (40 mg/ml) und für den Einsatz im IFN $\gamma$ -ELISpot-Assay in PBS weiterverdünnt (200  $\mu$ g/ml).

Für den Erkennungstest im IFN $\gamma$ -ELISpot-Assay wurden die Peptide am Tag der Belegung in 96-Well-Rundbodenplatten mit RPMI<sub>FCS</sub> auf 1  $\mu$ g/30  $\mu$ l (Endkonzentration im Test: 10  $\mu$ g/ml) eingestellt und 30  $\mu$ l je Testeinheit ausplattiert. Gegebenenfalls wurden dem IFN $\gamma$ -ELISpot-Assay je  $1 \times 10^4$  durch Elektroporation mit den einzelnen HLA-Klasse I-Allelen der Patientin transfizierte 293T-Zellen (insgesamt  $6 \times 10^4$  Zellen) oder  $5 \times 10^4$  INTH-EBV-B-Zellen in je 30  $\mu$ l/well als peptidpräsentierende Zellen hinzugefügt. Andernfalls dienten die T-Zellen selbst als peptidpräsentierende Zellen. Je nach Stärke der Reaktivität der Effektorzellen wurden  $1-5 \times 10^4$  Zellen *ad* 100  $\mu$ l RPMI<sub>FCS</sub> pro Testeinheit eingesetzt.

#### 2.2.1.14.2 *Ex-vivo*-Test

Zur Bestimmung der Frequenz peptidreaktiver T-Zellen in unstimulierten autologen PBMCs wurden *Ex-vivo*-IFN $\gamma$ -ELISpot-Assays durchgeführt. Hierfür wurden die kryokonservierten PBMCs am Tag vor der Testbelegung aufgetaut (siehe Abschnitt 2.2.1.3) und mit  $2 \times 10^6$  Zellen/ml in serumfreiem AIM-V in Kultur genommen.

Die Membranen der ELISpot-Platten wurden wie in Abschnitt 2.2.1.14 beschrieben vorbereitet und mit dem Erstantikörper beschichtet. Zum Blockieren der noch freien Bindestellen der Plattenmembranen wurden 100  $\mu$ l/well AIM-V<sub>komplett</sub> (siehe Abschnitt 2.1.18) ausplattiert. Nach 30-minütiger Inkubation bei 37 °C wurde das Medium verworfen und die Platten

wurden dreimal mit PBS gewaschen. Anschließend erfolgte die Belegung. Die PBMCs wurden einmal gewaschen, eingestellt auf  $5 \times 10^5$  Zellen/50  $\mu$ l serumfreiem AIM-V und ausplattiert. Die Peptidpools bzw. Einzelpeptidpräparationen wurden in serumfreiem AIM-V eingestellt auf 2  $\mu$ g/ml je Peptid und jeweils 50  $\mu$ l der vorverdünnten Peptide bzw. Peptidpools wurden zu den PBMCs gegeben. Als Positivkontrolle wurde PHA (Phytohämagglutinin, 10  $\mu$ g/ml Endkonzentration) verwendet. Die Entwicklung des Tests nach Inkubation über Nacht bei 37 °C, 5% CO<sub>2</sub> und 95% Luftfeuchtigkeit ist in Abschnitt 2.2.1.14 beschrieben.

### 2.2.2 Molekularbiologische Methoden

#### 2.2.2.1 Kultivierung und Kryokonservierung von Bakterien

Die Anzucht von Bakterien erfolgte standardmäßig bei 37 °C in LB-Medium bzw. auf LB-Agarplatten versetzt mit Selektionsantibiotika. Für manche transformierten Bakterienstämme wurde die Temperatur auf 30 °C reduziert, um die Aktivität rekombinierender Enzyme zu reduzieren.

Zur Kryokonservierung wurde die Bakterienflüssigkultur 1:4 mit Bakterieneinfriermedium (siehe Abschnitt 2.1.17, Endkonzentration: 15% Glycerol) gemischt und in Kryoröhrchen bei -80 °C gelagert.

#### 2.2.2.2 Isolierung rekombinanter Plasmid-DNA

Die Isolierung von Plasmid-DNA aus rekombinanten Bakterien wurden mithilfe der molekularbiologischen Kits der Firma QIAGEN (*HiSpeed<sup>®</sup> Plasmid Maxi Kit* und *QIAprep<sup>®</sup> Spin Miniprep Kit*) entsprechend den Herstellerprotokollen durchgeführt. Die anschließende Bestimmung der Konzentration der Plasmid-DNA erfolgte im Spektralphotometer (siehe Abschnitt 2.2.2.3.1).

#### 2.2.2.3 Photometrische Konzentrationsbestimmungen von Nucleinsäuren

##### 2.2.2.3.1 Spektralphotometer

Die photometrische Konzentrationsbestimmung von Nucleinsäuren im Spektralphotometer erfolgte durch Messung der Extinktion bei einer Wellenlänge von 260 nm. Hierbei entsprach eine Absorptionseinheit 50  $\mu$ g DNA bzw. 40  $\mu$ g RNA. Das Verhältnis der Extinktionen bei 260 nm und 280 nm diente als Maß für die Reinheit der Proben.

##### 2.2.2.3.2 Fluorometer

Die Konzentrationsmessung von Nucleinsäuren mithilfe des Fluorometers *Qubit 2.0* (Invitrogen) basiert auf einer Fluoreszenzmarkierung einzelner Nucleinsäurespezies (doppelsträngige bzw. einzelsträngige DNA, RNA) mit spezifischen Sonden. Innerhalb eines



Messansatzes wird jeweils nur eine Art von Nukleinsäuren detektiert, sodass Messfehler durch kontaminierende Proteine oder andere Nukleinsäurearten ausgeschlossen werden konnten. Die Bestimmung der Reinheit der Proben erforderte eine Messung der gleichen Probe mit anderen Sonden.

Die Vorbereitung der Proben für die Messung erfolgte nach Herstellerangabe mit den Kits für dsDNA (doppelsträngige DNA) und RNA (siehe Abschnitt 2.1.10). Das *Qubit 2.0* Fluorometer wurde vor jeder Messung mithilfe der Kit-Standards kalibriert.

#### **2.2.2.4 Isolation genomischer DNA**

Zur Extraktion genomischer DNA (gDNA) aus eukaryotischen Zellen wurde das Kit *QIAamp<sup>®</sup> DNA Mini* (QIAGEN) verwendet.  $2\text{-}5 \times 10^6$  Zellen pro Ansatz wurden für 10 min bei  $56^\circ\text{C}$  im Puffer AL mit Proteinase K und RNase A lysiert. Nach Zugabe von Ethanol (100%) wurde die Probe über die *QIAamp Mini Spin*-Säule zentrifugiert und die an die Kieselgelmembran gebundene gDNA mit den Puffern AW1 und AW2 gewaschen. Die gDNA wurde mit  $\text{dH}_2\text{O}$  (*ad iniectabilia*) eluiert und bei  $-20^\circ\text{C}$  gelagert. Die Konzentrationsbestimmung erfolgte mit dem Spektralphotometer (siehe Abschnitt 2.2.2.3.1).

#### **2.2.2.5 Isolation von Gesamt-RNA**

Für die Isolierung der Gesamt-RNA eukaryotischer Zellen wurde das *RNeasy<sup>®</sup> Mini Kit* (QIAGEN) verwendet. Hierfür wurden  $2\text{-}5 \times 10^6$  Zellen je Ansatz in PBS gewaschen und anschließend zur Lyse in RLT-Puffer mit 1%  $\beta$ -Mercaptoethanol aufgenommen. Gegebenenfalls wurden die Lysate bei  $-80^\circ\text{C}$  gelagert. Zur Homogenisierung wurden die Lysate durch *QIAshredder<sup>™</sup>*-Säulen zentrifugiert. Nach Zugabe von Ethanol (70%) wurde die Probe auf eine *RNeasy Spin*-Säule geladen und zur Bindung der RNA an die Quarzmembran abzentrifugiert. Für die Entfernung von kontaminierender gDNA wurde die Säule anschließend zunächst mit RW1-Puffer gewaschen und dann für 15 min bei Raumtemperatur mit dem DNase I-Inkubationsmix des *RNase-free DNase Set* (QIAGEN) inkubiert. Nach erneutem Waschen mit dem Puffer RW1 und zweimaligem Waschen mit Puffer RPE, wurde die RNA in RNase-freiem Wasser eluiert und bei  $-80^\circ\text{C}$  gelagert. Die Bestimmung der Konzentration erfolgte mit dem Spektralphotometer und dem Fluorometer *Qubit 2.0* (siehe Abschnitt 2.2.2.3.1 bzw. Abschnitt 2.2.2.3.2).

#### **2.2.2.6 cDNA-Synthese**

Zur Herstellung von cDNA aus Gesamt-RNA wurde das *SuperScript<sup>®</sup> III First Strand Synthesis System* (Invitrogen) verwendet. Bis zu 5  $\mu\text{g}$  RNA wurden mit 50 pmol Oligo(dT)<sub>20</sub>-Primern und 10 nmol dNTPs für 5 min bei  $65^\circ\text{C}$  vorinkubiert. Für die Synthese der komplementären cDNA wurden 200 U *SuperScript<sup>®</sup> III Reverse Transkriptase* und 40 U

RNase-Inhibitor *RNaseOUT*<sup>TM</sup> in RT-Puffer mit 10 µM MgCl<sub>2</sub> und 5 mM DTT hinzugefügt. Nach Inkubation für 50 min bei 50 °C wurden die Enzyme bei 85 °C für 5 min hitzeinaktiviert. Zum Abbau der RNA-Matrizen wurden die Reaktionsansätze mit 2 U *E. coli* RNase H für 20 min bei 37 °C inkubiert. Die synthetisierte cDNA wurde gegebenenfalls 1:5 mit Nuklease-freiem Wasser verdünnt und bei -20 °C gelagert.

### 2.2.2.7 Polymerase-Kettenreaktionen

Die Polymerase-Kettenreaktion (engl. *polymerase chain reaction*, PCR) dient der Amplifikation eines definierten DNA-Abschnitts ausgehend von einer DNA-Matrize unter Verwendung sequenzspezifischer Primer, die den gewünschten Bereich flankieren.

Je nach Ziel der Amplifikationsreaktion wurden verschiedene Polymerasen (vgl. Abschnitt 2.1.7) verwendet:

- Für den Nachweis definierter Sequenzabschnitte ohne Anspruch an eine fehlerfreie Amplifikation wurde der *Perfect Taq Plus Master Mix* (5prime) bzw. die *VWR-Taq*-Polymerase (VWR) eingesetzt.
- Sollte das Amplifikat anschließend kloniert werden, wurde die *Q5*-Polymerase (Invitrogen) verwendet.
- Zum Anhängen von 3'-A-Überhängen und bei kürzeren Amplifikaten (< 1 kb) wurde die *EasyA*- (Agilent) oder die *Advantage2*-Polymerase (Clontech) verwendet.

Für einen PCR-Ansatz wurden Primer (*sense* und *antisense*), DNA-Matrix, Polymerase mit dazugehörigem Reaktionspuffer und dNTPs in 25 bzw. 50 µl Endvolumen gemischt. Je nach verwendeter Polymerase, Länge des Amplifikats und Sequenz der Primer wurde das PCR-Programm angepasst (siehe **Tab. 2.5**). Der Erfolg der PCR wurde mittels DNA-Gelelektrophorese (siehe Abschnitt 2.2.2.8.1) überprüft.

**Tab. 2.5:** Übersicht der genutzten PCR-Programme. Um die PCR unter optimalen Bedingungen durchführen zu können, wurde das Thermocycler-Programm entsprechend der verwendeten Polymerase und Primer und entsprechend der Länge des Amplifikats gewählt.

Zyklen	Polymerase PCR-Schritt	Perfect Taq bzw. VWR		Q5		EasyA		Advantage2	
		1	Initiale Denaturierung	94 °C	3 min	98 °C	30 s	95 °C	2 min
25-30	Denaturierung	94 °C	30 s	98 °C	10 s	95 °C	40 s	95 °C	30 s
	Annealing der Primer	50-65 °C*	30 s	T <sub>m</sub> **+3 °C	30 s	T <sub>m</sub> ** -5 °C	30 s	50-65 °C*	30 s
	Elongation	72 °C	1 min/kb	72 °C	30 s/kb	72 °C	1 min/kb	68 °C	1 min/kb
1	Finale Elongation	72 °C	10 min	72 °C	2 min	72 °C	7 min	68 °C	3 min

\* Empirische Bestimmung der optimalen Annealingtemperatur

\*\* T<sub>m</sub>: Schmelztemperatur laut Synthesereport

### 2.2.2.8 Gelelektrophoretische Auftrennung von DNA und RNA

#### 2.2.2.8.1 Agarosegele

Für die gelelektrophoretische Auftrennung wurden entsprechend der zu analysierenden DNA-Längen 0,8%- bis 1,2%-ige Agarosegele mit  $0,5\times$  TBE-Puffer (siehe Abschnitt 2.1.11) und dem in DNA interkalierenden Farbstoff *GelRed* (1:50.000) gegossen. Die Proben wurden vor dem Beladen des Gels mit DNA-Ladepuffer (siehe Abschnitt 2.1.3) versetzt. Als Größenmarker wurden eine 100 bp DNA-Leiter bzw. für längere DNA eine 1 kb DNA-Leiter (siehe Abschnitt 2.1.5) verwendet. Die Auftrennung erfolgte in  $0,5\times$  TBE-Puffer bei einer Spannung von  $5-10\text{ V/cm}$  Gel. Die Auftrennung der DNA wurde mit UV-Licht visualisiert und mit dem *BioDocAnalyze*-System der Firma Biometra digitalisiert.

#### 2.2.2.8.2 Denaturierende Formaldehydgele

Für die Gelelektrophorese von RNA wurde bevorzugt mit denaturierenden Formaldehydgelen gearbeitet, da Sekundärstrukturen der RNA durch die Zugabe von Formaldehyd aufgelöst werden, so dass die Größenauftrennung nicht durch diese gestört wird.

Für ein 1% Agarose-Formaldehydgel wurden 0,54 g Agarose in 70 ml DEPC-behandeltem Wasser (siehe Abschnitt 2.1.11) geschmolzen. Der verbleibenden Mischung (ca. 40 ml) wurden 7 ml  $10\times$  MOPS-Puffer (siehe Abschnitt 2.1.11) und 7 ml 37% Formaldehyd hinzugefügt. Zur Visualisierung der RNA wurde ebenfalls der interkalierende Farbstoff *GelRed* (1:27.000) eingesetzt.

Vor dem Probenauftrag erfolgte ein Vorlauf des Gels für 10 min bei 60 V. Jeweils 5  $\mu\text{l}$  RNA bzw. 2  $\mu\text{l}$  des Größenmarkers (*RiboRuler<sup>TM</sup> High Range RNA Ladder*, Fermentas) wurden vor dem Beladen mit je 7,5  $\mu\text{l}$  Formaldehyd-Ladepuffer (Fermentas) versetzt und für 5 min bei  $72^\circ\text{C}$  inkubiert. Die Auftrennung der RNA erfolgte in  $1\times$  MOPS-Puffer für 1,5-2 h bei 85 V und wurde ebenso mit dem *BioDocAnalyze*-System (Biometra) dokumentiert.

#### 2.2.2.9 Restriktionsverdau von doppelsträngiger DNA

Zur Charakterisierung von transformierter Plasmid-DNA und zur Generierung spezifischer DNA-Fragmente für Klonierungsexperimente wurde die Restriktionshydrolyse eingesetzt. Hierfür wurde die zu verdauende DNA entsprechend den Protokollen der Enzymhersteller (siehe Abschnitt 2.1.7) im jeweiligen Puffer, gegebenenfalls unter Zugabe von acetyliertem BSA ( $100\text{ ng}/\mu\text{l}$ , Promega) für mindestens 1 h bzw. über Nacht mit 1 U pro  $10\text{ }\mu\text{g}$  DNA bei  $37^\circ\text{C}$  oder  $25^\circ\text{C}$  (je nach Herstellerspezifikation) inkubiert. Nach Hitzeinaktivierung für 20 min bei  $85^\circ\text{C}$  wurde der Verdau mittels Gelelektrophorese analysiert (siehe Abschnitt 2.2.2.8.1) und die verdaute DNA wurde gegebenenfalls aus dem Gel aufgereinigt (siehe Abschnitt 2.2.2.10.1).

### **2.2.2.10 Aufreinigung und -konzentrierung von Nukleinsäuren**

#### **2.2.2.10.1 Aufreinigung von DNA-Fragmenten aus Agarosegelen**

Mithilfe des *QIAquick Gel Extraction Kits* der Firma QIAGEN wurden DNA-Fragmente nach Restriktionsverdau bzw. PCR-Amplifikate aus Agarosegelen aufgereinigt. Die Durchführung erfolgte entsprechend dem Herstellerprotokoll.

#### **2.2.2.10.2 Aufreinigung von DNA über Säulen**

PCR-Amplifikate, die nach gelelektrophoretischer Analyse keine Verunreinigung durch unspezifische PCR-Produkte zeigten, wurden zur Entfernung der übrigen Bestandteile der Amplifizierungsreaktion entsprechend den Herstelleranweisungen über die Glasfasersäulen des *High Pure PCR Product Purification Kits* der Firma Roche bzw. die Silicamembransäulen des *QIAquick<sup>®</sup> PCR Purification Kits* der Firma QIAGEN aufgereinigt.

#### **2.2.2.10.3 Phenol-Chloroform-Extraktion und Ethanol-fällung**

Zur Aufreinigung von RNA wurde eine Phenol-Chloroform-Extraktion mit anschließender Ethanol-fällung durchgeführt. Dazu wurde die RNA auf 100 µl mit RNase-freiem H<sub>2</sub>O aufgefüllt und mit einem Volumen Phenol-Chloroform-Isoamylalkohol-Gemisch (25:24:1) durch Vortexen für 30 s gemischt. Nach Zentrifugation für 3 min bei 13.000 rpm wurde die obere wässrige Phase in einem neuen Reaktionsgefäß mit einem Volumen des Chloroform-Isoamylalkohol-Gemisches (24:1) durch Vortexen für 30 s gemischt. Nach erneuter Zentrifugation wurde die obere wässrige Phase wiederum in ein frisches Reaktionsgefäß überführt und zur Abtrennung eventueller Chloroform-Reste ein weiteres Mal zentrifugiert.

Zur Fällung wurde die RNA-Probe mit 0,1 Volumen 3 M Natriumacetat (AppliChem) und 2,5 Volumen kaltem Ethanol (100%) versetzt und über Nacht bei -80 °C inkubiert. Nach Zentrifugation für 30 min bei 4 °C und 13.000 rpm wurde das RNA-Pellet zweimal mit kaltem 70% Ethanol gewaschen, anschließend getrocknet und in 50 µl RNase-freiem H<sub>2</sub>O aufgenommen. Die Konzentration der aufgereinigten RNA wurde mit dem Spektralphotometer und mit dem Fluorometer (siehe Abschnitt 2.2.2.3.1 bzw. Abschnitt 2.2.2.3.2) bestimmt und die Integrität auf einem denaturierendem Formaldehydgel (siehe Abschnitt 2.2.2.8.2) überprüft.

### **2.2.2.11 Einbringen von cDNA oder cDNA-Fragmenten in Vektoren**

#### **2.2.2.11.1 TOPO-TA-Klonierung**

Die Klonierung von kurzen PCR-Produkten ( $\leq 1,5$  kb) erfolgte unter Verwendung des *pcDNA<sup>TM</sup> 3.1/V5-His TOPO<sup>®</sup> TA Expression Kits* der Firma Invitrogen. Für die Amplifikation des gewünschten *Inserts* wurde eine *Taq*-Polymerase (*EasyA* [Agilent]) oder

*Advantage2* [Clontech], siehe Abschnitt 2.1.7 und Tab. 2.5) verwendet, welche die für die TA-Ligation benötigten 3'-A-Überhänge macht.

Je Ligationsansatz wurden 2 µl des aufgereinigten PCR-Produktes (siehe Abschnitt 2.2.2.10.1 bzw. Abschnitt 2.2.2.10.2) mit 0,5 µl Vektor (*pcDNA<sup>TM</sup>3.1/V5-His TOPO<sup>®</sup>*) und 0,5 µl Salzlösung (200 mM NaCl, 10 mM MgCl<sub>2</sub>) gemischt und für 30 min bei Raumtemperatur inkubiert. Alle weiteren Schritte erfolgten nach Angaben des Herstellers.

### 2.2.2.11.2 Gateway-Klonierung

Im Falle längerer *Inserts* (>1,5 kb) wurde das *Gateway*-Klonierungssystem der Firma Invitrogen verwendet, welches auf zwei Rekombinationsreaktionen basiert. Hierfür wurden Primer verwendet, die das gewünschte *Insert* amplifizierten und dem *Insert* die für die Rekombination benötigten *attB*-Sequenzen anfügten. Die PCR wurde mit der *Q5*-Polymerase (NEB, siehe Abschnitt 2.1.7 und Tab. 2.5) durchgeführt. Die Annealingtemperatur wurde für die ersten fünf Zyklen basierend auf dem an das DNA-*Template* bindenden Teil der Primer gewählt. Für die restlichen 25 Zyklen wurde die Annealingtemperatur auf 72 °C erhöht. Das *attB*-PCR-Produkt wurde über ein Agarosegel aufgereinigt (siehe Abschnitt 2.2.2.10.1).

Die erste Rekombinationsreaktion, die BP-Reaktion, führte zum Austausch der *attP*-Sequenz-flankierten *ccdB*-Kassette des Vektors *pDONR<sup>TM</sup>221* (siehe Abschnitt 2.1.14) gegen das *attB*-PCR-Produkt, wodurch aus den *attB*-Sequenzen *attL*-Sequenzen wurden. Hierzu wurden 50-150 ng des *attB*-PCR-Produkts mit 150 ng des Vektors in TE-Puffer für 16-18 h mit der *BP Clonase II* (Invitrogen) bei 25 °C inkubiert. Die Reaktion wurde durch Zugabe von 2 µg Proteinase K und Inkubation für 10 min bei 37 °C gestoppt. 2 µl des Reaktionsansatzes wurden für die Transformation von chemisch kompetenten NEB 10-beta *E. coli* (NEB) verwendet (siehe Abschnitt 2.1.15). Als Selektionsantibiotikum wurde 50 µg/ml Kanamycin (Roth) eingesetzt.

Die entstandenen, sogenannten *Entry*-Klone wurden durch Sequenzierung, PCR oder Restriktionsverdau auf Korrektheit überprüft und dann in der zweiten Rekombinationsreaktion, der LR-Reaktion, zum Einbringen des *Inserts* in einen sogenannten *Destination*-Vektor verwendet. Für die LR-Reaktion wurden 10 fmol des *Entry*-Klons mit 20 fmol des *Destination*-Vektors in TE-Puffer für 18 h bei 25 °C mit der *LR Clonase II Plus* (Invitrogen) inkubiert. Nach Zugabe von 2 µg Proteinase K und Inkubation für 10 min bei 37 °C wurden 2 µl des Reaktionsansatzes zur Transformation von chemisch kompetenten NEB 10-beta *E. coli* (NEB) eingesetzt. Für den in dieser Arbeit verwendeten *Destination*-Vektor *pcDNA3.1 DEST #6* (siehe Abschnitt 2.1.14) wurde das Selektionsmedium mit 100 µg/ml Ampicillin (Sigma) versetzt.

### 2.2.2.12 Sequenzierverfahren

#### 2.2.2.12.1 Sanger-Sequenzierung

Die Sequenzierung von bis zu ca. 1,2kb langen DNA-Abschnitten erfolgte nach dem Sanger-Verfahren in den Firmen StarSEQ bzw. GENterprise Genomics (Mainz).

Für die Sequenzierung zirkulärer Plasmid-DNA wurden ca. 500 ng DNA mit 10 pmol des Sequenzierprimers *ad* 7 µl dH<sub>2</sub>O als „U-mix“-Probe zur Firma StarSEQ geschickt, die die Sequenzierungsreaktion und die Auftrennung der Sequenzierfragmente auf dem *ABI 3730 capillary sequencer* (Applied Biosystems) durchführte.

PCR-Amplifikate wurden vor der Sequenzierung über eine Säule aufgereinigt (siehe Abschnitt 2.2.2.10.2) und im Rahmen des *BYOS-K-Services* der Firma GENterprise Genomics sequenziert. Für eine Sequenzierreaktion wurden bis zu 5 µl der DNA-Probe mit 2 µl 5× *BigDye*-Puffer, 10 pmol Sequenzierprimer und 2 µl *BigDye*-Mix (Version 3.1) *ad* 10 µl dH<sub>2</sub>O (*ad iniectabilia*) gemischt. Die Sequenzierreaktion erfolgte über 30 Zyklen bei 96 °C für 10 s und 55 °C für 4 min.

#### 2.2.2.12.2 Hochdurchsatz-Sequenzierung

Die Generierung der Sequenzierbibliotheken aus den isolierten gDNA- bzw. Gesamt-RNA-Proben (siehe Abschnitt 2.2.2.4 bzw. Abschnitt 2.2.2.5) wurde von B. Ebner (Firma GENterprise Genomics, Mainz) unter Verwendung der Kits *TruSeq<sup>TM</sup> DNA Sample Preparation* bzw. *TruSeq<sup>TM</sup> RNA Sample Preparation v2* (Illumina) durchgeführt. Für die Anreicherung von exonbeinhaltenden DNA-Fragmenten für die Exomsequenzierung wurde das Illumina-Kit *TruSeq<sup>TM</sup> Exome Enrichment* verwendet. Die Sequenzierung (100 bp, *paired-end*) erfolgte auf dem *HiSeq 2500* von Illumina im Institut für Molekulargenetik (Johannes Gutenberg-Universität Mainz).

### 2.2.3 Bioinformatische Methoden

Zur Lokalisierung potentiell immunogener, somatischer Punktmutationen wurde ein Workflow aufgebaut, der im Anhang dargestellt ist (vgl. Abb. A.3). Die einzelnen Prozessschritte sowie alle weiteren verwendeten bioinformatischen Methoden sind im Folgenden beschrieben.

#### 2.2.3.1 Qualitätsbewertung der Illumina-Daten

Zur Bewertung der Qualität der Sequenzierrohdaten bzw. der nachbearbeiteten Daten wurden die einzelnen Datenpakete in das Programm *FastQC* (v.0.10.1, siehe Abschnitt 2.1.20) geladen. Die von der Software generierten Graphen gaben Aufschluss über

- die Sequenzqualität pro Base und gemittelt über die Sequenz,

- die Basenverteilung bzw. den GC-Gehalt je Position und gemittelt über die Sequenz,
- das Vorkommen von nicht-definierten Basen (Ns) an jeder Position,
- die Längenverteilung der Sequenzen und
- das Vorkommen von Duplikaten bzw. überrepräsentierten Sequenzen und  $k$ -meren.

Auf der Grundlage der Qualitätsbewertung mit *FastQC* wurden die Parameter für die Nachbearbeitung der Rohdaten (siehe Abschnitt 2.2.3.2) gewählt bzw. angepasst. Der Erfolg der Nachbearbeitung wurde ebenfalls mit *FastQC* überprüft.

### 2.2.3.2 Nachbearbeitung der Illumina-Rohdaten

Zur Nachbearbeitung der Rohdaten wurde das Tool *Trimmomatic* (v.0.22, siehe Abschnitt 2.1.20, Bolger *et al.*, 2014) verwendet. Die durchgeführte Nachbearbeitung der Rohdaten umfasste das Herausfiltern von Adaptersequenzen (ILLUMINACLIP), das Entfernen von Sequenzanfängen (LEADING) oder -enden (TRAILING) mit schlechter Qualität, das Überprüfen der Sequenz auf ein starkes Absinken der Qualität (SLIDINGWINDOW) und das Verwerfen von Sequenzen, die durch die Nachbearbeitung zu kurz geworden waren (MINLEN). Hierfür wurden folgende Parameter im *paired-end* Modus gewählt:

- ILLUMINACLIP:
  - Maximale Anzahl an *Mismatches*: 2
  - Genauigkeit des *Matches* zwischen den beiden adapterligierten Sequenzen für *paired-end* Palindrom-*Read-Alignments*: 40
  - Genauigkeit des *Matches* zwischen einer Adaptersequenz und *Read*: 15
- LEADING:
  - Minimale Qualität: 13 (entspricht einer Fehlerwahrscheinlichkeit von  $\leq 5\%$ )
- TRAILING:
  - Minimale Qualität: 13
- SLIDINGWINDOW:
  - Fenstergröße: 4
  - Minimale Qualität gemittelt über Fenster: 15 (entspricht einer Fehlerwahrscheinlichkeit von  $\leq 3\%$ )
- MINLEN:
  - Mindestlänge der verbleibenden *Reads*: 30

Für die Entfernung der Illumina-Adaptersequenzen wurden die von S. Rapp (Institut für Molekulargenetik, Johannes Gutenberg-Universität Mainz) zur Verfügung gestellten Sequenzen (siehe Anhang, Abschnitt A.2.2) für jeden Datensatz in einer FASTA-Datei zusammengestellt. Das Ergebnis der Nachbearbeitung der Rohdaten wurde mit *FastQC* (vgl. Abschnitt 2.2.3.1) überprüft.

### 2.2.3.3 Mapping

Die Zuordnung der sequenzierten *Reads* an eine Referenz, das *Mapping*, erfolgte sowohl für die Exom- als auch für die Transkriptomdaten jeweils an die genomische Referenz hg19 (*H. sapiens*, UCSC, Februar 2009). Für das *Mapping* der Exomdaten an hg19 wurden das Programm *Bowtie2* (Langmead & Salzberg, 2012) verwendet. Für das *Mapping* der Transkriptomdaten wurde das Programm *TopHat* (Trapnell *et al.*, 2012) eingesetzt. Die gewählten Parameter werden in den folgenden Abschnitten beschrieben.

#### 2.2.3.3.1 Bowtie2

Für das *Mapping* der Exomsequenzierdaten mittels *Bowtie2* (Langmead & Salzberg, 2012) wurden alle Standardparameter beibehalten. Einzige Ausnahme bildete die Option `-k`, die auf 1 gesetzt wurde, wodurch zu jedem *Read* bzw. *Read*-Paar nur ein *Mapping*-Ergebnis ausgegeben wurde. Als Referenz wurde der für *Bowtie2* vordefinierte Index<sup>2</sup> für das Humangenom hg19 (UCSC) von *Illumina iGenomes* verwendet. Nach dem *Mapping* wurden die generierten SAM-Dateien in das binäre BAM-Format umgewandelt (`samtools view`). Außerdem wurden die alignierten *Reads* nach der genomischen Koordinaten sortiert (`samtools sort`) und PCR-Duplikate entfernt (`samtools rmdup`). Die BAM-Dateien wurden abschließend indexiert (`samtools index`).

#### 2.2.3.3.2 TopHat

Das synchrone *Mapping* von gepaarten und ungepaarten *Reads* mittels *TopHat* (Trapnell *et al.*, 2012) wird nicht empfohlen. Daher wurde zunächst das *Mapping* der ungepaarten *Reads* durchgeführt und durch die Option `-j` mit dem *Mapping* der gepaarten *Reads* verknüpft. In beiden *Mapping*-Durchgängen wurden jeweils folgende Parameter gewählt:

- Anzahl der Mismatches pro Read (`-N`): 2 (entspricht Standardwert)
- Art der Bibliothek (`--library-type`): fr-unstranded
- Erwarteter durchschnittlicher Abstand zwischen gepaarten *Reads* (`-r`): 200
- Standardabweichung für den Abstand (`--mate-std-dev`): 50

---

<sup>2</sup>[ftp://igenome:G3nom3s4u@uscd-ftp.illumina.com/Homo\\_sapiens/UCSC/hg19/Homo\\_sapiens\\_UCSC\\_hg19.tar.gz](ftp://igenome:G3nom3s4u@uscd-ftp.illumina.com/Homo_sapiens/UCSC/hg19/Homo_sapiens_UCSC_hg19.tar.gz) (heruntergeladen am 04.01.2013)



Die letzten beiden Parameter wurden basierend auf der Größe der bei der Generierung der Sequenzierbibliothek entstandenen Insertlängen (90-380 bp, Mittelwert: 200 bp, persönliche Mitteilung von B. Ebner, GENterprise Genomics, Mainz) abgeschätzt (siehe auch Abschnitt 3.2).

Da *TopHat* das *Mapping* der *RNA-Seq-Reads* mit *Bowtie2* durchführt und dann auf durch Spleißen verbundene Exons überprüft, konnte auf den gleichen vordefinierten *Bowtie2*-Index als Referenz für das Humangenom zurückgegriffen werden wie für das *Mapping* der Exomdaten (siehe Abschnitt 2.2.3.3.1).

Nach dem *Mapping* wurden die alignierten *Reads* in den von *TopHat* generierten BAM-Dateien noch nach der genomischen Koordinate sortiert (`samtools sort`) und die prozessierten BAM-Dateien indexiert (`samtools index`). Abschließend wurden die BAM-Dateien der *Mappings* der gepaarten und der ungepaarten *Reads* einer Probe zusammengefasst (`samtools merge`) und ebenso sortiert und indexiert. Das Entfernen von PCR-Duplikaten (`samtools rmdup`, siehe Abschnitt 2.2.3.3.1) wurde im Falle der *RNA-Seq*-Daten nicht durchgeführt, da bei hochexprimierten Genen bei der Generierung der Sequenzierbibliothek identische Fragmente entstehen können. Die dazugehörigen *Reads* alignieren an der gleichen Stelle der Referenz, sind jedoch keine PCR-Duplikate.

#### 2.2.3.4 Qualitätsbewertung der *Mapping*-Ergebnisse

Zum Einschätzen des Erfolgs der Exom-*Mappings* wurde die Abdeckung der *Target-Exons* mithilfe von *bedtools* (`coverageBed`) überprüft. Die genomischen Koordinaten der *Target-Exons* wurden der *TruSeq*-Referenzdatei<sup>3</sup> entnommen. Weitere *Alignment*-Statistiken wurden mittels `samtools flagstat` berechnet. Zur Abschätzung der Abdeckung der Transkriptom-*Mappings* (Details im Ergebnisteil Abschnitt 3.3.3.2) wurde die Anzahl der alignierten Nukleotide mithilfe des eigenen Perl-Skripts `mappedReads_LengthDistribution.pl` (siehe Anhang, Abschnitt A.3.2.1) ermittelt. Die Anzahl eindeutig alignierter *Read*-Paare wurde in den mit *TopHat* alignierten Daten anhand des `FLAG` (`samtools view -f 2`) und des Tags `NH:i:1` bestimmt, welcher *Reads* mit nur einem *Alignment* markiert.

#### 2.2.3.5 Visualisierung der Daten

Zur Visualisierung der *Mappings* und der detektierten Mutationen wurden die generierten BAM-Dateien in den *Integrative Genome Viewer* (IGV, Thorvaldsdóttir *et al.*, 2013) mit hg19 (UCSC) als Referenz geladen. Dies ermöglichte die qualitative Überprüfung einzelner Gene bzw. Exons bezüglich ihrer Expression bzw. des Vorkommens von Mutationen.

---

<sup>3</sup>[http://support.illumina.com/downloads/truseq\\_exome\\_targeted\\_regions\\_bed\\_file.ilmm](http://support.illumina.com/downloads/truseq_exome_targeted_regions_bed_file.ilmm)  
(heruntergeladen am 04.05.2013)

### 2.2.3.6 Detektion und Filtern von Mutationen

#### 2.2.3.6.1 *Variant Calling*

Das Herausschreiben von Abweichungen der alignierten Daten zur Referenz, das sogenannte *Variant Calling*, erfolgte mithilfe der Tools der Programmpakete *SAMtools* (Version 0.1.18 (r982:295)) und *bcftools* (Version 0.1.17-dev (r973:277)) nach dem Protokoll<sup>4</sup> der Entwickler durch Kombinieren der Befehle `samtools mpileup` und `vcfutils.pl varFilter`.

Das *Variant Calling* wurde in den zuvor prozessierten BAM-Dateien gegebenenfalls nach Entfernen der PCR-Duplikate durchgeführt. Als Referenz diente die von *Illumina iGenomes* heruntergeladene Referenzsequenz (hg19, UCSC, siehe Abschnitt 2.2.3.3.1). Für die maximal erlaubte Abdeckung einer Mutation (Option `-D`) wurde für die Exomdaten abweichend vom Entwicklerprotokoll 200 gewählt, da aufgrund der Tiefe der durchgeführten Sequenzierung und der dadurch hohen Anzahl an alignierten *Reads* eine durchschnittlich höhere Abdeckung erwartet wurde. *Variants* mit überdurchschnittlich hoher Abdeckung können durch *Alignment*-Artefakte verursacht sein. Mutationen mit einer Abdeckung von  $>200$  *Reads* in INTH-EBV-B wurden in einer separaten Datei dokumentiert, um darin enthaltene mögliche Keimbahnmutationen herausfiltern zu können. Für die quantitativen Transkriptomdaten wurde keine obere Grenze gesetzt.

#### 2.2.3.6.2 Annotation der detektierten Mutationen

##### 2.2.3.6.2.1 rsIDs der *dbSNP*-Datenbank

In der NCBI-Datenbank *Database for Short Genetic Variations* (*dbSNP*) werden Einzelnukleotidvarianten, kleinere Insertionen und Deletionen, *Short Tandem Repeats* und Mikrosatelliten verschiedener Organismen katalogisiert. Um gefundene Mutationen direkt mit bereits in *dbSNP* annotierten Mutationen in Verbindung zu bringen, wurden die generierten Mutationslisten (VCF- bzw. BCF-Dateien) in der Spalte ID um die gegebenenfalls dazugehörige *rsID* der *dbSNP*-Datenbank ergänzt. Mutationen, die dort nicht annotiert waren, blieben unverändert.

Hierfür wurde der Befehl `vcf-annotate` (*VCFtools*) und die vom NCBI bereitgestellte VCF-Datei `00-A11.vcf`<sup>5</sup> mit den gelisteten humanen *dbSNP*-Einträgen (*Build* 137, 26.01.2012) verwendet. Um diese Datei kompatibel mit den generierten VCF-Dateien der Ma-Mel-86-Zelllinien zu machen, wurde die Nummerierung der Chromosomen von 1 bis 22 in `chr1` bis `chr22` umgewandelt. Entsprechend wurde für die Chromosomen X und Y verfahren. Für die Koordinaten mitochondrialer DNA wurde `MT` zu `chrM` umgeschrieben.

---

<sup>4</sup>Beschreibung der Befehle unter <http://samtools.sourceforge.net/mpileup.shtml> (zuletzt aufgerufen am 23.04.2014)

<sup>5</sup>[ftp://ftp.ncbi.nih.gov/snp/organisms/human\\_9606/VCF/00-A11.vcf.gz](ftp://ftp.ncbi.nih.gov/snp/organisms/human_9606/VCF/00-A11.vcf.gz) (heruntergeladen am 31.01.13)

### 2.2.3.6.2 Vorhersage des Effekts von Mutationen

Das Java-Programm *snpEff* (Version 3.1m, Cingolani *et al.*, 2012) wurde eingesetzt zur Annotation der detektierten Mutationen und zur Vorhersage des Effekts auf die betroffenen Gene. Die hierfür benötigte vordefinierte *snpEff*-Datenbank zur Referenz hg19 (UCSC) wurde durch Aufrufen des Kommandos `download hg19` heruntergeladen. Bei der Annotation der generierten VCF-Dateien der Ma-Mel-86-Zelllinien wurden alle Standardparameter des *snpEff*-Kommandos `eff` unverändert beibehalten, wodurch die Spalte `INFO` in den VCF-Dateien um den Eintrag `EFF` erweitert wurde. Dieser Eintrag enthielt nach der Annotation unter anderem folgende Informationen:

- Effekt der Mutation
- Stärke des Effekts (`HIGH/MODERATE/LOW/MODIFIER`)
- Funktionelle Klasse (`NONE/SILENT/MISSENSE/NONSENSE`)
- Veränderung des betroffenen Codons
- Aminosäure-Austausch
- Länge des Proteins
- Name des betroffenen Gens
- Transkript-ID (*RefSeq*)

Auf die so annotierten VCF-Dateien konnten später Textfilter zur Selektion von bestimmten Mutationsklassen angewandt werden (siehe Abschnitt 2.2.3.6.5).

### 2.2.3.6.3 Qualitätsfilter zur Identifizierung sicherer Mutationen

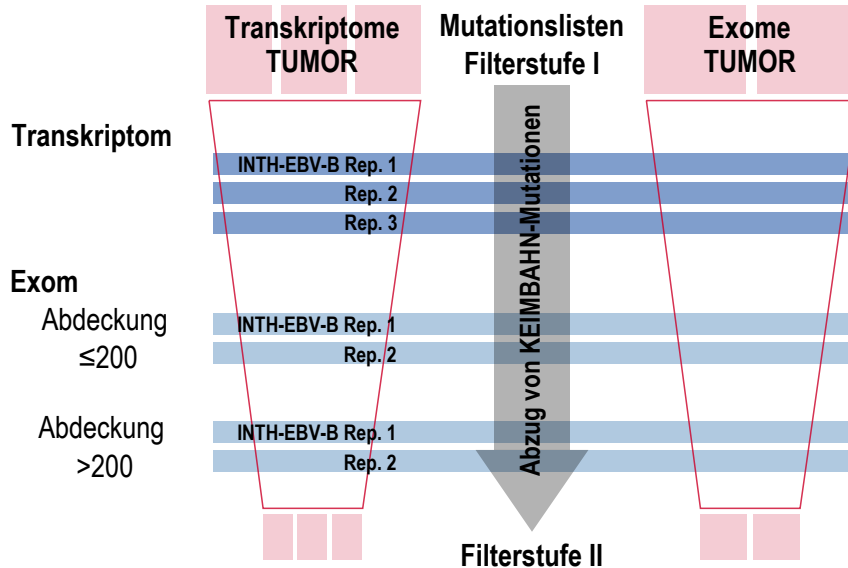
Zur Selektion von sicheren Mutationen mit guter Qualität und ausreichender Abdeckung wurden die VCF-Dateien mittels `vcf-annotate` (*VCFtools*) gefiltert. Als Filterkriterien wurden eine Qualität von mindestens 13 (entspricht einer Fehlerwahrscheinlichkeit von  $\leq 5\%$ ) und eine Mindestabdeckung von zehn *Reads* für die Exomdaten bzw. fünf *Reads* für die Transkriptomdaten gewählt.

### 2.2.3.6.4 Filtern somatischer Mutationen

Um die Keimbahnmutationen aus den vorgefilterten Mutationen mit guter Qualität und Abdeckung (siehe Abschnitt 2.2.3.6.3, Filterstufe I) herauszufiltern, wurden die Sequenzierungen der autologen EBV-immortalisierten B-Zelllinie INTH-EBV-B herangezogen. Dabei wurden in jedem Tumordatensatz alle Mutationen entfernt, die in einem der Exome oder Transkriptome der INTH-EBV-B-Zellen detektiert wurden. Zusätzlich wurden die Mutationen, die in den INTH-EBV-B-Exomen mit einer Abdeckung von mehr als 200 *Reads* vorlagen, ausgeschlossen (siehe **Abb. 2.1**). Es wurde angenommen, dass es sich vermutlich

nicht um *Alignment*-Artefakte handelte, wenn diese Mutationen mit niedrigerer Abdeckung in den Sequenzierungen der Tumorzelllinien detektiert wurden (vgl. Abschnitt 2.2.3.6.1).

Für den Abgleich der insgesamt fünf Mutationslisten je Tumorzelllinie mit der Keimbahnkontrolle INTH-EBV-B wurde `vcf-isec` (*VCFtools*) mit der Option `-c` verwendet. Somit wurden nur die Mutationen weitergeführt, die in keiner der INTH-EBV-B-Listen vorkamen (Filterstufe II).



**Abb. 2.1: Filtern somatischer Mutationen.** Mithilfe der Sequenzierungen der Keimbahnkontrolle INTH-EBV-B wurden nicht-somatische Mutationen herausgefiltert. Aus jeder Tumormutationsliste (3×Transkriptom, 2×Exom, rot) wurden alle Mutationen ausgeschlossen, die in mindestens einem INTH-EBV-B-Transkriptom (dunkelblau) oder mindestens einem INTH-EBV-B-Exom (hellblau) detektiert wurden. Zusätzlich wurden Exommutationen von INTH-EBV-B mit einer Abdeckung von mehr als 200 Reads entfernt.

### 2.2.3.6.5 Textfilter zur Selektion nicht-synonymer Punktmutationen

Zur weiteren Einschränkung der Mutationen wurden Textfilter auf die annotierten VCF-Dateien angewandt. Dazu wurden die Dateien zunächst in die Programmierumgebung *R* geladen (Version 2.15.1, Paket *VariantAnnotation*, Version 1.4.10). Die eingelesenen VCF-Dateien wurden als sogenannte VCF-Objekte angelegt, die durch ihre Struktur gezieltes Abfragen und Filtern ermöglichen, ohne dass verknüpfte Informationen verloren gingen.

Die Information darüber, ob eine Mutation einen kodierenden Sequenzabschnitt betraf, war im VCF-Objekt unter dem *data frame info* in der Spalte *EFF* abgelegt. Mittels eines Textfilters wurden alle Einträge, die den Begriff *CODING* enthielten, bestimmt und mit den jeweiligen genomischen Koordinaten (abgelegt im VCF-Objekt unter *rowData@ranges@NAMES*) in einem neuen *data frame* gesammelt. Die herausgefilterten Einträge bestanden ihrerseits wiederum je nach Anzahl der durch die Mutation betroffenen Transkriptionsvarianten aus mehreren Einträgen. Nicht jede Transkriptionsvariante war dabei in einem kodierenden Bereich mutiert. Solche Einträge wurden durch Umwandeln des *data frame* in einen Vektor und erneutes Filtern nach *CODING* entfernt. Die so herausgefilterten, einen proteinkodierenden Sequenzabschnitt betreffenden Mutationen wurden mittels eines weiteren Textfilters

auf die Keinnzeichnung **MODIFIER**, **LOW**, **MODERATE** und **HIGH** entsprechend des Effekts der vorliegenden Mutation in vier Gruppen eingeteilt. Die Gruppe **MODERATE** enthielt die nicht-synonymen *Missense*-Mutationen.

Abschließend wurden die gefilterten Mutationslisten zu allen vier Tumorzelllinien aus den Exom- und Transkriptomdaten in einzelne Tabellen umgewandelt und als kommaseparierte Textdateien zur Verwendung in Tabellenkalkulationsprogrammen, wie *Microsoft Excel*<sup>®</sup>, ausgegeben.

### 2.2.3.7 Vorhersage der Immunogenität von mutierten Peptiden

#### 2.2.3.7.1 Generierung mutierter Proteinabschnitte für Vorhersagealgorithmen

Mithilfe der weiter unten beschriebenen Vorhersagealgorithmen (siehe Abschnitt 2.2.3.7.2) sollte die Immunogenität von Peptiden, die die zuvor herausgefilterten Mutationen enthielten, vorhergesagt werden. Damit nur Vorhersagen zu Peptiden, die die Mutationen enthielten, gemacht wurden, sollte den Algorithmen jeweils nur der relevante Proteinabschnitt und nicht die komplette Proteinsequenz übergeben werden. Diese Proteinabschnitte hatten für die Vorhersage von Oktameren eine Länge von 15 Aminosäuren, für Nonamere eine Länge von 17 Aminosäuren und für Dekamere eine Länge von 19 Aminosäuren. Dabei lag die mutierte Aminosäure jeweils in der Mitte des Proteinabschnitts, so dass die Mutation in den daraus abgeleiteten Okta-, Nona- bzw. Dekamere genau einmal an jeder möglichen Peptidposition vorkommen konnte. Eine schematische Darstellung hierzu findet sich im Ergebnisteil (Abb. 3.16).

Zur Generierung dieser Proteinsequenzabschnitte wurden mehrere eigene *Perl*-Programme eingesetzt. Als Eingabe wurden die von UCSC und *RefSeq* heruntergeladenen Wildtypproteinsequenzen im FASTA-Format und eine tabulatorgetrennte Liste der Mutationen (Angabe von *RefSeq* Transkript-ID, Position der Mutation innerhalb der Proteinsequenz und vorliegender Aminosäure-Austausch) benötigt. Da Diskrepanzen zwischen den zwei Datenbanken und zwischen den Datenbanksequenzen und den Annotationen durch *snpEff* (siehe Unterabsatz 2.2.3.6.2.2) vorlagen, wurden die Datenbanksequenzen zunächst gegeneinander mithilfe des *Perl*-Skripts `compareRefSeqUSCS.pl` (siehe Anhang, Abschnitt A.3.2.2) abgeglichen und Redundanzen entfernt. Anschließend erfolgte das Einfügen der Mutationen in die Wildtypsequenzen und das Ausschneiden der relevanten Proteinabschnitte (`peptideCandidates_TAB.pl`, siehe Anhang, Abschnitt A.3.2.3). Dabei wurden alle Sequenzen verworfen, bei denen die Wildtypamino-säure an der zu mutierenden Position nicht mit der Wildtypamino-säure der Mutationsliste übereinstimmte (d. h. Diskrepanz zwischen Datenbank und *snpEff*-Annotation). Abschließend wurde für jede Mutation der oben beschriebene Sequenzabschnitt ins FASTA-Format konvertiert, wobei übereinstimmende Proteinabschnitte wiederum zusammengefasst wurden (`peptideCandidatesTAB_to_FASTA.pl`, siehe Anhang, Abschnitt A.3.2.4).

### 2.2.3.7.2 Vorhersage der Präsentation durch HLA-Klasse I-Moleküle

Zur Vorhersage der Wahrscheinlichkeit, dass ein mutiertes Peptid (Okta-, Nona- oder Dekamer) durch die HLA-Klasse I-Moleküle der Patientin INTH an der Tumorzelloberfläche präsentiert werden kann, wurden die vier in **Tab. 2.6** aufgelisteten Algorithmen verwendet. Bei allen Algorithmen wurden jeweils die Grundeinstellungen beibehalten und die in Abschnitt 2.2.3.7.1 generierten Proteinabschnitte für die Vorhersage von Okta-, Nona- bzw. Dekameren übergeben. Die gewählten Schwellenwerte sind ebenfalls **Tab. 2.6** zu entnehmen.

**Tab. 2.6:** Übersicht der verwendeten Algorithmen zur Vorhersage der HLA-Klasse I-Bindewahrscheinlichkeit von mutierten Okta-, Nona- und Dekameren. Die Auflistung zeigt die vier eingesetzten Webtools, die überprüften HLA-Klasse I-Moleküle, die jeweilige Art des ausgegebenen *Score* und welcher Schwellenwert gewählt wurde. Unterstrichen sind die HLA-Moleküle, die in der vierstelligen Allelbezeichnung mit den HLA-Klasse I-Allelen der Patientin INTH übereinstimmen.

Algorithmus	Unterstützte HLA-Klasse I-Allele			Score-Art	Gewählter Schwellenwert
	A	B	C		
SYFPEITHI <sup>a</sup>	A*01 <u>A*24:02</u>	B*08 <u>B*15:01</u>	-	Eigener <i>Score</i>	≥20
BIMAS <sup>b</sup>	A1 A24	B8 B62	Cw0301 Cw0702	Dissoziationszeit [s]	≥20
IEDB v2.9 <sup>c</sup>	<u>A*01:01</u> <u>A*24:02</u>	<u>B*08:01</u> <u>B*15:01</u>	<u>C*03:03</u> <u>C*07:01</u>	Prozentrang	≤10
NetMHC 3.4 Server <sup>d</sup>	<u>A*01:01</u> <u>A*24:02</u>	<u>B*08:01</u> <u>B*15:01</u>	<u>C*03:03</u> <u>C*07:01</u>	Bindeaffinität [nM]	≤500

<sup>a</sup> <http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm>

<sup>b</sup> [http://www.bimas.cit.nih.gov/molbio/hla\\_bind/](http://www.bimas.cit.nih.gov/molbio/hla_bind/)

<sup>c</sup> <http://tools.immuneepitope.org/mhci/>

<sup>d</sup> <http://www.cbs.dtu.dk/services/NetMHC/>

Da nur die Algorithmen NetMHC 3.4 Server (kurz: NetMHC) und von IEDB (*Immune Epitope Database*) die Vorhersage für alle sechs HLA-Klasse I-Allele der Patientin anboten, wurden nur diese beiden Algorithmen für die abschließende Auswahl der zu synthetisierenden und auf Immunogenität zu testenden Peptide herangezogen. Hierbei wurden nur Peptide gewählt, die eine vorhergesagte Bindeaffinität von ≤500 nM (NetMHC) und einen Prozentrang von ≤10 (IEDB) aufwiesen.

### 2.2.3.7.3 Vorhersage der proteasomalen Spaltung

Neben der Vorhersage der Bindung an HLA-Klasse I-Moleküle können potentielle proteasomale Schnittstellen in Proteinsequenzen vorhergesagt werden. Nach Identifikation der tatsächlich immunogenen Peptide in IFN $\gamma$ -ELISpot-Assays wurde die entsprechenden Proteinabschnitte mithilfe des Internettools *NetChop 3.1 Server*<sup>6</sup> auf solche Schnittstellen überprüft. Dabei wurde der Schwellenwert auf der Standardeinstellung (0,5) belassen und jeweils beide Vorhersagemethoden (*C term 3.0* bzw. *20S 3.0*) überprüft. Die Methoden unterscheiden sich bezüglich der Datensätze, die für das Training des neuronalen

<sup>6</sup><http://www.cbs.dtu.dk/services/NetChop/>

Netzwerks verwendet wurden. Die *C-term*-Methode basiert auf den C-Termini bekannter Epitope von cytotoxischen T-Zellen. Bei der *20S*-Methode wurden die Ergebnisse von *In-vitro*-Verdauexperimenten mit dem konstitutiven Proteasom als Trainingsdaten herangezogen.

### 2.2.3.8 Expressionsanalyse

Die Analyse von differenziell exprimierten Genen in den Transkriptomdaten wurde nach dem Verfahren *DESeq* von Anders & Huber (2010) durchgeführt. *DESeq* nimmt eine negative Binomialverteilung bei der Abschätzung der Varianz an, in der Mittelwert und Varianz durch lokale Regression miteinander verknüpft sind. Dadurch wird die Überdispersion biologischer Replikate berücksichtigt (Anders & Huber, 2010).

Zur Bestimmung der differentiellen Expression im Vergleich zu gesunden Melanozyten wurden öffentlich verfügbare *RNA-Seq*-Daten herangezogen. Die im *Sequence Read Archive* hinterlegten Daten (*Accession*: SRX101662) stammten von RFP (engl. *red fluorescent protein*)-transduzierten Melanozyten dreier Spender (Flockhart *et al.*, 2012). Die RNA-Isolation aus diesen Melanozyten und Generierung der Sequenzierbibliothek erfolgte mit den Kits, die auch für die eigenen Proben verwendet wurden (siehe Abschnitt 2.2.2.5 und Abschnitt 2.1.10). Ebenso wurden die Melanozytendaten auf einem Illumina *Hi-Seq 2000* (100 bp, *paired-end*) sequenziert und konnten somit parallel mit den eigenen Transkriptomdaten prozessiert werden.

Für die Expressionsanalyse der mit *TopHat* alignierten Transkriptomdaten wurde das Paket *DESeq* (Version 1.10.1) in der Programmierumgebung *R* (Version 2.15.1) verwendet (siehe Abschnitt 2.1.20). Die Analyse wurde wie in der Vignette (Anders & Huber, 2012) beschrieben durchgeführt. Zur Umwandlung der sortierten BAM-Dateien in eine sogenannte *count table*, die für jedes Gen die Anzahl der dort alignierten *Reads* (*Read-Counts*) angibt, wurde die Funktion `summarizeOverlaps` aus dem *R*-Paket *GenomicRanges* (Version 1.10.7) genutzt. Die Normalisierung der Daten wurde mit *DESeq* anhand der ermittelten Größen der Datensätze (`estimateSizeFactors`) durchgeführt. Hierfür wird ein Pseudoreferenzdatensatz aus den geometrischen Mittelwerten der *Read-Counts*  $k_{ij}$  für jedes Gen  $i$  über alle Datensätze  $j$  bestimmt. Der Normalisierungsfaktor  $\hat{s}_j$  entspricht dann dem Median der Verhältnisse zwischen den *Read-Counts*  $k_{ij}$  und dem entsprechenden Wert im Pseudoreferenzdatensatz über alle Gene (Anders & Huber, 2010):

$$\hat{s}_j = \text{Median}_i \frac{k_{ij}}{\sqrt[m]{\prod_{v=1}^m k_{iv}}}$$

wobei  $m$  die Anzahl der Datensätze ist.

Um die Expressionswerte zwischen Genen vergleichbar zu machen, wurden die mit *DESeq* normalisierten Daten zusätzlich anhand der ORF-Länge des längsten annotierten Transkripts jedes Gens in Kilobasen normalisiert. Zudem wurde die in

der *DESeq*-Vignette beschriebene Varianzstabilisierung der Daten mit der Funktion `getVarianceStabilizedData` des *R*-Pakets *vsn* (Version 3.26.0) durchgeführt (aufgerufen durch die *DESeq*-Funktion `varianceStabilizingTransformation`). Weiterhin wurden die Pakete *TxDb.Hsapiens.UCSC.hg19.knownGene* (Version 2.8.0), *gplots* (Version 2.11.0) und *RColorBrewer* (Version 1.0-5) verwendet, wobei *TxDb.Hsapiens.UCSC.hg19.knownGene* die Annotationen zur Referenz hg19 (UCSC) enthielt.

### 2.2.3.9 Phylogentische Analyse der Ma-Mel-86-Zelllinien

Die Berechnung der phylogenetischen Beziehung der Ma-Mel-86-Zelllinien wurde basierend auf den im Exom detektierten Einzelnukleotidvarianten (SNVs, engl. *single nucleotide variants*) erstellt. Für das hierbei benötigte *Alignment* wurde für jede Zelllinie eine artifizielle Sequenz generiert, die sich aus allen mutierten Positionen der fünf Zelllinien zusammensetzte. Lag eine Mutation in einer Zelllinie nicht vor, so wurde an dieser Stelle die Wildtypbase gesetzt.

Für die Generierung dieser Sequenzen wurden zunächst die alignierten Daten der Replikate einer Zelllinie nach Entfernen von PCR-Duplikaten (siehe Abschnitt 2.2.3.3.1) zusammengefasst (`samtools merge`). Anschließend wurden die Mutationen in den zusammengefassten BAM-Dateien wie in Abschnitt 2.2.3.6.1 beschrieben detektiert (`samtools mpileup`) und Mutationen mit guter Qualität ( $QUAL \geq 13$ ) und ausreichender Abdeckung ( $coverage \geq 10$ ) gefiltert. Da auch heterozygote Mutationen durch Verwendung des *ambiguity code* (IUBMB/IUPAC-Schreibweise, siehe auch Tab. A.1 im Anhang) bei der phylogenetischen Analyse berücksichtigt werden sollten, wurden die Mutationen diesbezüglich nochmals gefiltert und die Zygote folgendermaßen definiert:

- Heterozygot: 40-60% der alignierten Basen an dieser Koordinate wichen von der Referenz ab.
- Homozygot:  $\geq 90\%$  der alignierten Basen an dieser Koordinate wichen von der Referenz ab.
- Alle übrigen Mutationen wurden verworfen.

Hierzu wurde das eigene Perl-Programm `adjust_zygosity_stringent.pl` (siehe Anhang, Abschnitt A.3.2.5) verwendet. Das Programm passte in VCF-Dateien den Eintrag `GT`, welcher die Information über den Genotyp (1/1: homozygot mutiert bzw. 0/1: heterozygot mutiert) enthält, basierend auf der Anzahl der *Reads* unter `DP4` in der Spalte `INFO` an. `DP4` gibt durch Kommas getrennt die Anzahl der *Reads* an, die (1) in *forward*-Richtung mit der Referenz übereinstimmen, (2) in *reverse*-Richtung mit der Referenz übereinstimmen, (3) in *forward*-Richtung nicht übereinstimmen und (4) in *reverse*-Richtung nicht übereinstimmen. Das Perl-Programm berechnete hierfür den Quotienten aus der Anzahl der *Reads*, die in *forward*- und *reverse*-Richtung von der Referenz abwichen (Summe aus (3) und (4)), und der Gesamtzahl der *Reads* an dieser Position (Summe aus (1) bis (4)). Nach Vergleich



mit den vorgegebenen Schwellenwerten (Optionen `-t` und `-s`) wurde in der VCF-Datei die Zygote im Eintrag `GT` angepasst, beibehalten oder der komplette Mutationseintrag verworfen.

Die so bearbeiteten VCF-Dateien der fünf Ma-Mel-86-Zelllinien wurden dann zusammengefasst (`VCF-merge`), so dass die gesammelten Mutationen aller fünf Zelllinien in einer VCF-Datei vorlagen. Für jede mutierte genomische Koordinate wurde jeweils eine Zeile erzeugt, in der in der Spalte `GT` die Genotypen der fünf Zelllinien hintereinander angegeben wurden. Diese VCF-Datei wurde in eine tabulatorgetrennte Textdatei (TAB-Datei) umgewandelt (`VCF-to-tab`), in der alle in Ma-Mel-86 mutierten genomischen Koordinaten gelistet waren und für jede Zelllinie in einer eigenen Spalte angegeben wurde, ob und wie diese Position mutiert war. Die TAB-Datei wurde als Eingabe für das eigene Perl-Programm `VCF_tab_to_fasta_alignment_advanced.pl` (siehe Anhang, Abschnitt A.3.2.6) benötigt. Dieses Skript basierte auf `VCF_tab_to_fasta_alignment.pl`<sup>7</sup> von C. Bergey und sollte aus den Spalten der TAB-Datei für jede Linie eine Sequenz im FASTA-Format erstellen. Im Vergleich zum Originalskript musste der Umgang mit dem Symbol `./` undefiniert werden. In der mittels `VCF-to-tab` erstellten TAB-Datei wurde dieses Symbol gesetzt, wenn die betreffende genomische Koordinate in der VCF-Datei einer Zelllinie nicht aufgeführt und somit nicht mutiert war. Das Originalskript wandelte `./` in der FASTA-Sequenz zu einer Lücke („-“) um. Für das benötigte *Alignment* musste an dieser Stelle jedoch die Referenzbase stehen. Dies wurde im eigenen *Perl*-Skript `VCF_tab_to_fasta_alignment_advanced.pl` korrigiert. Durch die Anwendung dieses *Perl*-Programms wurden zudem auch alle Insertionen und Deletionen sowie Längenmutationen aus der „SNV-Sequenz“ entfernt.

Die so generierten Sequenzen wurden anschließend zur Berechnung eines phylogenetischen Baumes mit dem Programm *MEGA* (Version 5.1, Tamura *et al.*, 2011) am Institut für Molekulargenetik (A. Bicker, T. Hankeln, Johannes Gutenberg-Universität, Mainz) verwendet. Hierfür wurde die *Neighbor-Joining*-Methode (Modell: Kimura-2-Parameter) und die *Maximum-Likelihood*-Methode (Modell: GTR, engl. *general time reversible*) mit jeweils 1000 *Bootstrap*-Replikaten eingesetzt.

---

<sup>7</sup><http://code.google.com/p/VCF-tab-to-fasta/>, 2009

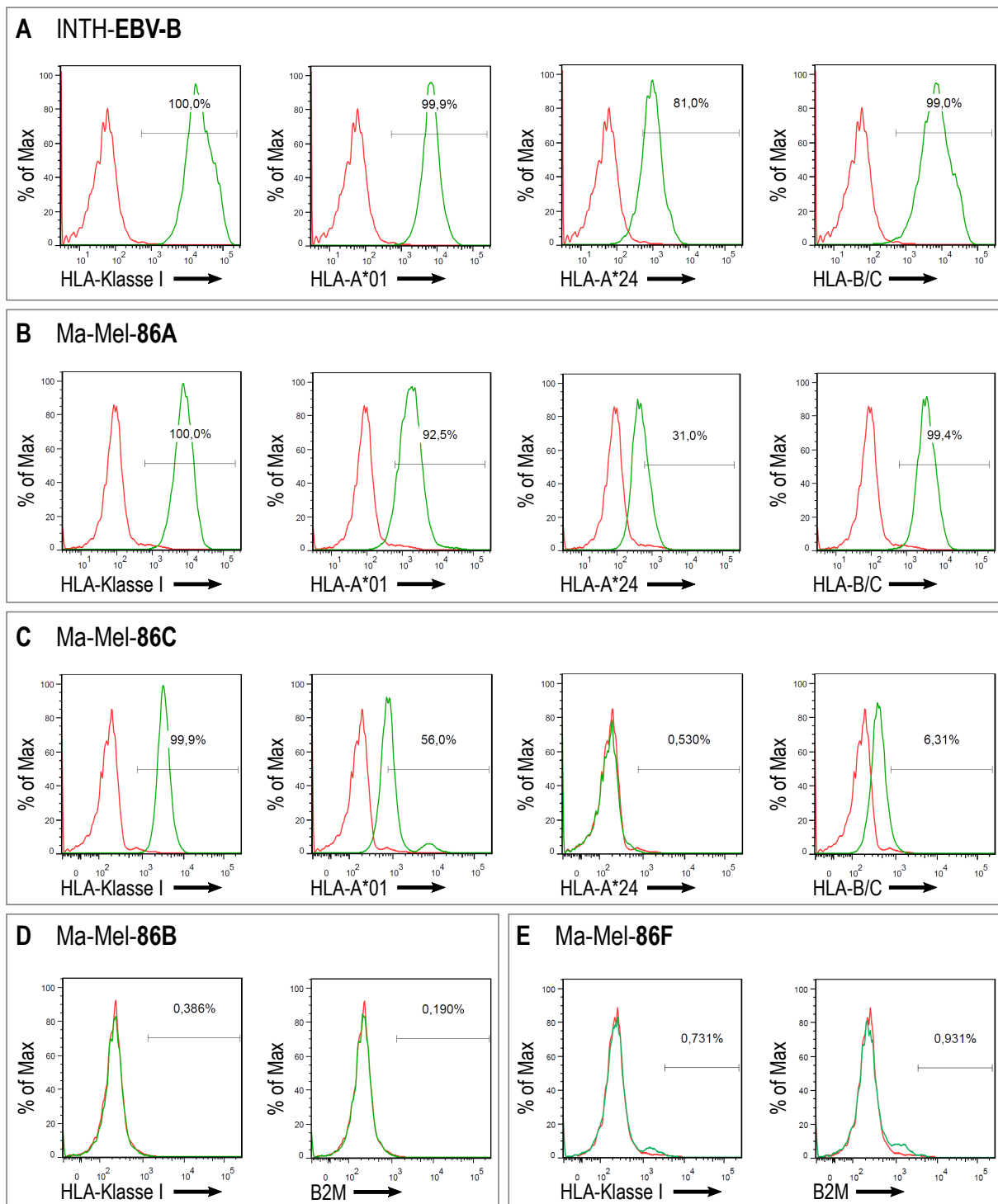


# 3 Ergebnisse

## 3.1 Vorbereitung der DNA- und RNA-Proben

### 3.1.1 Bestätigung der Identität der Ma-Mel-86-Zelllinien

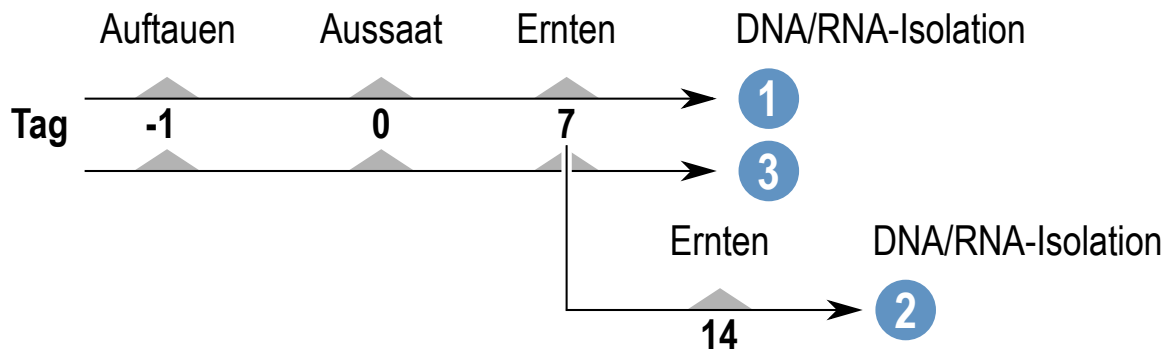
Vor der Isolation der genomischen DNA (gDNA) und Gesamt-RNA für die Generierung der Sequenzierbibliotheken wurde die Identität der kryokonservierten Zellen des Modells Ma-Mel-86 überprüft. Damit Zellen gleichen Ursprungs auch im nächsten Schritt für die Nukleinsäuregewinnung und für spätere Tests verwendet werden konnten, wurden Passagen gewählt, von denen noch mindestens fünf Aliquots des gleichen Einfrierdatums vorlagen. Außerdem sollten die Zellen zuvor mindestens zwanzigmal passagiert worden sein, um von einer möglichst reinen Population von Tumorzellen ausgehen zu können. Jeweils ein Aliquot wurde für die Validierung in Kultur genommen. Die vier adhärenenten Melanomzelllinien konnten morphologisch unter dem Auflichtmikroskop unterschieden werden (Daten nicht gezeigt). Zur Absicherung ihrer Identität wurde eine durchflusszytometrische Analyse der MHC-Klasse I-Expression durchgeführt (siehe **Abb. 3.1**). Alle Zelllinien wurden mit dem HLA-Klasse I-spezifischen Antikörper W6/32 untersucht. Bei den HLA-Klasse I-positiven Zelllinien Ma-Mel-86A, -86C und der autologen B-Zelllinie INTH-EBV-B wurde außerdem die Expression der HLA-A-Allele HLA-A\*01 und -A\*24 mithilfe von spezifischen Antikörpern und die Expression von HLA-B und -C durch gruppenspezifische Antikörper (B1.23.2) validiert. Die HLA-Klasse I-negativen Melanomzelllinien Ma-Mel-86B und -86F wurden hingegen auf die Präsenz von  $\beta$ 2-Mikroglobulin auf der Zelloberfläche überprüft. Die Zelllinien INTH-EBV-B und Ma-Mel-86A (**Abb. 3.1A** und **B**) exprimierten die beiden HLA-A-Allele HLA-A\*01 und HLA-A\*24 und wiesen eine hohe Expression von HLA-B und -C auf. Ma-Mel-86C (**Abb. 3.1C**) war hingegen negativ für HLA-A\*24 und zeigte eine schwächere Expression von HLA-B und -C. Bei den Linien Ma-Mel-86B und -86F (**Abb. 3.1D** und **E**) wurden weder HLA-Klasse I-Moleküle noch  $\beta$ 2-Mikroglobulin an der Zelloberfläche detektiert. Diese Ergebnisse bestätigten die bekannten HLA-Klasse I-Expressionsmuster der Modellzelllinien, wonach Ma-Mel-86A alle sechs HLA-Klasse I-Allele exprimierte, Ma-Mel-86C einen Haplotyp-Verlust aufwies und Ma-Mel-86B und -86F negativ für HLA-Klasse I-Expression waren (vgl. Einleitung, Abschnitt 1.1.1.3.3). Die Unterscheidung der beiden HLA-Klasse I-Verlustvarianten Ma-Mel-86B und -86F konnte anhand der Erkennung durch CSF2RA-reaktive T-Zellen erfolgen, da Ma-Mel-86F negativ für die Expression von CSF2RA war (Daten nicht gezeigt).



**Abb. 3.1: Validierung der Identität der Ma-Mel-86-Zelllinien.** Vor der Isolation von gDNA und Gesamt-RNA für die Sequenzierung der Zelllinien INTH-EBV-B (A), Ma-Mel-86A (B), -86C (C), -86B (D) und -86F (E) wurde die charakteristische Expression der HLA-Klasse I-Moleküle mittels Durchflusszytometrie überprüft. Die Beschriftungen der Histogramm-x-Achsen zeigen die Spezifität der Primärantikörper (HLA-Klasse I: W6/32, HLA-A\*01: HLA-A01/A11/A26+ (OneLamda, Canoga Park, USA), HLA-A\*24: C7709.A2.6, HLA-B/C: B1.23.2, B2M: B2M-02 (ThermoScientific, Schwerte)), die mithilfe FITC-markierter Sekundärantikörper detektiert wurden (grüne Kurven; *goat* F(ab')<sub>2</sub> Fragment Anti-mouse IgG + IgM (H+L)-FITC (Beckman Coulter GmbH, Krefeld)). Kontrollfärbungen ohne Primärantikörper sind in Rot dargestellt. Als Positivkontrolle wurden INTH-EBV-B-Zellen mit dem B2M-spezifischen Antikörper gefärbt (nicht gezeigt). Die Färbungen wurden durchgeführt, wie im Methodenteil in Abschnitt 2.2.1.11 beschrieben.

### 3.1.2 Isolation von genomischer DNA und Gesamt-RNA

Da für die geplanten Fragestellungen nur proteinkodierende Genomabschnitte relevant waren, sollten Exom- und Transkriptomsequenzierbibliotheken der Zelllinien des Modells generiert werden. Dabei wurden die Exome in Duplikaten und die Transkriptome in Triplikaten bearbeitet. Die Replikate stammten aus jeweils getrennten Kulturen bzw. Nukleinsäurepräparationen (vgl. **Abb. 3.2**), sodass im weiteren Sinn biologische Replikate generiert wurden. Hierfür wurden die kryokonservierten Zelllinien aufgetaut (Tag -1) und zunächst über Nacht in Kulturflaschen inkubiert. Am nächsten Tag (Tag 0) wurden die Zellen so eingestellt, dass innerhalb von sieben Tagen eine Konfluenz von 70-80% erreicht wurde. An Tag 7 wurden die Zellen geerntet und zum Teil für eine weitere Woche in Kultur ausgesät. Die restlichen Zellen wurden aufgeteilt und zur Isolation der gDNA bzw. Gesamt-RNA aufgeschlossen. Ebenso wurde an Tag 14 verfahren. Die Nukleinsäureisolate von Tag 7 und Tag 14 entsprachen den Replikaten 1 bzw. 2. Für das dritte Replikat der Transkriptomsequenzierung wurden neue Zellen (Passage und Einfrierdatum identisch) aufgetaut und das beschriebene Schema bis Tag 7 wiederholt. Im Gegensatz zur Verwendung von technischen Replikaten erlaubte es diese Vorgehensweise, biologische Varianz miteinzubeziehen. Die Präparationen der Nukleinsäuren wurden mit Standardkits durchgeführt (siehe Methodenteil, Abschnitt 2.2.2.4 und Abschnitt 2.2.2.5). Kontaminationen durch RNA in den DNA- bzw. DNA in den RNA-Isolaten wurden durch Behandlung mit RNase A bzw. DNase I minimiert.



**Abb. 3.2:** Zeitschema der Probengewinnung für die Exom- bzw. Transkriptomsequenzierungen der Ma-Mel-86-Zelllinien. Um Replikate aus unabhängigen Kulturen zu generieren, wurden die kryokonservierten Zelllinien aufgetaut (Tag 1) und ausgesät (Tag 0), so dass innerhalb von sieben Tagen eine 70-80%-ige Konfluenz erreicht wurde. An Tag 7 wurden die Zellen geerntet, um genomische DNA bzw. Gesamt-RNA zu isolieren (Replikat 1). Ein Teil der Zellen wurde weiterkultiviert, um an Tag 14 nach der ersten Aussaat weitere DNA- bzw. RNA-Proben zu präparieren (Replikat 2). Für RNA-Replikat 3 wurden neue Zellen aufgetaut, aus denen an Tag 7 RNA isoliert wurde, wie für das erste Replikat beschrieben.

Die Messung der DNA-Konzentration erfolgte am Spektralphotometer. **Tab. 3.1** zeigt die erzielten Ausbeuten. Normalisiert auf je  $5 \times 10^6$  Zellen wurden zwischen 30 und 58  $\mu\text{g}$  DNA isoliert.

Nach Rücksprache mit B. Ebner (Firma GENterprise GENOMICS, Mainz) wurden die RNA-Proben zusätzlich mittels Phenol-Chloroform-Extraktion und anschließender Ethanol-fällung aufgereinigt (siehe Methodenteil, Abschnitt 2.2.2.10.3), um RNase-Kontaminationen zu entfernen. Zur Bestimmung der RNA-Konzentration wurden die

### 3.1 Vorbereitung der DNA- und RNA-Proben

Proben nach der Aufreinigung mit dem Fluorometer und im Spektralphotometer gemessen (siehe Abschnitt 2.2.2.5). Pro  $1 \times 10^6$  Zellen konnten nach der Phenol-Chloroform-Extraktion zwischen 5 und 34  $\mu\text{g}$  Gesamt-RNA gemessen werden.

**Tab. 3.1: Isolation von genomischer DNA (gDNA) und Gesamt-RNA aus Ma-Mel-86-Zellen.** Entsprechend der Darstellung in Abb. 3.2 wurden aus allen fünf Zelllinien des Modells jeweils zwei getrennte gDNA- (Replikate 1 und 2) und drei getrennte RNA-Präparate (Replikate 1, 2 und 3) hergestellt. Die Messung der DNA-Konzentrationen erfolgte am Spektralphotometer. Gezeigt sind die gDNA-Ausbeuten je  $5 \times 10^6$  Zellen. Die RNA-Proben wurden zusätzlich durch Phenol-Chloroform-Extraktion aufgereinigt. Die Bestimmung der RNA-Konzentration erfolgte im Fluorometer und im Spektralphotometer mit jeweils zwei Verdünnungen. Die gezeigten Mittelwerte (normalisiert auf je  $1 \times 10^6$  Zellen) wurden aus den Einzelmessungen berechnet, die weniger als 15% voneinander abwichen.

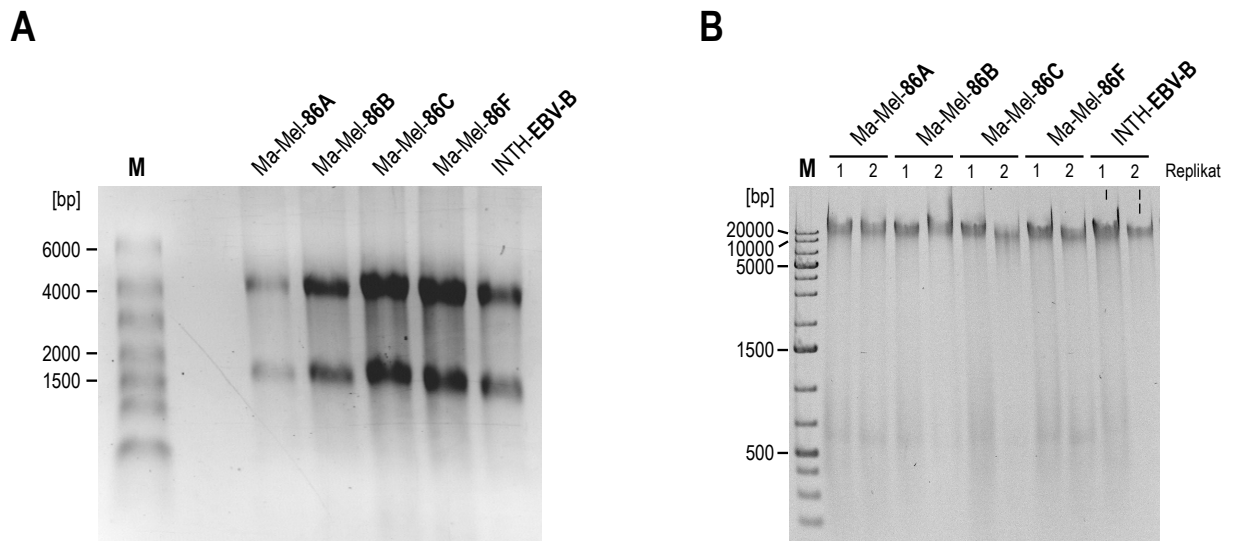
Zelllinie	Replikat	gDNA pro $5 \times 10^6$ Zellen [ $\mu\text{g}$ ]	Gesamt-RNA pro $1 \times 10^6$ Zellen [ $\mu\text{g}$ ]
Ma-Mel-86A	1	58,2	5,1
	2	55,1	14,1
	3	-	14,9
Ma-Mel-86B	1	38,7	11,7
	2	39,8	20,8
	3	-	11,1
Ma-Mel-86C	1	54,7	19,4
	2	30,7	26,1
	3	-	33,5
Ma-Mel-86F	1	33,3	11,1
	2	42,2	22,1
	3	-	28,1
INTH-EBV-B	1	34,0	15,3
	2	31,3	7,8
	3	-	11,4

#### 3.1.3 Überprüfung der Integrität der DNA- und RNA-Proben

Vor der Generierung der Sequenzierbibliotheken wurde die Integrität des isolierten Probenmaterials überprüft. Die durch Phenol-Chloroform-Extraktion aufgereinigten RNA-Proben wurden auf Formaldehydgelen (vgl. Methodenteil, Abschnitt 2.2.2.8.2) analysiert. Vor der Beladung des Gels wurden die RNA-Proben für 30 min auf Eis inkubiert und nach Zugabe des Ladepuffers für 5 min bei  $72^\circ\text{C}$  denaturiert (**Abb. 3.3A**). Im Falle einer Kontamination mit RNasen müsste nach der 30-minütigen Inkubation bereits eine Degradation der RNA zu erkennen sein. Die Analyse der gDNA-Proben erfolgte bei GENTERprise GENOMICS (Mainz, **Abb. 3.3B**).

Die elektrophoretische Auftrennung der Gesamt-RNA-Proben zeigt jeweils zwei deutliche Banden (**Abb. 3.3A**), die anhand ihrer Laufweiten als 28S- (ca. 5000 Nukleotide) bzw. 18S-rRNA (ca. 1.800 Nukleotide) zu identifizieren waren (Lodish *et al.*, 2000). Die kleineren rRNA-Untereinheiten (5S bzw. 5,8S) waren laut Herstellerangabe aufgrund des verwendeten Isolationsverfahrens nicht zu erwarten (*RNeasy® Mini Handbook*, QIAGEN, Juni 2012). Die Intensität der 28S-Bande war in allen Proben stärker als die der 18S-Bande. Somit lagen keine Anzeichen für eine Degradierung der RNA vor. Die isolierte gDNA zeigte sich

auf dem Agarosegel in allen Proben erwartungsgemäß als eine klare Bande mit einer Größe von mehr als 20.000 bp (**Abb. 3.3B**).



**Abb. 3.3:** Überprüfung der Probenintegrität. (A) Formaldehydgelelektrophorese der RNA-Proben nach Phenol-Chloroform-Extraktion. Exemplarisch dargestellt sind jeweils die ersten RNA-Präparationen der fünf Ma-Mel-86-Zelllinien (M: *RiboRuler™ High Range RNA Ladder*, Fermentas). Die RNA-Proben wurden vor der Zugabe des Ladepuffers für 30 min auf Eis inkubiert. (B) Agarosegelelektrophorese der gDNA-Proben. Gezeigt sind jeweils beide Replikate (1 und 2) der gDNA-Isolate aus den Ma-Mel-86-Zelllinien (M: *GeneRuler™ 1 kb Plus DNA Ladder*, Fermentas; Abb. zur Verfügung gestellt von B. Ebner, Firma GENTERprise GENOMICS, Mainz).

### Zusammenfassung

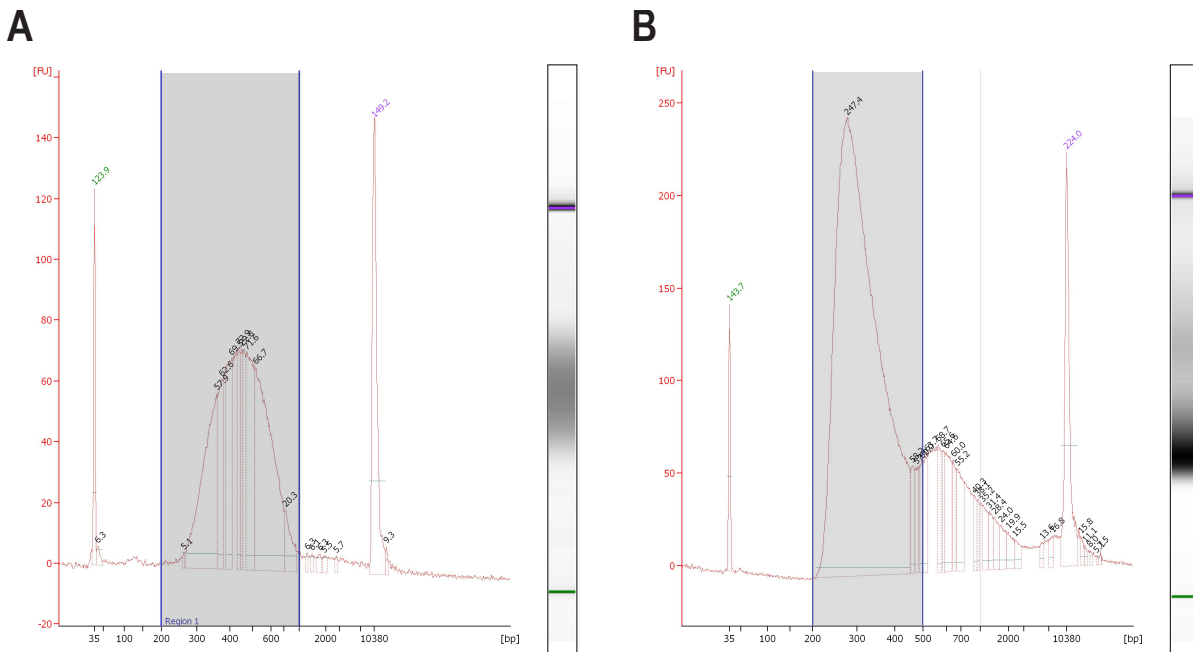
Nach der Verifizierung der Identität der fünf Zelllinien des Modells Ma-Mel-86 mittels Durchflusszytometrie wurden von jeder Zelllinie zwei gDNA- und drei Gesamt-RNA-Proben isoliert. Die Integrität des Probenmaterials wurde durch Gelelektrophorese bestätigt, so dass die Proben für die Generierung der Sequenzierbibliotheken freigegeben werden konnten.

## 3.2 Herstellung der Sequenzierbibliotheken und Qualitätskontrollen

Die Generierung der Sequenzierbibliotheken wurde von B. Ebner (Firma GENTERprise GENOMICS, Mainz) unter Verwendung der Kits *TruSeq™ DNA Sample Preparation* bzw. *TruSeq™ RNA Sample Preparation v2* von Illumina durchgeführt. Für die Exomanreicherung in der gDNA-Bibliothek wurde das Illumina-Kit *TruSeq™ Exome Enrichment* verwendet.

**Abb. 3.4** zeigt die Ergebnisse der Bioanalyserläufe der gepoolten Sequenzierbibliotheken. Die Fragmente der DNA-Bibliotheken lagen nach der Exomanreicherung in einem Größenbereich von 280 bis 800 bp (grau markiert in **Abb. 3.4A**). Die durchschnittliche Länge betrug 465 bp. Durch Abzug der Adaptersequenzen ( $2 \times 60$  bp) ergaben sich Insertlängen von 160 bis 680 bp.

Das Elektropherogramm der Transkriptombibliotheken (**Abb. 3.4B**) zeigte neben dem Hauptpeak einen zweiten kleineren Peak. Dieser entsprach RNA-Molekülen, die aufgrund von Sekundärstrukturen langsamer liefen als denaturierte RNA-Moleküle gleicher Größe (B. Ebner, pers. Mitteilung). Dieser Peak wurde bei der Größenbestimmung vernachlässigt. Somit ergab sich bei den Fragmenten der Transkriptombibliotheken eine durchschnittliche Länge von 320 bp (Bereich: 210 – 500 bp, grau markiert). Die Inserts wiesen dementsprechend Größen von 90 bis 380 bp auf.



**Abb. 3.4: Bioanalyzerrläufe zur Überprüfung der Sequenzierbibliotheken.** Die fertigen Sequenzierbibliotheken für die Exom- (**A**) und die Transkriptomsequenzierung (**B**) wurden gepoolt und auf dem Bioanalyzer analysiert. Gezeigt sind jeweils die Elektropherogramme (x-Achse: Molekülgröße in bp, y-Achse: Fluoreszenz in *arbitrary fluorescence units* (FU)) und das dazugehörige virtuelle Gelbild. Die grün bzw. violett beschrifteten Peaks entsprechen den Molekülgrößenmarkern. Die grauen Felder zeigen die für die Berechnung der durchschnittlichen Fragmentgrößen herangezogenen Bereiche. Die Exombibliotheken wiesen eine mittlere Länge von 465 bp auf. Bei den Transkriptombibliotheken lag die durchschnittliche Größe bei 320 bp. Abb. zur Verfügung gestellt von B. Ebner (Firma GENTERprise GENOMICS, Mainz).

## 3.3 Statistiken der Sequenzierungen

### 3.3.1 Sequenzierrohdaten

Die Sequenzierung der Exom- und Transkriptomproben erfolgte auf dem *HiSeq 2500* der Firma Illumina im Institut für Molekulargenetik (Johannes Gutenberg-Universität, Mainz). **Abb. 3.5** zeigt eine Übersicht der erzielten 100 bp *paired-end Reads* der einzelnen Proben.

In der Exomsequenzierung wurden im Durchschnitt  $65,8 \times 10^6$  *paired-end Reads* (Standardabweichung:  $11,8 \times 10^6$ ) generiert. Die Sequenzierung der Transkriptombibliotheken ergab im Mittel  $69,4 \times 10^6$  *paired-end Reads* (Standardabweichung:  $25,5 \times 10^6$ ). Um Expressionsanalysen der Melanomzelllinien im Vergleich zu gesunden Melanozyten durchführen



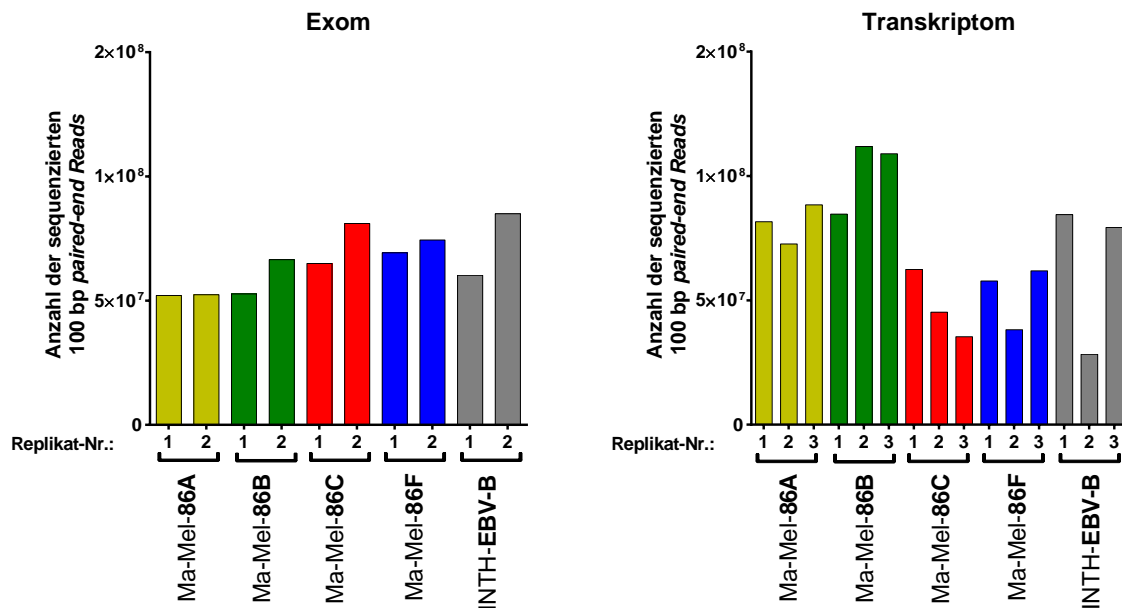


Abb. 3.5: Ergebnis der Exom- und Transkriptomsequenzierung. Die Sequenzierung der Exome (links) und Transkriptome (rechts) der Ma-Mel-86-Linien wurde auf einem *HiSeq 2500* (Illumina) am Institut für Molekulargenetik (Johannes Gutenberg-Universität, Mainz) durchgeführt. Es wurden durchschnittlich ca.  $68 \times 10^6$  *paired-end Reads* mit einer Länge von 100 bp generiert (durchschnittliche Anzahl *Reads* für Exome:  $65,8 \times 10^6$ , durchschnittliche Anzahl *Reads* für Transkriptome:  $69,4 \times 10^6$ ).

zu können, wurde ein im *Sequence Read Archive* (SRA) hinterlegter *RNA-Seq*-Datensatz hinzugezogen. Die Daten (*Accession*: SRX101662) stammten von RFP (*red fluorescent protein*)-transduzierten Melanozyten dreier Spender (Flockhart *et al.*, 2012). Die Auswahl dieser Daten erfolgte basierend auf der technischen Ähnlichkeit zu den eigenen Proben. So wurden die RNA-Isolation und die Generierung der Sequenzierbibliothek mit Kits durchgeführt, die auch für die Ma-Mel-86-Proben verwendet wurden. Die Sequenzierung erfolgte auf einem *HiSeq 2000* (Illumina). Es wurden ca.  $140 \times 10^6$  *paired-end Reads* mit einer Länge von 100 bp generiert. Dadurch war es möglich, die Melanozytendaten parallel zu den eigenen Transkriptomdaten zu prozessieren.

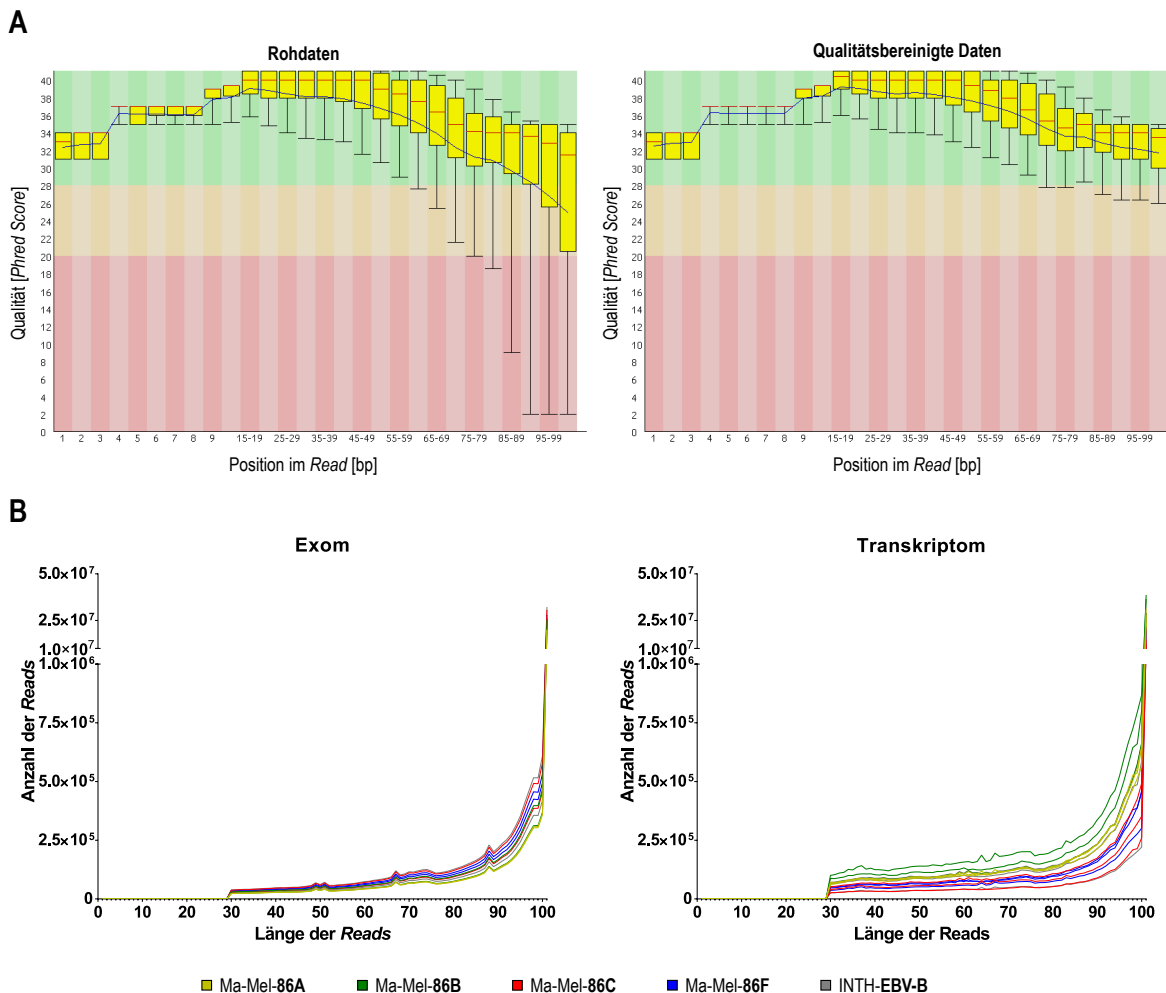
### 3.3.2 Aufarbeitung der Rohdaten

Da *Read*-Abschnitte mit schlechter Qualität und die Präsenz von Adaptersequenzen zu suboptimalen Ergebnissen bei den weiteren Analysen führen konnten, wurden die Rohdaten zunächst nachbearbeitet. Die Nachbearbeitung mit dem Programm *Trimmomatic* (Lohse *et al.*, 2012; Bolger *et al.*, 2014) umfasste

- das Herausfiltern von Adaptersequenzen,
- das Entfernen von Sequenzanfängen bzw. -enden mit schlechter Qualität,
- das Überprüfen der Sequenz auf ein starkes Absinken der Qualität und
- das Verwerfen von Sequenzen, die durch die Nachbearbeitung zu kurz geworden waren.

### 3.3 Statistiken der Sequenzierungen

Die verwendeten Parameter sind dem Methodenteil (Abschnitt 2.2.3.2) zu entnehmen. Der Erfolg der Qualitätsbereinigung wurde mit *FastQC* bzw. *fastx-toolkit* überprüft. **Abb. 3.6A** zeigt einen beispielhaften Vergleich der Qualitätswerte pro Base vor und nach der Aufarbeitung von 4 Millionen *Reads* des ersten Exomreplikats von Ma-Mel-86A. In den Rohdaten war ein Abfall der Qualität am 3'-Ende der *Reads* zu beobachten (blau: Mittelwerte, rot: Mediane). Dies konnte durch die Qualitätsbereinigung behoben werden, so dass die Mittelwerte in den grünen *Phred-Score*-Bereich verschoben wurden. Dieser Bereich markiert Qualitätswerte von über 28, was einer Fehlerwahrscheinlichkeit von weniger als 0,16% entspricht. Die *Whisker* umfassen den Wertebereich zwischen dem 10%- und dem 90%-Quantil der Daten. Auch hier konnte die Qualitätsbereinigung eine Anhebung des unteren *Whiskers* in den orangefarbenen Bereich bewirken, so dass hier mit einer Fehlerwahrscheinlichkeit von weniger als 1% (*Phred-Score* 20) zu rechnen war.



**Abb. 3.6: Qualitätsbereinigung der Rohdaten. (A) Qualitätswerte je Base.** Die Kastengrafiken (generiert mit *FastQC*) zeigen die Qualitätswerte je Position im *Read* exemplarisch für 4 Millionen *Reads* des ersten Exomreplikats von Ma-Mel-86A vor (links) und nach (rechts) der Bearbeitung mit *Trimmomatic* (Mittelwerte: blau, Mediane: rot, Interquartilumfang: gelbe Kästen, 10%- bzw. 90%-Quantil: unterer bzw. oberer *Whisker*). Die *Phred-Score*-Skala (y-Achse) ist eingeteilt in schlechte (rot), gute (orange) und sehr gute (grün) Qualitätswerte. Durch die Qualitätsbereinigung wurden *Read*-Abschnitte mit zu niedriger Qualität entfernt. **(B) Verteilung der Read-Längen.** Das Entfernen von Adaptersequenzen und *Read*-Abschnitten mit schlechter Qualität führte zu einer Verkürzung der *Reads*. Zu kurze *Reads* (<30 bp) wurden verworfen. Die Graphen zeigen die Längenverteilung für die kompletten Datensätze aller Exom- (links, jeweils zwei Replikate) und Transkriptomsequenzierungen (rechts, jeweils drei Replikate).

Insgesamt blieben im Mittel 99,9% der Exom-*Reads* und 95,4% der Transkriptom-*Reads* nach der Nachbearbeitung erhalten. Davon waren etwa 1,8% der Exom- und 4,3% der Transkriptom-*Reads* nicht mehr gepaart. Diese *unpaired Reads* entstanden, wenn die Qualität einer der beiden *Reads* nicht hoch genug war oder die ursprünglich gepaarten *Reads* vollständig überlappten und somit sequenzidentisch waren. Im letzteren Fall wurde einer der beiden *Reads* verworfen, da sich durch den zweiten *Read* kein Informationsgewinn ergab.

Anhand der mit *fastx-toolkit* berechneten Qualitätsstatistiken wurde außerdem die Verteilung der *Read*-Längen nach der Bearbeitung mit *Trimmomatic* bestimmt (**Abb. 3.6B**). *Reads*, deren Länge durch die Aufbereitung unter 30 bp fiel, wurden automatisch verworfen. Durchschnittlich 87,4% der bereinigten *Reads* aus der Exomsequenzierung (**Abb. 3.6B**, links) hatten noch eine Länge zwischen 90 und 100 bp. Bei der Transkriptomsequenzierung (**Abb. 3.6B**, rechts) lagen im Mittel 82,5% der *Reads* in diesem Längenbereich. Auf ein sogenanntes *Hard-Clipping*, bei dem in allen *Reads* eine definierte Anzahl von Basen am Anfang oder Ende der Sequenz abgeschnitten werden, wurde verzichtet, da das anschließende *Mapping* zufriedenstellende Ergebnisse lieferte (siehe Abschnitt 3.3.3) und somit nicht weitere Sequenzinformationen verworfen werden mussten.

### 3.3.3 Mapping an die Referenz

#### 3.3.3.1 Exomsequenzierung

Das *Mapping* der aufgearbeiteten *Reads* an die Referenzsequenz des Humangenoms hg19 (Februar 2009, UCSC) erfolgte für die Exomsequenzierung mithilfe des Programms *Bowtie2* (Langmead & Salzberg, 2012). Die Referenz konnte als bereits angelegter *Bowtie2*-Index von der Webseite der *Mapping*-Software heruntergeladen werden (siehe Methodenteil, Abschnitt 2.2.3.3.1). Das *Mapping* wurde für alle Datensätze (fünf Zelllinien mit je zwei Replikaten) separat durchgeführt, wobei jeweils gepaarte und ungepaarte *Reads* gleichzeitig aligniert wurden. Alle Standardparameter wurden beibehalten, außer der Vorgabe, dass je *Read* nur ein *Alignment* an die Referenzsequenz dokumentiert werden sollte. Somit erfolgte das *Mapping* im sensitiven *end-to-end*-Modus. Hierbei ist der Schwellenwert für ein erfolgreiches *Alignment* eine Funktion der *Read*-Länge  $L$ :

$$\text{Schwellenwert} = -0,6 + (-0,6) * L$$

Für jedes *Alignment* werden bei Abweichungen zur Referenz negative Strafpunkte (*Mismatch*: 6, *Gap*: -5, *Gap*-Verlängerung: -3) vergeben und aufsummiert. Nur wenn die Summe den obigen Schwellenwert übersteigt, wird das *Alignment* als erfolgreich eingestuft und dokumentiert. Somit wären für einen *Read* mit einer Länge von 100 bp beispielsweise zehn Fehlpaarungen denkbar, wodurch gegebenenfalls das nicht durchgeführte *Hard-Clipping* von 10 bp kompensiert werden würde.

Die *Mappings* wurden anschließend für die weiteren Analysen prozessiert (siehe Abschnitt 2.2.3.3.1) und ihre Qualität mithilfe von *SAMtools* und *bedtools* überprüft.

**Tab. 3.2** zeigt eine Übersicht zur Einschätzung des *Mapping*-Erfolgs.

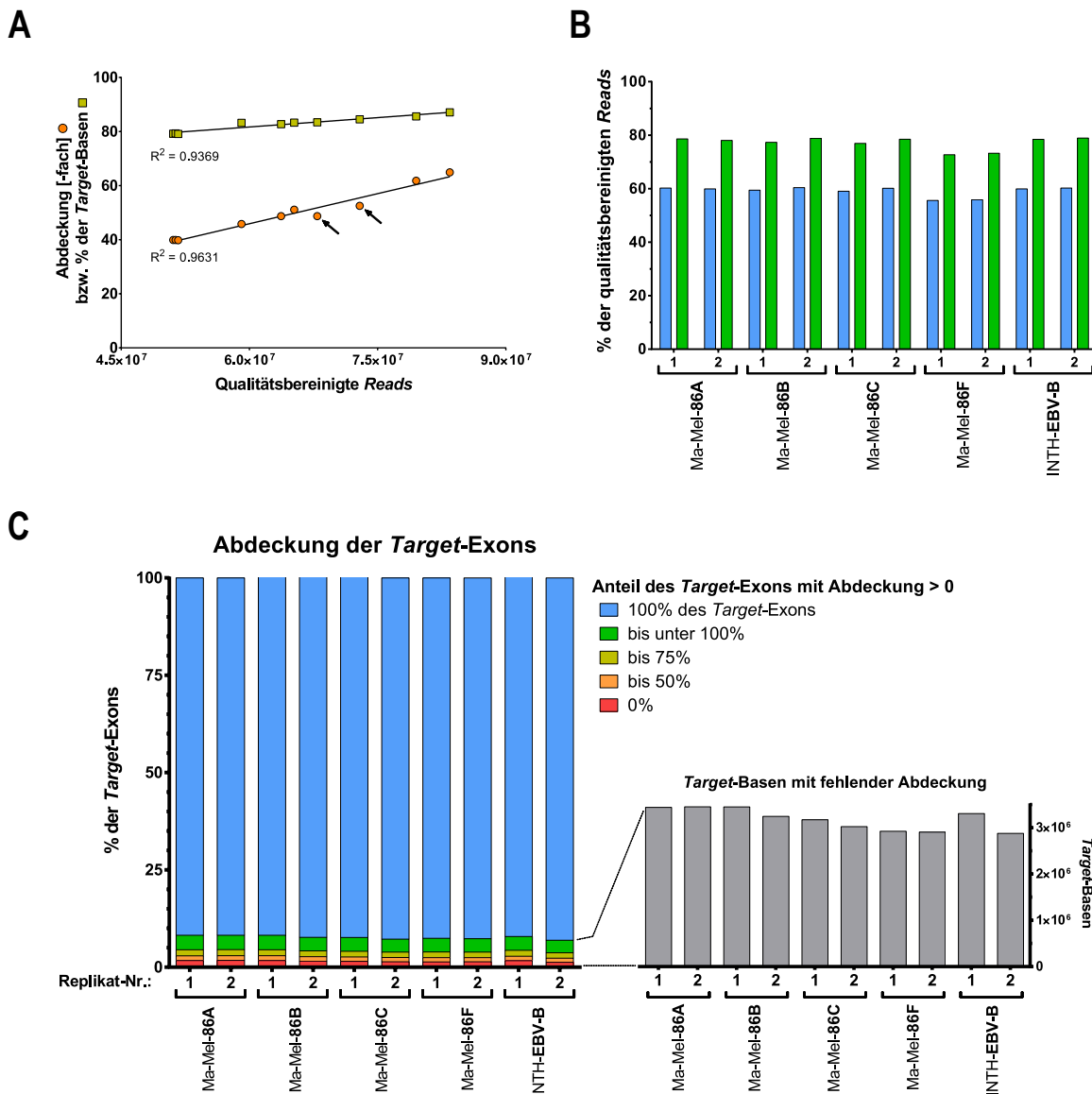
**Tab. 3.2:** Übersicht der *Mapping*-Statistiken der Exomsequenzierung. Das *Alignment* der qualitätsbereinigten *Reads* der Exomsequenzierung an die genomische Referenz hg19 (*H. sapiens*, UCSC) erfolgte mit *Bowtie2*. Die Tabelle zeigt eine Reihe von Kennziffern zur Beurteilung des *Mapping*-Erfolgs. Die gelisteten Angaben wurden aus den bereits prozessierten BAM-Dateien berechnet, die für die anschließenden Analysen verwendet wurden.

Zelllinie	Replikat	Anzahl der <i>Reads</i> nach Aufarbeitung ( <i>Input</i> )	Zugeordnete <i>Reads</i> [% von <i>Input</i> ]	Zugeordnete <i>Read</i> -Paare [% von zugeordnet]	Abdeckung [-fach]	Anteil der Basen mit $\geq 15$ -facher Abdeckung [%]
Ma-Mel-86A	1	51.085.483	99,2	97,3	39,9	79,2
	2	51.368.377	99,1	97,2	39,9	79,3
Ma-Mel-86B	1	51.665.217	99,3	97,4	39,8	79,1
	2	65.238.045	99,3	97,5	51,1	83,3
Ma-Mel-86C	1	63.704.508	99,2	97,4	48,7	82,7
	2	79.512.089	99,4	97,5	61,8	85,6
Ma-Mel-86F	1	67.931.163	99,0	97,1	48,7	83,4
	2	72.904.674	99,0	97,0	52,5	84,5
INTH-EBV-B	1	59.068.661	99,2	97,4	45,8	83,2
	2	83.437.818	99,3	97,5	64,9	87,1
<b>Durchschnitt</b>		<b>64.591.604</b>	<b>99,2</b>	<b>97,3</b>	<b>49,3</b>	<b>82,7</b>

Im Durchschnitt konnten 99,2% aller qualitätsbereinigten *Reads* zugeordnet werden. Bei den zugeordneten *Reads* handelte es sich bei durchschnittlich 97,3% um *Read*-Paare, von denen beide *Reads* zugeordnet wurden. Die mittlere Abdeckung der Exomsequenzen, die angereichert werden sollten (*Target*-Exons), lag zwischen 39,9 und 64,9 *Reads* (Mittelwert: 49,3), wobei durchschnittlich 82,7% der Basen der *Target*-Exons mit mindestens 15 *Reads* abgedeckt waren. Die Anzahl der qualitätsbereinigten *Reads*, die an *Bowtie2* für das *Mapping* übergeben wurden, korrelierte mit den erzielten Werten für die durchschnittliche Abdeckung (**Abb. 3.7A**, orange,  $R^2 = 0,9369$ ) bzw. für den Anteil der Basen mit mindestens 15-facher Abdeckung (**Abb. 3.7A**, gelb,  $R^2 = 0,9631$ ). Nur die beiden Replikate von Ma-Mel-86F wiesen trotz der relativ hohen Anzahl an *Input*-*Reads* im Vergleich eine marginal niedrigere Abdeckung auf (markiert mit Pfeilen in **Abb. 3.7A**). Insgesamt lagen jedoch alle Datensätze über dem von Stratton *et al.* (2009) abgegebenen Schätzwert, dass zur Detektion aller somatischen Mutationen eine mehr als 30-fache Abdeckung des Tumorgenoms und des normalen Genoms des gleichen Individuums nötig sei.

Der Erfolg der Exomanreicherung ließ sich anhand des Anteils der qualitätsbereinigten *Reads* einschätzen, denen *Target*-Exons zugeordnet werden konnten. So lag die prozentuale Anreicherung zwischen 56 und 60% (blaue Säulen in **Abb. 3.7B**). Die Sonden der Exomanreicherung sind jedoch so konstruiert, dass laut Herstellerangabe auch angrenzende Regionen angereichert werden. Bezieht man daher um 150 bp vor und nach den *Target*-Exons verlängerte Bereiche mit ein, lag die prozentuale Anreicherung bei 73-79% (grüne Säulen in **Abb. 3.7B**). Diese Anreicherungseffizienzen waren etwas niedriger als die vom Kit-Hersteller angestrebten Zielwerte ( $>65\%$  bzw.  $>80\%$ ). Daher wurde die Abdeckung

der einzelnen *Target*-Exons überprüft, woraus hervorging, dass nicht alle *Target*-Exons vollständig bzw. überhaupt sequenziert wurden (**Abb. 3.7C**). Da es sich um weibliches Patientenmaterial handelte, wurden die *Target*-Exons auf dem Chromosom Y hier nicht berücksichtigt. Von den restlichen *Target*-Exons (200.786 Exons auf Chromosom 1-22 und X) waren nur ca. 92% über die komplette Exonlänge abgedeckt (**Abb. 3.7C** links, blau). Für durchschnittlich 3,2 Millionen (5,1%) der ca. 62 Millionen *Target*-Basen lagen im Gegensatz dazu gar keine Sequenzinformationen vor (**Abb. 3.7C** rechts).



**Abb. 3.7: Erfolg des Exom-Mappings.** (A) Die Anzahl der qualitätsbereinigten *Reads* korrelierten linear mit der jeweils erzielten mittleren Abdeckung (orange) bzw. mit dem Anteil der *Target*-Basen mit mindestens 15-facher Abdeckung (gelb). Mit Pfeilen markiert sind die beiden Ma-Mel-86F-Replikate, bei denen trotz hoher *Read*-Anzahl eine im Vergleich geringere Abdeckung erzielt wurde. (B) Gezeigt ist die prozentuale Anreicherung der *Target*-Exons (blau) und der *Target*-Exons erweitert um  $\pm 150$  bp (grün). (C) Der linke Graph gibt an, welcher Anteil der *Target*-Exons in den Exomsequenzierungen vollständig (blau), nicht vollständig (grün, gelb und orange) bzw. gar nicht (rot) sequenziert wurden. Im rechten Graphen (grau) ist die Anzahl der *Target*-Basen mit fehlender Abdeckung für die einzelnen Datensätze dargestellt.

Bei Betrachtung aller zehn Datensätze fehlte für insgesamt 5.478 verschiedene *Target*-Exons die Abdeckung komplett (durchschnittlich etwa 3.100 Exons je Sequenzierung), wobei

1.052 Exons in keinem der zehn Datensätze sequenziert worden waren. Möglicherweise war die Effizienz mancher Anreicherungs sonden nicht optimal. Diese Beobachtungen wurden insbesondere bei der Detektion der somatischen Mutationen berücksichtigt. Für 15 *Target-Exons* wurden nur in den Replikaten der Keimbahnkontrolle INTH-EBV-B *Alignments* dokumentiert. Hierbei handelte es sich möglicherweise um somatische Deletionen, die allerdings nicht näher untersucht wurden.

#### 3.3.3.2 Transkriptomsequenzierung

Die *Reads* der Transkriptomsequenzierung sollten ebenfalls an die genomische Referenz aligniert werden. Um dabei Spleißstellen berücksichtigen zu können, wurde für das *Mapping TopHat* (Trapnell *et al.*, 2009) verwendet. Diese *Mapping*-Software basiert bei der Zuordnung der *Reads* auch auf *Bowtie2*, weswegen auf denselben Referenzindex (hg19, siehe Abschnitt 2.2.3.3.1) zurückgegriffen werden konnte. Zusätzlich werden von *TopHat* jedoch auch Spleißstellen erkannt und dokumentiert. Anders als beim *Mapping* der Exomdaten wurden gepaarte und ungepaarte *Reads* getrennt aligniert und die Ergebnisse später zusammengefasst (siehe Methodenteil, Abschnitt 2.2.3.3.2). Abweichungen von den Standardparametern wurden nur für die erwarteten Abstände zwischen den gepaarten *Reads* vorgenommen. Die Werte hierfür wurden anhand der Insertlängen der Sequenzierbibliotheken abgeschätzt (vgl. Abschnitt 3.2) und sind dem Methodenteil zu entnehmen. Aufgrund des internen Aufrufens von *Bowtie2* während des *Mappings* mit *TopHat* werden erfolgreiche *Alignments* bestimmt, wie zuvor beschrieben (siehe Abschnitt 3.3.3.1). *TopHat* verwirft abschließend allerdings *Alignments* mit mehr als zwei *Mismatches* (Standardwert).

Nach Prozessieren der generierten Daten für die weiteren Analysen (siehe Methodenteil, Abschnitt 2.2.3.3.2) wurde ebenso der Erfolg der *Mappings* überprüft (vgl. **Tab. 3.3**).

Mittels *TopHat* konnten im Mittel 96,6% der qualitätsbereinigten Transkriptom-*Reads* an die genomische Referenz aligniert werden. Durchschnittlich 91,2% der zugeordneten *Reads* konnten dabei eindeutig zugeordnet werden und 78,3% der alignierten *Reads* gehörten zu eindeutig zugeordneten *Read*-Paaren.

Die Bestimmung der mittleren Abdeckung im Fall der Transkriptomsequenzierung ist nur bedingt aussagekräftig, da durch unterschiedliche Expressionslevel keine gleichmäßige Abdeckung zu erwarten ist. Das ENCODE-Konsortium gibt jedoch eine Empfehlung<sup>8</sup> für die Abschätzung der Abdeckung pro Nukleotid einer mRNA, die mit nur einer Kopie pro Zelle vorliegt und eine Länge von 1.500 Nukleotiden aufweist:

$$\text{Abdeckung} = \frac{\text{Anzahl der sequenzierten Nukleotide}}{\frac{\text{Anzahl der mRNA-Moleküle pro Zelle}}{1.500 \text{ Nukleotide pro mRNA}}}$$

---

<sup>8</sup>ENCODE Consortium: *Standards, Guidelines and Best Practices for RNA-Seq*. V1.0. (2011), [http://encodeproject.org/ENCODE/protocols/dataStandards/ENCODE\\_RNAseq\\_Standards\\_V1.0.pdf](http://encodeproject.org/ENCODE/protocols/dataStandards/ENCODE_RNAseq_Standards_V1.0.pdf) (zuletzt aufgerufen am 01.07.2014)

Um die Abdeckung durch die tatsächlich zugeordneten *Reads* zu bestimmen, wurde die obige Formel in folgender abgewandelter Form angewandt:

$$\text{Abdeckung} = \frac{\frac{\text{Anzahl der zugeordneten Nukleotide}}{\text{Anzahl der mRNA-Moleküle pro Zelle}}}{1.500 \text{ Nukleotide pro mRNA}}$$

Hierzu wurde die Anzahl der tatsächlich zugeordneten Basen aus den *Mappings* mithilfe des eigenen *Perl*-Skripts `mappedRead_LengthDistribution.pl` (siehe Anhang, Abschnitt A.3.2.1) ausgelesen und aufsummiert. Ebenso wurde für die eindeutig zugeordneten Basen verfahren. Für die Anzahl der mRNA-Moleküle pro Zelle wurde ca. 300.000 Moleküle (Velculescu *et al.*, 1999) angenommen. Somit ergab sich für die Modellzelllinien eine durchschnittliche Abdeckung von 14,1 durch alle zugeordneten *Reads* und eine Abdeckung von 12,2 durch die eindeutig zugeordneten *Reads*. Der Vergleichsdatensatz des SRA wurde etwas tiefer sequenziert und wies eine 25,7- bzw. 24,4-fache Abdeckung durch die zugeordneten bzw. eindeutig zugeordneten *Reads* auf.

**Tab. 3.3: Übersicht der Mapping-Statistiken der Transkriptomsequenzierungen.** Das *Alignment* der qualitätsbereinigten *Reads* der Transkriptomsequenzierung an die genomische Referenz hg19 (*H. sapiens*, UCSC) erfolgte mit *TopHat*. Die gelisteten Kennziffern wurden aus den bereits prozessierten BAM-Dateien kalkuliert und dienen der Beurteilung des *Mapping*-Erfolgs. Die Abschätzung der Abdeckung erfolgte in Anlehnung an die Empfehlung des ENCODE-Konsortiums (Details im Text).

Zelllinie	Replikate	Anzahl der <i>Reads</i> nach Aufarbeitung ( <i>Input</i> )	Zugeordnete <i>Reads</i> [% von <i>Input</i> ]	Eindeutig zugeordnete <i>Reads</i> [% von zugeordnet]	Eindeutig zugeordnete <i>Read</i> -Paare [% von zugeordnet]	Abgeschätzte Abdeckung [-fach]	
						insgesamt	eindeutig
Ma-Mel-86A	1	77.792.234	96,8	92,0	79,9	16,5	14,5
	2	68.766.926	96,9	91,5	77,5	14,5	12,7
	3	84.137.830	96,9	92,2	79,8	18,0	15,9
Ma-Mel-86B	1	80.666.672	97,0	91,7	79,6	17,4	15,1
	2	105.562.290	96,9	91,3	78,1	22,4	19,5
	3	104.192.306	97,2	91,6	79,5	22,5	19,6
Ma-Mel-86C	1	59.468.122	97,2	91,9	79,4	12,7	11,1
	2	42.750.918	97,3	91,6	77,6	9,1	7,9
	3	33.877.483	97,2	91,9	79,7	7,3	6,4
Ma-Mel-86F	1	55.071.615	96,7	91,8	79,2	11,8	10,3
	2	36.056.083	96,8	91,4	78,1	7,6	6,6
	3	59.223.897	96,9	91,7	79,9	12,8	11,1
INTH-EBV-B	1	80.424.119	95,5	89,5	76,2	17,4	14,4
	2	26.741.378	95,0	88,8	74,4	5,7	4,7
	3	75.696.328	95,1	89,0	76,1	16,4	13,5
<b>Durchschnitt</b>		66.028.547	96,6	91,2	78,3	14,1	12,2
<b>Primäre Melanozyten</b> ( <i>Accession</i> : SRX101662)		136.128.698	95,4	94,3	77,5	25,7	24,4

### Zusammenfassung

Bei der Sequenzierung der Exome (jeweils Duplikate) und Transkriptome (jeweils Triplikate) der fünf Zelllinien des Modells wurden zwischen 28 und 112 Millionen *paired-end Reads* mit einer Länge von 100 bp generiert. Nach der Qualitätsbereinigung der Rohdaten wurden die *Reads* an die genomische Referenz hg19 (*H. sapiens*, UCSC) aligniert. Der Erfolg des *Mappings* konnte anhand des Anteils der alignierten *Reads*, der erzielten Abdeckung und der Anreicherungseffizienzen (Exombibliotheken) bestätigt werden. Die Exome wurden im Durchschnitt mit einer 49,3-fachen Abdeckung sequenziert, wobei etwa 82,7% der *Target*-Basen mindestens 15-fach abgedeckt waren. Es fiel jedoch auf, dass einige *Target*-Exons nicht sequenziert wurden. In den Transkriptom-*Mappings* wurden im Mittel 91,2% der alignierten *Reads* mit einer durchschnittlichen abgeschätzten Abdeckung von 12,2 Reads je Base eindeutig zugeordnet. Für die weiteren Auswertungen musste allerdings insbesondere auf die großen Unterschiede in der Anzahl der *Reads* der einzelnen Transkriptomdatensätze Rücksicht genommen werden.

## 3.4 Bioinformatische Auswertungen

Der Fokus der bioinformatischen Auswertungen dieser Arbeit lag auf der Etablierung einer Strategie zur Identifizierung von potentiell immunogenen Peptiden, die somatische nicht-synonyme Punktmutationen enthalten. Aufgrund dessen wurde in den Sequenzierdaten der Zelllinien des Modells primär nach Mutationen gesucht. Die gefundenen Mutationen wurden dabei auch zur Kalkulation der phylogenetischen Beziehung zwischen den Zelllinien herangezogen. Außerdem wurde die fehlende Expression von  $\beta$ 2-Mikroglobulin in Ma-Mel-86B und -86F näher analysiert. Weitere Statistiken dienten der Einschätzung der Eignung der generierten Daten und der gewählten Parameter.

### 3.4.1 Vergleich der Replikate

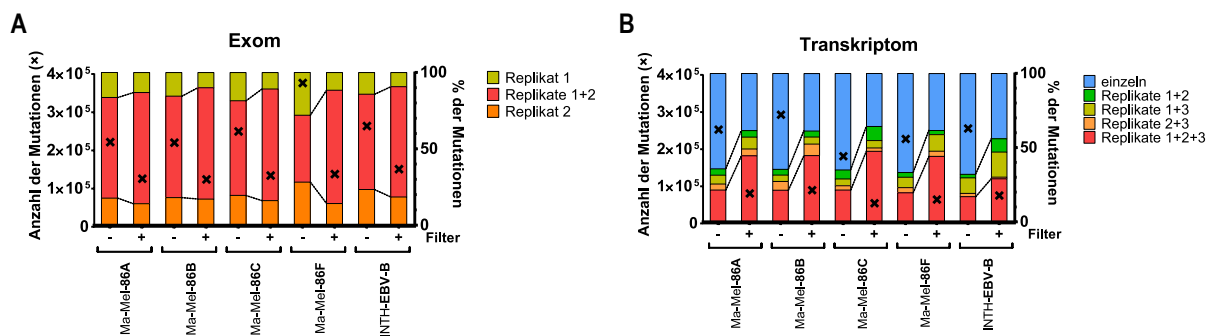
#### 3.4.1.1 Übereinstimmung der gefundenen Mutationen

Die Detektion von Abweichungen in den Sequenzierdaten der Ma-Mel-86-Zelllinien zur Referenz hg19 (UCSC), das *Variant Calling*, erfolgte ausgehend von den prozessierten *Mappings* (siehe Methodenteil, Abschnitt 2.2.3.3.1 und Abschnitt 2.2.3.3.2) mithilfe der Programmpakete *SAMtools* und *bcftools* (vgl. Methodenteil, Abschnitt 2.2.3.6.1). Die einzelnen Replikate wurden weiterhin getrennt bearbeitet. Zur Selektion von sicheren Mutationen wurden die detektierten *Variants* anschließend basierend auf der Qualität und der Abdeckung der Referenzabweichung gefiltert. Als Schwellenwerte wurden hierbei eine Fehlerwahrscheinlichkeit von  $\leq 5\%$  (entspricht einem *Phred-Score* von  $\geq 13$ ) und eine Abdeckung durch mindestens fünf (Transkriptom) bzw. zehn Reads (Exom) gewählt. Für die



*RNA-Seq*-Daten wurde eine niedrigere Soll-Abdeckung eingesetzt, um die Übereinstimmung zwischen den Transkriptomreplikaten zu verbessern und um die insgesamt weniger tiefe Sequenzierung einzelner Datensätze zu kompensieren (vgl. Abschnitt 3.3.3.2). Betrachtet man die relative Überlappung (**Abb. 3.8**) der gefundenen Mutationen vor (-) und nach (+) Anwenden des beschriebenen Filters, zeigte sich für alle Linien in den qualitätsgefilterten *Variants* eine größere Übereinstimmung zwischen den Replikaten. Bis zu 74% der Mutationen der Exomsequenzierungen waren in beiden Replikaten (rot) detektierbar.

Bei den Transkriptomsequenzierungen wurden durchschnittlich etwa 61% der Mutationen in mindestens zwei Replikaten (grün, gelb, orange und rot) gefunden. In drei der fünf Linien war die Ähnlichkeit zwischen den RNA-Replikaten 1 und 3 (gelb), welche jeweils an Tag 7 der Kultur (vgl. **Abb. 3.2**) generiert wurden, größer als die Übereinstimmungen mit dem Replikate 2 (Tag 14, grün bzw. orange). Im Durchschnitt wurden in den Exomen ca. 135.000 und in den Transkriptomen ca. 73.000 verschiedene Mutationen mit guter Qualität und Abdeckung gefunden (**Abb. 3.8**, linke y-Achsen). Erwartungsgemäß war die Anzahl der Mutationen in den Transkriptomen niedriger als in den Exomen, da nicht alle proteinkodierenden Bereiche mit Mutationen tatsächlich exprimiert werden.

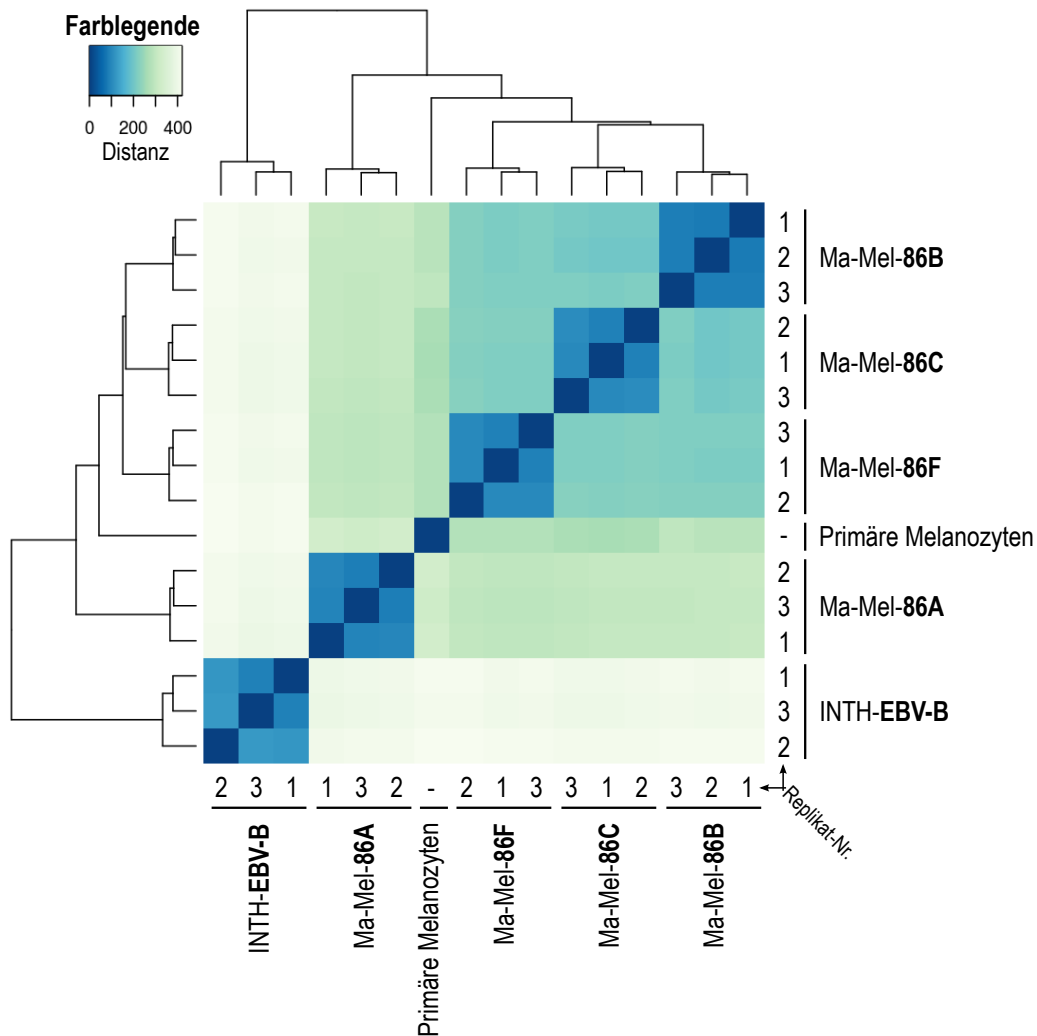


**Abb. 3.8: Übereinstimmung der gefundenen Mutationen in den Replikaten.** Nach der Detektion der Abweichungen zur Referenzsequenz hg19 (UCSC) wurden die gefundenen Mutationen basierend auf der Qualität und der Abdeckung gefiltert. Die Säulen zeigen die prozentuale Überlappung der Mutationen im Exom (**A**) und Transkriptom (**B**) vor (-) und nach (+) dem Filtern (rechte y-Achsen). Auf den linken y-Achsen sind die absoluten Mutationszahlen (x) für jede Zelllinie aufgetragen. Durch Anwenden des Filters konnte die Übereinstimmung zwischen den Replikaten gesteigert werden (gestrichelte Verbindungslinien).

### 3.4.1.2 Ähnlichkeit der Expressionsmuster

Um die Ähnlichkeit der Replikate der Transkriptomsequenzierungen auch anhand der Genexpressionsmuster einschätzen zu können, wurde mit den *RNA-Seq*-Daten, wie im Handbuch zum *Bioconductor*-Paket *DESeq* (Anders & Huber, 2012) beschrieben, eine Distanzmatrix berechnet. Dazu wurde aus den Transkriptom-*Mappings* mithilfe des *Bioconductor*-Pakets `summarizeOverlaps` eine Tabelle generiert, die für jeden Datensatz Auskunft gab über die Anzahl der *Reads*, die jedem Gen zugeordnet werden konnten. Diese Werte wurden anschließend nach Anders & Huber (2010) normalisiert, varianzstabilisiert (Huber *et al.*, 2003) und dann für die Berechnung des euklidischen Abstands zwischen den einzelnen Datensätzen verwendet. In der Darstellung der Distanzmatrix als *Heatmap* (**Abb. 3.9**) ist die Ähnlichkeit der Datensätze zueinander durch die Farbskalierung

(blau: hohe Ähnlichkeit, weiß: geringe Ähnlichkeit) und durch die hierarchische Clusterbildung in den Dendrogrammen visualisiert. Im Vergleich wurden auch die Distanzen zum Expressionsprofil der Melanozytendaten (*Accession*: SRX101662) berechnet.



**Abb. 3.9: Distanzmatrix der Genexpressionsprofile.** Die Anzahl der *Reads*, die einem Gen zugeordnet werden konnten, diente als Maß für die Höhe der Expression dieses Gens. Nach der Normalisierung und Varianzstabilisierung dieser Expressionswerte wurde der euklidische Abstand zwischen den einzelnen Datensätzen ermittelt. Die *Heatmap* visualisiert die Ähnlichkeiten der Expressionsprofile der Zelllinien bzw. der Replikate anhand der Farbskalierung (blau: hohe Ähnlichkeit, weiß: geringe Ähnlichkeit) und anhand der Dendrogramme. Als Kontrolle wurden die *RNA-Seq*-Daten der primären Melanozyten (SRA-Accession: *SRX101662*) miteinbezogen.

Bei allen fünf Modellzelllinien war die Ähnlichkeit zwischen den Replikaten deutlich höher als die Ähnlichkeit zu einer anderen Zelllinie des Modells. Hierbei wurde keine konstant niedrigere Distanz zwischen den an Tag 7 der Kultur isolierten RNA-Proben (Replikate 1 und 3) im Vergleich zu den RNA-Proben von Tag 14 (Replikate 2) beobachtet. Das Expressionsprofil von INTH-EBV-B unterschied sich klar von den Profilen der Melanomzelllinien und der Melanozyten. Dementsprechend wurden die Datensätze auch verschiedenen Clustern zugeordnet. Innerhalb des Clusters mit den Melanomzelllinien wies Ma-Mel-86A wiederum einen größeren Abstand zu den anderen Zelllinien auf als diese untereinander. Dies könnte mit dem Verlust der Expression von Differenzierungsantigenen in Ma-Mel-86A zusammenhängen (vgl. Einleitung, Abschnitt 1.1.1.3.3). Die Melanozyten

waren ebenfalls dem Cluster zugeordnet, das die Melanomzelllinien enthielt. Somit scheint hier, trotz der unterschiedlichen Expressionsmuster von gesunden Zellen und Tumorzellen, ein „melanozytäres Grundprofil“ erhalten geblieben zu sein.

### Zusammenfassung

Zur Beurteilung der Ähnlichkeit der generierten Replikate der Exom- und Transkriptomsequenzierungen der Modellzelllinien wurde die Übereinstimmung der gefundenen Mutationen betrachtet. Durch Filtern der detektierten Mutationen basierend auf Qualität und Abdeckung konnte der Anteil der Mutationen, die in mindestens zwei Replikaten gefunden wurden, gesteigert werden. Die Übereinstimmung in den Transkriptomdaten war etwas geringer als in den Exomdaten. Anhand der Expressionsprofile konnte jedoch die Ähnlichkeit der Transkriptomreplikate gezeigt werden.

### 3.4.2 Phylogenetische Beziehung der Ma-Mel-86-Zelllinien

Die Sequenzierung des Melanommodells Ma-Mel-86 bot auch die Möglichkeit, eine phylogenetische Analyse der Melanomzelllinien durchzuführen. Als Grundlage hierfür wurden die in den Exomdaten gefundenen Einzelnukleotidvarianten (SNVs, engl. *single nucleotide variants*) gewählt. Um einen robusten Stammbaum berechnen zu können, wurden die *Mappings* der Replikate zunächst zusammengefasst und erst anschließend wurde das *Variant Calling* durchgeführt. Mutationen mit schlechter Qualität oder zu geringer Abdeckung wurden verworfen (siehe Methodenteil, Abschnitt 2.2.3.6.3). Innerhalb der verbliebenen Mutationen wurde die Zygotie folgendermaßen definiert:

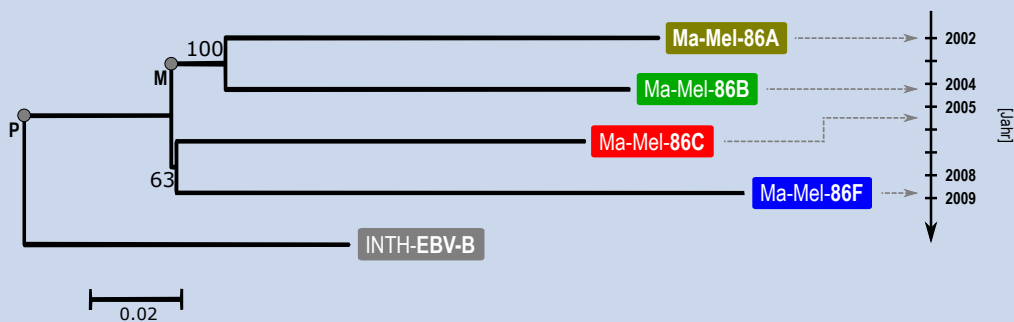
- Heterozygot: 40-60% der Basen an dieser genomischen Koordinate wichen von der Referenz ab.
- Homozygot: Mindestens 90% der Basen an dieser Koordinate wichen von der Referenz ab.
- Alle übrigen Mutationen wurden verworfen.

Durch Aneinanderreihen der nach den genomischen Koordinaten geordneten SNVs konnte unter Verwendung des *ambiguity code* (IUBMB/IUPAC-Schreibweise, siehe auch Tab. A.1) für jede Zelllinie eine artifizielle Sequenz erzeugt werden, die zur Erstellung des *Alignments* für die Stammbaumberechnung verwendet wurde. Die phylogenetische Analyse wurde am Institut für Molekulargenetik (A. Bicker, T. Hankeln, Johannes Gutenberg-Universität, Mainz) mithilfe der *Neighbor-Joining*-Methode (Modell: Kimura-2-Parameter) mit 1.000 *Bootstrap*-Replikaten durchgeführt. Bei der Berechnung wurde INTH-EBV-B als Wurzel definiert. Zur Validierung der Topologie wurde die Analyse auch mit der *Maximum-Likelihood*-Methode (Modell: GTR, engl. *general time reversible*) wiederholt. Durch beide Methoden konnte die in **Abb. 3.10** gezeigte Verwandtschaft ermittelt werden.

Auf der Grundlage der gefilterten Exom-SNVs ließen sich die aus Metastasen generierten Melanomzelllinien auf einen gemeinsamen Vorläufer „P“, eventuell den Primärtumor, zurückführen. Danach war eine Aufspaltung zu beobachten. Zum Einen leitete sich „M“ ab, der Vorläufer der Linien Ma-Mel-86A und Ma-Mel-86B, welche somit eine geschlossene Abstammungsgemeinschaft (Klade) bildeten. Dies wurde durch die hohe *Bootstrap*-Wahrscheinlichkeit von 100% unterstützt. Zum Anderen führte ein zweiter sich aufteilender Ast zu Ma-Mel-86C und -86F. Aufgrund der kurzen Astlänge und des niedrigen *Bootstrap*-Wertes (63%) ist davon auszugehen, dass sich Ma-Mel-86C und -86F parallel direkt vom gemeinsamen Vorläufer „P“ abspalteten.

#### Zusammenfassung

Insgesamt zeigte die phylogenetische Analyse der verschiedenen Ma-Mel-86-Zelllinien einen sich früh verzweigenden Aufbau. Auf der Grundlage dieses Stammbaums kann angenommen werden, dass die Metastasen nicht auseinander hervorgegangen sind, sondern sich genetisch früh voneinander getrennt hatten und wahrscheinlich erst nacheinander zu verschiedenen Zeitpunkten klinisch evident wurden (siehe Zeitstrahl rechts in **Abb. 3.10**).



**Abb. 3.10: Phylogenetische Beziehungen der Ma-Mel-86-Linien.** Zur Erstellung eines Alignments für die Stammbaumberechnung wurden die SNVs der Exomsequenzierungen herangezogen. Mit INTH-EBV-B als Wurzel wurde die phylogenetische Analyse mit der *Neighbor-Joining*-Methode (Modell: Kimura-2-Parameter) durchgeführt. Die Werte an den Knotenpunkten geben die *Bootstrap*-Wahrscheinlichkeiten (1.000 Replikate) an. Basierend auf dem so berechneten Stammbaum hatten alle Melanomzelllinien des Modells einen gemeinsamen Vorfahren „P“, von dem sich der Vorläufer von Ma-Mel-86A und -86B („M“) ableitete, und von dem sich parallel Ma-Mel-86C und -86F abspalteten. Der Maßstab gibt die evolutionäre Distanz in Mutationen pro Position an. Auf der Zeitachse (rechts) ist angegeben, in welchem Jahr die Melanomzelllinien aus Metastasen etabliert wurden.

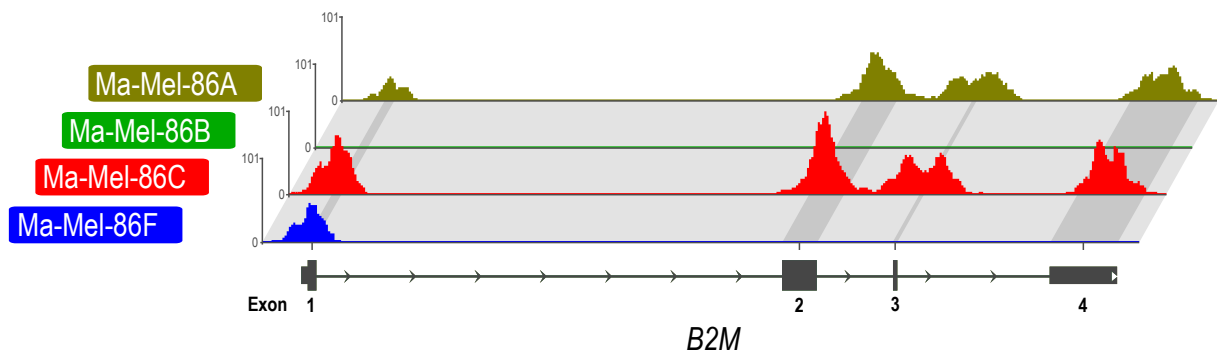
#### 3.4.3 Untersuchung des Defekts im $\beta$ 2-Mikroglobulinen

Schwinn *et al.* (2009) führten den Verlust der HLA-Klasse I-Moleküle auf der Zelloberfläche von Ma-Mel-86B auf eine fehlende  $\beta$ 2-Mikroglobulin-Expression zurück. In einer weiteren Publikation wurde der *B2M*-Defekt als genomische Deletion in Exon 1 und LOH (engl. *loss of heterozygosity*) in Chromosom 15 beschrieben (del Campo *et al.*, 2009). Im Zuge der hier durchgeführten Überprüfung der Abdeckung der *Target*-Exons (vgl. Abschnitt 3.3.3.1) fiel auf, dass alle vier *B2M*-Exons unter den *Target*-Exons mit fehlender Abdeckung gelistet waren (siehe **Tab. 3.4**).

**Tab. 3.4: Abdeckung der *B2M*-Exons in den Exomsequenzierungen der Ma-Mel-86-Zelllinien.** Die Überprüfung der Exomsequenzierungen und -*Mappings* umfasste auch die Dokumentation der *Target*-Exons mit fehlender Abdeckung (vgl. Abschnitt 3.3.3.1). Die Tabelle zeigt die Ergebnisse für die vier *B2M*-Exons (+: Abdeckung, -: keine Abdeckung). Im Gegensatz zu Ma-Mel-86A, -86C und INTH-EBV-B wurden in den Exom-*Mappings* von Ma-Mel-86B und -86F keine *Reads* an die *B2M*-Exons 1 bis 4 bzw. an die *B2M*-Exons 2 bis 4 aligniert.

Zelllinie Replikat		Ma-Mel-86A		Ma-Mel-86B		Ma-Mel-86C		Ma-Mel-86F		INTH-EBV-B	
		1	2	1	2	1	2	1	2	1	2
B2M-Lokus	chr15: 45003685-45003811 <b>Exon1</b>	+	+	-	-	+	+	+	+	+	+
	chr15: 45007621-45007899 <b>Exon2</b>	+	+	-	-	+	+	-	-	+	+
	chr15: 45008527-45008554 <b>Exon3</b>	+	+	-	-	+	+	-	-	+	+
	chr15: 45009805-45010356 <b>Exon4</b>	+	+	-	-	+	+	-	-	+	+

Dies konnte anhand der *Coverage*-Profile der Exom-*Mappings* am *B2M*-Lokus bestätigt werden (**Abb. 3.11**). So fehlte in Ma-Mel-86B die Abdeckung über den kompletten *B2M*-Lokus und in Ma-Mel-86F wurde kein *Alignment* für die *B2M*-Exons 2 bis 4 gefunden. Die Betrachtung der benachbarten Exons (Daten nicht gezeigt) ergab, dass sich die Deletion in Ma-Mel-86F nur auf diese drei *B2M*-Exons beschränkte. In Ma-Mel-86B schien der verlorene Genomabschnitt nicht nur den *B2M*-Lokus zu umfassen. Aufgrund des kompletten Fehlens von *Alignments* in diesen Regionen musste es sich jeweils um biallelische Deletionen handeln, die somit einen vollständigen Verlust von  $\beta$ 2-Mikroglobulin zur Folge hatten.



**Abb. 3.11: Coverage-Profil der Exomsequenzierungen am *B2M*-Lokus.** Zur Visualisierung der genomischen Deletion am *B2M*-Lokus wurden mithilfe des *Integrative Genomics Viewer* (IGV, Thorvaldsdóttir *et al.*, 2013) *Coverage*-Profile dieses Lokus erstellt (x-Achsen: Position innerhalb des Lokus, y-Achsen: Anzahl der alignierten Reads). Die Grafik unterhalb der Profile zeigt die Exon-Intron-Struktur von *B2M* (Exon: Rechteck, Intron: Linie, Lokus-Orientierung in Pfeilrichtung).

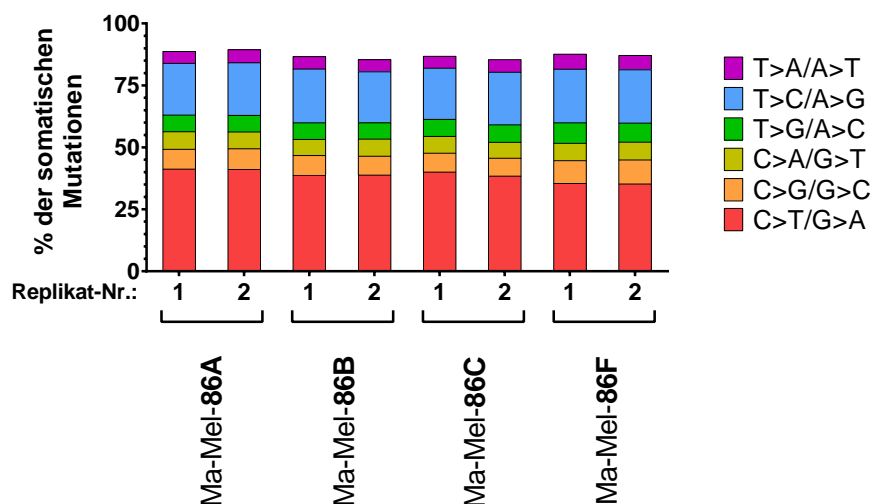
### Zusammenfassung

Anhand der Exomsequenzierungen der Zelllinien Ma-Mel-86B und -86F konnten die *B2M*-Defekte, die zum Verlust der HLA-Klasse I-Moleküle auf der Zelloberfläche der beiden Tumorzelllinien führte, näher beschrieben werden. Demnach ist *B2M* in Ma-Mel-86B komplett deletiert, wohingegen in Ma-Mel-86F das erste Exon noch erhalten ist. Somit handelte es sich um zwei unterschiedliche Mechanismen, die zum Verlust der HLA-Klasse I-Expression und somit zur Reduktion der Immuno-

genität der beiden Tumorzelllinien führte. Eine exakte Bestimmung der Grenzen der deletierten Genomabschnitte war jedoch mit den generierten Exomdaten nicht möglich.

### 3.4.4 Übersicht der somatischen Mutationen

In den Exomen der vier Tumorzelllinien des Modells Ma-Mel-86 wurden nach Abzug der Mutationen, die in der Keimbahnkontrolle vorlagen (Details siehe Abschnitt 3.4.5.1), durchschnittlich ca. 4.200 somatische Mutationen (Substitutionen und kurze Insertionen und Deletionen) detektiert. Über 85% dieser Mutationen waren jeweils Einzelnukleotidsubstitutionen (siehe **Abb. 3.12**). Bei den restlichen Mutationen handelte es sich um Insertionen oder Deletionen von einzelnen oder mehreren Nukleotiden. Am häufigsten wurde der Austausch von Cytosin gegen Thymin (C>T) bzw. von Guanin gegen Adenin (G>A) beobachtet, was für Melanome, für die UV-Strahlung als ursächlich erachtet wird, typisch ist (Hill *et al.*, 2013). Im Mittel wurden 265 nicht-synonyme somatische Mutationen in den vier Tumorexomen detektiert.



**Abb. 3.12: Einzelnukleotidaustausche in den somatischen Mutationen der Ma-Mel-86-Exome.** Zwischen 85 und 89% der somatischen Mutationen, die in den Exomen der Tumorzelllinien des Modells detektiert werden konnten, waren Substitutionen einzelner Nukleotide (A: Adenin, C: Cytosin, G: Guanin, T: Thymin). Der hohe Anteil von C>T/G>A-Austauschen implizierte die Dominanz UV-induzierter Mutationen.

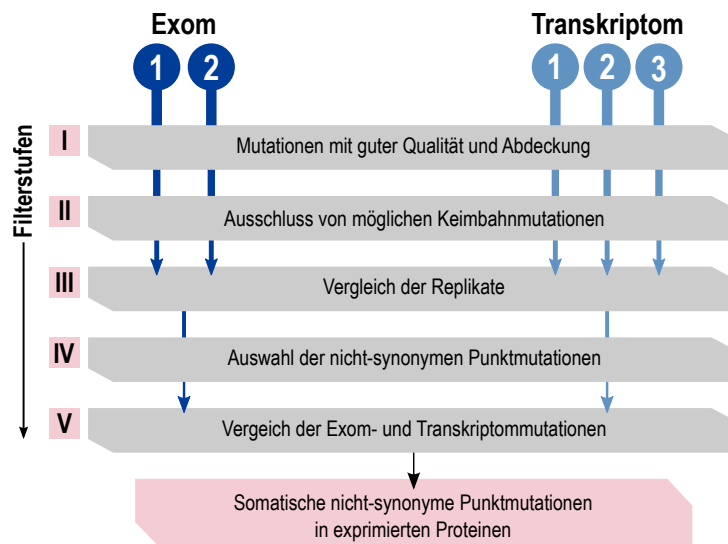
Die von Hodis *et al.* (2012) in Ma-Mel-86A detektierten (potentiellen) *Driver*-Mutationen in *BRAF*, *TP53*, *PTEN* und *MAP2K1* konnten in den hier durchgeführten Exomsequenzierungen bestätigt werden. Außerdem lagen in den Exomsequenzierungen der Tumorzelllinien des Modells auch Gene mutiert vor, in denen bereits Mutationen in anderen Melanomen publiziert worden waren (Kunz *et al.*, 2013). Dabei handelte es sich um *MST1* (engl. *macrophage stimulating 1*), *GPR98* (engl. *G protein-coupled receptor 98*), *FAT2* (engl. *FAT atypical cadherin 2*), *TRRAP* (engl. *transformation/transcription domain-associated protein*), *GRIN2A* (engl. *glutamate receptor, ionotropic, N-methyl D-aspartate 2A*), *DCC* (engl. *deleted in colorectal carcinoma*), *ERBB4* (engl. *v-erb-b2 avian erythroblastic leukemia*

*viral oncogene homolog 4*), *LRRFIP1* (engl. *leucine rich repeat (in FLII) interacting protein 1*), *PDGFRA* (engl. *platelet-derived growth factor receptor, alpha polypeptide*) und *PTK2B* (engl. *protein tyrosine kinase 2 beta*). So waren beispielsweise *GRIN2A* (Wei *et al.*, 2011) und *MST1* (Berger *et al.*, 2010) in mehr als 30% der jeweils untersuchten Melanome mutiert. Da jedoch nicht in allen Publikationen die genaue Position der detektierten Mutationen innerhalb der jeweiligen Gene angegeben waren, wurde kein systematischer Vergleich der Mutationen durchgeführt.

### 3.4.5 Detektion von potentiell immunogenen Mutationen

#### 3.4.5.1 Auswahl von mutierten Proteinen

Die Grundlage einer Vorhersage von potentiell immunogenen Peptiden ist zunächst die Auswahl von Proteinen bzw. Proteinabschnitten, innerhalb derer nach HLA-Bindemotiven oder Proteasomschnittstellen gesucht wird. In der vorliegenden Arbeit sollten Proteinabschnitte um die in den Melanomzelllinien des Modells gefundenen somatischen Mutationen auf das Vorkommen potentiell HLA-bindender Peptide überprüft werden. Wie in Abschnitt 3.4.1.1 gezeigt, wurden in den Exomen der Zelllinien des Modells durchschnittlich ca. 135.000 und in den Transkriptomen ca. 73.000 verschiedene Mutationen mit ausreichender Qualität und Abdeckung detektiert (Filterstufe I, vgl. **Abb. 3.13**). Damit eine Mutation jedoch im Kontext eines Peptids spezifisch auf Tumorzellen von T-Zellen erkannt werden kann, muss es sich um eine somatische Mutation handeln, die zur Veränderung der Aminosäuresequenz eines exprimierten Proteins führt. Daher wurden die detektierten Mutationen weiter gefiltert (Filterstufen II-V, **Abb. 3.13**). Dazu wurden zunächst die Mutationen, die auch in der Keimbahnkontrolle INTH-EBV-B detektiert wurden (Filterstufe II, Details siehe Methodenteil, Abb. 2.1) ausgeschlossen. Anschließend wurden die Replikate der Exomsequenzierungen und die der Transkriptomsequenzierungen untereinander verglichen. Dabei wurden alle Mutationen verworfen (Filterstufe III), die nicht in beiden Exomreplikaten gefunden bzw. die nicht durch mindestens ein zweites Transkriptomreplikat bestätigt wurden. Innerhalb der verbliebenen Mutationen wurden anschließend *Missense*-Mutationen gesucht (Filterstufe IV). Im letzten Schritt (Filterstufe V) wurden die vorgefilterten Exom- und Transkriptommutationen jeder Zelllinie miteinander verglichen. Bei Mutationen, die nur in der Transkriptomliste vorkamen, wurde überprüft, ob sie aufgrund einer zu schwachen Abdeckung in den entsprechenden Exomsequenzierungen (vgl. Abschnitt 3.3.3.1) nicht in der Exomliste zu finden waren. In diesem Fall wurden diese „Nur Transkriptom“-Mutationen in die finalen Listen mit aufgenommen. Mutationen, die nur in den Exomen detektiert wurden, wurden als nicht exprimiert angesehen und daher nicht weiter berücksichtigt. Im Gegensatz zu bereits publizierten Strategien der Priorisierung von Mutationen (Löwer *et al.*, 2012; Castle *et al.*, 2012) wurden die Transkriptomdaten dadurch explizit zur Detektion relevanter Punktmutationen genutzt und nicht nur zur Validierung der Expression des mutierten Proteins bzw. der Mutation an sich verwendet.



**Abb. 3.13: Strategie zur Auswahl mutierter Proteine für die Vorhersage von potentiell immunogenen Peptiden.** Ausgehend von den in den Exom- und Transkriptomsequenzierungen detektierten Mutationen wurden in fünf Filterstufen (I-V) somatische nicht-synonyme Punktmutationen in exprimierten Proteinen lokalisiert. Bis zum dritten Filterschritt wurden die Replikate getrennt prozessiert (blaue Pfeile), danach wurden nur Mutationen weitergeführt, die in beiden Exomreplikaten oder in mindestens zwei Transkriptomreplikaten gefunden wurden. Im letzten Schritt wurden auch die Exom- und Transkriptomlisten gegeneinander abgeglichen.

Um die tatsächliche Verteilung der Mutationen auf die vier Tumorzelllinien zu verifizieren, wurden die Mappings aller Datensätze zum Vergleich an den jeweiligen genomischen Koordinaten im *Integrative Genomics Viewer* (IGV, Thorvaldsdóttir *et al.*, 2013) aufgerufen und die Listen dementsprechend angepasst. Die Zuordnung der Mutationen zu den Tumorzelllinien erfolgte basierend auf ihrer Expression und nicht aufgrund des Vorkommens in den Exomdaten.

Im *Circos-Plot* in **Abb. 3.14** ist dargestellt, wie viele Mutationen nach jeder Filterstufe übrig blieben. Hier fiel auf, dass nach der zweiten und dritten Filterstufe in fast allen Datensätzen mehr Mutationen in den Transkriptomen weitergeführt wurden als in den Exomen. Begründet werden kann dies teilweise durch die gewählten Parameter bei der Mutationsdetektion. Einerseits wurde für die *RNA-Seq*-Daten eine niedrigere Mindestabdeckung gefordert (Transkriptom: fünf *Reads*, Exom: zehn *Reads*) und andererseits wurden in den Exomen nur Mutationen mit einer maximalen Abdeckung von 200 *Reads* (siehe Methodenteil, Abschnitt 2.2.3.6.1) dokumentiert, um *Alignment*-Artefakte auszuschließen. Somit war die Detektion der Mutationen in den Transkriptomen weniger eingeschränkt. Die gute Übereinstimmung zwischen den Transkriptomreplikaten (siehe Filterstufe III) rechtfertigte jedoch diese Parameterwahl.

Weiterhin wurde eine Mutation, die sich in einem Sequenzabschnitt befand, der in den Exomsequenzierungen von INTH-EBV-B nicht ausreichend abgedeckt war und gleichzeitig aber auch nicht von INTH-EBV-B exprimiert wurde, in den Mutationslisten der Tumortranskriptome als somatische Mutation eingeordnet. Hierbei entstandene falsch positive Einträge konnten durch die weitere Filterung, spätestens jedoch durch die finale Überprüfung der „Nur Transkriptom“-Mutationen entfernt werden.



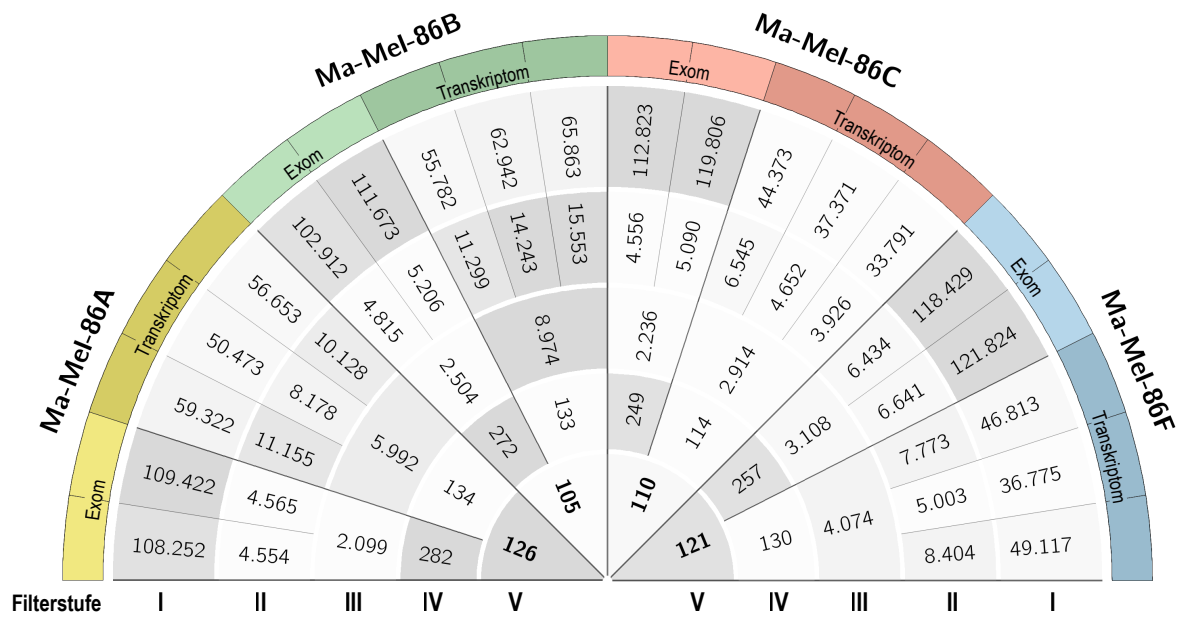
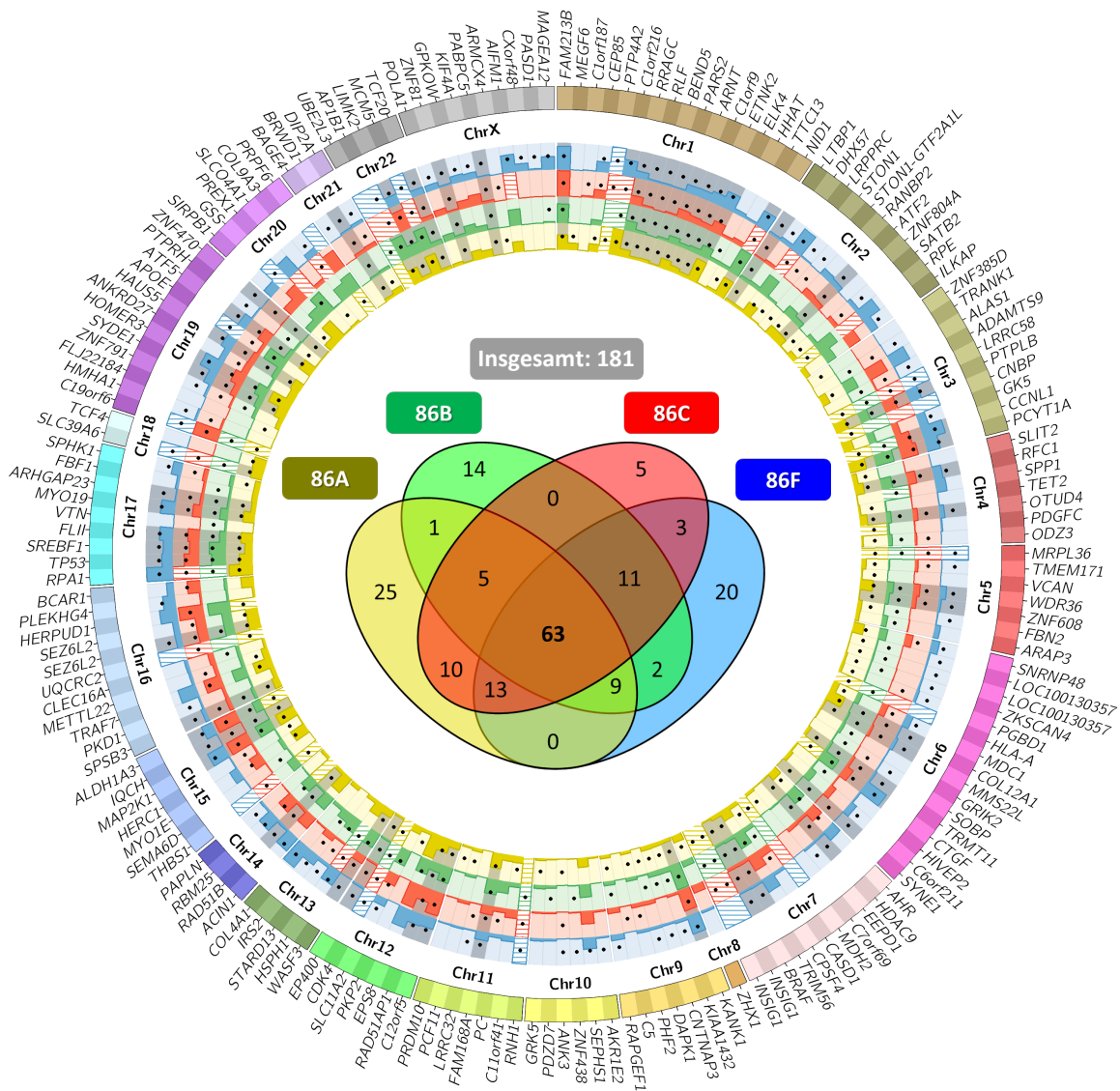


Abb. 3.14: Auswahl mutierter Proteine für die Vorhersage von potentiell HLA-bindenden Peptiden. Wie in Abb. 3.13 dargestellt, wurden die in den Sequenzierdaten der Zelllinien des Modells detektierten Mutationen gefiltert, um Proteine zu selektieren, die somatische nicht-synonyme Punktmutationen aufwiesen. Der *Circos-Plot* zeigt von außen nach innen die Anzahl der Mutationen, die in den Replikaten bzw. zusammengefassten Mutationslisten während der Filterstufen I bis V gefunden wurden. Die Intensität der Graufärbung ist für jeden Halbkreis proportional zur Anzahl der Mutationen gewählt.

Nach dem fünften Filterschritt verblieben insgesamt 181 somatische nicht-synonyme Punktmutationen, deren Verteilung auf die vier Tumorzelllinien im Venn-Diagramm in **Abb. 3.15** dargestellt ist. 63 dieser Mutationen konnten in allen vier Tumorzelllinien detektiert werden (grau hinterlegt in **Abb. 3.15**). Insgesamt waren 178 verschiedene Gene von den Mutationen betroffen. Die Gene *LOC100130357* und *SEZ6L2* wiesen jeweils zwei Mutationen auf, und die *INSIG1*-Mutation an der genomischen Koordinate *chr7:155094465* hatte je nach Transkriptionsvariante zwei unterschiedliche Aminosäureaustausche (P > S bzw. S > F) zur Folge. Die Expressionslevel der einzelnen Gene (Säulendiagramme in **Abb. 3.15**) wurden mit dem *Bioconductor*-Paket *DESeq* aus den Transkriptomsequenzierungen ermittelt (vgl. Methodenteil, Abschnitt 2.2.3.8).

Da nur die Tumorzelllinien Ma-Mel-86A und -86C HLA-Klasse I-Moleküle exprimierten, wurden nur die 145 Mutationen, die in mindestens einer dieser beiden Tumorzelllinien (gelb und rot in **Abb. 3.15**) detektiert wurden, für die Suche nach potentiell HLA-bindenden Peptiden verwendet. Aufgrund der gewählten Filterstrategie und der Absicherung durch Replikate zweierlei Ursprungs wurden die Mutationen als sicher klassifiziert, und es wurde auf eine zusätzliche Validierung durch Sanger-Sequenzierungen verzichtet.

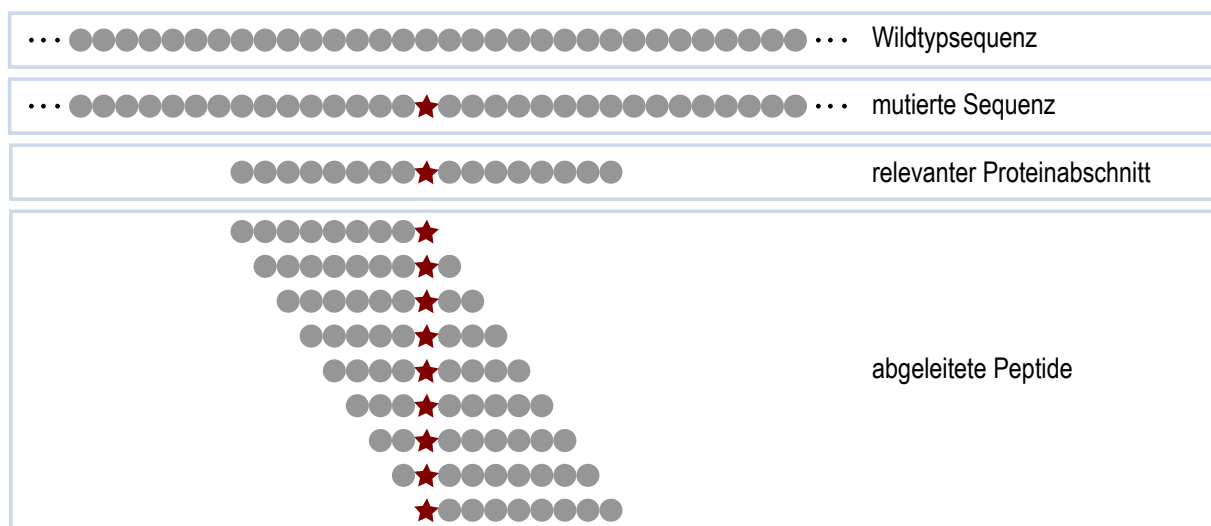


**Abb. 3.15: Somatische nicht-synonyme Punktmutationen exprimierter Gene in Ma-Mel-86.** Durch Filtern entsprechend des Schemas in Abb. 3.13 wurden 181 exprimierte somatische nicht-synonyme Punktmutationen lokalisiert. Das Venn-Diagramm (Mitte) gibt an, wie diese Mutationen auf die vier Tumorzelllinien des Modells verteilt waren. Im *Circos-Plot* sind auf dem äußersten Ring die 178 betroffenen Gene sortiert nach der genomischen Koordinate ihrer Mutation dargestellt. Gene, die doppelt aufgeführt sind, trugen entweder zwei verschiedene Mutationen (*LOC100130357* und *SEZ6L2*) oder eine einzelne Mutation mit zwei unterschiedlichen Effekten auf die Aminosäuresequenz (*INSIG1*). Die Chromosomen-Blöcke (Chr1-22 und ChrX) sind entsprechend der Anzahl der jeweils detektierten Mutationen skaliert. Die Säulendiagramme zeigen die auf die Größe der Datensätze (*DESeq, size factors*) und auf die Länge der Gene (in kb) normalisierten Expressionswerte (Mittelwerte aus je drei Replikaten) der Transkriptomsequenzierungen von Ma-Mel-86A (gelb), Ma-Mel-86B (grün), Ma-Mel-86C (rot) und Ma-Mel-86F (blau). Zur besseren Übersicht sind die y-Achsen auf Expressionswerte von maximal 2.000 beschränkt. Höhere Expressionslevel sind schraffiert gekennzeichnet. Die schwarzen Punkte markieren, in welcher Zelllinie die Mutationen exprimiert wurden. Die 63 Mutationen, die in allen vier Tumorzelllinien gefunden wurden, sind grau hinterlegt.

### 3.4.5.2 Vorhersage und Selektion von Peptidkandidaten

Laut Annotation von *SnpEff* hatten die 145 gefilterten Mutationen der Tumorzelllinien Ma-Mel-86A und -86C eine Auswirkung auf 264 *RefSeq*-Transkriptionsvarianten von 142 verschiedenen Genen. Da für das *Mapping* jedoch die UCSC-Referenz hg19 verwendet wurde und für einzelne Transkripte Diskrepanzen zwischen den Datenbanken

vorlagen, wurden die Wildtypproteinsequenzen dieser Transkripte aus beiden Datenbanken (*RefSeq* und UCSC) heruntergeladen und zur Absicherung mithilfe des *Perl*-Skripts **compareRefSeqUCSC.pl** (siehe Anhang, Abschnitt A.3.2.2) miteinander verglichen. Anschließend wurden die Mutationen automatisiert in die abgeglichenen Wildtypsequenzen eingefügt und die Sequenzen auf jeweils um die Mutation zentrierte Abschnitte reduziert (*Perl*-Skripte **peptideCandidates\_TAB.pl** und **peptideCandidatesTAB\_to\_FASTA.pl**, siehe Methodenteil, Abschnitt 2.2.3.7.1 und Anhang, Abschnitt A.3.2.3 und Abschnitt A.3.2.4). Da potentiell HLA-bindende Okta-, Nona- und Dekamere mit den detektierten Mutationen vorhergesagt werden sollten, wiesen die relevanten Proteinabschnitte eine Länge von 15, 17 bzw. 19 Aminosäuren auf. **Abb. 3.16** zeigt dies exemplarisch für Nonamere. Dadurch wurde von den Algorithmen nur zu Peptiden, die die Mutation tatsächlich enthielten, eine Vorhersage gemacht.



**Abb. 3.16: Schematische Darstellung der relevanten Proteinabschnitte für die Peptidvorhersagen.** In der Proteinwildtypsequenz wurde die zuvor bestimmte Mutation eingefügt (●: Wildtyp-Aminosäure, ★: mutierte Aminosäure). Für die Vorhersage der Bindewahrscheinlichkeit von Peptiden, die diese Mutation enthielten, wurde nur ein um die Mutation zentrierter Proteinabschnitt benötigt. Dieser hatte für Nonamere eine Länge von 17 Aminosäuren, sodass neun verschiedene, die Mutation beinhaltende Peptide abgeleitet werden konnten. Entsprechend wurden für die Vorhersage von Okta- bzw. Dekameren Proteinabschnitte mit 15 bzw. 19 Aminosäuren ausgewählt.

Für die Vorhersage von Peptiden, die sich aus diesen Proteinabschnitten ableiteten und an die HLA-Klasse I-Moleküle der Patientin INTH binden konnten, wurden die matrixbasierten Algorithmen SYFPEITHI (Rammensee *et al.*, 1999) und BIMAS (Parker *et al.*, 1994), der ANN-basierte Algorithmus des NetMHC Server (Lundegaard *et al.*, 2008) und das Vorhersageprogramm der *Immune Epitope Database* (IEDB; Zhang *et al.*, 2008) gewählt (siehe **Tab. 3.5**, weitere Details im Methodenteil, Abschnitt 2.2.3.7.2). Sofern möglich, wurden mit den vier Algorithmen in allen generierten Proteinabschnitten nach Okta-, Nona- und Dekameren für die Bindung an die sechs HLA-Klasse I-Allele des Modells gesucht. Einschränkungen gab es beim SYFPEITHI-Algorithmus, der nicht für alle Peptidlängen und nicht für HLA-C-Allele verfügbar ist, und bei BIMAS, wo nur die Allelgruppen, nicht aber die benötigten spezifischen HLA-Allele ausgewählt werden konnten (z. B. HLA-B62 für HLA-B\*15:01). Im Gegensatz dazu konnten die Bindewahrscheinlichkeiten mithilfe

von NetMHC und IEDB uneingeschränkt vorhergesagt werden.

**Tab. 3.5: Übersicht der verwendeten Algorithmen zur Peptidbindungsvorhersage.** Für die Lokalisation von Peptiden, die die ausgewählten exprimierten somatischen nicht-synonymen Punktmutationen enthielten und die an die HLA-Klasse I-Allele des Modells (HLA-A\*01:01, HLA-A\*24:02, HLA-B\*08:01, HLA-B\*15:01, HLA-C\*03:03 und HLA-C\*07:01) binden konnten, wurden die vier gelisteten Vorhersagealgorithmen eingesetzt. Diese unterschieden sich unter anderem in der Art der Bewertung der Peptidkandidaten. Die gewählten Schwellenwerte sind angegeben. Für SYFPEITHI und BIMAS ergaben sich Einschränkungen, da Vorhersagen nicht für alle HLA-Allele der Patientin zur Verfügung standen (siehe Bemerkungen).

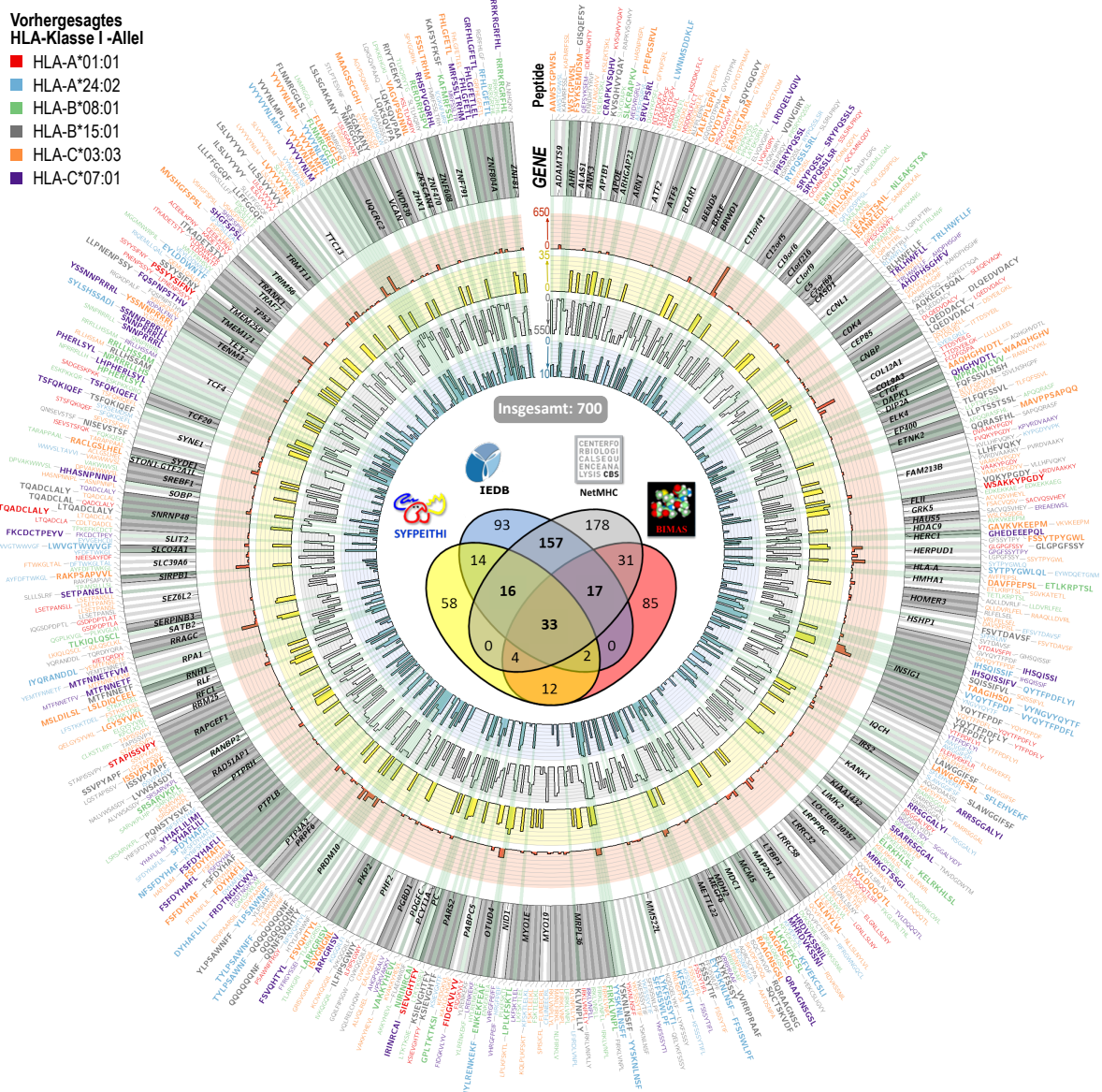
Algorithmus	Bewertung	Gewählter Schwellenwert	Bemerkungen
<b>SYFPEITHI</b>	Eigener <i>Score</i>	$\geq 20$	Keine Vorhersage für HLA-C-Allele, nicht alle Peptidlängen für alle Allele verfügbar
<b>BIMAS</b>	Dissoziationszeit [s]	$\geq 20$	Nur HLA-Gruppen verfügbar
<b>IEDB</b>	Prozentrang	$\leq 10$	Alle spezifischen Modell-HLA-Allele verfügbar
<b>NetMHC 3.4 Server</b>	Bindeaffinität [nM]	$\leq 500$	Alle spezifischen Modell-HLA-Allele verfügbar

700 verschiedene HLA-Peptid-Kombinationen erfüllten mindestens einen der in **Tab. 3.5** gelisteten Schwellenwerte. Das Venn-Diagramm in **Abb. 3.17** verdeutlicht, inwiefern die Vorhersagen der vier Algorithmen übereinstimmen. Die IEDB-Vorhersage kombiniert in Abhängigkeit des zu überprüfenden HLA-Allels verschiedene Vorhersagemethoden. Eine Methode beinhaltet auch den ANN-Algorithmus von NetMHC. Dementsprechend war die Übereinstimmung dieser beiden Programme erwartungsgemäß am größten. Dennoch erfüllten nur 223 (41%) der 545 HLA-Peptid-Kombinationen, für die ein Prozentrang von  $\leq 10$  (IEDB) und/oder eine Bindeaffinität von  $\leq 500$  nM (NetMHC) vorhergesagt wurde, beide Schwellenwerte.

Den 700 HLA-Peptid-Kombinationen lagen 562 verschiedene Peptide zugrunde (**Abb. 3.17**, äußerster Ring), die sich aus 125 mutierten Genen (grau-weißer Ring in **Abb. 3.17**) ableiteten. Für einzelne Gene (z. B. *MMS22L*) wurden bis zu 28 potentiell immunogene HLA-Peptid-Kombinationen innerhalb des relevanten Proteinabschnittes bestimmt. Die meisten Peptide wurden für HLA-C\*03:03 vorhergesagt (198 Peptide, Sequenzen in Orange in **Abb. 3.17**). Danach folgten HLA-B\*15:01 mit 154 Peptiden (grau), HLA-B\*08:01 mit 108 Peptiden (grün), HLA-C\*07:01 mit 85 Peptiden (violett) und HLA-A\*24:02 mit 84 Peptiden (blau). Für die Bindung an HLA-A\*01:01 wurden nur 71 passende Peptide (gelb) gefunden. Die Ergebnisse der Vorhersagen der vier Algorithmen sind in den Säulendiagrammen der **Abb. 3.17** dargestellt (roter, gelber, grauer und blauer Ring).

Da nur mittels IEDB und NetMHC Peptide für alle sechs HLA-Klasse I-Allele des Modells vorhergesagt werden konnten, wurden nur die Ergebnisse dieser beiden Algorithmen für die Auswahl der auf Immunogenität zu testenden Peptide berücksichtigt. Wie oben beschrieben, erfüllten 223 verschiedene HLA-Peptid-Kombinationen (fett gedruckt in **Abb. 3.17**, dazugehörige Vorhersageresultate grün hinterlegt) die für die IEDB- und NetMHC-Vorhersagen gesetzten Schwellenwerte. Sechs Peptide wurden für die Bindung an HLA-A\*01:01, 32 Peptide für HLA-A\*24:02, 24 Peptide für HLA-B\*08:01, 67 Peptide für HLA-B\*15:01, 44 Peptide für HLA-C\*03:03 und 50 Peptide für HLA-C\*07:01 vorhergesagt.

Da für manche Peptide potentielle Bindungen an mehrere HLA-Allele vorhergesagt wurden, lagen insgesamt 207 verschiedene Peptide vor. Nach einem manuellen Abgleich dieser Peptidkandidaten in den alignierten Sequenzierdaten der Ma-Mel-86-Zelllinien mithilfe des IGV wurden 174 der mutierten Peptide (41 Okta-, 61 Nona- und 72 Dekamere) nochmals als sicher exprimiert bestätigt und im nächsten Schritt auf die Erkennung durch autologe T-Zellen getestet (siehe Abschnitt 3.5.5). Eine Auflistung dieser Peptide ist im Methodenteil (Abschnitt 2.1.13) zu finden.



**Abb. 3.17: Ergebnisse der Vorhersagealgorithmen.** Von außen nach innen zeigt der *Circos-Plot* (i) die Peptidsequenzen der 700 Peptid-Restriktionsmolekül-Kombinationen, die mindestens einen der geforderten Schwellenwerte (vgl. Tab. 3.5) erfüllten, (ii) die Gene, von denen sich die vorgeschagten Peptide ableiteten, und (iii) als Säulendiagramme die Ergebnisse der Vorhersagealgorithmen BIMAS (rot), SYFPEITHI (gelb), NetMHC (grau) und IEDB (blau), wenn die gesetzten Schwellenwerte erfüllt wurden. Die Farben der Peptidsequenzen kodieren die vorhergesagten HLA-Klasse I-Allele (siehe Legende). Zur besseren Übersicht sind die y-Achsen der Säulendiagramme jeweils so orientiert, dass den besten HLA-Bindern die höchsten Säulen zugeordnet sind (BIMAS: 0 → 650 [s], SYFPEITHI: 0 → 35 (SYFPEITHI-Score), NetMHC: 550 → 0 [nM], IEDB: 10 → 0 (Prozentrang)). Die Peptidsequenzen der 223 Vorhersagen, bei denen eine Bindeaffinität von ≤500 nM (NetMHC) und ein Prozentrang von ≤10 (IEDB) bestimmt wurde, sind fett gedruckt und die dazugehörigen Vorhersageergebnisse grün markiert. Das Venn-Diagramm in der Mitte zeigt die Übereinstimmung der vier Vorhersagealgorithmen. Die einzelnen Vorhersageergebnisse sowie eine Version dieser Abbildung mit hoher Auflösung sind im elektronischen Anhang zu finden.

#### **Zusammenfassung**

In den Sequenzierungen der Exome und Transkriptome der Zelllinien des Modells Ma-Mel-86 wurden durch schrittweises Filtern 181 somatische nicht-synonyme Punktmutationen in exprimierten Proteinen lokalisiert. In den HLA-Klasse I-exprimierenden Zelllinien Ma-Mel-86A und -86C wurden 145 dieser Mutationen detektiert. Mithilfe von Algorithmen zur Vorhersage von HLA-Bindewahrscheinlichkeiten wurden die relevanten Proteinabschnitte um diese 145 Mutationen auf das Vorkommen von Peptiden untersucht, die an die HLA-Klasse I-Allele der Patientin INTH binden könnten. Auf der Grundlage der Vorhersagen der Algorithmen von NetMHC und IEDB wurden 174 Peptide als potentielle HLA-Binder eingestuft. Die ausgewählten Peptide hatten eine vorhergesagte Bindeaffinität von  $\leq 500$  nM (NetMHC) und erzielten einen Prozentrang von  $\leq 10$  (IEDB). Die Erkennung dieser Peptide durch autologe T-Zellen sollte im nächsten Schritt *in vitro* überprüft werden.

## 3.5 Identifizierung immunogener Peptide

### 3.5.1 Auswahl der T-Zellen für die Analyse der Peptidimmunogenität

Für die Analyse der Immunogenität der Peptidkandidaten mittels IFN $\gamma$ -ELISpot-Assays standen aus den Vorarbeiten im Modell Ma-Mel-86 (siehe Einleitung, Abschnitt 1.1.1.3.3) 15 verschiedene, kryokonservierte MLTC-Responderpopulationen mit mindestens je  $25 \times 10^6$  Zellen zur Verfügung (vgl. **Tab. 3.6**).

**Tab. 3.6: Übersicht der verfügbaren MLTCs im Modell Ma-Mel-86.** Fünfzehn verschiedene MLTCs, die in Vorarbeiten aus PBMCs oder aus PBMCs isolierten CD8<sup>+</sup> T-Zellen drei verschiedener Blutabnahmen generiert wurden, standen für die Peptidkennungstests zur Verfügung. In der zweiten Spalte ist das Alter der MLTCs in Tagen (dx+y) angegeben, wobei y die Anzahl der Tage bis zur Kryokonservierung seit der letzten Restimulation an Tag x angibt. Wo y nicht angegeben ist, wurden die MLTCs nach dem Auftauen zunächst mit Ma-Mel-86A bzw. -86C restimuliert. Bei der Generierung der MLTCs 7C bis 12C wurde versucht, die Immunogenität der Tumorzellen oder die Reaktivität der in den CD8-negativen *Feeder*-Zellen noch vorhandenen antigenpräsentierenden Zellen durch Zugabe von Immunstimulantien (in Klammern) zu erhöhen. Die Reagenzien wurden freundlicherweise von Prof. G. Hartmann (Bonn) zur Verfügung gestellt. Weiterhin wurde in der MLTC 15C der Effekt des Braf-Inhibitors PLX4032 untersucht.

MLTC	eingefroren an	generiert aus	Datum der Abnahme	Stimulatorzellen
<b>1A.1</b>	d23	PBMCs	25.08.2004	Ma-Mel- <b>86A</b>
<b>1A.2</b>	d38+4	MLTC 1A.1 CD8 <sup>+</sup> *	-	Ma-Mel- <b>86C</b>
<b>3A</b>	d28+4 d49+3	CD8 <sup>+</sup>	07.04.2004	Ma-Mel- <b>86A</b>
<b>4A</b>	d56+4	PBMCs	07.05.2002	Ma-Mel- <b>86A</b>
<b>1C</b>	d23+4	PBMCs	25.08.2004	Ma-Mel- <b>86C</b>
<b>2C</b>	d23+4	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b>
<b>4C</b>	d35+4 d63+4	PBMCs	07.05.2002	Ma-Mel- <b>86C</b>
<b>5C</b>	d29+5	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b>
<b>7C</b>	d29+5	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b> (APC + 3p-RNA)
<b>9C</b>	d21	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b>
<b>10C</b>	d21	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b> (APC + 3p-RNA)
<b>11C</b>	d21	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b> (APC + Poly(I:C))
<b>12C</b>	d21	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b> (APC + 3p-RNA + Poly(I:C))
<b>13C</b>	d21+5	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b>
<b>15C</b>	d21+5	CD8 <sup>+</sup>	25.08.2004	Ma-Mel- <b>86C</b> (+ PLX4032)

\* MLTC 1A.2 wurde aus CD8<sup>+</sup> T-Zellen generiert, die an Tag 23 aus MLTC 1A.1 isoliert und mit Ma-Mel-86C weiterstimuliert wurden.

Gegebenenfalls wurden die MLTCs einmal mit bestrahlten autologen Tumorzellen restimuliert, um die Zellen nochmals zu expandieren oder zu aktivieren. Aus MLTC 3A und 4C konnten jeweils zwei Populationen verschiedenen Alters getestet werden. Zur Einschätzung der Reaktivität der verfügbaren MLTC-Responderpopulationen und zur Ermittlung der Anteile HLA-Klasse I-abhängiger T-Zellantworten wurde zunächst die Erkennung der autologen Tumorzelllinien mit und ohne Zugabe des an ein gemeinsames

### 3.5 Identifizierung immunogener Peptide

Epitop von HLA-Klasse I-Molekülen bindenden Antikörpers W6/32 getestet (**Abb. 3.18**).

		Zielzellen										Auswahl
		-		Ma-Mel-86A		Ma-Mel-86B		Ma-Mel-86C		Ma-Mel-86F		
α-HLA I	-	-		+		-		+		-		
	MLTC	1A.1 (d23+6)										
1A.2 (d38+6)												
3A (d28+6)												*
3A (d49+5)												
4A (d56+6)												*
1C (d23+6)												*
2C (d23+6)												*
4C (d35+6)												
4C (d63+6)												
5C (d29+7)												
7C (d29+7)												*
9C (d21+5)												
10C (d21+5)												
11C (d21+5)												
12C (d21+5)												
13C (d21+7)												
15C (d21+7)											*	

**Abb. 3.18: Vortest der Ma-Mel-86-MLTCs.** Die Erkennung der autologen Melanomzelllinien (jeweils 50.000 Zellen/Well) durch 15 verschiedene MLTCs (jeweils 30.000 Zellen/Well; MLTC 4A: 20.000 Zellen/Well) wurde in IFN $\gamma$ -ELISpot-Assays überprüft. Die unterschiedliche Grundfärbung der Wells ist durch die Umstellung der ELISpot-Platten bedingt. Die MLTCs 1A.1 bis 4C und 7C wurden auf *MultiScreen<sub>HTS</sub>*-Filterplatten (MSIPS4510, Millipore) und die MLTCs 5C und 9C bis 15C wurden auf *MultiScreen*-Filterplatten (S2EM004M99, Millipore) getestet. In Klammern ist das Alter der MLTCs (dx+y) am Tag des Tests angegeben, wobei y die Anzahl der Tage nach der letzten Restimulation an Tag x angibt. Von den MLTCs 3A und 4C wurden jeweils zwei Populationen verschiedenen Alters getestet. Durch eine Blockade mit dem Antikörper W6/32 ( $\alpha$ -HLA I) sollte der Anteil der HLA-Klasse I-abhängigen Antworten überprüft werden. Die letzte Spalte (Auswahl) gibt an, mit welchen MLTC-Responderpopulationen weitergearbeitet wurde.

Anhand dieses Vortests wurden basierend auf dem Erkennungsmuster, der Stärke der



Reaktivität, der spontanen IFN $\gamma$ -Sekretion (d. h. Hintergrund) und der Blockierbarkeit durch  $\alpha$ -HLA-Klasse I-Antikörper sieben MLTCs für die weiteren Tests ausgewählt (gekennzeichnet in **Abb. 3.18**).

### 3.5.2 Generierung der MLTC 16C

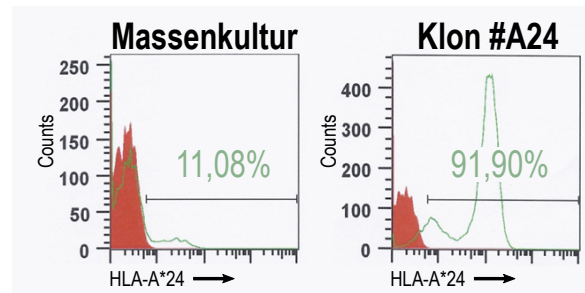
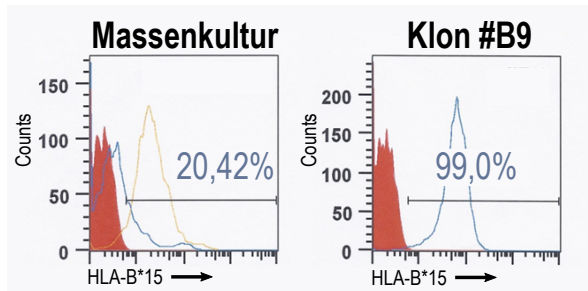
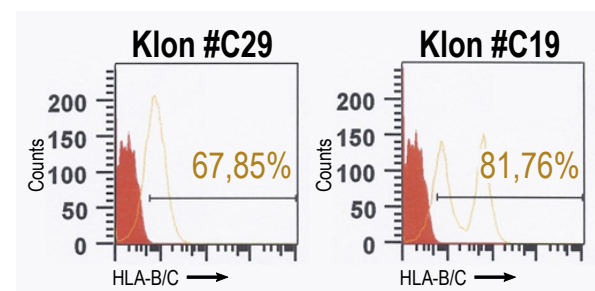
Um die Immunogenität der elf mutierten Peptide (#164 bis #174) zu testen, die in Ma-Mel-86C und nicht in Ma-Mel-86A exprimiert und laut der Vorhersagen nur über die in Ma-Mel-86C fehlenden HLA-Klasse I-Allele präsentiert werden, wurden stabile Ma-Mel-86C-HLA-Transfektanten hergestellt, in denen die Expression der drei Allele rekonstituiert wurde. Diese Transfektanten wurden für die Stimulation einer neuen MLTC verwendet, bezeichnet als MLTC 16C.

#### 3.5.2.1 Generierung der Ma-Mel-86C-HLA-Transfektanten

Die Ausstattung von Ma-Mel-86C mit HLA-A\*24:02, -B\*15:01 und -C\*03:03 erfolgte in drei getrennten Ansätzen mit dem Ziel drei stabile Transfektanten mit je einem zusätzlichen HLA-Klasse I-Allel zu erzeugen. Die cDNAs der HLA-Klasse I-Moleküle lagen im Expressionsvektor *pcDNA<sup>TM</sup>3.1/V5-His TOPO<sup>®</sup>* vor. Die Transfektion wurde per Elektroporation von je  $3 \times 10^6$  Zellen Ma-Mel-86C mit jeweils 10  $\mu$ g Plasmid-DNA durchgeführt, wie zuvor beschrieben (siehe Methodenteil, Abschnitt 2.2.1.13). Jeweils die Hälfte der Zellen wurde als Massenkultur in RPMI mit 10% FCS (RPMI<sub>FCS</sub>) unter Zugabe von 1 mg/ml Geneticin kultiviert. Die restlichen Zellen wurden mit 5.000 Zellen pro Testeinheit in 96-Well-Flachbodenplatten in RPMI<sub>FCS</sub> und 1 mg/ml Geneticin zur Generierung klonaler Zellpopulationen ausplattiert. An Tag 28 nach der Elektroporation wurden die bis dahin gewachsenen Klone gezählt (vgl. **Tab. 3.7**) und die Expression der transfizierten HLA-Klasse I-Moleküle in den ausreichend expandierten Transfektanten durch Durchflusszytometrie überprüft (siehe **Abb. 3.19**). Für jeden Transfektionsansatz konnte mindestens ein positiver Klon etabliert werden.

**Tab. 3.7: Klonierung der HLA-Transfektanten von Ma-Mel-86C.** Die Melanomzelllinie Ma-Mel-86C sollte stabil mit den angegebenen HLA-Klasse I-Allelen transfiziert werden. Ein Teil der gewachsenen Klone (Spalte 2) wurde an Tag 28 der Kultur mittels Durchflusszytometrie auf die Expression der transfizierten HLA-Allele überprüft (Spalte 3). In bis zu drei der getesteten Klone jeder Transfektante wurde die Expression des jeweiligen Transgens bestätigt (Spalte 4). Ergebnisse der Durchflusszytometrie sind exemplarisch in **Abb. 3.19** gezeigt.

Ma-Mel-86C transfiziert mit	Anzahl der gewachsenen Klone	Anzahl der getesteten Klone	Anzahl der positiven Klone
HLA-A*24:02	26	8	3
HLA-B*15:01	14	5	1
HLA-C*03:03	39	14	3

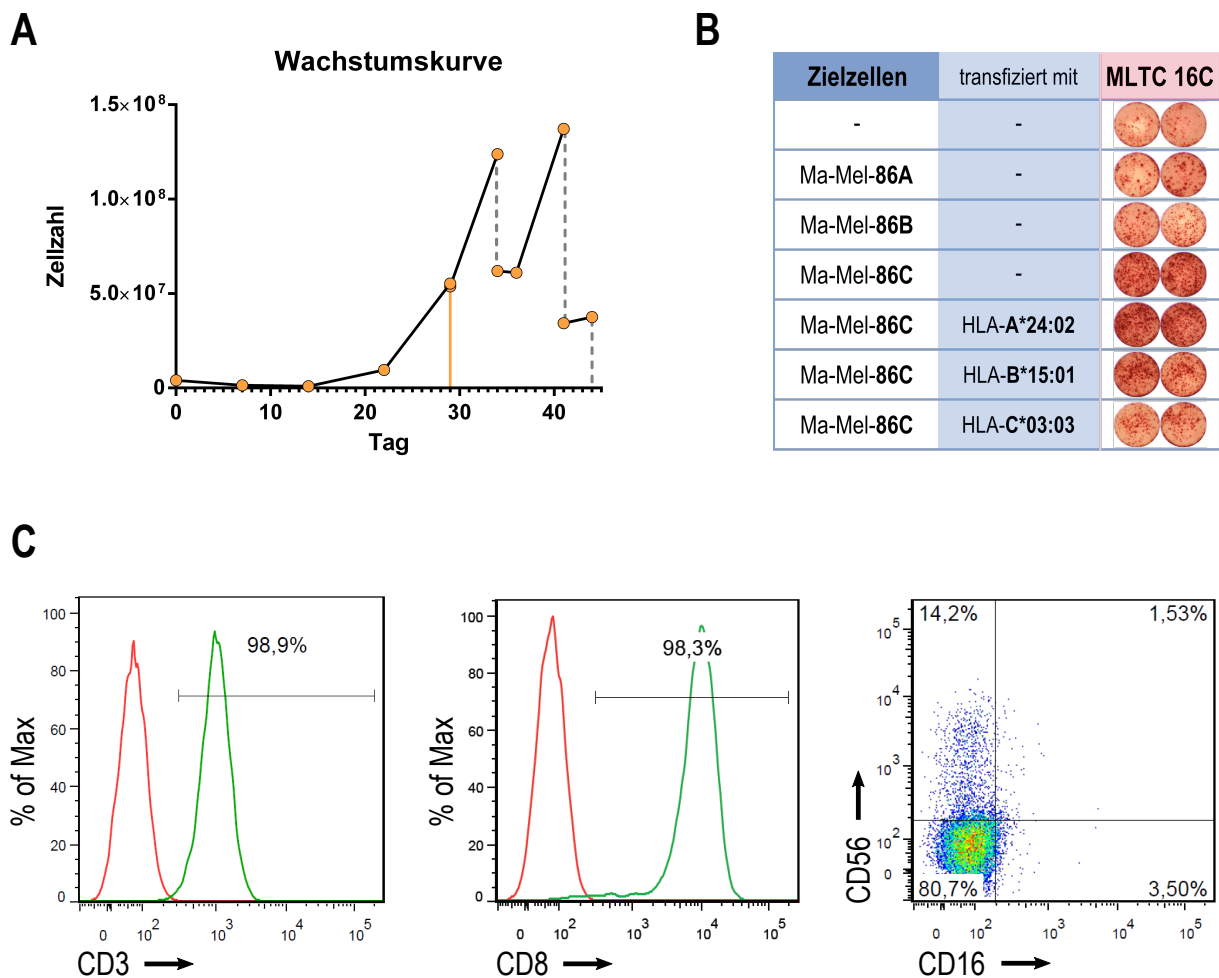
**A** Ma-Mel-86C + HLA-A\*24:02**B** Ma-Mel-86C + HLA-B\*15:01**C** Ma-Mel-86C + HLA-C\*03:03

**Abb. 3.19: Validierung der stabilen Expression der in Ma-Mel-86C transfizierten HLA-Klasse I-Moleküle.** Für die Transfektion von HLA-A\*24:02 (A), HLA-B\*15:01 (B) bzw. HLA-C\*03:03 (C) in Ma-Mel-86C wurden jeweils eine Massenkultur und eine Klonierung angelegt, die an Tag 28 der Kultur durchflusszytometrisch auf die Expression der transfizierten HLA-Allele überprüft wurden. Für jeden Transfektionsansatz ist ein positiver Klon (#A24, #B9 bzw. #C19) gezeigt. Die FACS-Färbung des negativen Klon #C29 ist zum Vergleich dargestellt. Die Beschriftungen der Histogramm-x-Achsen geben die Spezifitäten der Primärintikörper (grün: HLA-A\*24, blau: HLA-B\*15, gelb: HLA-B/C) an, die mithilfe eines FITC-markierten Sekundärintikörpers detektiert wurden. Die Färbungen mit den Isotyp-Kontrollen sind in Rot dargestellt. Für den Klon #C19 wurde die Expression von HLA-C\*03:03 zusätzlich durch eine HLA-C-spezifische PCR und Sequenzierung (nicht gezeigt) validiert (Details siehe Text).

Da für die durchflusszytometrische Bestätigung der Expression von HLA-C\*03:03 im Klon #C19 keine allelspezifischen Antikörper zur Verfügung standen, wurden die Ergebnisse der Durchflusszytometrie mit dem gegen HLA-B/C gerichteten Antikörper (B1.23.2) zusätzlich durch PCR und Sanger-Sequenzierung validiert. Hierfür wurde RNA aus der Transfektante Klon #C19 isoliert und cDNA synthetisiert. Mit HLA-C-spezifischen Primern wurde die Expression der HLA-C-mRNA in einer PCR überprüft. Die Sequenzierung des Amplifikats ergab Übereinstimmungen mit HLA-C\*03:03 (oder HLA-C\*03:01), wodurch die Expression von HLA-C\*03:03 in Klon #C19 bestätigt wurde (nicht gezeigt).

**3.5.2.2 Kultur der MLTC 16C**

Die MLTC 16C wurde aus bereits vorisolierten CD8<sup>+</sup> T-Zellen generiert, welche aus PBMCs vom 07.04.2004 stammten, wie im Methodenteil beschrieben (Abschnitt 2.2.1.7). Zur wöchentlichen Restimulation der MLTC wurden die in **Abb. 3.19** analysierten HLA-Transfektanten von Ma-Mel-86C in gleichen Anteilen nach Bestrahlung mit 10.000 rad (100 Gy) verwendet. Dafür wurden sie eine Passage vor der Stimulation ohne Selektionsantibiotikum kultiviert und in regelmäßigen Abständen auf die Expression der transfizierten HLA-Moleküle überprüft. Für die MLTC 16C wurde auf die Zugabe von *Feeder*-Zellen verzichtet.

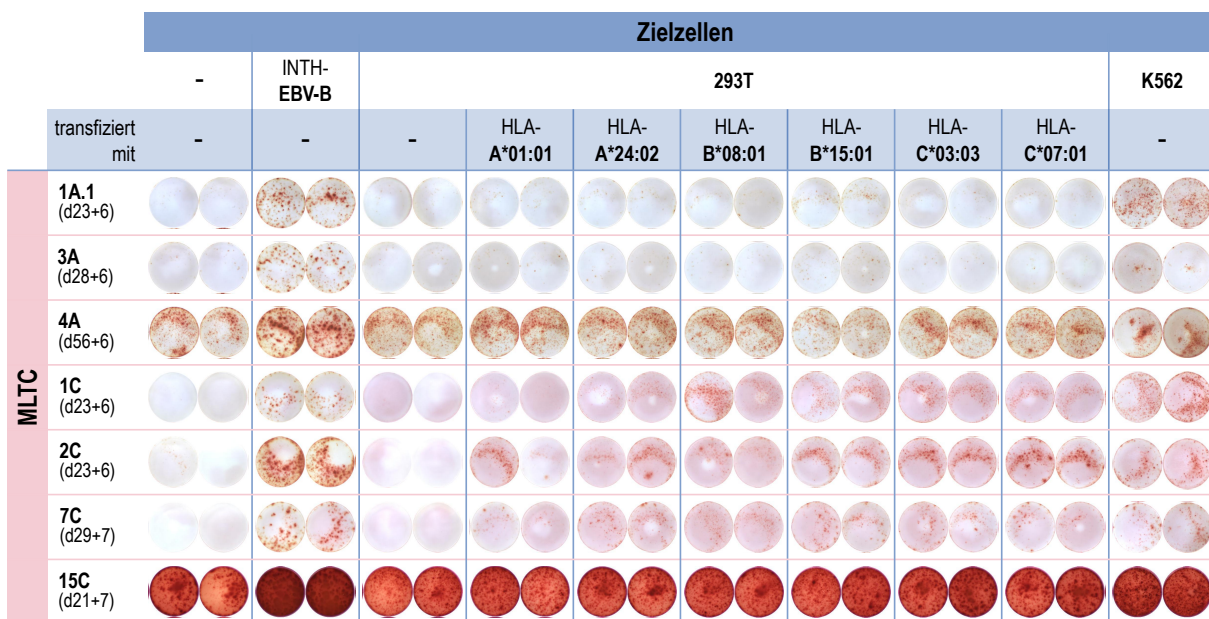


**Abb. 3.20: Wachstum und Charakterisierung der MLTC 16C.** (A) Nach Restimulation der MLTC 16C an Tag 14 wurde ein exponentielles Wachstum der *Responder*-Zellen beobachtet. Ab Tag 34 wurde die MLTC kryokonserviert (graue gestrichelte Linien). Die orangefarbene Linie markiert den Zeitpunkt des IFN $\gamma$ -ELISpot-Assays (B) und der durchflusszytometrischen Analyse der Zellen (C). (B) Die Reaktivität der MLTC 16C gegen die autologen Tumorzelllinien Ma-Mel-86A, -86B und -86C sowie die HLA-Transfektanten von Ma-Mel-86C wurde an Tag 29 der Kultur in einem IFN $\gamma$ -ELISpot-Assay bestimmt (MLTC 16C: 30.000 Zellen/Well, Tumorzelllinien: 50.000 Zellen/Well). (C) Mittels Durchflusszytometrie wurde an Tag 29 der Anteil CD8<sup>+</sup> T-Zellen in der MLTC 16C überprüft. Die x-Achsen der Histogramme sind mit den Spezifitäten der FITC-gekoppelten Antikörper beschriftet (grüne Kurven). Die roten Kurven zeigen jeweils die Isotyp-Kontrollen. Im *Dotplot* ist die Doppelfärbung der Oberflächenmoleküle CD16 (FITC-gekoppelt) und CD56 (PE-gekoppelt) zur Visualisierung natürlicher Killerzellen (CD16<sup>+</sup>CD56<sup>+</sup>) dargestellt, welche nicht nachgewiesen wurden.

Die Wachstumskurve in **Abb. 3.20A** zeigt, dass erst ab der zweiten Woche in Kultur ein Anstieg der Zellzahlen festgestellt werden konnte. An Tag 29 (orangefarbene Linie in **Abb. 3.20A**) wurde die Erkennung der autologen Tumorzelllinien Ma-Mel-86A, -86B und -86C sowie der HLA-Transfektanten von Ma-Mel-86C in einem IFN $\gamma$ -ELISpot-Assay überprüft (**Abb. 3.20B**). Aufgrund des starken Hintergrundes konnte die Reaktivität der MLTC 16C gegen Ma-Mel-86A und -86B nicht eindeutig geklärt werden. Ma-Mel-86C-Zellen sowie die HLA-Transfektanten wurden von der MLTC 16C erkannt, wobei die Expression der HLA-Allele keine gesteigerte Erkennung gegenüber untransfizierten Ma-Mel-86C zur Folge hatte. Zur Reduktion der spontanen IFN $\gamma$ -Sekretion wurden die Testbedingungen nochmals optimiert (siehe Abschnitt 3.5.4). Die T-Zellen exprimierten zum Teil den Aktivierungsmarker CD56, waren aber negativ für CD16. Somit waren in der MLTC 16C keine natürlichen Killerzellen expandiert worden.

### 3.5.3 Auswahl antigenpräsentierender Zellen für die Peptidbeladung

Um die Immunogenität der vorhergesagten Peptidkandidaten in  $\text{IFN}\gamma$ -ELISpot-Assays zu überprüfen, sollten die synthetisierten Peptide im Komplex mit HLA-Molekülen präsentiert werden. Da nicht genügend Patientenmaterial zur Verfügung stand, um autologe dendritische Zellen als professionelle antigenpräsentierende Zellen (APCs, engl. *antigen-presenting cells*) zu generieren, wurde der mögliche Einsatz der autologen B-Zelllinie INTH-EBV-B bzw. von 293T- oder K562-Zellen, die mit den HLA-Klasse I-Allelen der Patientin transfiziert waren, getestet (**Abb. 3.21**). Wichtig hierbei war, dass die EBV- oder Allo-Reaktivität und damit die Hintergrundaktivität der ausgewählten MLTCs möglichst gering war, um auch eine schwache Peptiderkennung erfassen zu können.

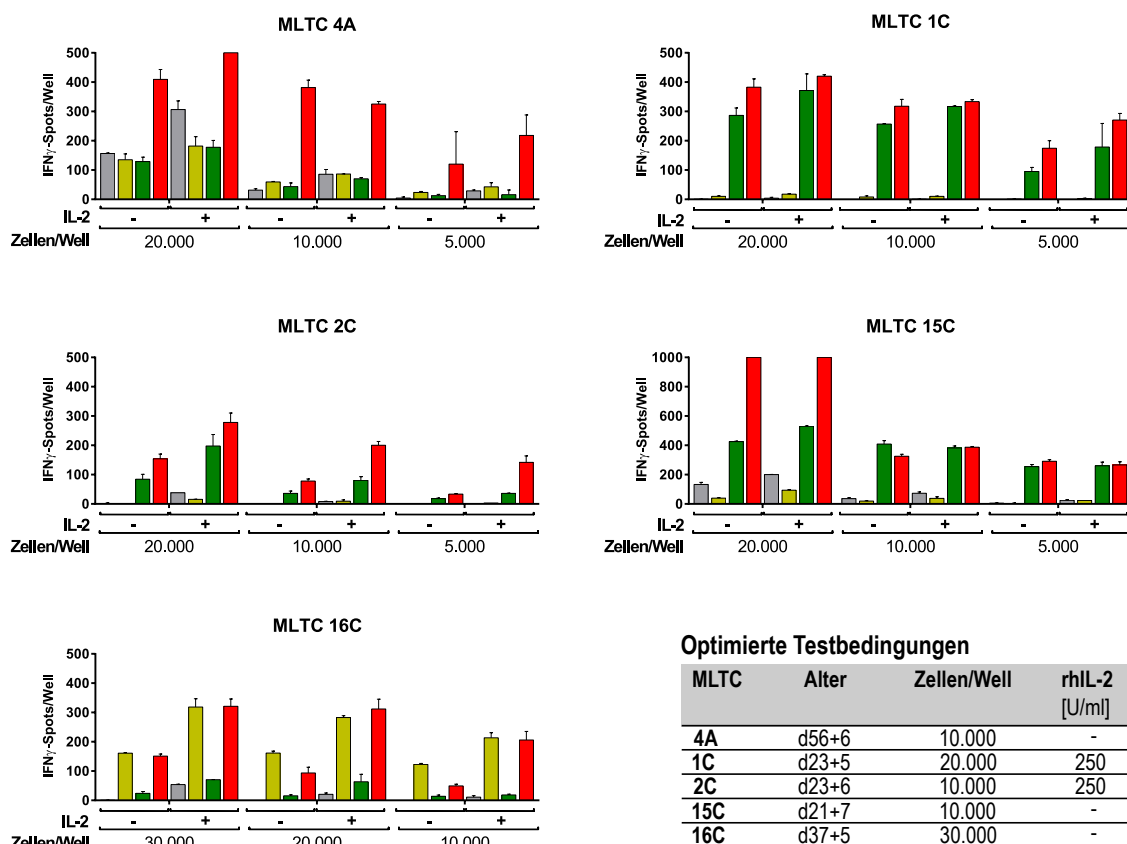


**Abb. 3.21: Auswahl der APCs für die Peptidbeladung.** Die Sekretion von  $\text{IFN}\gamma$  in der Anwesenheit von INTH-EBV-B (75.000 Zellen/Well), 293T- bzw. K562-Zellen (jeweils 30.000 Zellen/Well) durch die zuvor ausgewählten MLTCs (jeweils 30.000 Zellen/Well; MLTC 4A: 20.000 Zellen/Well) wurde in  $\text{IFN}\gamma$ -ELISpot-Assays getestet. 293T-Zellen wurden transient mit den einzelnen HLA-Klasse I-Allelen der Patientin INTH transfiziert und der Hintergrund gegen die Transfektanten (jeweils 30.000 Zellen/Well) ermittelt. Die MLTC 15C wurde auf *MultiScreen*-ELISpot-Platten (S2EM004M99, Millipore) getestet. Für alle anderen MLTCs wurden *MultiScreen*<sub>HTS</sub>-Filterplatten (MSIPS4510, Millipore) verwendet. Die MLTCs 1A.1 und 3A zeigten keine bzw. kaum  $\text{IFN}\gamma$ -Sekretion nach Koinkubation mit 293T-Zellen. Alle anderen MLTCs wiesen spätestens nach Transfektion der HLA-Klasse I-Allele einen Hintergrund auf. Der gezeigte Test wurde parallel zum ELISpot-Assay in Abb. 3.18 durchgeführt. Die Reaktivität gegen die Modellzelllinien (Positivkontrollen) ist dort dargestellt.

Die Koinkubation mit INTH-EBV-B bzw. K562-Zellen führte bei allen getesteten MLTCs zu einer starken  $\text{IFN}\gamma$ -Sekretion. Sie eigneten sich daher nicht für die geplanten Tests. 293T-Zellen eigneten sich uneingeschränkt nur für die MLTCs 1A.1 und 3A. Alle anderen MLTCs zeigten gegen die transfizierten Zellen einen zu hohen Hintergrund. Außerdem wiesen die MLTCs 4A und 15C starke spontane  $\text{IFN}\gamma$ -Sekretion auf. Aufgrund dessen und aufgrund der Umstellung auf sensitivere ELISpot-Platten wurden die Testbedingungen nochmals angepasst (siehe nächster Abschnitt).

### 3.5.4 Optimierung der Bedingungen zur Reduktion spontaner IFN $\gamma$ -Sekretion

Die MLTCs 4A, 1C, 2C, 15C und 16C zeigten unter Standardtestbedingungen (30.000 Zellen/Well, Zugabe von 250 U/ml rhIL-2) mit den sensitiveren *MultiScreen*-ELISpot-Platten (S2EM004M99, Millipore) eine zu starke spontane IFN $\gamma$ -Sekretion für die geplanten Peptiderkennungstests. Deshalb wurde die eingesetzte Zellzahl pro Testeinheit nochmals titriert und der Einfluss der Zugabe von rhIL-2 überprüft (**Abb. 3.22**). Mithilfe dieser Vortests wurden für alle MLTCs Testbedingungen definiert, die dazu geeignet waren, auch schwach erkannte Peptide detektieren zu können. Die Tabelle in **Abb. 3.22** fasst die wichtigsten Parameter zusammen.



**Abb. 3.22: Vortests der MLTCs zur Minimierung spontaner IFN $\gamma$ -Sekretion.** Durch Titration der Zellzahl im IFN $\gamma$ -ELISpot-Assay (*MultiScreen*-Filterplatten, S2EM004M99, Millipore) zwischen 30.000 und 10.000 Zellen/Well und durch Vergleich der Reaktivität mit und ohne die Zugabe von rhIL-2 (250 U/ml) sollten für die gezeigten MLTCs Testbedingungen bestimmt werden, bei denen die spontane IFN $\gamma$ -Sekretion (■: ohne *Target*-Zellen) möglichst gering, und gleichzeitig die Erkennung der autologen Tumorzelllinien (jeweils 50.000 Zellen/Well, ■: Ma-Mel-86A, ■: Ma-Mel-86B, ■: Ma-Mel-86C) noch stark genug war. Dargestellt sind jeweils die Mittelwerte aus Doppelwerten (Fehlerbalken: Standardfehler der Mittelwerte). Die kryokonservierten MLTCs wurden jeweils zwei Tage vor der Testbelegung in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Die fortan gewählten Testkonditionen sind in der Tabelle unten rechts zusammengefasst.

#### Zusammenfassung

Aus den zur Verfügung stehenden MLTCs des Modells wurden sieben Responderpopulationen aufgrund ihres Erkennungsmusters, der Stärke ihrer Reaktivität

und des Anteils HLA-abhängiger T-Zellantworten für die geplanten Peptiderkennungstests ausgewählt. Mit HLA-transfizierten Ma-Mel-86C-Zellen wurde außerdem die MLTC 16C generiert, um T-Zellen gegen Antigene zu expandieren, die in Ma-Mel-86C exprimiert und nur auf den verlorenen HLA-Klasse I-Allelen präsentiert wurden. Wenn nötig, wurden die Testbedingungen für den IFN $\gamma$ -ELISpot-Assay optimiert, um die spontane IFN $\gamma$ -Freisetzung mancher MLTCs zu minimieren und die Peptiderkennungstests zu ermöglichen.

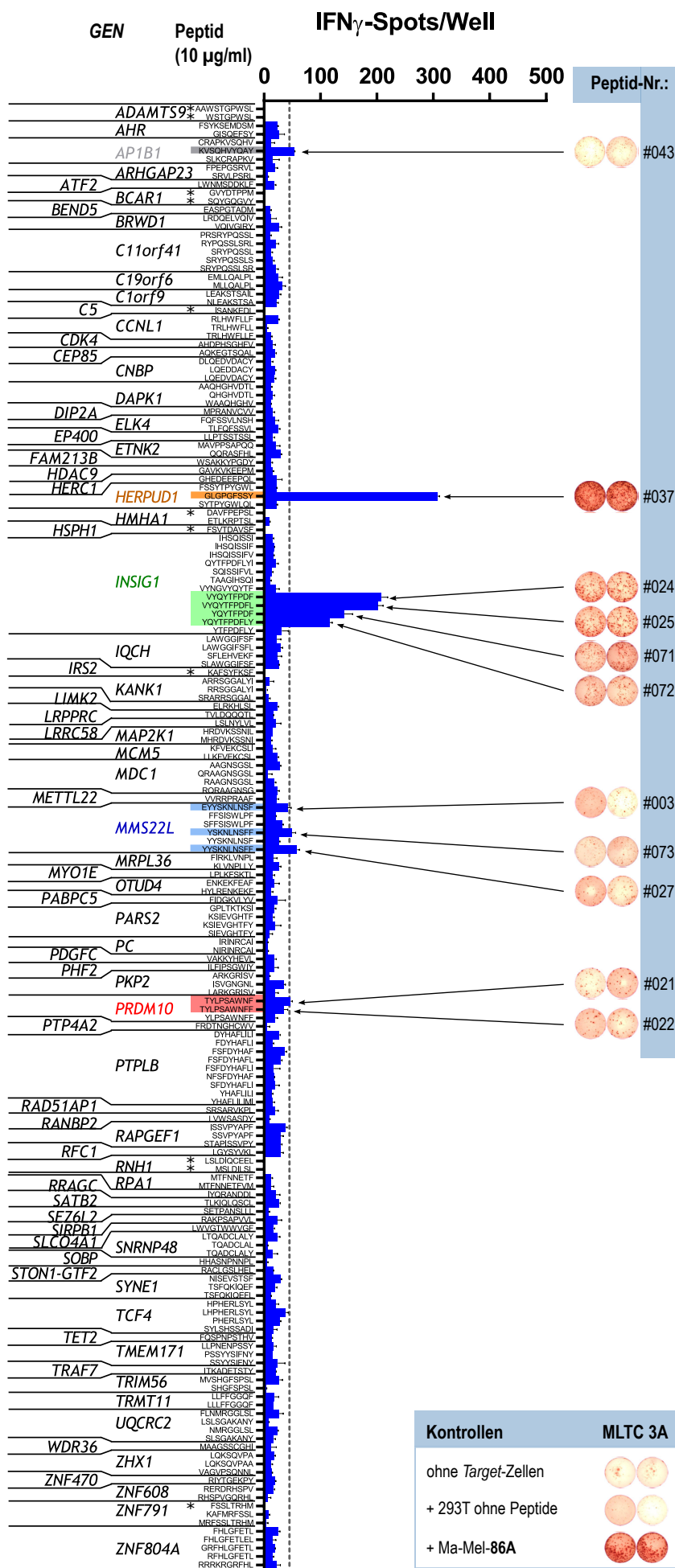
#### 3.5.5 Test der Erkennung der mutierten Peptide

Für die Überprüfung der Immunogenität der 174 Peptidkandidaten in IFN $\gamma$ -ELISpot-Assays wurden die ausgewählten Peptide durch den *High-Throughput Peptide Synthesis-Service* der Firma Peptide 2.0 (Chantilly, Virginia, USA) mit einer Reinheit von 20-80% synthetisiert. Im Erkennungstest wurden die Peptidpräparationen in einer Endkonzentration von 10  $\mu\text{g}/\text{ml}$  (entspricht ca. 10  $\mu\text{g}$ ) eingesetzt. Elf der Peptide waren abgeleitet von Proteinen, die in Ma-Mel-86C und nicht in Ma-Mel-86A mutiert waren, und sollten laut der Vorhersagealgorithmen nur über die HLA-Klasse I-Allele des in Ma-Mel-86C verlorenen Haplotyps (HLA-A\*24:02, -B\*15:01 und -C\*03:03) präsentiert werden. Daher wurde die Erkennung dieser elf Peptide nur mit der MLTC 16C, die mit HLA-rekonstituierten Ma-Mel-86C-Zellen stimuliert wurde, überprüft. Die übrigen 163 Peptide wurden mit allen ausgewählten MLTCs getestet.

Aufgrund der Ergebnisse der Vortests wurde zunächst die Peptidreaktivität der MLTC 3A analysiert. Als peptidpräsentierende Zellen wurden HLA-transfizierte 293T-Zellen verwendet. Diese wurden in sechs getrennten Ansätzen mit je einem der HLA-Klasse I-Allele der Patientin elektroporiert. Nach 24 h wurden im IFN $\gamma$ -ELISpot-Assays jeweils 10.000 Zellen jedes Transfektionsansatzes pro Testeinheit gemischt, so dass je Testeinheit insgesamt 60.000 antigenpräsentierende Zellen, die die sechs Ma-Mel-86-HLA-Klasse I-Allele exprimierten, eingesetzt wurden. Die Effizienz der Transfektion wurde mittels Durchflusszytometrie überprüft (nicht gezeigt) und lag je nach transfiziertem HLA-Allel zwischen 10 und 54%.

Für die in Doppelwerten getesteten Peptide wurden jeweils die Mittelwerte der gezählten Spots ( $\bar{x}_{\text{Erkanntes Peptid}}$ ) ermittelt und mit den gemittelten Spot-Counts der Hintergrundaktivität gegen die transfizierten, aber nicht mit Peptid beladenen 293T-Zellen verglichen ( $\bar{x}_{\text{Hintergrund}}$ ). Als Grenzwert für eine Peptiderkennung wurde der Mittelwert des Hintergrundes plus dessen dreifache Standardabweichung ( $\sigma_{\text{Hintergrund}}$ ) festgelegt:

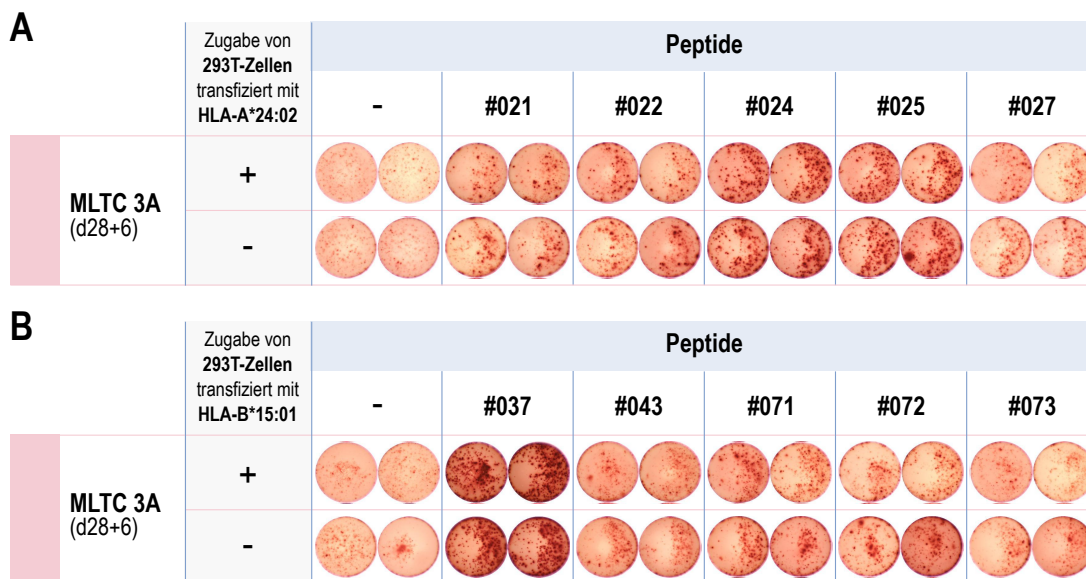
$$\bar{x}_{\text{Erkanntes Peptid}} \geq \bar{x}_{\text{Hintergrund}} + 3 \times \sigma_{\text{Hintergrund}}$$



**Abb. 3.23: Peptiderkennung durch MLTC 3A.** 163 der 174 ausgewählten mutierten Peptide mit einer vorhergesagten HLA-Bindeaffinität von  $\leq 500$  nM (NetMHC) und einem Prozentsatz von  $\leq 10$  (IEDB) wurden auf die Erkennung durch die MLTC 3A (23.000 Zellen/Well) in einem IFN<sub>γ</sub>-ELISpot-Assay getestet (Peptidkonzentration: 10 µg/ml, \*: nicht getestet). Als peptidpräsentierende Zellen wurden 293T-Zellen verwendet, die in sechs getrennten Ansätzen mit den HLA-Klasse I-Allelen der Patientin INTH transient transfiziert worden waren. Jeweils 10.000 Zellen/Well wurden von jeder Transfektante im Test eingesetzt. Das Balkendiagramm zeigt die Mittelwerte der gemessenen Spot-Counts (n=2, Fehlerbalken: Standardfehler des Mittelwerts). Farblich markiert sind die elf Peptide, die entsprechend den vordefinierten Bedingungen (gestrichelte Linie) Reaktivität induzierten sowie die fünf Gene, von denen die erkannten Peptide abgeleitet waren. Kontrollen: Testeinheiten ohne Target-Zellen, mit HLA-transfizierten 293T-Zellen bzw. mit Ma-Mel-86A.

Elf der 163 getesteten Peptide wurden von den T-Zellen erkannt (siehe **Abb. 3.23**, farblich gekennzeichnet). Unter ihnen war das zuvor identifizierte Peptid aus dem durch cDNA-Bank-Screening identifizierten Antigen HERPUD1<sup>mutt</sup> (Peptidsequenz: GLGPGFSSY, siehe Einleitung, Abschnitt 1.1.1.3.3). Die anderen erkannten Peptide waren abgeleitet von den vier mutierten Genen *AP1B1* (engl. *adaptor-related protein complex 1, beta 1 subunit*), *INSIG1* (engl. *insulin induced gene 1*), *MMS22L* (engl. *MMS22-like, DNA repair protein*) und *PRDM10* (engl. *PR domain containing 10*). Für die erkannten Peptide war eine Präsentation durch HLA-A\*24:02 bzw. HLA-B\*15:01 vorhergesagt worden.

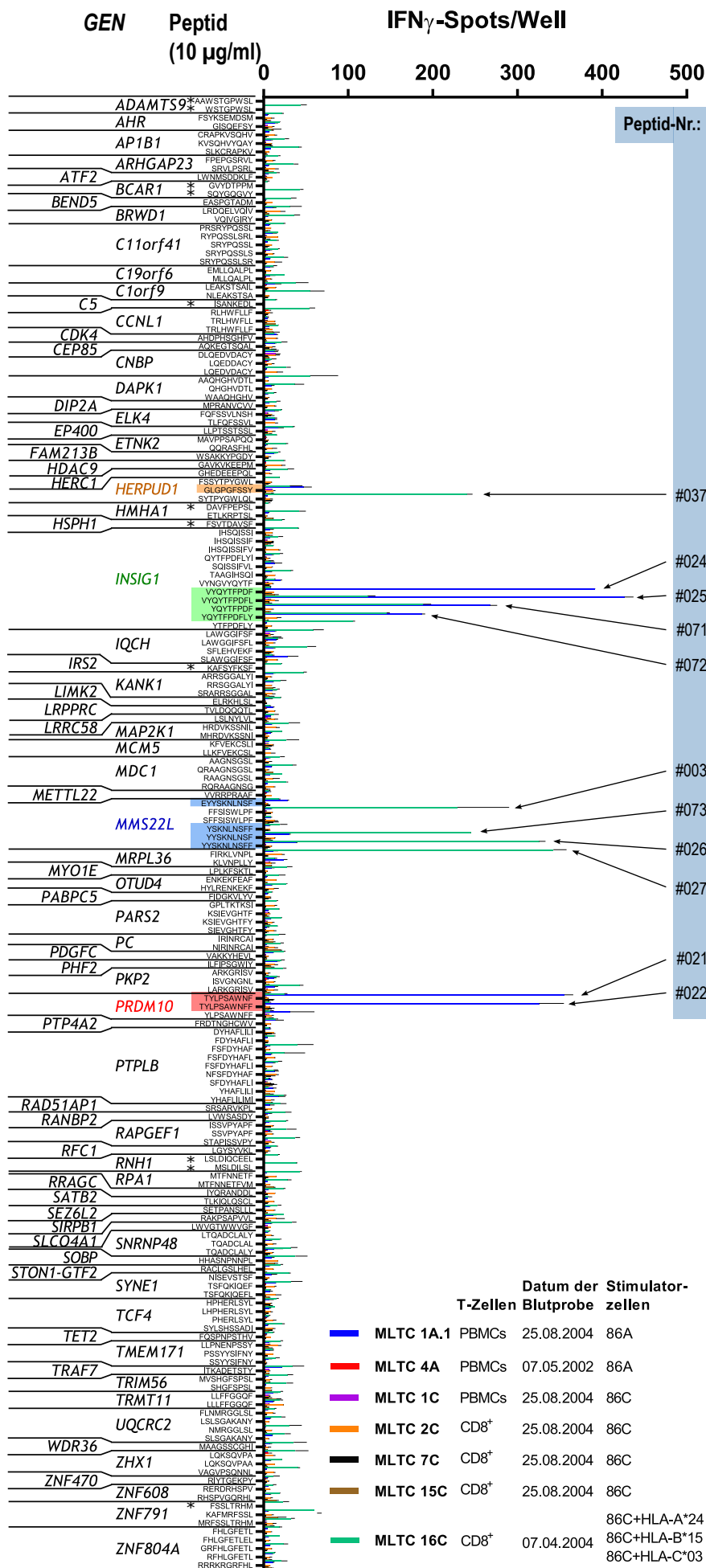
Aufgrund der teilweise schwachen Reaktivität sollte die Erkennung der Peptide in einem zweiten IFN $\gamma$ -ELISpot-Assay validiert werden (**Abb. 3.24**). Abgesehen von den Reaktivitäten gegen das AP1B1<sup>mutt</sup>-Peptid #043 (Peptidsequenz: KVSQHVYQAY) und das MMS22L<sup>mutt</sup>-Peptid #073 (Peptidsequenz: YSKNLNSFF), die nicht sicher bestätigt werden konnten, wurden alle anderen Peptide erneut von den T-Zellen erkannt. Dabei stellte sich heraus, dass eine Zugabe von HLA-transfizierten peptidpräsentierenden Zellen für die Peptiderkennungstests nicht notwendig war. Allein die Zugabe der Peptide zur MLTC 3A führte zu einer Erkennung, die mit der gegen HLA-transfizierte und mit den Peptiden beladenen 293T-Zellen vergleichbar war (**Abb. 3.24**). Daraus wurde geschlossen, dass die T-Zellen sich die Peptide gegenseitig präsentierten.



**Abb. 3.24: Validierung der Peptidreaktivität der MLTC 3A in An- und Abwesenheit HLA-transfizierter peptidpräsentierender 293T-Zellen.** Die Erkennung der im ersten Peptidtest (siehe Abb. 3.23) von der MLTC 3A erkannten Peptide sollte in einem zweiten IFN $\gamma$ -ELISpot-Assay bestätigt werden (Peptidkonzentration: 10  $\mu$ g/ml). Die an Tag 28+4 kryokonservierte MLTC wurde zwei Tage vor der Testbelegung in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen und im Test mit 25.000 Zellen/Well eingesetzt. Als peptidpräsentierende Zellen wurden 293T-Zellen entweder mit HLA-A\*24:02 (A) oder mit HLA-B\*15:01 (B) transfiziert und mit 10.000 Zellen/Well eingesetzt (+). Zum Vergleich wurde die Erkennung ohne HLA-transfizierte 293T-Zellen getestet (-).

Aufgrund dieser Beobachtung wurden die weiteren Peptiderkennungstests mit den MLTCs 1A.1, 4A, 1C, 2C, 7C, 15C und 16C (**Abb. 3.25**) ohne Zugabe von peptidpräsentierenden Zellen durchgeführt.





**Abb. 3.25: Peptid-erkennungstests mit weiteren MLTCs.** Die zuvor ausgewählten mutierten Peptide wurden auf die Erkennung durch die MLTCs 1A.1 (d30+7, 30.000 Zellen/Well), 4A (d56+6, 10.000 Zellen/Well), 1C (d23+5, 20.000 Zellen/Well), 2C (d23+6, 10.000 Zellen/Well), 7C (d34+6, 20.000 Zellen/Well), 15C (d21+7, 10.000 Zellen/Well) und 16C (d37+6, 30.000 Zellen/Well) in IFN<sub>γ</sub>-ELISpot-Assays getestet (Peptidkonzentration: 10 µg/ml, \*: nur mit MLTC 16C getestet). In der Legende sind die Charakteristika der sieben MLTCs zusammengefasst. Alle MLTCs außer der MLTC 7C wurden ein bis zwei Tage vor der Testbelegung in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Die MLTC 7C wurde aufgetaut, mit bestrahlten Ma-Mel-86C restimuliert und sechs Tage später im Test eingesetzt. Um den Hintergrund zu reduzieren, wurde in den Tests der MLTCs 4A, 15C und 16C kein rhIL-2 verwendet. Für die Tests wurden die Peptide direkt in die MLTCs gegeben. Das Balkendiagramm zeigt die Mittelwerte der gemessenen Spotzahlen (n=2, schwarze Fehlerbalken: Standardfehler des Mittelwerts). Farblich markiert sind die elf Peptide, die spezifisch von mindestens einer MLTC erkannt wurden sowie die Gene, von denen die erkannten Peptide abgeleitet waren.

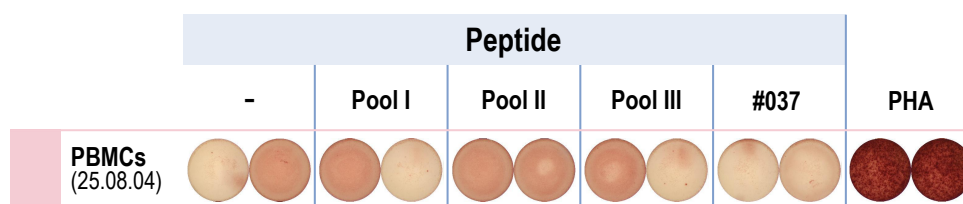
Wie in **Abb. 3.25** gezeigt, wurde nur in den MLTCs 1A.1 und 16C (blaue bzw. grüne Balken) eine Peptidreaktivität oberhalb des gesetzten Grenzwertes detektiert. Die Immunogenität der bereits im Test mit der MLTC 3A erkannten Peptide konnte mit einer Ausnahme bestätigt werden. Das Peptid #043 (AP1B1<sup>mut</sup>) wurde durch keine weitere MLTC erkannt. Im Gegensatz dazu induzierte ein weiteres MMS22L<sup>mut</sup>-Peptid (Peptid #026, Sequenz: YSKNLNSF) in der MLTC 16C eine starke IFN $\gamma$ -Freisetzung. Die zuvor fragliche Erkennung des MMS22L<sup>mut</sup>-Peptids #073 (Sequenz: YSKNLNSFF) wurde durch MLTC 16C bestätigt. Außerdem wurde keines der elf Peptide erkannt, die nur mittels MLTC 16C getestet wurden.

Anhand der Balkendiagramme wurde auch ersichtlich, dass in den einzelnen MLTCs unterschiedliche Peptidantworten dominierten. So war die Reaktivität gegen HERPUD1<sup>mut</sup> und INSIG1<sup>mut</sup> in der MLTC 3A am stärksten. Im Gegensatz dazu wurde HERPUD1<sup>mut</sup> von der MLTC 1A.1 deutlich weniger stark erkannt. Hier lagen insbesondere T-Zellantworten gegen INSIG1<sup>mut</sup> und PRDM10<sup>mut</sup> vor. In der MLTC 16C dominierte die Reaktivität gegen die MMS22L<sup>mut</sup>-Peptide.

### 3.5.5.1 *Ex-vivo*-Test

Nach Abschließen der Peptiderkennungstests mit den vorstimulierten MLTC-Responderpopulationen sollte überprüft werden, ob T-Zellantworten gegen die immunogenen Peptide auch *ex vivo* in unstimulierten PBMCs detektiert werden konnten. Dies würde einen Hinweis darauf geben, ob für vergleichbare Peptid-Screenings bei anderen Patienten stabile Tumorzelllinien zur Stimulation von autologen MLTCs benötigt würden. Als Detektionsgrenze des IFN $\gamma$ -ELISpot-Assays wurde 20 in 10<sup>5</sup> CD8<sup>+</sup> T-Zellen angenommen (Lennerz *et al.*, 2005).

Der *Ex-vivo*-Test wurde mit kryokonservierten PBMCs vom 25.08.2004 durchgeführt. Aus PBMCs bzw. CD8<sup>+</sup> T-Zellen dieser Blutabnahme waren die MLTCs 1A.1, 1C, 2C, 7C und 15C generiert worden. Zumindest in der MLTC 1A.1 konnte Reaktivität gegen die mutierten Peptide nachgewiesen werden. Der *Ex-vivo*-Assay wurde, wie im Methodenteil (Abschnitt 2.2.1.14.2) beschrieben, mit 500.000 PBMCs pro Testeinheit in AIM-V-Medium ohne Humanserum und ohne rhIL-2 durchgeführt. Standardmäßig wurde 1  $\mu$ g/ml von jedem Peptid eingesetzt. Der Anteil der CD8<sup>+</sup> T-Zellen wurde mangels ausreichender Zellmenge nicht durchflusszytometrisch ermittelt.

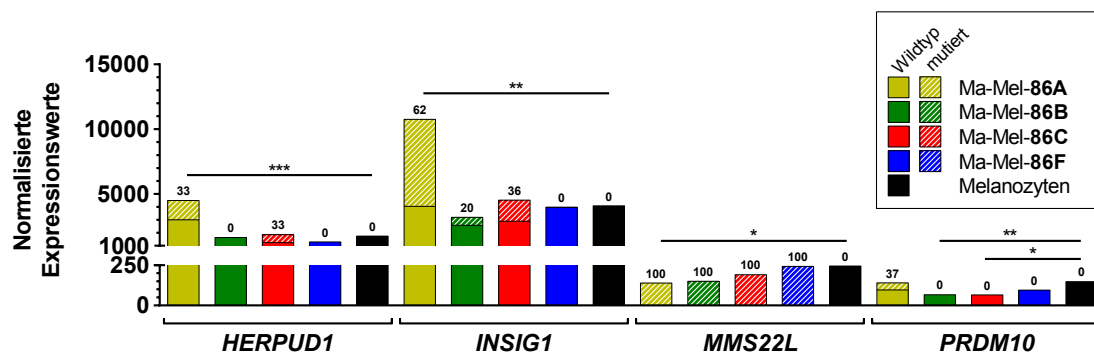


**Abb. 3.26:** Testung der *Ex-vivo*-Reaktivität gegen mutierte Peptide. Die Reaktivität von unstimulierten PBMCs (Blutabnahme vom 25.08.2004, 500.000 Zellen/Well) gegen die zuvor als immunogen identifizierten Peptide (je 1  $\mu$ g/ml; Pool I: #021, #022; Pool II: #024, #025, #071, #072; Pool III: #003, #026, #027, #073; Peptid #037) wurde in einem IFN $\gamma$ -ELISpot-Assay bestimmt. Mit Phytohämagglutinin (PHA, 10  $\mu$ g/ml) aktivierte PBMC dienten als Positivkontrolle.

Während eine starke Reaktion der T-Zellen auf PHA deren generelle Fähigkeit zeigte,  $\text{IFN}\gamma$  zu produzieren, konnten keine spezifischen T-Zellantworten gegen die mutierten Peptide nachgewiesen werden (**Abb. 3.26**). Die Frequenz der peptidreaktiven T-Zellen im peripheren Blut der Patientin war somit zu gering, um ohne eine vorherige Anreicherung in einer MLTC detektiert werden zu können.

### 3.5.6 Expressionsanalyse der immunogenen Mutationen auf mRNA-Ebene

Die Analyse der Expressionsstärke der Gene mit den immunogenen Mutationen sollte Aufschluss darüber geben, wie stark sie auf mRNA-Ebene im Vergleich zu den Wildtyp-homologen exprimiert waren. In den Transkriptomsequenzierungen der vier Tumorzelllinien wurden daher die Expressionslevel von *HERPUD1*, *INSIG1*, *MMS22L* und *PRDM10* anhand der Zahl der alignierten *Reads* (siehe Methodenteil, Abschnitt 2.2.3.8) bestimmt und jeweils der Anteil des Wildtyp- bzw. des mutierten Allels ermittelt (**Abb. 3.27**). Zum Vergleich wurden wiederum die Melanozytendaten aus dem SRA hinzugezogen. Die Analyse differentieller Expression erfolgte mithilfe des R-Pakets *DESeq*.



**Abb. 3.27: mRNA-Expressionslevel der Gene mit immunogenen Mutationen.** Das Säulendiagramm zeigt die auf die Größe der Datensätze (*DESeq, size factors*) und auf die Länge des offenen Leserahmens (in kb) normalisierten Expressionswerte (Mittelwerte aus je drei Replikaten; Melanozyten als Einzelwert) der Gene *HERPUD1*, *INSIG1*, *MMS22L* und *PRDM10*, die aus den *RNA-Seq*-Daten der Tumorzelllinien des Modells und der primären Melanozyten (SRA-Accession: SRX101662) ermittelt wurden. Schraffiert ist jeweils der Anteil des mutierten Allels dargestellt. Die Zahlen oberhalb der Säulen geben die dazugehörigen Prozentwerte an. Im Falle differentieller Expressionen im Vergleich zu den Melanozyten ist das Signifikanzniveau der adjustierten *p*-Werte der *DESeq*-Expressionsanalyse angegeben ( $p < 0,005$ : \*,  $p < 0,0005$ : \*\*,  $p < 0,00005$ : \*\*\*).

Die allgemeine Expressionsanalyse zeigte, dass *HERPUD1* und *INSIG1* insgesamt (Wildtyp und mutiert) deutlich stärker exprimiert waren als *MMS22L* und *PRDM10*. In Ma-Mel-86A war die Expression von *HERPUD1* und *INSIG1* nochmals höher als in den übrigen Datensätzen. Im Vergleich zu den Melanozyten waren diese Gene in Ma-Mel-86A ca. 2,6-fach hochreguliert (adjustierte *p*-Werte:  $1,83 \times 10^{-6}$  bzw.  $1,26 \times 10^{-4}$ ). Die Expression von *MMS22L* in Ma-Mel-86A war im Gegensatz dazu herabreguliert (*Fold Change*: 0,57; adjustierter *p*-Wert: 0,00378). In Ma-Mel-86B und -86C war eine signifikante Herabregulation von *PRDM10* zu beobachten (*Fold change*: 0,45 bzw. 0,44; adjustierte *p*-Werte:  $8,90 \times 10^{-5}$

bzw.  $5,84 \times 10^{-4}$ ), wobei die immunogene Mutation in keiner der beiden Zelllinien vorlag. Abgesehen davon waren die Expressionslevel der vier analysierten Gene in den Melanozyten ähnlich hoch wie in den Tumorzelllinien.

Der Vergleich zwischen den Wildtypallelen und den mutierten Allelen ergab, dass das mutierte Allel bei *HERPUD1* und *PRDM10* unterrepräsentiert war (33% bzw. 37%), wohingegen es bei *INSIG1* zumindest in Ma-Mel-86A (62%) überwog. *MMS22L* hingegen lag in allen vier Melanomzelllinien nur mutiert vor. Trotz der teilweise schwächeren mRNA-Expression der mutierten Allele konnte durch die Stimulation der MLTCs mit Ma-Mel-86A mutationsspezifische T-Zellen expandiert werden.

### Zusammenfassung

Die 174 Peptide, die laut den Vorhersagen von IEDB und NetMHC an die HLA-Klasse I-Moleküle der Patientin INTH binden konnten, wurden mit einer Reinheit von 20-80% synthetisiert. 163 dieser Peptide wurden auf die Erkennung durch sieben verschiedene autologe MLTCs getestet. Die Immunogenität weiterer elf Peptide wurde nur in der MLTC 16C überprüft. Von den 174 wurden elf Peptide von bis zu drei MLTCs erkannt. **Tab. 3.8** gibt hierzu eine Zusammenfassung.

**Tab. 3.8: Übersicht zur Erkennung der Peptide.** Die gelisteten elf Peptide wurden von den MLTCs 1A.1, 3A bzw. 16C erkannt (×: Erkennung nach den in 3.5.5 formulierten Kriterien). Die mutierte Aminosäure ist in den Peptidsequenzen jeweils rot hervorgehoben. Laut der Vorhersagen von NetMHC und IEDB wurden die Peptide entweder durch HLA-A\*24:02 oder HLA-B\*15:01 präsentiert. In der Spalte Codonaustausch sind das jeweilige Wildtypcodon und das zugehörige mutierte Codon angegeben, wobei die mutierten Basen in Großbuchstaben geschrieben wurden. Die Ergebnisse aller vier genutzten Algorithmen sind angegeben (/: keine Vorhersage für an HLA-B\*15-bindende Oktamere möglich).

Peptid-Nr.	Sequenz	Bindung an HLA- (lt. Vorhersage)	Gen (Aminosäure-austausch)	Codon-austausch	Peptid erkannt von MLTC			SYFPEITHI	BIMAS [s]	IEDB	NetMHC [nM]
					1A.1	3A	16C				
021	TYLPSAWN <b>F</b>	A*24	<i>PRDM10</i> (S1050F)	tCc/tTc	×	×		24	180	0.526	17
022	TYLPSAWN <b>F</b>	A*24			×	×		24	216	0.526	11
024	VYQYT <b>F</b> PDF	A*24	<i>INSIG1</i> (S238F)	tCc/tTc	×	×	×	23	180	0.526	15
025	VYQYT <b>F</b> PDFL	A*24			×	×	×	23	300	2.105	84
071	YQYT <b>F</b> PDF	B*15			×	×	×	/	/	0.526	30
072	YQYT <b>F</b> PDFLY	B*15			×	×	×	21	40	0.789	14
003	EYYSKNLNS <b>F</b>	A*24	<i>MMS22L</i> (S437F)	tCc/tTc	×	×	×	20	120	0.526	80
026	YYSKNLNS <b>F</b>	A*24				×	×	21	120	0.263	10
027	YYSKNLNS <b>F</b> F	A*24				×	×	21	120	0.526	12
073	YSKNLNS <b>F</b> F	B*15				×	×	9	6	4.211	35
037	GLGPGFSS <b>Y</b>	B*15	<i>HERPUD1</i> (G161S)	Ggt/Agt	×	×	×	23	192	2.105	31

Die immunogenen Peptide leiteten sich von den vier mutierten Genen *PRDM10*, *INSIG1*, *MMS22L* und *HERPUD1* ab, wobei das *HERPUD1*<sup>mut</sup>-Peptid bereits durch cDNA-Bank-Screening identifiziert worden war. Für die immunogenen Peptide

war jeweils eine Präsentation über HLA-A\*24:02 oder HLA-B\*15:01 vorhergesagt worden. Diese HLA-Klasse I-Allele wurden nur auf Ma-Mel-86A und den jeweiligen HLA-Transfektanten von Ma-Mel-86C exprimiert.

## 3.6 Generierung mutationsspezifischer T-Zellklone aus MLTC-Responderpopulationen

Zur Isolation von peptidreaktiven T-Zellklonen wurden die MLTC 1A.1, 3A und 16C mittels des Grenzverdünnungsverfahrens kloniert (vgl. Methodenteil, Abschnitt 2.2.1.8). Die T-Zellklone wurden zunächst wöchentlich mit bestrahlten Tumorzellen restimuliert und nach ca. drei Wochen auf die Erkennung der immunogenen Peptide (jeweils 5 µg/ml Endkonzentration) getestet. Um die Tests möglichst früh mit wenigen T-Zellen durchführen zu können, wurden die Peptide folgendermaßen gepoolt:

- Pool I: Peptide #021 und #022 (jeweils PRDM10<sup>mut</sup>)
- Pool II: Peptide #024, #025, #071 und #072 (jeweils INSIG1<sup>mut</sup>)
- Pool III: Peptide #003, #026, #027 und #073 (jeweils MMS22L<sup>mut</sup>)

Da für das Antigen HERPUD1<sup>mut</sup> bereits T-Zellklone vorhanden waren, wurde auf die Testung des HERPUD1<sup>mut</sup>-Peptids verzichtet. T-Zellklone, die die neu identifizierten mutierten Peptide erkannten, wurden anschließend durch unspezifische Restimulation mit OKT-3 expandiert, die Reaktivität wurde nochmals überprüft und die T-Zellen wurden kryokonserviert.

### 3.6.1 Klonierung der MLTC 1A.1

An Tag 35 kryokonservierte CD8<sup>+</sup> T-Zellen der MLTC 1A.1 wurden aufgetaut und einmal mit Ma-Mel-86A restimuliert. Eine Woche nach der Restimulation wurde die MLTC auf fünf 96-Well-Rundbodenplatten mit je 1 Zelle/Well kloniert. Nach 20 Tagen wurden Aliquots von 67 gewachsenen T-Zellklonen (Durchschnitt: 13,4 pro Platte) ohne Bestimmung der Zellzahl in einem IFN $\gamma$ -ELISpot-Assay auf die Erkennung des Peptid-Pools I und einer Mischung der Pools II und III überprüft. Bei der erzielten Klonierungseffizienz von 14,0% lag die statistische Mindestwahrscheinlichkeit für Klonalität bei 92,7% (Berechnung siehe Abschnitt 2.2.1.8). 14 der 67 getesteten T-Zellklone erkannten Pool I und elf dieser PRDM10<sup>mut</sup>-reaktiven T-Zellklone konnten mittels OKT-3-Restimulation expandiert werden. Außerdem erkannten 32 T-Zellklone die gemischten Pools II und III. Die Spezifität dieser T-Zellklone wurde in einem zweiten IFN $\gamma$ -ELISpot-Assay ermittelt. Hierfür wurden die Pools II und III getrennt getestet und die T-Zellklone zu 10.000 Zellen pro Testeinheit eingesetzt. Alle T-Zellklone erkannten INSIG1<sup>mut</sup> (Pool II). Die acht T-Zellklone mit der stärksten INSIG1<sup>mut</sup>-Reaktivität wurden für die OKT-3-Restimulation ausgewählt.

### 3.6.2 Klonierung der MLTC 3A

Die Klonierung der MLTC 3A erfolgte auf neun 96-Well-Rundbodenplatten mit je 1 Zelle/Well drei Tage nach dem Auftauen der an Tag 32 kryokonservierten CD8<sup>+</sup> T-Zellen. Nach 20 Tagen konnten 226 gewachsene T-Zellklone auf die Erkennung der Peptid-Pools I bis III getestet werden. Nach weiteren zwei Wochen wurden weitere 50 T-Zellklone nachgetestet (Durchschnitt insgesamt: 30,7 pro Platte). Aufgrund des hohen Anteils positiver Kultureinheiten (etwa 31,9%) betrug die statistische Wahrscheinlichkeit für Klonalität nur mindestens 82,0%. Insgesamt erkannten sechs der 276 mutmaßlichen T-Zellklone die von INSIG1<sup>mut</sup> abgeleiteten Peptide (Pool II). Vier von ihnen wurden erfolgreich durch OKT-3-Restimulation expandiert. Für den in Abschnitt 3.7.2 eingesetzten T-Zellklon 3A/115 wurde mittels Durchflusszytometrie ermittelt, dass die unspezifisch expandierten T-Zellen ausschließlich die TCR-Kette V $\beta$ 2 exprimierten (nicht gezeigt), sodass von einer klonalen T-Zellpopulation auszugehen war.

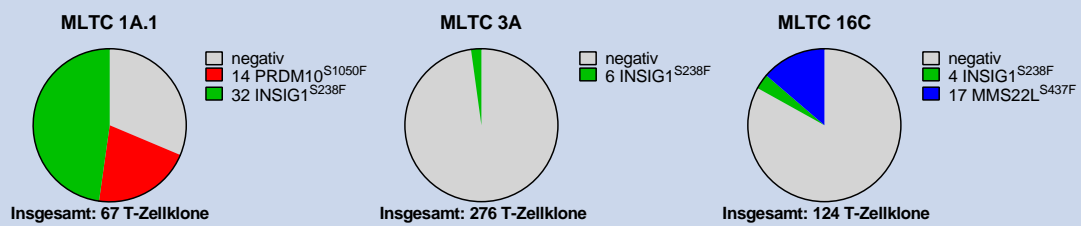
### 3.6.3 Klonierung der MLTC 16C

Für die Klonierung der MLTC 16C wurden an Tag 41 eingefrorene CD8<sup>+</sup> T-Zellen aufgetaut und nach vier Tagen mit je 1 Zelle/Well auf zehn 96-Well-Rundbodenplatten ausplattiert. Für die wöchentliche Restimulation wurden die Transfektanten von Ma-Mel-86C, die mit HLA-A\*24:02 bzw. -B\*15:01 ausgestattet waren, zu gleichen Anteilen verwendet, da die immunogenen Peptide laut Vorhersage entweder von HLA-A\*24:02 oder von HLA-B\*15:01 präsentiert werden sollten. An Tag 19 wurden 124 gewachsene T-Zellklone (Durchschnitt: 12,4 pro Platte) auf die Erkennung der Pools II und III überprüft. Pool I wurde nicht getestet, da *PRDM10* in Ma-Mel-86C nicht mutiert ist. Aus der MLTC 16C wurden so vier weitere INSIG1<sup>mut</sup>-reaktive (Pool II) und 17 MMS22L<sup>mut</sup>-reaktive (Pool III) T-Zellklone isoliert. Ein gegen INSIG1<sup>mut</sup> gerichteter und 14 MMS22L<sup>mut</sup>-reaktive T-Zellklone wurden mittels OKT-3-Restimulation expandiert. Bei der Klonierung der MLTC 16C ist basierend auf der Klonierungseffizienz von 12,9% eine statistische Klonalitätswahrscheinlichkeit von mindestens 93,2% anzunehmen.

#### Zusammenfassung

Mithilfe des Grenzverdünnungsverfahrens wurden aus den MLTCs 1A.1, 3A und 16C T-Zellklone isoliert, die gegen die drei neu identifizierten Neoantigene gerichtet waren. Dabei waren jeweils unterschiedliche Frequenzen peptidspezifischer T-Zellklone zu beobachten (**Abb. 3.28**). In der Klonierung der MLTC 1A.1 dominierten INSIG1<sup>mut</sup>- und PRDM10<sup>mut</sup>-reaktive T-Zellen. Im Gegensatz dazu waren fast 98% der T-Zellen der 3A-Klonierung negativ für die getesteten Peptid-Pools. Durch die Klonierung der MLTC 16C konnten insbesondere MMS22L<sup>mut</sup>-spezifische

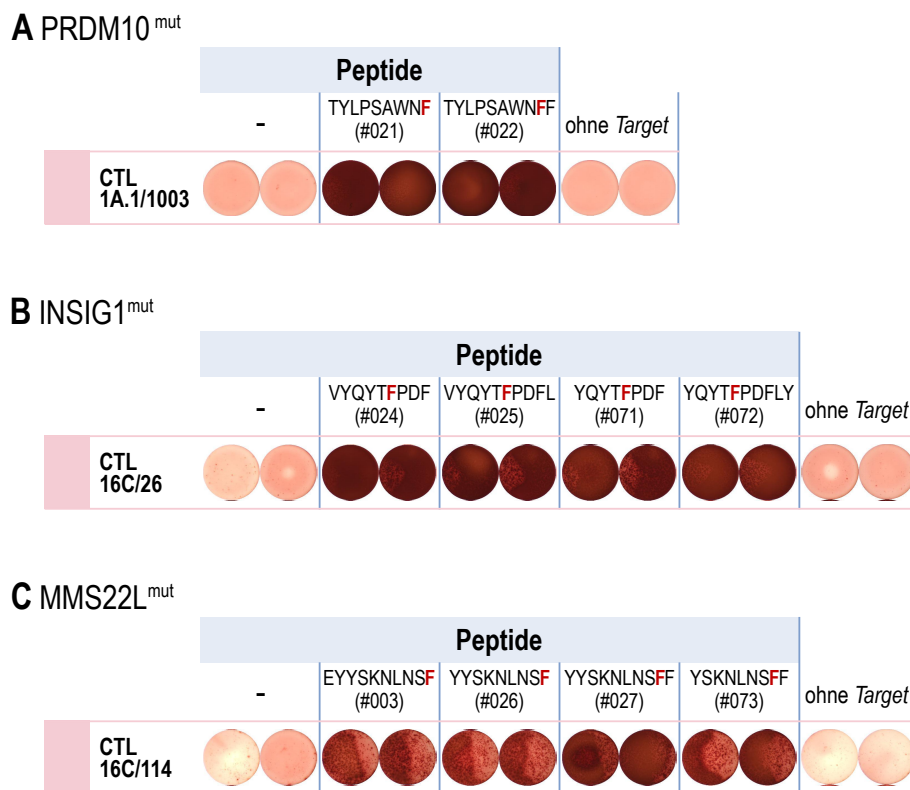
T-Zellklone expandiert werden. Die isolierten T-Zellklone ermöglichten im Folgenden die genauere Charakterisierung der Peptiderkennung.



**Abb. 3.28:** Generierung peptidreaktiver T-Zellklone aus den MLTCs 1A.1, 3A und 16C. Die drei MLTCs wurden kloniert, wie im Text beschrieben, um T-Zellklone gegen die neu identifizierten Neoantigene zu isolieren. Die Tortendiagramme zeigen, zu welchen Anteilen die peptidreaktiven T-Zellen in den MLTCs vorlagen. Für jede MLTC ist angegeben, wie viele T-Zellklone insgesamt auf Peptiderkennung getestet werden konnten. In den Legenden ist die absolute Anzahl der gegen das jeweilige mutierte Antigen gerichteten T-Zellklone wiedergegeben.

### 3.7 Charakterisierung der Peptiderkennung

Da die Erkennung der immunogenen Peptide durch die generierten T-Zellklone nur in Pools getestet worden war, musste noch geklärt werden, ob alle T-Zellklone, die gegen einen Pool reagierten, das gleiche Peptid erkannten bzw. welches der Peptide unterschiedlicher Länge am besten erkannt wurde. Aufgrund dessen wurden nach der OKT-3-Restimulation jeweils zwei T-Zellklone für jeden Pool ausgewählt und auf die Erkennung der einzelnen Peptide des jeweiligen Pools überprüft. Hierbei wurden die autologen INTH-EBV-B-Zellen als antigenpräsentierende Zellen verwendet. Bei der Testung der Einzelpeptide war jedoch kein signifikanter Unterschied in der IFN $\gamma$ -Freisetzung durch die peptidreaktiven T-Zellklone zu beobachten (**Abb. 3.29**). Die Sequenzen der Peptide eines Pools unterschieden sich jeweils nur um einzelne Aminosäuren am N- oder C-Terminus. Außerdem wiesen die Peptide nur eine Reinheit von 20-80% auf. Daher war anzunehmen, dass auch kürzere Peptide in den Peptidpräparationen vorhanden waren.



**Abb. 3.29: Erkennung der einzelnen Peptide aus den Peptidpools.** Jeweils zwei der gegen PRDM10<sup>mut</sup> (A), INSIG1<sup>mut</sup> (B) bzw. MMS22L<sup>mut</sup> (C) gerichteten T-Zellklone wurden gegen die einzelnen Peptide (Peptidnummern in Klammern) mittels des IFN $\gamma$ -ELISpot-Assays getestet (exemplarisch dargestellt für jeweils einen T-Zellklon; T-Zellklon 1A.1/1003: 20.000 Zellen/Well, T-Zellklon 16C/26: 28.000 Zellen/Well, T-Zellklon 16C/114: 44.000 Zellen/Well). Als peptidpräsentierende Zellen wurden 50.000 INTH-EBV-B-Zellen pro Testeinheit eingesetzt. Kontrollen waren die Erkennung der INTH-EBV-B-Zellen ohne Peptidzusatz (-) und die spontane IFN $\gamma$ -Freisetzung der T-Zellklone (ohne Target).

Die abschließende Charakterisierung der drei Neoantigene INSIG1<sup>mut</sup>, MMS22L<sup>mut</sup> und PRDM10<sup>mut</sup> sollte die Bestimmung ihrer HLA-Restriktion, die Validierung der intrazellulären Prozessierung und Präsentation, sowie die Bestimmung des C-terminalen Peptidendes umfassen. Letzteres diente der Identifizierung des tatsächlich erkannten Peptids (bei



PRDM10<sup>mut</sup>) oder zumindest der weiteren Eingrenzung der tatsächlich erkannten Peptide von INSIG1<sup>mut</sup> und MMS22L<sup>mut</sup>. Hierfür wurden zunächst die cDNAs der mutierten Antigene und die entsprechenden Wildtyp-cDNAs in Expressionsvektoren kloniert. Dazu wurden die aus Ma-Mel-86A (Replikat 3) bzw. aus INTH-EBV-B (Replikat 1) isolierte RNA in cDNA revers transkribiert und die antigenkodierenden cDNAs spezifisch amplifiziert (Primer siehe Abschnitt 2.1.12). Außerdem wurden 3'-cDNA-Fragmente generiert, um den für die Erkennung benötigten C-Terminus der Peptide zu ermitteln. Dieser wurde durch die Erkennungstests mit den synthetischen Peptiden bereits eingeschränkt. Als Startpunkt für die Fragmentierung der drei Neoantigene wurde die C-terminale Aminosäure des Peptids gewählt, das von den erkannten Peptiden am weitesten entfernt vom C-Terminus des mutierten Proteins endete. Die antigenkodierende cDNA wurde entsprechend nach diesem Codon trunziert und schrittweise um drei Nukleotide verkürzt bzw. verlängert. Dies ist beispielhaft für INSIG1<sup>mut</sup> in **Abb. 3.30** gezeigt. Am 3'-Ende der cDNA-Fragmente wurde bei der Amplifikation außerdem ein Stopcodon angehängt (Primer siehe Abschnitt 2.1.12), um das Ende der Translation sicherzustellen.

### INSIG1<sup>mut</sup>

Codon	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245		
cDNA	...	AATGGTGTCTATCAGTATA	CA	T	AT	CC	AG	AT	TT	CT	CT	TAT	AT	TC	GT	...	
Protein	...	N	G	V	Y	Q	Y	T	<b>F</b>	P	D	F	L	Y	I	R	...
Peptide																	

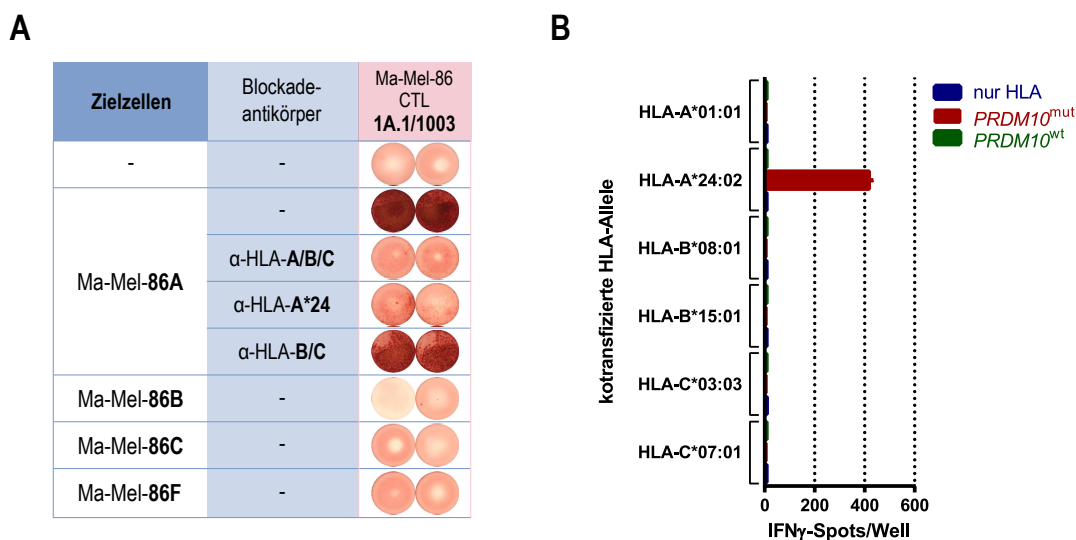
**Abb. 3.30: Fragmentierung der antigenkodierenden cDNA.** Als Ausgangspunkt für die Herstellung von 3'-Fragmenten der antigenkodierenden cDNA wurden die C-Termini der erkannten Peptide verwendet. Die Abbildung zeigt dies am Beispiel von INSIG1<sup>mut</sup>. Die mutierte Base bzw. Aminosäure ist in Rot dargestellt. Die Peptide #024 und #071 (dunkelblaue Balken) endeten an Codon 241 der INSIG1-Isoform 1, wohingegen die Peptide #025 und #072 (hellblaue Balken) weiter C-terminal endeten. Die cDNA-Sequenz wurde daher zunächst hinter dem Codon 241 trunziert (gestrichelte Linie) und dann schrittweise um drei Nukleotide verkürzt bzw. verlängert. Analog wurde bei der Fragmentierung der MMS22L<sup>mut</sup>- bzw. PRDM10<sup>mut</sup>-cDNA verfahren.

### 3.7.1 PRDM10<sup>mut</sup>

Da der offene Leserahmen (ORF, engl. *open reading frame*) von PRDM10 (3.471 bp) für eine TOPO-TA-Klonierung zu lang war, wurde die cDNA mit dem Gateway-Verfahren kloniert. Die Wildtyp- und die mutierte cDNA wurden mit PRDM10-spezifischen Primern aus Ma-Mel-86A-cDNA amplifiziert. Das Amplifikat wurde mit Gateway-kompatiblen attB-Primern und der Q5-Polymerase (NEB) reamplifiziert, und das attB-PCR-Produkt wurde nach einer Gelextraktion in den Gateway-Entry-Vektor pDONR221<sup>TM</sup> kloniert. Nach einem zweiten Rekombinationsschritt lagen Wildtyp- bzw. mutiertes PRDM10 im Destination-Vektor pcDNA3.1 DEST #6 vor. Die Sequenzierung der Plasmid-DNA zeigte, dass es sich jeweils um die PRDM10-Isoform 2 (RefSeq ID: NM\_199437.1) handelte.

## 3.7.1.1 Bestimmung der HLA-Restriktion und Validierung der Peptidprozessierung

Um das restringierende HLA-Klasse I-Molekül der PRDM10<sup>mut</sup>-reaktiven T-Zellklone zu definieren, wurde zunächst ein Antikörperblockadetest durchgeführt (**Abb. 3.31A**). Dabei wurde gezeigt, dass die IFN $\gamma$ -Sekretion durch die Zugabe von einem Gesamt-HLA-Klasse I-spezifischen Antikörper (W6/32) und durch einen HLA-A\*24-spezifischen Antikörper inhibiert werden konnte. Ein Antikörper gegen eine gemeinsame Determinante von HLA-B- und HLA-C-Molekülen (B1.23.2) hatten keinen sicher blockierenden Effekt. Aus den Ergebnissen in **Abb. 3.31A** ist auch zu entnehmen, dass der T-Zellklon 1A.1/1003 nur Ma-Mel-86A erkannte. Ma-Mel-86B, -86C und -86F trugen auf ihrer Oberfläche kein HLA-A\*24:02, exprimierten darüber hinaus kein PRDM10<sup>mut</sup> (siehe Abb. 3.27 in Abschnitt 3.5.6) und wurden deswegen nicht erkannt.

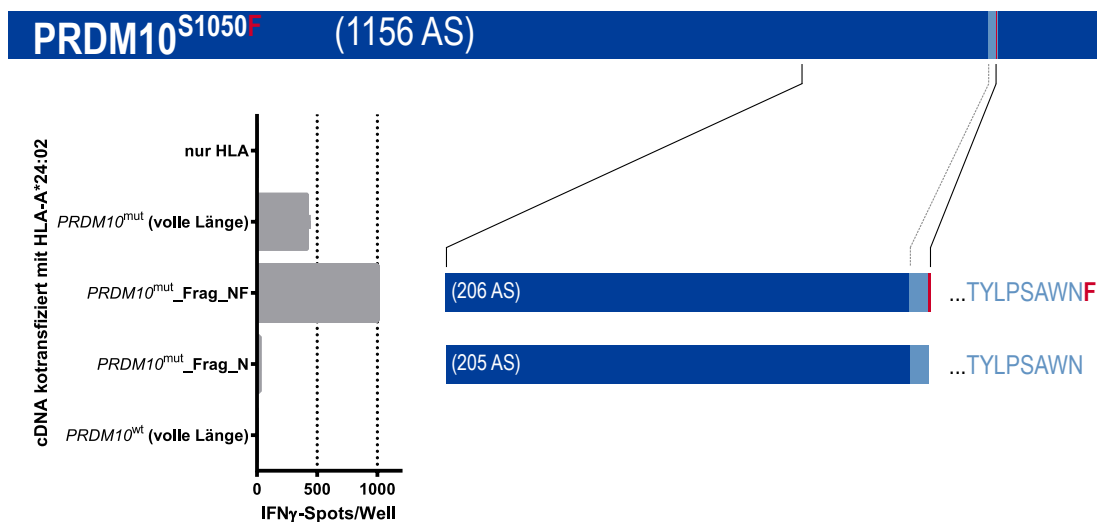


**Abb. 3.31: Bestimmung des präsentierenden HLA-Klasse I-Moleküls und Validierung der intrazellulären Prozessierung des Antigens PRDM10<sup>mut</sup>.** (A) Die Erkennung der Ma-Mel-86A (50.000 Zellen/Well) durch den PRDM10<sup>mut</sup>-spezifischen T-Zellklon 1A.1/1003 (20.000 Zellen/Well) wurde in einem IFN $\gamma$ -ELISpot-Assay in der Anwesenheit HLA-spezifischer Antikörper ( $\alpha$ -HLA-A/B/C: 200  $\mu$ g/ml W6/32,  $\alpha$ -HLA-A\*24: 100  $\mu$ g/ml C7709.A2.6 und  $\alpha$ -HLA-B/C: 85  $\mu$ g/ml B1.23.2) getestet. Der OKT-3-restimulierte T-Zellklon 1A.1/1003 (d15) wurde einen Tag vor der Testbelegung aufgetaut und in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Als Kontrollen wurden die spontane IFN $\gamma$ -Freisetzung ohne Zielzellen und die Erkennung der HLA-Verlustvarianten Ma-Mel-86B, -86C und -86F (jeweils 50.000 Zellen/Well) überprüft. (B) Anhand der Erkennung von COS-7-Zellen (20.000 Zellen/Well), die am Tag vor der Testbelegung mit mutierter bzw. nicht mutierter PRDM10-cDNA (PRDM10<sup>mut</sup> bzw. PRDM10<sup>wt</sup>, 300 ng Plasmid-DNA je Ansatz) und mit cDNAs, die für die HLA-Klasse I-Allele der Patientin kodierten (jeweils 100 ng Plasmid-DNA je Ansatz), kotransfiziert wurden, sollten das restringierende HLA-Klasse I-Molekül bestätigt sowie die intrazelluläre Prozessierung und die Präsentation des mutierten Peptids validiert werden. Der T-Zellklon 1A.1/1003 wurde eingesetzt wie in (A) beschrieben. Nur mit HLA transfizierte COS-7-Zellen dienten als Kontroll-Targetzellen. Das Balkendiagramm gibt die Mittelwerte der gemessenen Spotzahlen wieder (n=2, Fehlerbalken: Standardfehler des Mittelwerts).

In einem zweiten IFN $\gamma$ -ELISpot-Assay (**Abb. 3.31B**) wurden PRDM10<sup>mut</sup>- und PRDM10<sup>wt</sup>-kodierende cDNAs in COS-7-Zellen kotransfiziert mit cDNAs, die für die HLA-Klasse I-Allele der Patientin kodierten. Nur nach Kotransfektion von mutierter cDNA mit HLA-A\*24:02-cDNA wurden die COS-7-Zellen durch den T-Zellklon 1A.1/1003 erkannt. Dieser Test validierte die HLA-A\*24:02-Restriktion des PRDM10<sup>mut</sup>-spezifischen T-Zellklons und bestätigte die intrazelluläre Prozessierung und die Präsentation des mutierten Peptids.

### 3.7.1.2 Lokalisation der peptidkodierenden Region

Zur Bestimmung des C-terminalen Peptidendes wurden Fragmente der mutierten *PRDM10*-cDNA hergestellt. Diese wurden ebenfalls mittels des *Gateway*-Systems in *pcDNA3.1 DEST #6* kloniert. Als PCR-*Template* wurde das *attB*-PCR-Produkt der mutierten *PRDM10*-cDNA voller Länge verwendet (siehe Abschnitt 3.7.1). Da sich die immunogene Mutation am Ende des sehr großen ORF befindet, wurde ein interner ORF, beginnend mit dem für Methionin kodierenden Codon 845 der *PRDM10*-Isoform 2, zur Klonierung der Fragmente genutzt. Als Translationsinitiationssequenz wurde das Startcodon in eine Kozak-Sequenz (ACCAUGG; Startcodon unterstrichen, wobei das folgende Guanin bereits zum nächsten Codon des internen ORF gehörte) integriert (Kozak, 1986), um die Effizienz der Translation in COS-7 zu optimieren. Das 3'-Ende des ersten Fragments wurde durch den C-Terminus des Peptids #021 festgelegt, welches an Codon 1050 von *PRDM10*<sup>mut</sup> endete. Die trunkierten cDNAs sowie die cDNAs voller Länge wurden zusammen mit HLA-A\*24:02-kodierender cDNA in COS-7-Zellen transfiziert und deren Erkennung durch *PRDM10*<sup>mut</sup>-spezifische T-Zellen getestet (**Abb. 3.32**).



**Abb. 3.32: Bestimmung des C-terminalen Peptidendes des Neoantigens *PRDM10*<sup>mut</sup>.** Die angegebenen cDNAs kompletter Länge (*PRDM10*<sup>mut</sup> und *PRDM10*<sup>wt</sup>) sowie die cDNA-Fragmente von *PRDM10*<sup>mut</sup> (Frag\_NF und Frag\_N) wurden zusammen mit HLA-A\*24:02-kodierender cDNA in COS-7-Zellen (20.000 Zellen/Well) transfiziert. Die Erkennung der Transfektanten durch den T-Zellklon 1A.1/1003 (20.000 Zellen/Well) wurde in einem IFN $\gamma$ -ELISpot-Assay getestet. Aufgrund der Länge des ORF und der Position der Mutation (rot markiert) wurde ein interner ORF (Start: Codon 845) für die Generierung der cDNA-Fragmente genutzt. Für den Test wurde der nach OKT-3-Stimulation kryokonservierte T-Zellklon am Tag vor der Belegung aufgetaut und in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Der durch die synthetischen Peptide abgedeckte Proteinabschnitt ist hellblau gefärbt. Rechts sind jeweils die Aminosäuresequenzen an den C-terminalen Fragmentenden angegeben. In Klammern stehen die Längen der erwarteten Translationsprodukte [Anzahl Aminosäuren (AS)]. Zur Kontrolle wurde die Erkennung HLA-A\*24:02-transfizierter COS-7-Zellen (nur HLA) überprüft. Im Balkendiagramm sind die Mittelwerte der gemessenen Spotzahlen dargestellt (n=2, Fehlerbalken: Standardfehler des Mittelwerts).

Das Fragment *PRDM10*<sup>mut</sup>\_Frag\_NF, das mit der mutierten Codonposition 1050 endete, wurde nach Transfektion stärker erkannt als die cDNA voller Länge. Dies bestätigte auch die Vorhersage potentieller Proteasomschnittstellen mit *NetChop 3.1 Server* (kurz: *NetChop 3.1*), wonach die wahrscheinlichste proteasomale Spaltung in diesem Proteinabschnitt nach Aminosäure 1050 erfolgen sollte. Ein mit der Codonposition 1049 endendes Fragment, bei

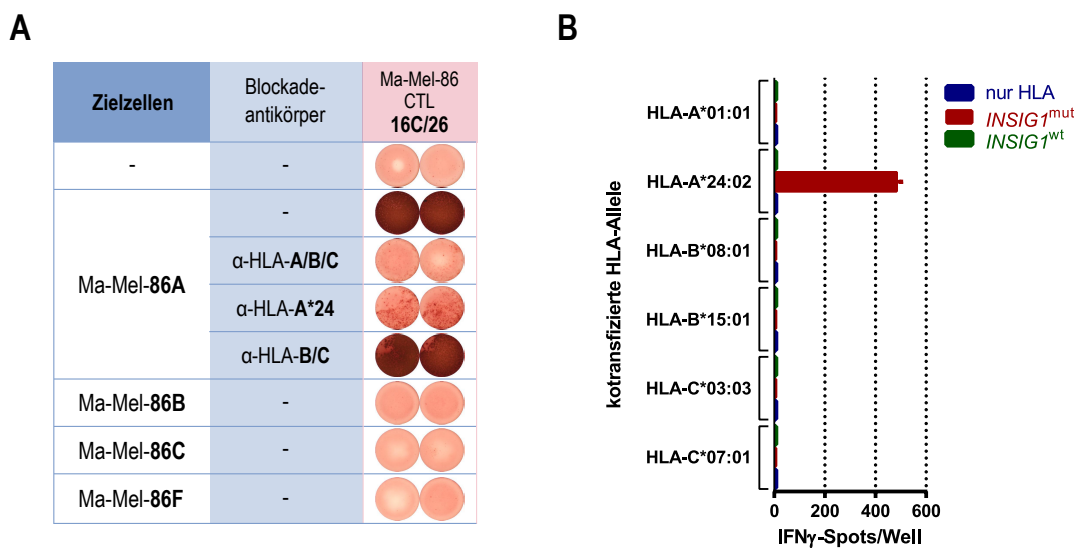
dem das Codon 1050 durch ein Stopcodon ersetzt worden war, führte nach Transfektion hingegen zu keiner Erkennung durch den T-Zellklon 1A.1/1003. Die mutierte Aminosäure bildete somit mit hoher Wahrscheinlichkeit den C-Terminus des immunogenen Peptids.

### 3.7.2 INSIG1<sup>mut</sup>

Wildtypische *INSIG1*-cDNA wurde aus INTH-EBV-B-cDNA und mutierte *INSIG1*-cDNA wurde aus Ma-Mel-86A-cDNA mit der Polymerase *EasyA* (Agilent Technologies) amplifiziert. Beide cDNAs wurden mithilfe des *TOPO TA Expression Kit*<sup>®</sup> in den Vektor *pcDNA<sup>TM</sup>3.1/V5-His TOPO*<sup>®</sup> (Invitrogen/Life Technologies) kloniert.

#### 3.7.2.1 Bestimmung der HLA-Restriktion und Validierung der Peptidprozessierung

Mittels eines Antikörperblockadetests sollte auch das Restriktionsmolekül der INSIG1<sup>mut</sup>-reaktiven T-Zellen bestimmt werden (**Abb. 3.32A**). Das Blockademuster war identisch zu dem des PRDM10<sup>mut</sup>-spezifischen T-Zellklons (vgl. Abb. 3.31). Ebenso war keine Erkennung der HLA-A\*24:02-negativen Tumorzelllinien Ma-Mel-86B, -86C und -86F zu beobachten.

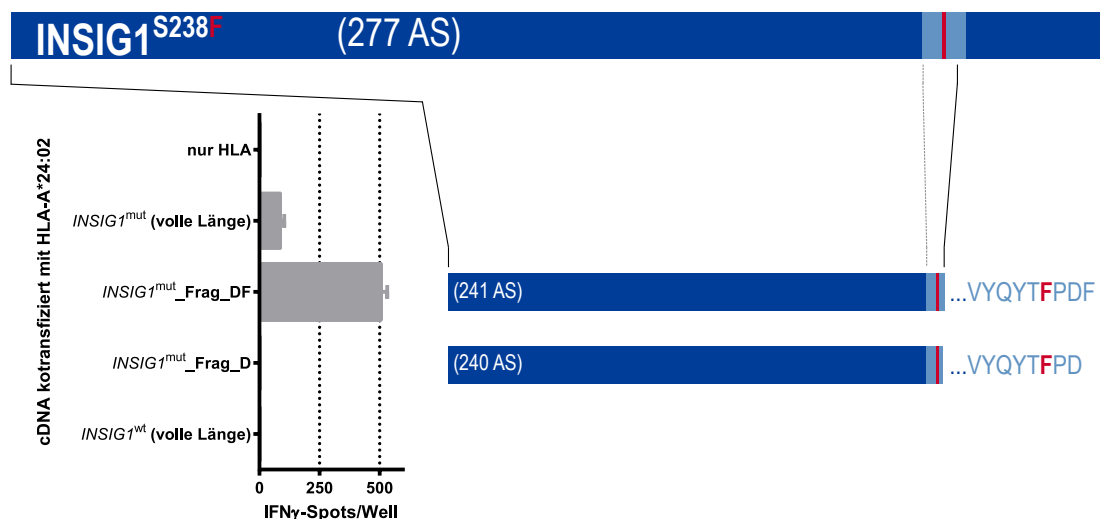


**Abb. 3.33: Bestimmung des präsentierenden HLA-Klasse I-Moleküls und Validierung der intrazellulären Prozessierung des Antigenes INSIG1<sup>mut</sup>.** (A) In einem Antikörperblockadetest ( $\alpha$ -HLA-A/B/C: 200  $\mu$ g/ml W6/32,  $\alpha$ -HLA-A\*24: 100  $\mu$ g/ml C7709.A2.6 und  $\alpha$ -HLA-B/C: 85  $\mu$ g/ml B1.23.2) wurde die Erkennung der autologen Ma-Mel-86A (50.000 Zellen/Well) durch den INSIG1<sup>mut</sup>-spezifischen T-Zellklon 16C/26 (28.000 Zellen/Well) getestet. Der OKT-3-restimulierte T-Zellklon 16C/26 (d15) wurde einen Tag vor der Testbelegung aufgetaut und in AIM-V-Medium mit 10% HS und 250 U/ml rIL-2 in Kultur genommen. Als Kontrollen wurden die spontane IFN $\gamma$ -Freisetzung ohne Zielzellen und die Erkennung der HLA-Verlustvarianten Ma-Mel-86B, -86C und -86F (jeweils 50.000 Zellen/Well) überprüft. (B) Der Test der Erkennung von COS-7-Zellen (20.000 Zellen/Well), die am Tag vor der Testbelegung mit mutierter bzw. nicht mutierter *INSIG1*-cDNA (*INSIG1*<sup>mut</sup> bzw. *INSIG1*<sup>wt</sup>, 300 ng Plasmid-DNA je Ansatz) und mit cDNAs, die für die HLA-Klasse I-Allele der Patientin kodierten (jeweils 100 ng Plasmid-DNA je Ansatz), kotransfiziert wurden, diente der Bestätigung des restringierenden HLA-Klasse I-Moleküls sowie der Validierung der intrazellulären Prozessierung und der Präsentation des mutierten Peptids. Der nach der vierzehntägigen OKT-3-Restimulation kryokonservierte T-Zellklon 3A/115 wurde zwei Tage vor der Testbelegung aufgetaut, in AIM-V-Medium mit 10% HS und 250 U/ml rIL-2 in Kultur genommen und mit 20.000 Zellen/Well im Test eingesetzt. Nur mit HLA transfizierte COS-7-Zellen dienen als Kontroll-Targetzellen. Im Balkendiagramm dargestellt sind die Mittelwerte der gemessenen Spotzahlen (n=2, Fehlerbalken: Standardfehler des Mittelwerts).

**Abb. 3.32B** zeigt die Reaktivität des ebenso für  $INSIG1^{mut}$ -spezifischen T-Zellklons 3A/115 gegen COS-7-Zellen, die mit  $INSIG1^{mut}$ - bzw.  $INSIG1^{wt}$ -cDNA und HLA-kodierender cDNA kotransfiziert waren. Für die Erkennung der Transfektanten war eine Kotransfektion der mutierten cDNA mit HLA-A\*24:02-kodierender cDNA notwendig. Somit wurde das Ergebnis des Antikörperblockadetests bestätigt und die interne Prozessierung und Präsentation des  $INSIG1^{mut}$ -Peptids demonstriert.

### 3.7.2.2 Lokalisation der peptidkodierenden Region

Durch 3'-Fragmentierung der  $INSIG1^{mut}$ -cDNA voller Länge sollte abschließend der C-Terminus des immunogenen  $INSIG1^{mut}$ -Peptids definiert werden. Zur Herstellung der *Template* für die Amplifizierung der cDNA-Fragmente wurde die  $INSIG1^{mut}$ -kodierende Plasmid-DNA mit EcoRV und BstXI verdaut und das  $INSIG1^{mut}$ -Insert (866 bp) durch Gelextraktion vom Vektor-*Backbone* getrennt. Durch entsprechende 3'-Primer wurden die trunkierten cDNAs amplifiziert und wie die cDNAs voller Länge mittels TOPO-TA-Klonierung in den Vektor *pcDNA<sup>TM</sup>3.1/V5-His TOPO<sup>®</sup>* eingebracht. Wie zuvor beschrieben (vgl. Abb. 3.30), wurde die Fragmentierung hinter dem Codon 241 begonnen.



**Abb. 3.34: Bestimmung des C-terminalen Peptids des Neoantigens  $INSIG1^{mut}$ .** Die gezeigten cDNAs voller Länge ( $INSIG1^{mut}$  bzw.  $INSIG1^{wt}$ ) sowie die cDNA-Fragmente von  $INSIG1^{mut}$  (Frag\_DF und Frag\_D) wurden zusammen mit HLA-A\*24:02-kodierender cDNA in COS-7-Zellen (20.000 Zellen/Well) transfiziert. In einem IFN $\gamma$ -ELISpot-Assay wurde die Reaktivität des T-Zellklons 3A/115 (20.000 Zellen/Well) gegen die Transfektanten bestimmt. Die OKT-3-restimulierten und kryokonservierten T-Zellen (d14) wurden am Tag vor der Testbelegung aufgetaut und in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Der durch die synthetischen Peptide abgedeckte Proteinabschnitt ist hellblau gefärbt. Die Position der Mutation ist rot markiert. Rechts sind jeweils die C-terminalen Aminosäuresequenzen angegeben. In Klammern stehen die Längen der erwarteten Translationsprodukte [Aminosäuren (AS)]. Zur Kontrolle wurde die Erkennung HLA-A\*24:02-transfizierter COS-7-Zellen (nur HLA) überprüft. Im Balkendiagramm sind die Mittelwerte der gemessenen Spotzahlen gezeigt (n=2, Fehlerbalken: Standardfehler des Mittelwerts).

In **Abb. 3.34** ist die Erkennung von COS-7-Zellen dargestellt, die mit den generierten Fragmenten bzw. mit cDNA voller Länge und mit HLA-A\*24:02-kodierender cDNA kotransfiziert waren. Während das mit Codon 241 endende 3'-Fragment nach Transfektion erkannt wurde, induzierte das mit Codon 240 endende 3'-Fragment keine Erkennung durch den  $INSIG1^{mut}$ -spezifischen T-Zellklon 3A/115. Somit war davon auszugehen, dass

Phenylalanin (F) an Position 241 den C-Terminus des immunogenen Peptides bildete, und dass die mutierte Aminosäureposition 238 im immunogenen Peptid an vierter Stelle vor dem C-Terminus lag.

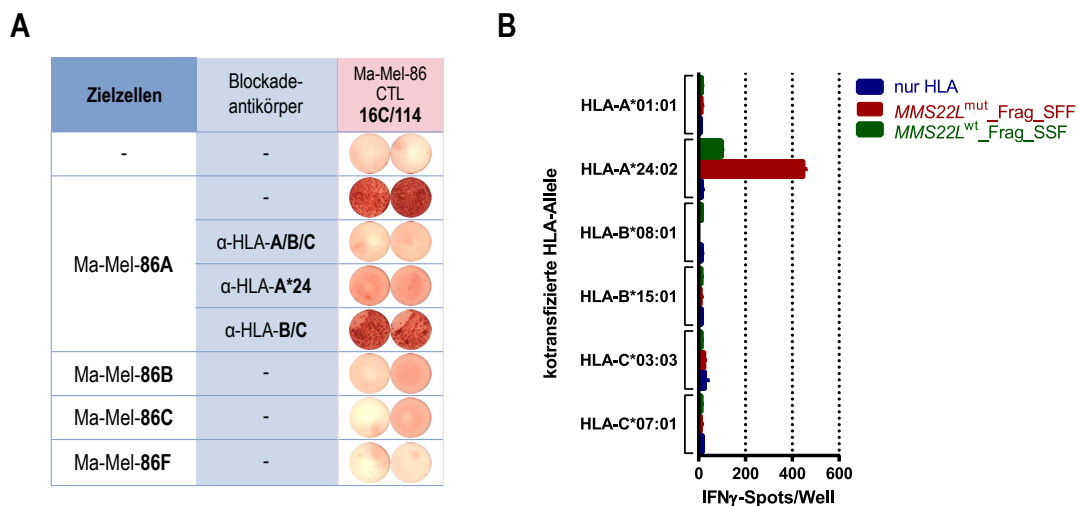
Die wahrscheinlichste Proteasomschnittstelle laut *NetChop 3.1* (Methode: *C-term 3.0*) läge zwischen Codon 243 und 244. Eine Spaltung nach dem durch Codon 241 kodierten Phenylalanin wurde mit dieser Methode nicht vorhergesagt. Basierend auf den Trainingsdaten des konstitutiven 20S-Proteasoms (*NetChop 3.1*, Methode: *20S 3.0*) wurde hier zumindest eine suboptimale Schnittstelle gefunden.

#### 3.7.3 MMS22L<sup>mut</sup>

Die cDNA von *MMS22L* sollte, ebenso wie für *PRDM10*, aufgrund der Länge des ORF (3.732 bp) mithilfe des *Gateway*-Systems in einen Expressionsvektor kloniert werden. Obwohl das *attB*-PCR-Produkt laut Sequenzierung in voller Länge vorlag, wurden bereits bei der Rekombination in den *Entry*-Vektor *pDONR221*<sup>TM</sup> nur Klone isoliert, denen verschiedene Genabschnitte fehlten. Dabei war entweder ein Teil der peptidkodierenden Sequenz verloren gegangen oder es waren Sequenzabschnitte deletiert, wodurch sich ein Rasterschub ergab. Als Ursache hierfür wurden fehlerhafte Rekombinationen angenommen. Aufgrund dessen wurden von *MMS22L* nur cDNA-Fragmente mit einer Länge von bis zu 1.320 bp hergestellt, die mittels TOPO-TA-Ligation in *pcDNA*<sup>TM</sup>3.1/*V5-His TOPO*<sup>®</sup> inseriert werden konnten. Hierbei wurde das 3'-Ende (Codonposition 437) zunächst durch den C-Terminus der Peptide #003 und #026 vorgegeben (vgl. Beschreibung für INSIG1<sup>mut</sup> in Abb. 3.30). Zudem wurden Fragmente generiert, die die *MMS22L*-Sequenz bis Codon 436, 438 bzw. 439 kodierten.

##### 3.7.3.1 Bestimmung der HLA-Restriktion und Validierung der Peptidprozessierung

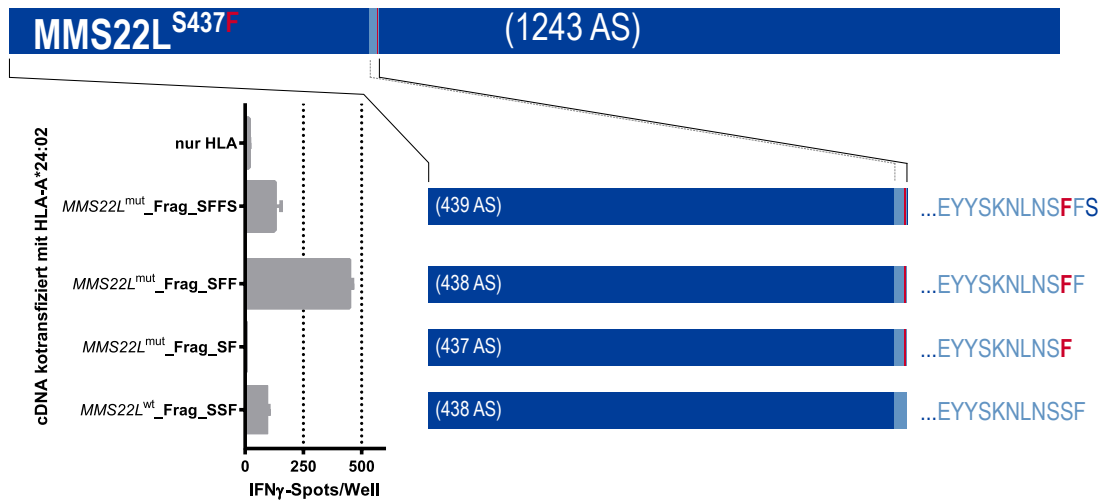
Wie in **Abb. 3.35A** gezeigt, war das Blockierungsmuster des MMS22L<sup>mut</sup>-spezifischen T-Zellklons 16C/114 identisch zu den Blockierungsmustern der T-Zellen gegen PRDM10<sup>mut</sup> und INSIG1<sup>mut</sup> (vgl. Abb. 3.31A und Abb. 3.33A). Die Transfektion der einzelnen HLA-cDNAs der Patientin in COS-7-Zellen zusammen mit den generierten mutierten und wildtypischen cDNA-Fragmenten (*MMS22L*<sup>mut</sup>\_Frag\_SFF bzw. *MMS22L*<sup>wt</sup>\_Frag\_SSF), die jeweils mit dem Codon 438 endeten, zeigte, dass die Erkennung von der Koexpression des Allels HLA-A\*24:02 abhängig war (**Abb. 3.35B**). Somit war auch dieses Antigen durch HLA-A\*24:02 restringiert. Gleichzeitig wurde die Prozessierung und Präsentation des MMS22L<sup>mut</sup>-Peptids bestätigt. Im Gegensatz zu den Antigenen PRDM10<sup>mut</sup> und INSIG1<sup>mut</sup> lag jedoch auch eine schwache Reaktivität des T-Zellklons 16C/114 gegen das Wildtyp-Fragment vor.



**Abb. 3.35: Bestimmung des präsentierenden HLA-Klasse I-Moleküls und Validierung der intrazellulären Prozessierung des Antigen MMS22L<sup>mut</sup>.** (A) Nach Zugabe von HLA-Klasse I-spezifischen Antikörpern ( $\alpha$ -HLA-A/B/C: 200  $\mu$ g/ml W6/32,  $\alpha$ -HLA-A\*24: 100  $\mu$ g/ml C7709.A2.6 und  $\alpha$ -HLA-B/C: 85  $\mu$ g/ml B1.23.2) wurde die Erkennung der autologen Ma-Mel-86A (50.000 Zellen/Well) durch den MMS22L<sup>mut</sup>-spezifischen T-Zellklon 16C/114 (44.000 Zellen/Well) getestet. Der T-Zellklon 16C/114 (d14) wurde nach der OKT-3-Restimulation kryokonserviert, für den IFN $\gamma$ -ELISpot-Assay am Tag vor der Belegung aufgetaut und in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Als Kontrollen wurden die spontane IFN $\gamma$ -Freisetzung ohne Zielzellen und die Erkennung der HLA-Verlustvarianten Ma-Mel-86B, -86C und 86F (jeweils 50.000 Zellen/Well) überprüft. (B) Die Erkennung von COS-7-Zellen (20.000 Zellen/Well), die am Tag vor der Testbelegung mit mutierter bzw. nicht mutierter MMS22L-cDNA (MMS22L<sup>mut</sup>\_Frag\_SFF bzw. MMS22L<sup>wt</sup>\_Frag\_SSF, 300 ng Plasmid-DNA je Ansatz) und mit cDNAs, die für die HLA-Klasse I-Allele der Patientin kodierten (jeweils 100 ng Plasmid-DNA je Ansatz), kotransfiziert wurden, wurde getestet, um das restringierende HLA-Klasse I-Molekül zu bestätigen und die intrazelluläre Prozessierung und die Präsentation des mutierten Peptids zu validieren. Nach OKT-3-Restimulation wurde der kryokonservierte T-Zellklon 16C/114 zwei Tage vor der Testbelegung aufgetaut und eingesetzt wie in (A) beschrieben. Nur mit HLA transfizierte COS-7-Zellen dienten als Kontroll-Targetzellen. Das Balkendiagramm zeigt die Mittelwerte der gemessenen Spotzahlen (n=2, Fehlerbalken: Standardfehler des Mittelwerts).

### 3.7.3.2 Lokalisation der peptidkodierenden Region

Im Gegensatz zu den mit den Codons 439 und 438 endenden, an Codon 437 mutierten cDNA-Fragmenten wurde das mit dem mutierten Codon 437 endende Fragment (MMS22L<sup>mut</sup>\_Frag\_SF) nicht mehr von dem T-Zellklon 16C/114 erkannt (**Abb. 3.36**). Daraus war zu folgern, dass Phenylalanin (F) an Position 438 den C-Terminus des immunogenen Peptides bildete, und dass die mutierte Aminosäureposition 437 im immunogenen Peptid an der vorletzten Position enthalten war. Dieses C-terminale Peptid wurde auch mit der Methode 20S 3.0 durch NetChop 3.1 vorhergesagt. Laut der Methode C-term 3.0, die mit C-Termini bekannter T-Zellantigene trainiert wurde, wäre jedoch eine Proteasomschnittstelle nach Codon 437 wahrscheinlicher als nach Codon 438. Aufgrund der Überlappung der immunogenen MMS22L<sup>mut</sup>-Peptide #003, #026, #027 und #073 wurden die elf C-terminalen Aminosäuren EYYSKNLNSFF des erkannten cDNA-Fragments MMS22L<sup>mut</sup>\_Frag\_SFF von den Peptiden abgedeckt. An Codon 438 endeten nur die Peptide #027 (YYSKNLNSFF) und #073 (YSKNLNSFF). Da für das Peptid #027 eine größere Übereinstimmung mit dem Bindemotiv von HLA-A\*24:02 vorlag, wurde angenommen, dass am ehesten das Dekamer YYSKNLNSFF (MMS22L<sup>S437F</sup><sub>429-438</sub>) auf HLA-A\*24:02 präsentiert und von MMS22L<sup>mut</sup>-reaktiven T-Zellen erkannt wurde.



**Abb. 3.36: Bestimmung des C-terminalen Peptidendes des Neoantigens MMS22L<sup>mut</sup>.** Die gezeigten cDNA-Fragmente von MMS22L<sup>mut</sup> (Frag\_SFFS, Frag\_SFF und Frag\_SF) bzw. MMS22L<sup>wt</sup> (Frag\_SSF) wurden zusammen mit HLA-A\*24:02-kodierender cDNA in COS-7-Zellen (20.000 Zellen/Well) transfiziert, um die Erkennung der Transfektanten durch den T-Zellklon 16C/114 (30.000 Zellen/Well) in einem IFN $\gamma$ -ELISpot-Assay zu bestimmen. Die nach der vierzehntägigen OKT-3-Stimulation kryokonservierten T-Zellen wurden zwei Tage vor der Testbelegung in AIM-V-Medium mit 10% HS und 250 U/ml rhIL-2 in Kultur genommen. Der durch die synthetischen Peptide #003, #026, #027 und #073 abgedeckte Proteinabschnitt ist hellblau gefärbt. Die Position der Mutation ist rot markiert. Rechts sind jeweils die Aminosäuresequenzen an den C-terminalen Fragmentenden angegeben. In Klammern stehen die Längen der erwarteten Translationsprodukte [Aminosäuren (AS)]. Zur Kontrolle wurde die Erkennung HLA-A\*24:02-transfizierter COS-7-Zellen (nur HLA) überprüft. Im Balkendiagramm sind die Mittelwerte der gemessenen Spotzahlen gezeigt (n=2, Fehlerbalken: Standardfehler des Mittelwerts).

### Zusammenfassung

Durch die abschließenden Tests zur Bestimmung der HLA-Klasse I-Restriktion wurde für die drei neu identifizierten Neoantigene übereinstimmend eine Präsentation durch HLA-A\*24:02 gezeigt. Durch Kotransfektion der HLA-cDNA und peptidkodierender cDNA in COS-7-Zellen wurde die intrazelluläre Prozessierung und die Präsentation der Peptide bestätigt. Mithilfe von 3'-Fragmenten der Antigen-cDNAs wurden die C-Termini der immunogenen Peptide bestimmt, so dass im Rahmen dieser Arbeit die drei in **Tab. 3.9** gelisteten, bisher unbekanntenen Neoantigene identifiziert wurden.

**Tab. 3.9: Übersicht der mittels NGS gefundenen Neoantigene im Melanommodell Ma-Mel-86.** Durch die Sequenzierung der Melanomzelllinien des Modells konnten drei bisher unbekanntene mutierte Antigene identifiziert werden. Anhand der abschließenden Tests mit peptidreaktiven T-Zellklonen wurden die peptidkodierenden Regionen definiert und die vorhergesagte Präsentation durch HLA-A\*24:02 für die drei Neoantigene bestätigt. Das Vorkommen der vier zugrundeliegenden Mutationen in den Transkriptomen der Tumorzelllinien des Modells ist gekennzeichnet (+: Mutation exprimiert, -: Mutation nicht exprimiert).

Name	Peptidsequenz*	Präsentation durch	Mutation exprimiert in Ma-Mel-86			
			A	B	C	F
PRDM10 <sup>S1050F</sup> <sub>1042-1050</sub>	TYLPSAWN <b>F</b>	HLA-A*24:02	+	-	-	-
INSIG1 <sup>S238F</sup> <sub>233-241</sub>	VYQY <b>T</b> FPDF	HLA-A*24:02	+	+	+	-
MMS22L <sup>S437F</sup> <sub>429-438</sub>	YYSKNLNS <b>FF</b>	HLA-A*24:02	+	+	+	+

\* vorhergesagt nach Testung synthetischer Peptide und von cDNA-Fragmenten



## 4 Diskussion

Das Ziel der vorgelegten Arbeit bestand in der systematischen Identifizierung von Neoantigenen im humanen Melanommodell Ma-Mel-86, die auf somatische, nicht-synonyme Punktmutationen zurückzuführen waren. Die Grundlage für die Detektion immunogener Mutationen bildeten die Exom- und Transkriptomsequenzierungen der vier Tumorzelllinien des Modells und der autologen EBV-immortalisierten B-Zelllinie. Der bioinformatische Teil des Projekts konzentrierte sich auf die Detektion und das Filtern der Mutationen in den Exomen und Transkriptomen und die anschließende Vorhersage von potentiell HLA-Klasse I-bindenden Peptiden, die die durch die jeweilige Mutation ausgetauschte Aminosäure enthielten. Die tatsächliche Immunogenität der aufgrund der Vorhersagen ausgewählten Peptide wurde anschließend mit autologen MLTCs *in vitro* überprüft. Dabei wurden zum einen T-Zellantworten gegen das bereits durch cDNA-Bank-Screening identifizierte Neoantigen HERPUD1<sup>G161S</sup> und zum anderen gegen die drei bisher unbekannt Antigene INSIG1<sup>S238F</sup>, MMS22L<sup>S437F</sup> und PRDM10<sup>S1050F</sup> gefunden. Abschließend wurden T-Zellklone gegen die neu identifizierten Neoantigene generiert, um die vorhergesagte HLA-Restriktion zu bestätigen. Zudem wurden die T-Zellklone auf die Erkennung von Antigen-cDNA voller Länge bzw. von Antigen-cDNA-Fragmenten getestet, um die intrazelluläre Prozessierung der Peptide nachzuweisen und deren C-Termini zu bestimmen.

### 4.1 Bewertung der generierten Sequenzierdaten

Für die Lokalisierung von exprimierten, somatischen, nicht-synonymen Punktmutationen wurden die Exome und Transkriptome der fünf Zelllinien des Melanommodells Ma-Mel-86 sequenziert. Somit wurden gezielt proteinkodierende Sequenzabschnitte, die Quelle potentiell immunogener Peptide, sequenziert und in dem für die geplante Fragestellung relevanten Anteil des Humangenoms (ca. 1%) eine möglichst hohe Abdeckung erzielt. Die Sequenzierung der Transkriptome (*RNA-Seq*) lieferte zusätzlich korrespondierende Expressionsdaten und deckte außerdem exprimierte Genomabschnitte ab, die in der Exombibliothek nicht ausreichend angereichert wurden (vgl. auch Abschnitt 4.3.1).

Mit einer durchschnittlich 49,3-fachen Abdeckung in den Exomen waren die Sequenzierungen der Ma-Mel-86-Zelllinien hinsichtlich der Sequenziertiefe vergleichbar mit anderen Studien, die mittels Exomsequenzierung somatische Mutationen analysierten (Robbins *et al.*, 2013). In anderen Arbeiten wurden jedoch auch deutlich niedrigere (Castle *et al.*, 2012, ca.

20-fach) bzw. höhere Abdeckungen (Rajasagi *et al.*, 2014; Wang *et al.*, 2011, >130-fach) genutzt. Die Abdeckungen der Transkriptomsequenzierungen wurden in Anlehnung an die Empfehlungen des ENCODE-Konsortiums geschätzt (siehe Abschnitt 3.3.3.2). Da in der Literatur noch keine einheitliche Angabe der Abdeckung von *RNA-Seq*-Daten anhand dieser Leitlinien existierte, dienten die geschätzten Werte nur dem Vergleich der Datensätze untereinander. Durch die Berechnung auf der Grundlage der Anzahl der alignierten anstelle der sequenzierten Nukleotide konnte berücksichtigt werden, dass nicht alle sequenzierten *Reads* zugeordnet, und dass manche *Reads* bei der Aufarbeitung (siehe Abschnitt 3.3.2) gekürzt wurden. Die so ermittelten Abdeckungen (5,7- bis 22,5-fach, Durchschnitt: 14,1-fach) spiegelten jeweils die zur Verfügung stehende *Read*-Anzahl je Rohdatensatz wider. Die Qualität der einzelnen Datensätze bzw. *Mappings* waren somit in Anbetracht der Anzahl der *Reads*, die für jedes *Mapping* genutzt werden konnten, sehr ähnlich.

Ein Vergleich der Transkriptomsequenzierungen mit denen anderer Publikationen war vorwiegend anhand der Anzahl der sequenzierten *Reads* möglich. Auch hier fanden sich sehr unterschiedlich groß angelegte Projekte mit beispielsweise durchschnittlich 8,3 Millionen (Xu *et al.*, 2013), 30 Millionen (Castle *et al.*, 2012) oder 110 Millionen *Reads* (Atak *et al.*, 2013). Diese großen Unterschiede ergaben sich unter anderem dadurch, dass mit einer Sequenzierung mehrere Fragestellungen adressiert wurden, die unterschiedliche Sequenziertiefen voraussetzten. Somit war die Größe der im Modell Ma-Mel-86 generierten *RNA-Seq*-Daten mit ca. 28 bis 111 Millionen *Reads* je Datensatz vergleichbar mit anderen Studien, die SNVs in Transkriptomen detektierten.

In einem kürzlich erschienenen Übersichtsartikel von Robasky *et al.* (2014) wurde der Einsatz von Replikaten zur Minimierung von Sequenzierfehlern in NGS-Daten thematisiert. Die Autoren bevorzugten die Verwendung von Replikaten gegenüber einer tieferen Sequenzierung, da Fehler nicht nur während der Sequenzierung, sondern auch während der Probengewinnung bzw. der Herstellung der Bibliotheken entstehen können. Die Fehlerquote wird mithilfe von Replikaten verringert, da in Replikaten identische Fehler nicht auftreten sollten. In der hier durchgeführten Untersuchung wurden von jedem Exom Duplikate und von jedem Transkriptom Triplikate generiert. Da die DNA- bzw. RNA-Präparate für die einzelnen Replikate aus getrennten Kulturen verschiedener Zeitpunkte stammten (vgl. Abb. 3.2), handelte es sich um biologische Replikate im weiteren Sinn (Robasky *et al.*, 2014). Durch Sequenzieren von Exom- und Transkriptombibliotheken der gleichen Proben wurden, zumindest für exprimierte Target-Exons, Sequenzierungen zweierlei biologischen Ursprungs durchgeführt. Außerdem wurden die Exome zusätzlich jeweils auf zwei *Lanes* der *Flow Cell* aufgeteilt, so dass hier technische Replikate generiert wurden, die ebenso den Effekt von technischen Sequenzierfehlern reduzierten.

Der Vergleich der in den Exomen und Transkriptomen gefundenen Mutationen vor und nach Qualitätsfilterung zeigte, dass sich Mutationen anreichern ließen, die in mindestens zwei Replikaten nachweisbar waren (vgl. Abschnitt 3.4.1.1). Der Anteil der verlässlichen Mutationen mit einer ausreichend guten Qualität und Abdeckung war somit innerhalb der

durch mindestens ein Replikat bestätigten Mutationen am größten. Es verblieb jedoch für jede Zelllinie noch ein bedeutender Anteil an Mutationen (durchschnittlich 27% [Exome] bzw. 39% [Transkriptome]), die nur in einem der zwei bzw. drei Replikate detektiert wurden. Solche Mutationen wurden nicht weiter prozessiert. Allein durch einen Filter würden diese potentiell falschpositiven Mutationen nur auf Kosten tatsächlich positiver Mutationen eliminiert werden. Der beschriebene, im Vergleich zu den Exomdaten höhere Anteil an Mutationen, die auch nach dem Filtern nur in einer der drei Transkriptomsequenzierungen je Zelllinie gefunden wurden, ergab sich vermutlich aus der natürlichen Dynamik des Transkriptoms. Die Betrachtung der Expressionsprofile der einzelnen Datensätze bestätigte die Ähnlichkeit der Replikate (vgl. Abschnitt 3.4.1.2). Die Kultivierung der Zellen für eine weitere Woche zur Generierung des zweiten Replikates ergab weder signifikante Änderungen der Expressionsprofile, noch wurden Unterschiede in der Detektion der Mutationen in den Exomen oder Transkriptomen gefunden. Durch dieses Vorgehen konnte somit in geeigneter Weise biologische Varianz miteinbezogen werden. Mit dieser Grundlage konnten wiederholt detektierte, relevante Mutationen verlässlich identifiziert werden, weswegen auf eine kosten- und zeitintensive Validierung der Einzelmutationen durch Sanger-Sequenzierung verzichtet wurde. Der oben genannte Artikel (Robasky *et al.*, 2014) unterstützt diese Entscheidung nachträglich.

## 4.2 Bioinformatische Analysen

### 4.2.1 Stammbaum der Zelllinien

Aus dem Krankheitsverlauf der Patientin INTH war das chronologische Auftreten der vier Lymphknotenmetastasen bekannt, aus denen die stabilen Tumorzelllinien des Modells generiert wurden (vgl. Abb. 1.2). Um die (phylo-)genetischen Zusammenhänge zwischen den Tumorzelllinien zu ermitteln, wurde mithilfe der in den Exomsequenzierungen entdeckten SNVs ein Stammbaum der Ma-Mel-86-Zelllinien berechnet. Für die Analyse wurden nur Mutationen mit guter Qualität und Abdeckung verwendet, deren Zygotität aus den Sequenzierdaten sicher abgeleitet werden konnte (siehe Abschnitt 2.2.3.9). Der ermittelte phylogenetische Baum (vgl. Abb. 3.10) basierte auf einem artifiziellen *Alignment* aus 170.015 aneinandergereihten SNVs. Nur 22.512 der 120.484 somatischen Mutationen (18,7%) lagen in allen vier Tumorzelllinien vor, sodass anzunehmen war, dass die einzelnen Metastasen sich genetisch früh voneinander getrennt hatten. Die berechnete Topologie bestätigte diese Annahme. Da vom Primärtumor kein Material vorlag, wurde die autologe B-Zelllinie INTH-EBV-B als Wurzel festgelegt, wodurch gezeigt werden konnte, dass sich die vier Tumorzelllinien auf einen gemeinsamen Vorläufer („P“ in Abb. 3.10), möglicherweise den Primärtumor, zurückführen ließen. Ausgehend von „P“ ergab sich ein verzweigter Aufbau als Hinweis darauf, dass die Metastasen nicht sequenziell auseinander hervorgegangen sind. Die gezeigte Topologie wurde unter Verwendung zweier verschiedener Methoden

(*Neighbor-Joining* und *Maximum-Likelihood*) ermittelt und erschien somit robust. Demzufolge sollte die gewählte Strategie der Stammbaumberechnung einen adäquaten Hinweis auf die phylogenetischen Verhältnisse ergeben haben. Durch eine detailliertere Analyse und Vergleiche mit Bäumen, die mit veränderten Parametern berechnet wurden, könnte dies noch validiert werden.

Für die Berechnung dieses Stammbaums wurden jedoch beispielsweise Längenmutationen, wie der Verlust von *B2M* in Ma-Mel-86B und -86F (vgl. Abb. 3.11), vernachlässigt. Das unterschiedliche Ausmaß dieses Defekts ließ eine evolutionäre Trennung der beiden Zelllinien vor dem Verlust von *B2M* vermuten. Tatsächlich schloss die ermittelte Verwandtschaftsbeziehung zwischen den Tumorzelllinien auch unter Vernachlässigung dieser Information in der phylogenetischen Analyse eine Trennung nach Auftreten des Defekts aus, da sonst alle vier Zelllinien betroffen gewesen wären. Im Gegensatz zu der scheinbaren Ähnlichkeit von Ma-Mel-86B und -86F bezüglich des kompletten HLA-Klasse I-Verlustes unterschied sich Ma-Mel-86A hinsichtlich der Expression der HLA-Klasse I-Moleküle und melanozytärer Differenzierungsantigene (MDA) deutlich von Ma-Mel-86B. Im berechneten Stammbaum bildeten die beiden Zelllinien jedoch eine gemeinsame Klade und waren somit näher miteinander verwandt als ihr Phänotyp erwarten ließ. Allerdings zeigte der vom gemeinsamen Vorfahren „M“ ausgehende Ast eine rasche Verzweigung an. Insgesamt deutete diese verzweigte Tumorevolution, die ähnlich auch in anderen Patienten beschrieben wurde (Campbell *et al.*, 2008; Gerlinger *et al.*, 2012, 2014), auf eine fortwährende parallele Entwicklung von Subpopulationen innerhalb eines Tumors und folgender Metastasen hin.

### 4.2.2 Analyse des Mutationsprofils von Ma-Mel-86

Melanome gehören zu den Tumorarten, deren Pathogenese insbesondere von der mutagenen UV-Strahlung beeinflusst wird. Daher weisen sie in der Regel mehr Mutationen auf als viele andere solide Tumore oder als Leukämien. So wurden beispielsweise im Mittel  $202,9 \pm 182,0$  nicht-synonyme Mutationen pro Melanom beschrieben, während in Pankreaskarzinomen  $48,1 \pm 23,2$  nicht-synonyme Mutationen detektiert wurden (Vogelstein *et al.*, 2013). Die Annotation der in den Tumorexomen der Ma-Mel-86-Zelllinien gefundenen somatischen Mutationen ergab, dass im Mittel  $265,0 \pm 14,8$  Mutationen zu einem Aminosäureaustausch führten. Die Studien von Champiat *et al.* (2014) und Snyder *et al.* (2014) liefern Evidenzen, dass die Mutationslast eines Tumorpatienten prognostisch wirksam im Zusammenhang mit Immun-*Checkpoint*-Inhibitoren ist. Möglicherweise war das lange Überleben der Patientin INTH auch die Folge einer überdurchschnittlich hohen Anzahl immunogener Mutationen. Bei etwa 46% der somatischen Basensubstitutionen handelte es sich um C>T- bzw. G>A-Austausche (vgl. Abb. 3.12), die typischerweise infolge UV-induzierter Pyrimidindimere entstehen (Bertolotto, 2013). Bei der Suche nach potentiell immunogenen Mutationen wurden 181 exprimierte, somatische Aminosäuresubstitutionen lokalisiert (vgl. Abb. 3.15), von denen 63 (34,8%) in allen vier Melanomzelllinien vorhanden waren. Diesen gemeinsamen

Mutationen lagen zu fast 75% Substitutionen mit UV-Signatur zugrunde. Robbins *et al.* (2013) beobachteten solche Transitionen bei sogar 80 bis 90% der nicht-synonymen Punktmutationen dreier Melanomzelllinien. Aufgrund der Anreicherung von C>T- bzw. G>A-Austauschen in den gemeinsamen Mutationen, die vermutlich bereits zu einem großen Anteil im Primärtumor vorlagen, war die Melanomerkkrankung der Patientin INTH wahrscheinlich UV-induziert. In den aus Lymphknotenmetastasen generierten Melanomzelllinien wurde dieser Mutationstyp vermutlich bei nicht mehr vorhandener UV-Exposition durch weitere und andersartige Mutationen zurückgedrängt, die beispielsweise durch Replikationsfehler und Reparaturdefekte verursacht wurden.

Neben den 63 gemeinsamen Mutationen waren weitere 64 der 181 erwähnten Mutationen (35,4%) in jeweils nur einer der vier Zelllinien exprimiert. Dies entsprach in etwa den Beobachtungen von Gerlinger *et al.* (2014, 2012), die, wie in Abschnitt 1.2.2.3 beschrieben, die Heterogenität innerhalb von Nierenzellkarzinomen und assoziierten Metastasen analysierten. Im Gegensatz dazu deutete die gezielte Sequenzierung einer Reihe von krebsassoziierten Genen sowohl in Brust- als auch Lungenkrebspatienten auf eine hohe genetische Übereinstimmung zwischen Primärtumor und Metastasen bzw. Rezidiven hin (Meric-Bernstam *et al.*, 2014; Vignot *et al.*, 2013). Viele der Mutationen in krebsassoziierten Genen führen allerdings auch erst zum Erhalt der malignen Eigenschaften, weswegen sie tendenziell früh auftreten und stabil sind. Dementsprechend fanden sich unter den 63 Genen, die in allen vier Melanomzelllinien mutiert waren, beispielsweise auch die tumor- bzw. melanomassoziierten Gene *CDK4*, *TP53*, *MAP2K1* und *BRAF* (Hodis *et al.*, 2012; Kunz *et al.*, 2013).

Die identifizierten immunogenen Mutationen des Melanommodells Ma-Mel-86 waren spezifisch für die Patientin INTH. Keiner der Einträge für die Gene *INSIG1*, *MMS22L*, *PRDM10* bzw. *HERPUD1* im *Catalogue of Somatic Mutations in Cancer* (COSMIC, Release v70) stimmte mit den Mutationen in Ma-Mel-86 überein. Lediglich in *MMS22L* war in einem Endometrialkarzinom eine stumme Punktmutation (COSM1082753) in demselben Codon gefunden worden, das auch in Ma-Mel-86 mutiert war. Mutationen in *INSIG1* und *HERPUD1* in malignen Melanomen waren bisher nicht gelistet. In *MMS22L* wurden hingegen vier verschiedene *Missense*-Mutationen (COSM223114, COSM1697673, COSM1697672, COSM229013) in Melanomen detektiert und in *PRDM10* wurden eine *Missense*- (COSM1704822) und eine *Nonsense*-Mutation (COSM1704821) nachgewiesen. Die Immunogenität dieser Mutationen wurde jedoch nicht untersucht (Krauthammer *et al.*, 2012; Berger *et al.*, 2012).

## 4.3 Identifizierung mutierter Neoantigene

### 4.3.1 Strategie der Lokalisierung relevanter Punktmutationen

Neben der Kombination zweier Sequenziermethoden bildete die gewählte Strategie zur Definition relevanter Punktmutationen die Basis für die Identifizierung weiterer mutierter Neoantigene im Melanommodell Ma-Mel-86. Um dabei möglichst effektiv von den generierten Replikaten zu profitieren (siehe Abschnitt 4.1), wurden diese so lange wie möglich parallel prozessiert, um dann zur Bestimmung sicherer Mutationen gegeneinander abgeglichen zu werden (vgl. Abb. 3.13). Außerdem wurden die Transkriptomdaten neben den Exomsequenzierungen explizit bei der Suche nach immunogenen Mutationen eingesetzt. Für exprimierte *Target*-Exons lagen dadurch für jede Zelllinie bis zu fünf Sequenzierungen vor. Bisher waren *RNA-Seq*-Daten primär für die Überprüfung der Expression der detektierten Mutation bzw. der mutierten Gene vorgesehen (Castle *et al.*, 2012; Heemskerk *et al.*, 2013). Es wurden allerdings auch erste Studien veröffentlicht, die allein Tumortranskriptom zur Detektion von somatischen Mutationen nutzten (Xu *et al.*, 2013; Atak *et al.*, 2013). Eine Kombination der generierten DNA- und RNA-Sequenzierungen zur Identifikation von mutierten Tumorantigenen, wie sie in der vorgelegten Arbeit beschrieben ist, wurde in der Literatur bisher nicht berichtet.

Vor- und Nachteile der jeweiligen Sequenzierart (Heemskerk *et al.*, 2013) machten sich auch bei der Filterung der somatischen Mutationen bemerkbar. So sollten zur Erfassung von Keimbahnmutationen die Sequenzierungen der autologen EBV-B-Zelllinie INTH-EBV-B herangezogen werden. Aufgrund der natürlicherweise sehr unterschiedlichen Expressionsprofile einer B-Zelllinie im Vergleich zu Melanomzellen (vgl. auch Abb. 3.9) waren die im Transkriptom von INTH-EBV-B gefundenen Mutationen nicht geeignet, um alle in den Tumorzelllinien exprimierten Keimbahnmutationen herauszufiltern. Daher wurden aus den Mutationslisten der Tumorsequenzierungen sowohl die in den INTH-EBV-B-Exom als auch die in den INTH-EBV-B-Transkriptomsequenzierungen detektierten Mutationen entfernt (Details siehe Abschnitt 2.2.3.6.4). Erwartungsgemäß hätten alle im Transkriptom detektierten Mutationen auch im Exom vorliegen müssen, wenn man Mechanismen, wie *RNA-Editing* (Avesson & Barry, 2014), vernachlässigt. Es verblieben jedoch je Tumorzelllinie zwischen 30 und 53 und insgesamt 117 verschiedene Mutationen, die in mindestens zwei Transkriptomreplikaten, aber nicht in den Exomsequenzierungen detektiert wurden. Diese Mutationen wurden im *Integrative Genomics Viewer* IGV (Thorvaldsdóttir *et al.*, 2013) überprüft, woraufhin 45 Mutationen verworfen werden mussten, da sie nicht sicher somatisch waren. Diese Keimbahnmutationen waren nicht herausgefiltert worden, da die Abdeckung in den INTH-EBV-B-Exomen zu gering war oder fehlte (vgl. auch Abb. 3.7). Für weitere neun Einträge konnte die Mutation im IGV nicht bestätigt werden. Im Gegensatz dazu erwiesen sich nur neun der 127 Mutationen, die sowohl in den Exomen als auch in den Transkriptomen detektiert wurden, als nicht somatisch. Insgesamt wurden

damit 181 Mutationen sicher identifiziert und nicht zusätzlich per Sanger-Sequenzierung validiert (vgl. auch Abschnitt 4.1). Allerdings war an dieser Stelle die beschriebene manuelle Überprüfung notwendig. Ohne die Transkriptomsequenzierung wären jedoch mehr als 60 Mutationen, und darunter die immunogene Mutation in *PRDM10*, nicht weiter untersucht worden.

Durch eine Keimbahnkontrolle, deren Expressionsmuster dem der Tumorzellen mehr ähnelt, könnte die Filterung von somatischen Mutationen verbessert werden. Dies ist jedoch dadurch limitiert, dass sich bereits die Expressionsprofile von malignen und den entsprechenden gesunden Zellen unterscheiden (Heemskerk *et al.*, 2013). Aufgrund dessen wurden in der Studie von Xu *et al.* (2013) alle Mutationen, die in der *dbSNP*-Datenbank hinterlegt waren, ausgeschlossen. Da in der Datenbank jedoch auch mit Erkrankungen assoziierte Mutationen gelistet sind, hätte dies im Fall von Ma-Mel-86 zum Beispiel den Ausschluss der Mutationen in *CDK4* und *BRAF* bedeutet. Um dies zu vermeiden, wurden in der Arbeit von Atak *et al.* (2013) auch unter anderem die *dbSNP*-geführten Mutationen herausgefiltert, ohne die gleichzeitig in COSMIC gelisteten Mutationen zu verwerfen.

### 4.3.2 Auswahl potentiell HLA-bindender Peptide

Für die Vorhersage der HLA-Bindewahrscheinlichkeit mutierter Peptide wurden nur die 145 Mutationen herangezogen, die zumindest in einer der beiden HLA-Klasse I-exprimierenden Melanomzelllinien, Ma-Mel-86A und -86C, exprimiert wurden. Für die finale Auswahl der Peptidkandidaten wurden der ANN-basierte Algorithmus NetMHC (Lundegaard *et al.*, 2008) und der Algorithmus von IEDB verwendet, in dem Vorhersagen von ANN (Nielsen *et al.*, 2003; Lundegaard *et al.*, 2008), SMM (engl. *stabilizing matrix method*) (Peters & Sette, 2005) und comblib (engl. *combinatorial peptide libraries*) (Sidney *et al.*, 2008) kombiniert wurden. Diese Algorithmen wurden in erster Linie deshalb gewählt, da sie Vorhersagen zu allen sechs HLA-Klasse I-Allelen des Modells liefern. Die einzelnen Algorithmen sagten 332 (IEDB) bzw. 436 (NetMHC) HLA-Peptid-Kombinationen vorher, die die gewählten Schwellenwerte erfüllten (NetMHC:  $\leq 500$  nM, IEDB: Prozentrang  $\leq 10$ ). Durch Kombinieren der beiden Algorithmen verblieben 223 Vorhersagen mit 174 verschiedenen Peptiden, für die eine Bindung an ein bis drei HLA-Klasse I-Allele der Patientin anzunehmen war. Elf der 174 vorhergesagten Peptide wurden von autologen MLTC-Responderlymphozyten erkannt (siehe Tab. 3.8). Dies entsprach einer Erfolgsquote von 6,3%. Durch Verschärfen der Schwellenwerte auf eine Bindeaffinität von  $\leq 200$  nM (NetMHC) und einem Prozentrang von maximal 2,5 (IEDB), wären 63 Peptide zu testen gewesen, von denen zehn (16%) erkannt worden wären, ohne dass eine der identifizierten immunogenen Mutationen übersehen worden wäre.

Da jedoch Peptide mit  $\leq 80\%$  Reinheitsgrad getestet wurden und die Peptidpräparationen somit auch kürzere Fragmente enthielten (siehe auch nächster Abschnitt), war nicht davon auszugehen, dass tatsächlich elf verschiedene Peptide erkannt wurden. Zur korrek-

teren Einschätzung des Vorhersageerfolgs sollte daher angegeben werden, dass die 174 vorhergesagten Okta-, Nona- und Dekamere von 80 verschiedenen *Missense*-Mutationen abgeleitet waren, von denen wiederum vier, also 5%, erkannt wurden. In einer Studie zur Untersuchung der T-Zellerkennung von somatischen Mutationen in Ovarialkarzinomen (Wick *et al.*, 2014) induzierte in einem Patienten eines von 37 ausgewählten mutierten Genen (2,7%) eine spezifische T-Zellantwort. Bei zwei untersuchten Patienten mit chronisch lymphatischer Leukämie wurden, nach vorheriger *In-vitro*-Bestätigung der Peptidbindung an die HLA-A- bzw. HLA-B-Allele der Patienten, zwei von dreizehn bzw. eine von zwölf *Missense*-Mutationen (7,7% bzw. 8,3%) von autologen T-Zellen aus dem peripheren Patientenblut erkannt (Rajasagi *et al.*, 2014). Robbins *et al.* (2013) detektierten somatische *Missense*-Mutationen in den Tumorexomen dreier Melanompatienten, die auf eine TIL-Therapie angesprochen hatten. Es wurden 46 bis 82 Mutationen anhand der vorhergesagten Bindeaffinität für die patientenspezifischen HLA-A-Allele ausgewählt und pro Patient wurden zwei bis drei der Mutationen (2,4%, 3,2% bzw. 6,5%) von CD8<sup>+</sup> T-Zellen der autologen TIL-Zelllinien erkannt. Somit konnte nur durch die zusätzliche Überprüfung der Bindung der Peptide an die HLA-A/B-Moleküle (Rajasagi *et al.*, 2014) die Effizienz der Antigen suche im Vergleich zur vorliegenden Arbeit um wenige Prozente angehoben werden. In Anbetracht des Umstands, dass bei der hier genutzten Strategie zur Identifizierung mutierter Antigene in Ma-Mel-86 keine Einschränkung vorgenommen wurde auf HLA-A- bzw. HLA-B-Allele, für die aufgrund der besser untersuchten Bindemotive genauere Vorhersagen möglich sind als für HLA-C-Allele (Walshe *et al.*, 2009), war die erzielte Ausbeute somit vergleichbar mit anderen Studien.

#### 4.3.3 Überprüfung der Peptidimmunogenität

Zur Bestimmung der Immunogenität der Peptidkandidaten wurde der IFN $\gamma$ -ELISpot-Assay als Testformat gewählt, da damit auch T-Zellspezifitäten, die nur in niedriger Frequenz vorhanden sind, mit hoher Sensitivität und Spezifität in Mischpopulationen nachgewiesen werden können. Daher war es möglich, für die initialen Peptiderkennungstests MLTC-Responder-Populationen einzusetzen. Die Verwendung von Peptidpräparationen mit einer Reinheit von nur 20 bis 80% bot die Möglichkeit, ein großes Peptidpanel parallel und dabei kosteneffizient zu testen. Dabei war es möglich, die Peptide (vgl. Abb. 3.24 und Abb. 3.25) ohne Zugabe peptidpräsentierender Zellen direkt in die MLTCs zu geben und im IFN $\gamma$ -ELISpot-Assay zu testen.

In der vorliegenden Arbeit war die Erkennung mehrerer ähnlicher Peptidpräparationen jedoch auch ein Indiz dafür, dass es sich tatsächlich um eine spezifische Reaktivität handelte. In Abb. 4.1 ist dargestellt, welche der von den immunogen mutierten Genen abgeleiteten Peptide erkannt bzw. nicht erkannt wurden. Trotz der „kruden“ Peptidpräparationen lag jeweils eine spezifische Reaktivität gegen eine Gruppe von Peptiden vor, die zumindest für PRDM10<sup>mut</sup> auch eindeutig auf den N-Terminus des erkannten Peptids schließen ließ.



	HERPUD1 <sup>mut</sup>	INSIG1 <sup>mut</sup>	MMS22L <sup>mut</sup>	PRDM10 <sup>mut</sup>
erkannte Peptide	GLGPGFSSY	YQYTFPDLFY VYQYTFPDLFL YQYTFPDF VYQYTFPDF	EYYSKNLNSF YSKNLNSFF YYSKNLNSF YYSKNLNSFF	TYLPSAWNFF TYLPSAWNFF
nicht erkannte Peptide	FSSYTPYGWL SYTPYGWLQL	QYTFPDLFYI YTFPDLFY VYNGVYQYTF IHSQISSI IHSQISSIF IHSQISSIFV SQISSIFVL TAAGIHSQI	FFSISWLPF SFFSISWLPF	YLPSAWNFF

Abb. 4.1: Übersicht der von den immunogen mutierten Proteinen HERPUD1, INSIG1, MMS22L und PRDM10 abgeleiteten, getesteten Peptide. Die gezeigten mutierten Peptide (mutierte Aminosäure in Rot dargestellt) wurden im IFN $\gamma$ -ELISpot-Assay auf die Erkennung durch MLTC-Responderlymphozyten der Patientin INTH getestet. Jeweils nur ein Teil der Peptidpräparationen (oberhalb der Linie) induzierte eine IFN $\gamma$ -Sekretion. Die nicht fettgedruckten INSIG1<sup>mut</sup>-Peptide leiteten sich von der INSIG1-Isoform 3 (NCBI Reference Sequence: NP\_938151, siehe Abschnitt 4.3.4.1) ab.

Mithilfe von MLTC-Responderpopulationen wurden im Rahmen dieser Arbeit vier Neoantigene in Peptiderkennungstests identifiziert. Eines der Antigene, HERPUD1<sup>mut</sup>, war bereits durch cDNA-Expressionsklonierung gefunden worden. T-Zellantworten gegen mutierte Peptide wurden zuvor auch *ex vivo* nachgewiesen (z.B.: Lennerz *et al.*, 2005). In PBMCs der Patientin INTH vom 25.08.2004 war in der *Ex-vivo*-Testung gegen keines der mutierten Peptide Reaktivität nachweisbar (vgl. Abb. 3.26), obwohl die MLTC 1A.1 mit darin angereicherter Reaktivität gegen INSIG1<sup>mut</sup> und PRDM10<sup>mut</sup> aus PBMCs dieses Blutentnahmedatums generiert worden war (vgl. Abb. 3.28). Es ist also anzunehmen, dass die Frequenz dieser T-Lymphozyten zu gering war für eine Detektion ohne eine vorherige Stimulation und eine damit verbundene spezifische Expansion. Möglicherweise waren die mutationsspezifischen T-Zellpopulationen, die aufgrund ihrer HLA-A\*24:02- bzw. -B\*15:01-Restriktion nur Ma-Mel-86A erkennen konnten, mehr als zweieinhalb Jahre nach Entfernung der entsprechenden Lymphknotenmetastase zu stark zurückgegangen. Da kein Material mehr von früheren Blutentnahmen zur Verfügung stand, konnte die Detektierbarkeit der T-Zellen *ex vivo* zu früheren Zeitpunkten nicht überprüft werden. Alternativ könnte bereits eine einwöchige *In-vitro*-Peptid- oder z.B. Antigen-RNA-Stimulation vor der Durchführung des IFN $\gamma$ -ELISpot-Assays genügen (McCutcheon *et al.*, 1997; Rezvani *et al.*, 2009), um T-Zellantworten in Kurzzeitkulturen ohne die Etablierung von patienteneigenen Tumorzelllinien zu detektieren. Außerdem könnte die Identifizierung von Biomarkern für die tumorreaktiven, mutationsspezifischen T-Lymphozyten eine „Vorauswahl“ der gesuchten T-Zellen ermöglichen, beispielsweise durch fluoreszenzaktivierte Zellsortierung (FACS). So kennzeichnete beispielsweise die Expression von PD-1 auf CD8<sup>+</sup> TILs das Repertoire klonal expandierter, melanomreaktiver T-Zellen, die unter anderem gegen ein mutiertes Neoantigen gerichtet waren (Gros *et al.*, 2014).

Als Alternative zum IFN $\gamma$ -ELISpot-Assay wurde beispielsweise von Robbins *et al.* (2013) ein ELISA (engl. *enzyme-linked immunosorbent assay*) eingesetzt. Die Sensitivität des ELISA ist jedoch 10- bis 200-mal geringer (Tanguay & Killian, 1994; Slota *et al.*, 2011). Weiterhin könnte zur reinen Detektion peptidspezifischer T-Zellen eine Markierung mit Multimeren von HLA-Molekülen, die das gewünschte Peptid präsentieren, für ei-

ne durchflusszytometrische Analyse genutzt werden (Altman *et al.*, 1996). Mithilfe eines UV-induzierten Ligandenaustausches können Panels von Peptid/HLA-Komplexen generiert werden (Rodenko *et al.*, 2006). Für die Messung von 100 Peptidkandidaten wären jedoch etwa 500 ml peripheres Blut notwendig. Das benötigte Patientenmaterial kann allerdings durch die Verwendung von Multimeren mit kombinierbarer Fluoreszenzmarkierung reduziert werden (Heemskerk *et al.*, 2013). So wurden bis zu 28 verschiedene T-Zellpopulationen in einer Probe analysiert (Hadrup *et al.*, 2009; Andersen *et al.*, 2012; Newell *et al.*, 2009). Auf diese Weise wurden beispielsweise auch CD8<sup>+</sup> T-Zellen gegen zwei Neoantigene in einem Melanompatienten nachgewiesen, der auf eine Behandlung mit Ipilimumab angesprochen hatte (van Rooij *et al.*, 2013).

### 4.3.4 Charakteristika der identifizierten Neoantigene

#### 4.3.4.1 INSIG1<sup>mut</sup>

In den Peptiderkennungstests wurde in den MLTCs 1A.1, 3A und 16C Peptidreaktivität gegen die Peptide #024 (VYQYTFPDF), #025 (VYQYTFPDFL), #071 (YQYTFPDF) und #072 (YQYTFPDFLY) nachgewiesen, welche sich von dem mutiertem Gen INSIG1 (engl. *insulin-induced gene 1*, Gene ID: 3638) ableiteten. INSIG1 ist auf dem langen Arm des Chromosoms 7 (Bande 7q36) lokalisiert (Peng *et al.*, 1997). Es sind drei INSIG1-Transkriptionsvarianten beschrieben. Die erste Isoform hat eine Länge von 277 Aminosäuren (NCBI Reference Sequence: NM\_005542/NP\_005533). Die zweite INSIG1-Isoform (NM\_198336/NP\_938150) nutzt eine alternative Spleißstelle in der 5'-Region des Gens. Außerdem kommt es durch Einfügen eines zusätzlichen Exons in der 3'-Region zu einem Rasterschub, weswegen die zweite Isoform gegenüber der ersten um 99 Aminosäuren kürzer ist und einen anderen C-Terminus aufweist. Der dritten beschriebenen Transkriptionsvariante von INSIG1 (NM\_198337/NP\_938151) fehlen in der Mitte des Gens zwei Exons, sodass ein Rasterschub ebenso zu einem alternativen C-Terminus und einem kürzeren, 164 Aminosäuren langen Translationsprodukt führt. Die immunogene Mutation befand sich an Position 713 (c.713C>T<sup>9</sup>) des ORF der Isoform 1. Aufgrund des veränderten Leserasters in den Isoformen 2 und 3 lagen die erkannten Peptide nur in der ersten Transkriptionsvariante vor. Hier führte die Basensubstitution c.713C>T zum Austausch von Serin nach Phenylalanin (p.S238F<sup>10</sup>). Die Analyse der Expression der einzelnen Transkriptionsvarianten mithilfe von *cufflinks* (Trapnell *et al.*, 2010) ergab, dass in den vier Melanomzelllinien jeweils die Isoform 1 am stärksten exprimiert wurde (nicht gezeigt).

Das Protein INSIG1 ist mit sechs Transmembrandomänen (TMD) und nur kurzen luminalen bzw. cytosolischen Schleifen größtenteils in der ER-Membran verankert (Feramisco *et al.*,

---

<sup>9</sup>Beschreibung der Punktmutation in kodierender DNA-Sequenz („c.“) entsprechend der Empfehlung der *Human Genome Variation Society* (den Dunnen & Antonarakis, 2000)

<sup>10</sup>Schreibweise entsprechend Fußnote 9 für Proteinsequenzen („p.“)

2004). Dort spielt INSIG1 durch Bindung an SCAP (engl. *SREBP cleavage-activating protein*) bzw. an die HMG-CoA-Reduktase (3-Hydroxy-3-Methylglutaryl-Coenzym-A-Reduktase) eine zentrale Rolle in der Regulation der Cholesterol- und Fettsäuresynthese. Die Bindung an diese Proteine erfolgt im Bereich der TMD (Yang *et al.*, 2002). Da sich die in Ma-Mel-86 detektierte Aminosäuresubstitution S238F in einer luminalen Schleife befand, hatte die Mutation vermutlich keinen Einfluss auf die Funktion von INSIG1. Es handelte sich höchstwahrscheinlich um eine *Passenger*-Mutation.

Laut den Expressionsanalysen von Krapivner *et al.* (2008) ist INSIG1 primär in der Leber und nur schwach in extrahepatischem Gewebe exprimiert. Eine signifikante Hochregulation der INSIG1-Expression wurde jedoch beispielsweise im Zusammenhang mit alternder Haut beschrieben (Lener *et al.*, 2006). Die Transfektion von mutierter bzw. wildtypischer *INSIG1*-cDNA in COS-7-Zellen (vgl. Abb. 3.33B) zeigte, dass die beobachtete T-Zellreaktivität nur gegen das mutierte INSIG1 gerichtet war. Somit wäre trotz der Expression von INSIG1 in gesundem Gewebe nicht mit einer T-Zellantwort gegen das nicht mutierte INSIG1 zu rechnen.

#### 4.3.4.2 PRDM10<sup>mut</sup>

Aus den mit Ma-Mel-86A stimulierten MLTCs 1A.1 und 3A konnten T-Zellen isoliert werden, die die Peptide #021 (TYLPSAWNF) und #022 (TYLPSAWNFF) erkannten. Diese Peptide entstanden durch die Prozessierung von mutiertem *PRDM10* (engl. *PR domain containing 10*, *Gene ID*: 56980). Das Gen liegt auf dem langen Arm des Chromosoms 11 (Bande 11q25). Es sind vier Transkriptionsvarianten von *PRDM10* annotiert. Die längste Variante, Isoform 1 (*NCBI Reference Sequence*: NM\_020228/NP\_064613), umfasst 1160 Aminosäuren. Die zweite Isoform (NM\_199437/NP\_955469) unterscheidet sich nur durch das Fehlen eines Exons, das für vier Aminosäuren kodiert. Die weiteren Transkriptionsvarianten (NM\_199438/NP\_955470 bzw. NM\_199439/NP\_955471) weisen im Vergleich zur Isoform 1 mehrere Unterschiede insbesondere am N-Terminus auf. Der peptidkodierende Bereich war jedoch in allen vier *PRDM10*-Varianten vorhanden. Im Modell Ma-Mel-86 wurde die cDNA der Isoform 2 kloniert (1156 Aminosäuren). Die in Ma-Mel-86 vorliegende Punktmutation c.3149C>T (bezogen auf Isoform 2) führte in Codon 1050 zu einem Aminosäureaustausch von Serin nach Phenylalanin (p.S1050F).

*PRDM10* ist eines der 17 Mitglieder der *PRDM*-Familie, die eine N-terminale PR (engl. *PRDI-BF1 [positive regulatory domain I-binding factor 1] and RIZ [retinoblastoma-interacting zinc finger protein 1] homology*)-Domäne und in der Regel mehrere Zinkfinger gemeinsam haben. *PRDM*-Proteine sind zelltyp- bzw. gewebsspezifische transkriptionelle Regulatoren, die entweder selbst als Histon-Methyltransferasen (HMT) oder durch die Rekrutierung anderer Enzyme an der epigenetischen Regulation der Genexpression beteiligt sind (Fog *et al.*, 2012). Für *PRDM10* wurde bisher keine HMT-Aktivität nachgewiesen (Di Zazzo *et al.*, 2013). Das Protein konnte jedoch in Assoziation mit der Deubiquitinase

BAP1 (engl. *BRCA1 associated protein 1*) aus Flag-HA-BAP1-transduzierten HeLa-Zellen aufgereinigt werden (Yu *et al.*, 2010). In Mäusen wurde die Expression von PRDM10 in verschiedenen Geweben auf mRNA- und Proteinebene gezeigt (Siegel *et al.*, 2002; Park & Kim, 2010). Aufgrund der ausschließlichen Erkennung des mutierten PRDM10 (vgl. Abb. 3.31B) wäre, auch im Falle ähnlicher Expressionsprofile im Menschen, keine T-Zellreaktivität gegen gesunde Zellen zu erwarten.

Von der immunogenen Mutation waren weder die PR-Domäne noch die Zinkfinger betroffen, so dass vermutlich keine funktionelle Veränderung induziert wurde. Trotz des beschriebenen onkogenen bzw. tumorsuppressiven Potentials mancher PRDM-Proteine (Fog *et al.*, 2012) war im hier untersuchten Melanommodell nicht von einer Beteiligung von mutiertem PRDM10 als *Driver* auszugehen, zumal die Mutation auch nur in Ma-Mel-86A vorlag.

#### 4.3.4.3 MMS22L<sup>mut</sup>

Weiterhin wurde im Modell Ma-Mel-86 in den MLTCs 3A und 16C die Erkennung der Peptide #003 (EYYSKNLNSF), #026 (YYSKNLNSF), #027 (YYSKNLNSFF) und #073 (YSKNLNSFF) beobachtet. Diese Peptide stammten von mutiertem *MMS22L* (engl. *methyl methanesulfonate-sensitivity protein 22-like*, Gene ID: 253714) ab. Das Gen *MMS22L* ist auf dem langen Arm des Chromosoms 6 (Bande 6q16.1) lokalisiert und es ist nur eine Isoform (NM\_198468/NP\_940870) mit einer Länge von 1243 Aminosäuren beschrieben. Die detektierte Punktmutation c.1310C>T betraf Codon 437 (p.S437F) und befand sich innerhalb eines konservierten Proteinabschnitts (Duro *et al.*, 2010).

Aufgrund der Sequenzähnlichkeit dieser konservierten Region (Codons 373-535) zu *Mms22* in *Schizosaccharomyces pombe* wurde das ursprünglich nur als C6ORF167 bezeichnete humane Protein mit unbekannter Funktion als *Mms22-like* (MMS22L) bezeichnet, die Beteiligung von MMS22L an der homologen Rekombination von Doppelstrangbrüchen untersucht und MMS22L als entferntes Ortholog des Hefe-*Mms22* bestätigt (Duro *et al.*, 2010). Im Komplex mit TONSL (engl. *Tonsoku-like*) nimmt MMS22L eine wichtige Rolle bei der Wiederherstellung der Replikationsgabel im Falle von DNA-Läsionen während der S-Phase ein (O'Donnell *et al.*, 2010; Piwko *et al.*, 2010; O'Connell *et al.*, 2010; Duro *et al.*, 2010). Somit könnten genetische Alterationen in MMS22L bzw. TONSL mit Krebs assoziiert sein (O'Donnell *et al.*, 2010). Ein Hinweis auf die onkogenen Eigenschaften von MMS22L lieferte eine Studie, wonach MMS22L in einer Reihe von Lungen- und Speiseröhrentumoren überexprimiert und auf Proteinebene im Normalgewebe, außer in den Hoden, kaum zu detektieren war (Nguyen *et al.*, 2012). Durch siRNAs (engl. *small interfering RNAs*) gegen *MMS22L* wurde das Wachstum von Tumorzellen inhibiert und umgekehrt in COS-7- und HEK293-Zellen nach Transfektion von *MMS22L* induziert. Die Studie zeigte außerdem Verbindungen zum NF $\kappa$ B (engl. *nuclear factor of  $\kappa$  light polypeptide gene enhancer in B-cells*)-Signalweg und zu den anti-apoptotischen Molekülen Bcl-XL (engl. *B-cell lymphoma extra-large*) und TRAF1 (engl. *tumor necrosis factor receptor-associated factor 1*).

In Ma-Mel-86 war *MMS22L* zumindest in Ma-Mel-86A signifikant herabreguliert (vgl. Abb. 3.27). Im Gegensatz zu den übrigen Neoantigenen, die im Rahmen dieser Arbeit identifiziert wurden, war *MMS22L* in allen vier Tumorzelllinien des Modells mutiert, wobei kein Wildtypallel mehr vorlag. Dies implizierte, dass die Mutation bereits früh aufgetreten war. CNV (engl. *copy number variation*)-Analysen deuteten auf eine Deletion des zweiten (wildtypischen) Allels hin. Die Resultate waren jedoch nicht eindeutig (nicht gezeigt). Ob sich die Aminosäuresubstitution p.S437F funktionell auswirkte, kann nur spekuliert werden, da entsprechende Untersuchungen fehlen und keine Ähnlichkeit zu annotierten Proteindomänen bestand. Für die Validierung der intrazellulären Prozessierung des identifizierten Peptids  $\text{MMS22L}_{429-438}^{\text{S437F}}$  wurden nur cDNA-Fragmente transfiziert, sodass kein Effekt dieser Transfektionen auf das Wachstum der Rezipientenzellen erwartet wurde. Da jedoch die konservierte Region betroffen war und da das Wildtypallel in allen Tumorzelllinien verloren war, liegt der Gedanke an das Vorliegen einer onkogenen *Driver*-Mutation nahe.

Bei der Charakterisierung der Erkennung des Neoantigens  $\text{MMS22L}^{\text{mut}}$  fiel auf, dass einerseits das cDNA-Fragment, welches mit dem Codon 437 endete, nicht erkannt wurde (vgl. Abb. 3.36), obwohl die Peptide #003 ( $\text{MMS22L}_{428-437}^{\text{S437F}}$ ) und #026 ( $\text{MMS22L}_{429-437}^{\text{S437F}}$ ) erkannt wurden (vgl. z. B. Abb. 3.29C). Andererseits wurde eine schwache Reaktivität gegen das wildtypische cDNA-Fragment beobachtet (vgl. Abb. 3.35B und Abb. 3.36). In Anbetracht der Funktion von *MMS22L* und der daraus zu erwartenden Folgen eines Expressionsverlustes ist das von Nguyen und Kollegen (2012) beschriebene Expressionsmuster (siehe oben) eher fraglich. So ist auch in Abb. 3.27 gezeigt, dass *MMS22L* auf mRNA-Ebene in den Referenzmelanozyten höher exprimiert war als in den Tumorzelllinien des Modells. Die „schwache“ Expression in Ma-Mel-86A genügte bereits für eine Erkennung durch spezifische T-Zellen (vgl. z. B. Abb. 3.35A). Da jedoch die  $\text{MMS22L}^{\text{wt}}$ -exprimierenden INT-EBV-B-Zellen nicht durch spezifische T-Zellklone erkannt wurden (vgl. Abb. 3.29C), sind vermutlich im Falle einer physiologischen Expression des wildtypischen *MMS22L* keine Autoimmunreaktionen zu erwarten.

Die Diskrepanz zwischen den erkannten Peptidpräparationen und den erkannten cDNA-Fragmenten wurde im Rahmen dieser Arbeit nicht im Detail untersucht. Da die transfizierten cDNA-Fragmente vollständig sequenziert worden waren, war bekannt, welche Aminosäure am C-Terminus vorlag. Dies war für die Peptidpräparationen mit einer Reinheit von 20-80% nicht exakt definiert. Daher wurde vermutet, dass die Peptide #003 und #026, die mit dem Codon 437 endeten, zwar von HLA-A\*24:02 präsentiert werden konnten, dass jedoch das entsprechende cDNA-Fragment aufgrund des Fehlens des Phenylalanins an Position 438 nicht intrazellulär prozessiert werden konnte. Aufgrund dessen wurde das Peptid  $\text{MMS22L}_{429-438}^{\text{S437F}}$  als das auf Ma-Mel-86A erkannte Neoantigen angenommen.

### 4.3.4.4 Eigenschaften der identifizierten Neoantigene

Den drei neu beschriebenen Antigenen war gemeinsam, dass sie durch HLA-A\*24:02 präsentiert wurden, und dass die Basensubstitution C>T jeweils zu einem Aminosäureaustausch von Serin zu Phenylalanin führte. In PRDM10<sup>S1050F</sup><sub>1042-1050</sub> befand sich das Phenylalanin am Carboxylende, in MMS22L<sup>S437F</sup><sub>429-438</sub> an vorletzter Position des Peptids und in INSIG1<sup>S238F</sup><sub>233-241</sub> zwischen den beiden primären Ankerpositionen, vermutlich an Position 6.

Die primären Ankerpositionen in Peptiden, die an HLA-A24 binden, sind Position 2 und der C-Terminus (Kondo *et al.*, 1995; Sidney *et al.*, 2005). Die drei identifizierten Neoantigene trugen an diesen Positionen übereinstimmend die bevorzugten Aminosäuren Tyrosin (Position 2) bzw. Phenylalanin (C-Terminus). Im Fall von PRDM10<sup>mut</sup> führte erst der detektierte Aminosäureaustausch p.S1050F zur Entstehung dieses Bindemotivs. Bei PRDM10<sup>S1050F</sup><sub>1042-1050</sub> kann daher von der Generierung eines Aggretops ausgegangen werden.

Die mRNA-Expressionsanalyse der vier immunogen mutierten Gene (vgl. Abb. 3.27) ergab, dass *HERPUD1* und *INSIG1* in Ma-Mel-86A im Vergleich zu den Referenzmelanozyten mehr als zweifach überexprimiert waren. *MMS22L* und *PRDM10* waren hingegen eher herabreguliert bzw. auf dem Expressionslevel der Melanozyten. Außerdem lagen das mutierte und das wildtypische Allel von *PRDM10* im Verhältnis von ca. 1:2 vor. Dennoch gehörte die Reaktivität der MLTC 1A.1 gegen die PRDM10<sup>mut</sup>-Peptide zu den dominanten T-Zellantworten dieser MLTC (vgl. Abb. 3.25, blaue Balken). Es wurden nur aus dieser MLTC PRDM10<sup>mut</sup>-reaktive T-Zellklone isoliert (vgl. Abb. 3.28). Die von MMS22L<sup>mut</sup> abgeleiteten Peptide wurden am stärksten von der MLTC 16C erkannt. Dies hing eventuell damit zusammen, dass *MMS22L* in Ma-Mel-86C stärker exprimiert war als in Ma-Mel-86A (vgl. Abb. 3.27), wenn angenommen wird, dass die für die Stimulation der MLTC verwendeten, mit HLA-A\*24:02 rekonstituierten Ma-Mel-86C-Zellen vergleichbare Expressionslevel aufwiesen. Auch für dieses Antigen wurden nur durch Klonieren der MLTC mit der stärksten MMS22L<sup>mut</sup>-Reaktivität spezifische T-Zellklone generiert. INSIG1<sup>mut</sup>-spezifische T-Zellen konnten hingegen aus allen drei peptidreaktiven MLTCs isoliert werden. Insgesamt erklärte die Präsentation der identifizierten Neoantigene über HLA-A\*24:02, weshalb keine Peptidreaktivität in den MLTC-Responderpopulationen detektiert wurde, die mit den HLA-A\*24:02-negativen Ma-Mel-86C-Zellen restimuliert worden waren.

Die schwache Expression der mutierten *MMS22L*- bzw. *PRDM10*-mRNA in Ma-Mel-86A könnte ein Grund dafür sein, dass die beiden Antigene bei den zuvor durchgeführten cDNA-Expressionsscreenings der Ma-Mel-86-MLTCs mit der Ma-Mel-86A-cDNA-Bank nicht identifiziert wurden (vgl. Abschnitt 1.1.1.3.3). Durch die Erkennungstests mit den mutierten Peptiden, die dabei im Überschuss eingesetzt wurden, wurde diese Einschränkung umgangen. Weiterhin wiesen die negativen *Ex-vivo*-Tests (vgl. Abb. 3.26) auf eine niedrige Frequenz der entsprechenden T-Zellpopulationen hin. Im Rahmen der Doktorarbeit von S. Lübcke isolierte HLA-A\*24:02-restringierte T-Zellklone konnten nicht ausreichend expandiert werden, um ein cDNA-Bank-Screening durchzuführen. Für das Screening

der hier getesteten 174 Peptide wurde im Vergleich zu einem vollständigen cDNA-Bank-Screening mit 100.000 bis 200.000 cDNA-Klonen nur etwa ein Achtel bis ein Viertel der T-Zellen benötigt.

### 4.3.5 Bedeutung mutierter Antigene

Mutierte Neoantigene repräsentieren tatsächlich tumorspezifische Antigene (vgl. Einleitung, Abschnitt 1.2.1), gegen die keine zentrale immunologische Toleranz vorliegt. Da die zugrundeliegenden Mutationen jedoch meist patientenspezifisch sind, sind mutierte Neoantigene in aller Regel nur Zielstrukturen für individualisierte Therapieverfahren. Dennoch bleibt im Vergleich zu überexprimierten oder selektiv exprimierten Antigenen der Vorteil der fehlenden Toleranz ohne das Risiko autoreaktiver Immunantworten (Rajasagi *et al.*, 2014). In den vergangenen 20 Jahren wurden etwa 50 mutierte, von T-Zellen erkannte Peptide in soliden Tumoren und Leukämien beschrieben (siehe <http://cancerimmunity.org/peptide/mutations/>, Stand: 13.08.2014). Durch die Entwicklung der Hochdurchsatzsequenzierung und die Verbesserung der Algorithmen zur Vorhersage der HLA-Bindewahrscheinlichkeiten können Neoantigene nun systematisch identifiziert werden. So wurden beispielsweise mithilfe von Exomsequenzierungen immunogene mutierte Peptide in CLLs (Rajasagi *et al.*, 2014), Melanomen (Robbins *et al.*, 2013) und Ovarialkarzinomen (Wick *et al.*, 2014) charakterisiert. Rajasagi *et al.* (2014) analysierten 2488 Exomsequenzierungen und postulierten im Mittel für jede Punktmutation ca. 1,5 potentiell HLA-bindende Peptide und für jede Rasterschubmutation etwa vier HLA-Binder vorhersagen ließen. Im Modell Ma-Mel-86 wurden durchschnittlich 1,2 Peptide pro somatische Punktmutation als Kandidaten ausgewählt.

Das immuntherapeutische Potential von Neoantigenen wurde bereits in Maus (Mandelboim *et al.*, 1995; Hanson *et al.*, 2000; Castle *et al.*, 2012) und Mensch (Zhou *et al.*, 2005; Lu *et al.*, 2013; Tran *et al.*, 2014) gezeigt. In manchen Patienten dominierten die T-Zellantworten gegen mutierte Antigene (Lennerz *et al.*, 2005; Novellino *et al.*, 2003). Auch im Melanommodell Ma-Mel-86 waren vier der fünf bisher aufgeklärten Ma-Mel-86A-reaktiven T-Zellantworten gegen Neoantigene gerichtet, während keines der MDA (Tyrosinase, gp100) und keines der HLA-unabhängigen Antigene (TRP-2, CSF2RA) in Ma-Mel-86A exprimiert war und somit auch nicht zur Induktion entsprechender T-Zellantworten führen konnte. Gerade für Tumorarten wie Melanome, Lungen- oder Kolorektalkarzinome mit hoher Mutationslast (Vogelstein *et al.*, 2013) könnten immunogene tumorspezifische Mutationen mit standardisierten Verfahren identifiziert und in einer personalisierten Therapie genutzt werden. Bestrebungen in diese Richtung werden bereits verfolgt (Rammensee & Singh-Jasuja, 2013; Castle *et al.*, 2012) und erste klinische Studien werden durchgeführt, wie z.B. IVAC MUTANOME oder GAPVAC (NCT02035956 und NCT02149225, <https://www.clinicaltrials.gov/>). Wenn sich außerdem die Hinweise weiter bestätigen lassen, dass eine hohe Anzahl an (immunogenen) Mutationen im Tumor

als molekulare Determinante für den Erfolg einer Behandlung mit *Checkpoint*-Inhibitoren anwendbar ist (Champiat *et al.*, 2014; Snyder *et al.*, 2014), gewinnen mutierte Neoantigene einerseits als prognostischer Biomarker und andererseits als Zielstrukturen in einer Kombinationstherapie an Wichtigkeit.

## 4.4 *Immune Escape* und *Immunoediting* in Ma-Mel-86

Wie bereits im Einleitungsteil (Abschnitt 1.1.1.3.2) beschrieben, stellten drei der vier Tumorzelllinien des Melanommodells Ma-Mel-86 HLA-Klasse I-Verlustvarianten dar. Die vierte Zelllinie, Ma-Mel-86A, war negativ für die Expression von MDA. Dies ermöglichte es den Tumorzellen, der Eliminierung durch HLA-Klasse I-abhängige und MDA-reaktive T-Zellen zu entgehen. Somit trugen alle Modellzelllinien *Immune Escape*-Merkmale. Im Fall der HLA-Klasse I-negativen Zelllinien Ma-Mel-86B und -86F wurden zwei unterschiedliche Defekte im *B2M*-Lokus identifiziert (vgl. Abb. 3.11). Es ist daher vorstellbar, dass sie unabhängig voneinander unter dem Selektionsdruck durch HLA-Klasse I-abhängige T-Zellen entstanden. Laut den Untersuchungen von Aptsiauri *et al.* (2008) können in Folge eines durch Immuntherapie verursachten Selektionsdrucks zwei Arten von Metastasen entstehen, die sich anhand der HLA-Expressionslevel unterscheiden lassen: regressive Metastasen, die eine erhöhte HLA-Expression aufweisen, und progressive Läsionen mit niedrigeren HLA-Leveln. Im Fall der Patientin INTH wurde im Jahr 2002 nach der operativen Entfernung der Lymphknotenmetastase, aus der Ma-Mel-86A generiert wurde, die erste immuntherapeutische Behandlung durchgeführt (vgl. Abb. 1.2), woraufhin im weiteren Krankheitsverlauf Metastasen mit HLA-Defekten isoliert werden konnten. Ob jedoch dafür tatsächlich ein kausaler Zusammenhang bestand, ist spekulativ.

Studien an Sarkomen im Mausmodell fanden Hinweise dafür, dass der Prozess des *Immunoediting* abhängig ist von der Erkennung dominanter tumorspezifischer Antigene durch T-Lymphozyten. Ein mutiertes Antigen war primär für die Abstoßung einer noch nicht durch das Immunsystem „editierten“ Sarkomzelllinie verantwortlich, die in immunkompetente Mäuse transplantiert wurde (Matsushita *et al.*, 2012). Aufgrund ihrer Präsentation durch HLA-A\*24:02 konnten auch die Neoantigene INSIG1<sup>mut</sup>, PRDM10<sup>mut</sup> und MMS22L<sup>mut</sup> nur auf der Melanomzelllinie Ma-Mel-86A erkannt werden. Gleiches galt für das durch HLA-B\*15:01 präsentierte Antigen HERPUD1<sup>mut</sup>. Im Gegensatz dazu konnte das überexprimierte Antigen CSPG4<sub>554-562</sub> im Komplex mit HLA-C\*07:01 nicht nur auf Ma-Mel-86A, sondern auch auf -86C präsentiert und erkannt werden. Das Vorkommen der neoantigenspezifischen T-Zellen, die die späteren HLA-Verlustvarianten des Modells nicht mehr erkennen konnten, unterstützt die Annahme, dass T-Zellen, die gegen tumorspezifische Mutationen gerichtet sind, eine bedeutende Rolle während des *Immunoediting* einnehmen. Aufgrund der fehlenden MDA-Expression bei gleichzeitiger Präsenz MDA-reaktiver T-Lymphozyten könnte Ma-Mel-86A allerdings selbst bereits aus einem Selektionsprozess heraus entstanden sein.



In der von Wick *et al.* (2014) publizierten Analyse zur Erkennung eines mutierten Antigens auf dem Primärtumor und zwei Rezidiven einer Ovarialkarzinompatientin wurde gezeigt, dass mit dem Anstieg der Frequenz des mutierten Allels nach der ersten Remission mutationsspezifische T-Zellen im Aszites der Patientin zu beobachten waren. Obwohl die betreffende Mutation weiterhin exprimiert wurde, waren beim zweiten Rezidiv keine spezifischen T-Zellen mehr detektierbar. Die Autoren vermuteten eine klonale Deletion der mutationsspezifischen T-Lymphozyten aufgrund chronischer Antigenexposition. Für eine Vakzinierungsstrategie gegen ein mutiertes Antigen müsste daher nicht nur eine transiente Expression des Antigens, sondern auch die Möglichkeit einer nur transient vorhandenen T-Zellreaktivität berücksichtigt werden. Eine Charakterisierung des T-Zellrepertoires im Verlauf der Krankheit der Patientin INTH mit T-Zellnachweisverfahren war aufgrund des geringen verbliebenen Probenmaterials nicht möglich. Durch eine gezielte Sequenzierung der TCRs im peripheren Blut, wie es beispielsweise von Linnemann *et al.* (2013, 2014) beschrieben wurde, könnten allerdings die Frequenzen der Ma-Mel-86-T-Zellpopulationen, deren Spezifitäten bisher aufgeklärt wurden bzw. deren TCR-Sequenzen bekannt sind, bestimmt und verglichen werden.



## 5 Fazit

Im Rahmen dieser Arbeit wurden mittels Exom- und Transkriptomsequenzierungen der Zelllinien des Modells Ma-Mel-86 vier mutierte Antigene identifiziert und charakterisiert, von denen drei bisher unbekannt waren. Drei der vier immunogenen Mutationen lagen in mehr als einer Melanomzelllinie vor. Allerdings war die Erkennung der mutierten Antigene aufgrund der fehlenden Expression der restringierenden HLA-Klasse I-Allele in den aus anderen Metastasen generierten Zelllinien auf Ma-Mel-86A beschränkt. Diese Daten lassen die Vermutung zu, dass gerade die T-Zellen, die gegen die mutierten Antigene reagierten, einen Selektionsdruck ausübten, der zur Etablierung der HLA-Klasse I-Verlustvarianten führte. Die Heterogenität der Mutationsprofile der einzelnen Metastasen und die Entstehung der HLA-Klasse I-Verlustvarianten zeigen auf, dass für eine effektive Immuntherapie des Melanoms mehrere Antigene kombiniert werden müssten. Dabei ist trotz des möglicherweise großen Potentials der neoantigenspezifischen T-Zellen auch die HLA-unabhängige T-Zellerkennung von Zielstrukturen auf der Tumorzelloberfläche von Bedeutung, um HLA-defiziente Tumorzellen targetieren zu können. Da es sich bei mutierten Neoantigenen jedoch meist um patientenspezifische Antigene handelt, werden die hier charakterisierten Antigene nicht für die Behandlung anderer Patienten genutzt werden können. Die vorgelegte Arbeit belegt dennoch die Anwendbarkeit und Effektivität der Methode und kann als Grundlage für die Suche immunogener Mutationen in anderen Patienten verwendet werden.



# A Anhang

## A.1 Vektorkarten

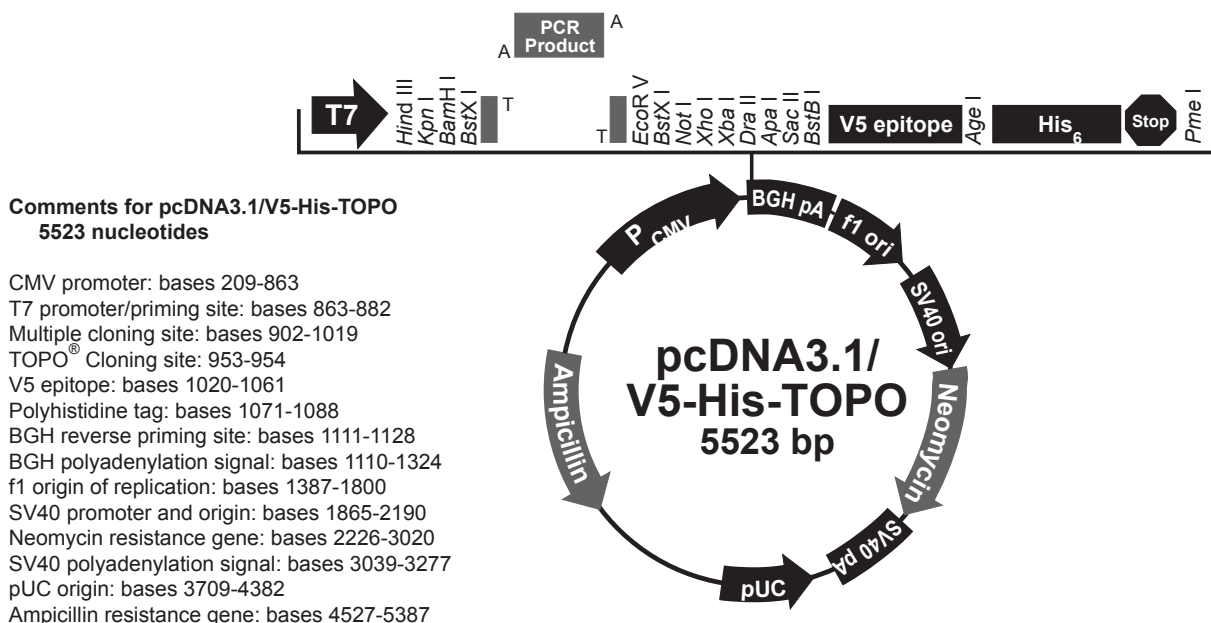


Abb. A.1: Vektorkarte von *pcDNA<sup>TM</sup>3.1/V5-His TOPO<sup>®</sup>*

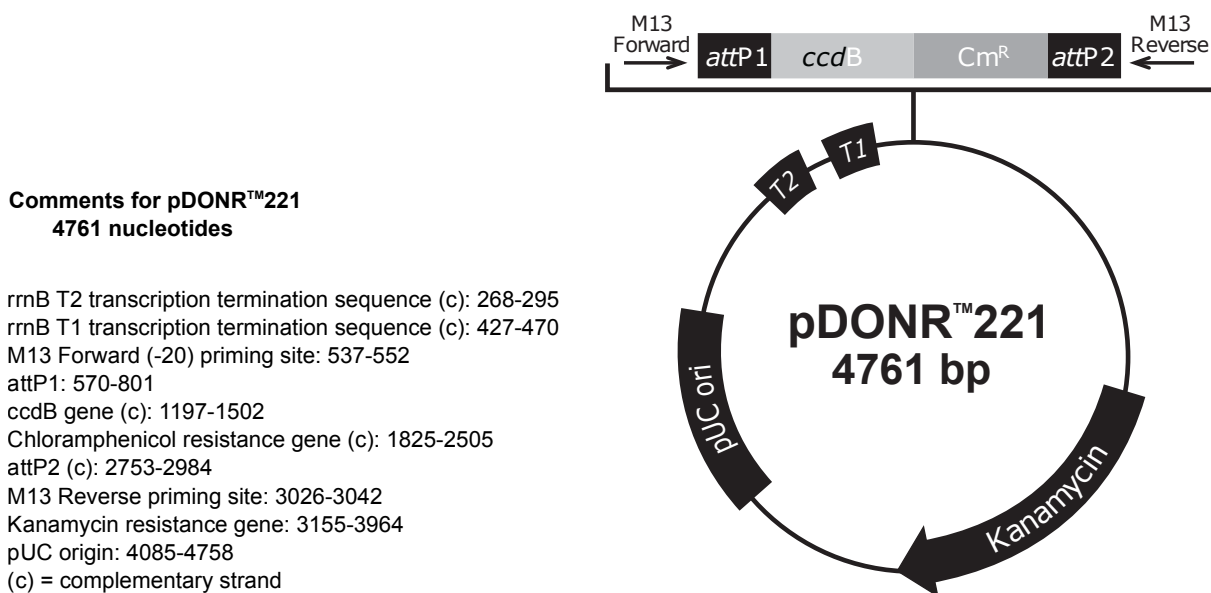


Abb. A.2: Vektorkarte von *pDONR<sup>TM</sup>221*

---

## A.2 Sequenzen

### A.2.1 IUPAC Ambiguity Code

Entsprechend der Nomenklatur von IUPAC (*International Union of Pure and Applied Chemistry*) und IUBMB (*International Union of Biochemistry and Molecular Biology*) können nicht eindeutig festgelegte Basen in Nukleinsäuresequenzen mithilfe des sogenannten *ambiguity code* bezeichnet werden (Cornish-Bowden, 1985). Die folgende Tabelle zeigt die empfohlene und in dieser Arbeit verwendete Schreibweise.

Tab. A.1: *Ambiguity code* zur Bezeichnung nicht eindeutig festgelegter Basen in Nukleinsäuresequenzen.

Symbol	Bedeutung	Herkunft der Bezeichnung
G	G	<i>Guanine</i>
A	A	<i>Adenine</i>
T	T	<i>Thymine</i>
C	C	<i>Cytosine</i>
R	G oder A	<i>puRine</i>
Y	T oder C	<i>pYrimidine</i>
M	A oder C	<i>aMino</i>
K	G oder T	<i>Keto</i>
W	A oder T	<i>Weak interaction</i>
H	A oder C oder T	<i>not G</i> (H folgt G im Alphabet)
B	G oder T oder C	<i>not A</i> (B folgt A im Alphabet)
V	G oder C oder A	<i>not T/U</i> (V folgt U im Alphabet)
D	G oder A oder T	<i>not C</i> (D folgt C im Alphabet)
N	G oder A oder T oder C	<i>aNy</i>

### A.2.2 Illumina-Adaptersequenzen

Zum Herausfiltern der Adaptersequenzen in den FASTQ-Rohdaten (vgl. Abschnitt 2.2.3.2) wurden dem Programm *Trimmomatic* die folgenden Sequenzen übergeben:

Uni: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

NEXTflex: GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-Index-ATCTCGTATGCCGTCTTCTGCTTG

Die Indices der einzelnen Proben lauteten:

#### Exome

→ Ma-Mel-86A

Replikat 1

ACAGTG

↔ 2

GCCAAT

→ Ma-Mel- <b>86B</b>	Replikat 1	CAGATC
	↔ 2	CTTGTA
→ Ma-Mel- <b>86C</b>	Replikat 1	AGTTCC
	↔ 2	ATGTCA
→ Ma-Mel- <b>86F</b>	Replikat 1	CCGTCC
	↔ 2	GTCCGC
→ INTH- <b>EBV-B</b>	Replikat 1	CGATGT
	↔ 2	TGACCA
<b>Transkriptome</b>		
→ Ma-Mel- <b>86A</b>	Replikat 1	TTAGGC
	↔ 2	ACAGTG
	↔ 3	ATCACG
→ Ma-Mel- <b>86B</b>	Replikat 1	GATCAG
	↔ 2	GCCAAT
	↔ 3	ACTTGA
→ Ma-Mel- <b>86C</b>	Replikat 1	CGTACG
	↔ 2	CTTGTA
	↔ 3	TAGCTT
→ Ma-Mel- <b>86F</b>	Replikat 1	ACTGAT
	↔ 2	CCGTCC
	↔ 3	GGCTAC
→ INTH- <b>EBV-B</b>	Replikat 1	ATTCCT
	↔ 2	GTGAAA
	↔ 3	GTTTCG

### A.2.3 Sequenzen der immunogen mutierten Gene und Proteine

Die mutierten Basen bzw. Aminosäuren in den cDNA- bzw. Proteinsequenzen sind jeweils in Rot dargestellt. Die peptidkodierenden cDNA-Abschnitte und die immunogenen Peptide sind in Blau hervorgehoben. Die Bezeichnungen der gezeigten mutierten Transkriptionsvarianten (v) bzw. Proteinisotypen (i) folgt der empfohlenen Nomenklatur der *Human Genome Variation Society* (den Dunnen & Antonarakis, 2000).

#### A.2.3.1 HERPUD1<sup>mut</sup>

##### cDNA-Sequenz NM\_014685 (*HERPUD1\_v1*):c.481G>T

ATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAGCCCCAACCAGCGCCACCGCGACTTGG  
AGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCCACCTGAGCCGCGTCTACCCCGAGCGTCC  
GCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCTGTTGTTGGATCACCAATGTCTCAGGGACTTGCTT  
CCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTGTGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCA  
ACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCTCAAG  
TGATGGTTTAAGGCAAAGGGAAGTTCTTCGGAACCTTTCTCCCTGGATGGGAAAACATCTCAAGGCCTGAA  
GCTGCCAGCAGGCATTCCAAGGCCTGGGTCTGGTTTCTCCAGTTACACACCCTATGGGTGGCTTCAGCTTT  
CCTGGTTCCAGCAGATATATGCACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTT

---

TGTTCCACCACCAAGTGCACAAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTT  
CCAGCTGAAAACCAGCCTGCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAATT  
TGCGGATGAATGCACAAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGAC  
CTATTCAGCAGCTACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCCTCATG  
GTCATGGGGGCCACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGTTTCCAG  
ACTTCCCAAATGATGGTCTCCTCCTGACGTTGTAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGA  
TCCTGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCC  
TTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCATCG  
CAAACCTGA

### Aminosäuresequenz NP\_055500 (HERPUD1\_i1):p.G161S

MESETEPEPVTLVVKSPNQRHRDLELSGDRGWSVGHKLAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLRDLL  
PKQEKRHVHLVLCNVKSPSKMPEINAKVAESTEPEAGSNRGQYPEDSSSDGLRQREVLRLNLSSPGWENISRPE  
AAQQAFQGLGPGFSSYTPYGLWQLSWFQQIYARQYYMQYLAATAASGAFVPPPSAQEIPVVSAPAPAPIHNQF  
PAENQPANQNAAPQVVVNPGANQNLRMNAQGPIVEEDEINRDWLDWTYSAATFSVFLSILYFYSSLSRFLM  
VMGATVVMYLHHVGFPRPRPVQNFNDGPPPDVVNQDPNNLQEGTDPETEDPNHLPDRDVLGEQTS  
FMSTAWLVFKTFASLLPEGPPAIAN

### A.2.3.2 INSIG1<sup>mut</sup>

#### cDNA-Sequenz NM\_005542 (INSIG1\_v1):c.713C>T

ATGCCAGATTGCACGACCACTTCTGGAGCTGCTCCTGTGCGCACAGCGGAGGCGCGAGGCCCGGAGCCCCCGCGAG  
CCAGCGCCGCGGGGCTGGCGGCCAAGGTTGGGAGATGATCAACGTTTCCGTGTCCGGGCCCTCCCTGCTGGC  
GGCCACGGTGCCCCGACGCTGACCCCGCGCCAGGGGCCGAGTGTGCGATGAGCGGCCCGAGCCCGGC  
AGCCCCTACCCCAACACCTGGCATCATCGCCTGTTGCAGAGGAGCCTCGTGCTCTTCTCGGTTGGGGTGGTCC  
TAGCCCTGGTGCTCAACCTGCTGCAGATCCAGAGGAATGTCCTCTTCCCCGAGGAGGTGATCGCCACCAT  
CTTTTCCCTCCGCCTGGTGGGTCCCTCCCTGCTGCGGGACAGCAGCTGCTGTTGTTGGCCTACTGTACCCCTGT  
ATCGACAGTCACCTCGGAGAACCCCAAAATTTAAGAGAGAATGGGCCAGTGTGATGCGCTGCATAGCAGTTT  
TTGTTGGCATTAAACCACGCCAGTGCTAAATTGGATTTTGCCAATAATGTCCAGCTGTCCCTGACTTTAGCAGC  
CCTATCTTTGGGCCTTTGGTGGACATTTGATCGTTCCAGAAGTGGCCTTGGGCTGGGGATCACCATAGCTTTT  
CTAGCTACGCTGATCACGCAGTTTCTCGTGTATAATGGTGTCTATCAGTATACATCCCAGATTTCTCTATA  
TTCGTTCTTGGCTCCCTTGTATATTTTTCTCAGGAGGCGTCACGGTGGGGAACATAGGACGACAGTTAGCTAT  
GGGTGTTCCCTGAAAAGCCCCATAGTGATTGA

### Aminosäuresequenz NP\_005533 (INSIG1\_i1):p.S238F

MPRLHDFWSCSCAHSARRRGPPRASAAGLAAKVGEMINVSVSGPSLLAAHGAPDADPAPRGRSAAMSGPEPG  
SPYPNTWHHRLQLRSLVLFVSVGVVLAALVNLQIQRNVTLFPEEVIATIFSSAWVPPCCGTAAAVVGLLYPC  
IDSHLGEPHKFKREWASVMRCIAVFGINHASAKLDFANNVQLSLTLAALSGLWTFDRSRSLGLGITIAF  
LATLITQFLVYNGVYQYTFPDLFIYRSWLPCIFFSGGVTVGNIGRQLAMGVPEKPHSD



A.2.3.3 MMS22L<sup>mut</sup>cDNA-Sequenz NM\_198468 (*MMS22L*):c.1310C>T

ATGGAGAACTGTTCTGCTGCATCGACGTTCTGACTGACAGCTTAGAGCTGGAGCTGGGGACGGAATGGTGCA  
AACCTCCTTACTTTTTCTGTGCTGTTGACAACAGAGGAGGAGGAAAACATTTTTCTGGAGAATCCTACCTCTG  
CAGCGGAGCCCTTAAGCGATTGATTTTTGAATCTTGACCCTTACCAACTAATTTTGAAGAAGATACCTTGAA  
ATATTTGGCATTTCAGTGGGTTACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGC  
AACAACTGTACAACCTTGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACACTGCAA  
AGCAGACAATATTAGGCAGCAGTGTGTAATAATTTCTCCATTATGTTAAAGTTTTTCATCTTCAGGTATCTGAAA  
GTACAGAAATGCTGAGAGTCATGTTCCCTGTCCATCCTTATGAGGCTTTGGAGGCTCAGCTTCCCTCAGTGTGTA  
TTGATGAGCTTCATGGATTACTCTTGTATATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATT  
TGTAATCAAACCAGATTAAGCTTTTTCCACCGTCATGGCATTATTACATCTCCACTTGGATATACATTGG  
CTGGTGCTAGAAATCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCAGTTTATGAATC  
TGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGAAACTCTCCTTTGTGATTTAATAAG  
CCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTCTGAATCATTAATGAGTGACCAGTGTCCATGTTTATGC  
ATTAAGAATTATGGGTTCTACTTATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTT  
GGAACTGGTTGAATAAACTACTTAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCCTGT  
AATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATCATTTTACAAGTTTGAT  
CGCCATGGAGTACCAGATGAAATGAGAAAAGTGAATCAAATTGGAACCTTGTAGAAGAAGTCTGAAAAAGT  
CCATCAGTGTTTCAGGGTGTCAATTCTAGAAGAACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGA  
TTTTCTGGGAGCCAAACATTGCAATTGTTACCATTTTATGGGAAATATTATAGTAAGAACCTGAATAGTTCTCTC  
AGTATTTCTTGGCTTCCTTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTGAAATGGTGA  
AGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGTAGTTATACTATTTTTCTTTGTAT  
TCTGGCAAAAAGTTGTTAAAAAAGCAATGAAGAGCAATGGCCCTCATCCTTGAAACAAGTCAAAGGAAGAATA  
TATTCAAAAATCCATCAAAAAAGAATGGAAGAATAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTTC  
TACTGTTAGCAGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCTCAAGCC  
TGCTTTTGTAAACGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCTTCTGATGTATGCCAGAAA  
AATCTGGACATTGGTGTTTTGGCTGAGAAAATTTTCATGTGCTTTCCGGGAGAAAGCAAAGGAATTCTTGGTGT  
CTAAGAATGAGGAAATGGTACAGAGACAGACTATCTGGACCCTTCTTCCATATACATTGATGGTGTTCAGA  
AGTGTTTGAGACCAGCTATTGCTTGTATCCTTCCCATGAAAACTGCTTAATGATGGATTTAGTATGCTTCTG  
CGAGCATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCTTACAAGCTGTTCTGGCCAGAATCAGGAGTA  
TGCATCAACAATTGTGTCAGGAACCTCAAAGGGACAATGTGGACCTATTTGTACAGTCTTCATTATCGGCTAA  
AGAGCGCCACCTTGCTGCAGTTGCCAGTGCAGTGTGGAGACATTTCTTTTCATTTTTGAAGAGTCAGAGAATG  
TCACAGGTAGTGCCTTTCTCACAACCTTGGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAA  
GCACAGCTCCATCAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCAT  
CTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTGAAGCACTTTCTCAT  
TCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTGTTGTGTTTTGCAAATGTATATTA  
AACCTCTCTGGCCCTGATGATTTGCTCATAGATAAAAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACA  
GTTGGTCAAACCTGACAAGATTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGTTGAA  
TATTTATCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTTACGGGA

ACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGAATACCTTGGTGAAGTATTTAAAATA  
TATTAAGCCTTATTTGGGAAAAAAGTTTTTCAGTGCAGGGCTGCAGCTGACTTATGGAATGATGGGAATTCTT  
GTGAAATCATGGGCACAAATCTTTGCCACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTAC  
TGCTGCCACATGCAGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT  
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAATCCGAATGCCTATTTGAATCAATTG  
CTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTCACCATATGTTTCAGATCTTGGACAAC  
ATCCTGTTTTGCTGGCATTGAGAAAACACAGCCACTATTCCACCAATATCATCTCTAAAGAAATGCATTGTGCA  
AGTCATAAGGAAATCCTACCTTGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTC  
ATCCTCCAACCTTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTAAAAT  
GCTTGGTGTAGTACAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAATACATGGTAAAAGCCTG  
CCAAGTGGGGTCAGAAGAAGAACCTTCTCCCAGCTGACTTCTGTGTTTAGGCAGTTTATCCAGGATTATGGT  
ATGAGGTACTATTACCAGGTTTACAGCATTTTAGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACT  
TGATTTCTACCCTTACTCAGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAG  
GGAAGCCTATAGCAAACCTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACTGGAAAATGAT  
AATACTTAA

#### **Aminosäuresequenz NP\_940870 (MMS22L):p.S437F**

MENCSAASTFLTDSLELELGTEWCKPPYFSCAVDNRGGGKHFSGESYLCSGALKRLILNLDPLPTNFEEEDTLE  
IFGIQWVTETALVNSSRELFLHFRQQLYNLETLLQSSCDFGKVSTLHCKADNIRQQCVLFLHYVKVIFRYYLK  
VQNAESHVPVHPYEALQPSVLIDELHGLLLYIGHLSELPSVNI GAFVNQNIKLFPPSWHLLHLHLDIHW  
LVLEILYMLGEKLVVYGHQFMNLASDNLTNISLFEHCETLLCDLISLSLNRDYDKVRSSSELSMSDQCPCLC  
IKELWVLLIHLLDHRKWFVSESFWNWLNKLLKTLLEKSSDRRRSSMPVIQSRDPLGFSWWIITHVASFYKFD  
RHGVPDEMVKVESNWNFVEELLKKSISVQGVILEEQLRMYLHCCLTLCDFWEPNIAIVTILWEYYSKNLNSFF  
SISWLPFKGLANTMKSPLSMLEMVKTCCCDKQDQELYKSSSSYTIIFLCILAKVVKKAMKSNPHPWKQVKGRI  
YSKFHQKRMEELTEVGLQNFFSLFLLAAVAEVEDVASHVLDLLNFKPAFVTSQRALIWKGHMAFLMYAQK  
NLDIGVLAEKFSCAFREKAKEFLVSKNEEMVQRQTIWTLLSIYIDGVQEVFETSICYLPSHEKLLNDGFSMLL  
RACRESELRTVLSFLQAVLARIRSMHQQLCQELQRDNVDFVQSSLSAKERHLAAVASALWRHFFSFLKSQRM  
SQVVPFSQLADAAADFTLLAMDMPSTAPSDFPQPVISIIQLFGWDDIICPQVVARYLSHVLQNSTLCEALSH  
SGYVSFQALTVRSWIRCVLQMYIKNLSGPDDLIDKNLEEAVEKEYMKQLVKLTRLLFNLSEVKSIFSKAQVE  
YLSISEDPKKALVRFVEAVGVTYGNVQTLSDKSAMVTKSLEYLGEVLKYIKPYLGKKVFSAGLQLTYGMMGIL  
VKSQAQIFATSKAQKLLFRIIDCLLLPHAVLQKEKELPAPMLSAIQKSLPLYLQGMCI VCCQSQNPAYLNQL  
LGNVIEQYIGRFLPASPYVSDLGQHPVLLALRNTATIPPISSLKKCIVQVIRKSYLEYKSSPPPRLASILAF  
ILQLFKETNTDIYEVELLLPGILKCLVLVSEPQVKRLATENLQYMKACQVGSEEEPPSSQLTSVFRQFIQDYG  
MRYYYQVYSILETVATLDQVVIHLISTLTQSLKDSEQKWGLGRNIAQREAYSKLLSHLGQMGQDEMQRLEND  
NT

**A.2.3.4 PRDM10<sup>mut</sup>****cDNA-Sequenz NM\_199437 (*PRDM10\_v2*):c.3149C>T**

ATGGATTCCAAAGATGAAAGCTCGCATGTGTGGCCGACATCTGCAGAGCATGAACAGAATGCCGCACAGGTGC  
ACTTTGTTCCGGACACAGGAACAGTGGCTCAGATTGTCTATACCGATGACCAGGTTCCGCCCCACAGCAGGT  
GGTGTACACGGCAGATGGTGCCTCCTACACATCAGTGGATGGTCCAGAGCACACGCTGGTGTACATCCACCCG  
GTGGAAGCTGCACAGACTCTGTTTACAGACCCCGCCAGGTAGCTTATGTCCAACAGGATGCTACAGCTCAGC  
AGGCCTCATTGCCAGTTCATAACCAGGTGCTGCCTTCCATCGAGAGTGTAGATGGGTCCGACCCTTTGGCAAC  
TCTGCAGACCCCTCTAGGCAGACTGGAGGCCAAAGAGGAAGAGGATGAGGATGAGGACGAGGACACTGAGGAA  
GATGAGGAAGAAGACGGTGAGGACACGGATCTGGATGACTGGGAGCCAGACCCGCCCCGGCCCTTCGACCCAC  
ACGACTTGTGGTGTGAGGAGTGAATAACGCGCATGCTTCAGTGTGTCCGAAGCACGGCCCTTGCACCCGAT  
CCCCAACCGGCCGGTGTCTACCCGGGCCAGGGCGAGCCTCCCCCTGGTGTCTACATAGACAGGTTTCTGGGC  
GGGGTGTCTCCAAGCGGCGCATCCCCAAGCGCACCCAGTTTGGCCCCGTGGAGGGGCCTCTCGTCAGGGGCT  
CGGAGCTGAAAGACTGTTACATTCACCTCAAGGTTTCTCTTGATAAAGGGGACAGGAAAGAAAGGGATTTACA  
TGAAGACCTATGGTTTGAGTTGTCTGATGAGACGCTTTGTAAGTGGATGATGTTTGTACGGCCAGCCAGAAT  
CACCTGGAGCAGAACCTGGTGGCTTACCAGTATGGCCACCATGTGTATTATACAACCATAAAAAATGTGGAGC  
CCAAGCAGGAACTGAAGGTGTGGTATGCCGCATCCTATGCTGAGTTCGTGAACCAGAAAATTCATGACATTTT  
TGAGGAAGAAAGGAAAGTTCTTCGAGAGCAAGAGAAGAATTGGCCCTGCTATGAATGTAACCGCCGATTTATA  
AGCTCGGAGCAGTTGCAACAGCATCTCAATTCTCATGATGAGAACTAGATGTGTTTAGCAGAACAAGAGGCA  
GAGGAAGGGGACGAGGCAAGAGGCGATTCCGGTCCAGGTCGACGGCCGGGGCGTCCTCCAAAATTTATCCGCCT  
GGAAATCACCAGCGAAAATGGGGAAAAGAGTGACGATGGGACACAGGACTTGCTACATTTTCCACAAAAGGAG  
CAATTTGATGAGGCTGAACCAGCCACTCTGAATGGGCTGGATCAACCAGAACAGACCACTATCCCAATCCCTC  
AGCTGCCACAGGAAACCCAGTCTTCCCTGGAACATGAACCAGAACTCACACCCTGCACCTGCAGCCGCAGCA  
TGAAGAGAGCGTGGTGGCCACCCAGAGCACGCTGACAGCCGACGACATGCGCAGAGCCAAGCGCATCCGAAAT  
GCAGCTCTTCAGCATCTGTTTATTTCGGAAGTCTTCCGGCCTTTTAAATGCTTGCAGTGTGGAAAGGCCTTCC  
GGGAAAAGGACAAACTGGACCAGCACTTACGCTTCCATGGGCGGGAGGGGAACTGCCACTGACCTGTGATCT  
CTGTAACAAGGGCTTCATCAGCAGCACATCCTTGGAGAGCCACATGAAGCTCCACTCAGACCAGAAGACTTAC  
TCTTGCATTTTTTGGCCAGAATCCTTTGACCGCCTTGATTTGTTGAAAGATCATGTGGCCATTCATATCAATG  
ATGGCTACTTCACCTGCCAACTTGTAAGAAACGGTTCCAGATTTTATCCAGGTGAAAAACACGTGCGCAG  
CTTCCACTCAGAAAAGATCTACCAGTGACAGAGTGTGACAAGGCCTTCTGTGCCCCGATAAACTGCGACTC  
CACATGCTCCGGCATTCCGACCGCAAAGACTTCCGTGTTCCACCTGTGGGAAGCAATTTAAGCGAAAAGACA  
AACTACGGGAACACATGCAGAGGATGCATAATCCTGAGAGGGAGGCCAAGAAAGCCGACCGCATCAGCCGCTC  
CAAGACGTTCAAGCCCCGCATCACGTCCACAGACTACGACAGCTTACGTTCAAGTGCCGCTGTGCATGATG  
GGCTTCCGGCGGCGCGGCATGCTGGTAAATCACTTATCGAAGAGACACCCAGACATGAAGATAGAAGAGGTGC  
CAGAGTTAACTTACCCATCATAAAACCCAATCGTGATTACTTTTTGTCAGTATTGCGATAAGGTTTATAAAAG  
TGCCAGCAAGCGCAAAGCCCACATTCTGAAGAACCACCCAGGAGCAGAGCTCCCACCGAGCATTCCGGAAGCTC  
CGACCCGCTGGTCTGGAGAGCCAGACCCCATGCTGAGCACACACACCCAGCTGACGGGCACCATCGCCACCC  
CTCCCGTCTGCTGTCCCCACTGCTCCAAGCAGTACAGCAGCAAGACCAAGATGGTCCAGCACATTCGAAAGAA  
GCATCCAGAGTTCCGCCAGCTCTCCAACACCATAACACACCACTGACGACAGCTGTGATCAGTGCCACCCCA  
GCGGTTTTGACTACAGACAGCGCCACTGGAGAGACTGTGGTGACGACGGACCTGCTCACCCAAGCAATGACAG

AACTGTCCCAGACCTTAACGACAGACTACCGAACGCCACAAGGGGATTACCAGAGAATTCAGTACATCCCTGT  
GTCGCAGTCGGCGTCTGGCCTCCAGCAGCCTCAGCACATACAGCTGCAAGTGGTTCAAGTGGCCTCGGCCACT  
TCCCCTACCAGTCACAGCAGTCCACTGTGGATGTTGGCCAGCTCCATGATCCTCAGCCCTACCCCAGCAGC  
CCATCCAGGTGCAGCACATCCAGGTGAGCGAGCCTACCGCCTCGGCCCCGTCTCCGCCAGGTATCTGGGCA  
GCCGTTGAGTCCCTCAGCCCAGCAGGCTCAGCAGGGGCTCAGCCCCTCCCACATCCAGGGCAGTTCTTCCACA  
CAGGGGCAGGCTCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGAATTCCTCTGTGCAGCACACGTACCTGCCAGTG  
CTTGGAAATTCCTTCGTTGGCTATTTCATCTGAGATTCAAATGATGACGCTTCCCTCCGGGTCAGTTTGTGATTAC  
AGACAGTGGTGTGGCAACTCCAGTTACTACTGGCCAGGTGAAGGCGGTTACTTCGGGTCATTATGTGTTATCA  
GAAAGTCAATCAGAATTGGAAGAAAAGCAAACCTTCTGCCCTCTCTGGTGGAGTCCAGGTCGAGCCACCTGCAC  
ACAGTGAAGTCCCTGGACCCCCAGACCAACAGCCAACAGCAGACCACACAGTACATCATCACCACCACCACCAA  
CGGGAACGGAAGCAGCGAAGTGCATATCACCAAACCATGA

**Aminosäuresequenz NP\_955469 (PRDM10\_i2):p.S1050F**

MDSKDESSHVWPTS AEHEQNAAQVHFV PDTGTVAQIVYTDQVRPPQVVYTDAGASYTSVDGPEHTLVYIHP  
VEAAQTLFTDPGQVAVYVQQDATAQQASLPVHNQVLP SIESVDGSDPLATLQTPLGRLEAKEEED EDEDTEE  
DEEEDGEDTDLDDWEPDPPRPFDPHDLWCEECNNAHASVCPKHG PLHP IPNRPVLTRARASLPVLVYIDRFLG  
GVFSKRRIPKRTQFGPVEGPLVRGSELKDCYIHLKVS LDKGDRKERDLHEDLWFELSD ETL CNWMMFVRPAQN  
HLEQNLVAYQYGHVYYTTIKNVEPKQELKVWYAASYAEFVNQKI HDI SEEERKVLREQKNWPCYECNRRFI  
SSEQLQHLNSHDEKLDVFSRTRGRGRGRGRKRRFGPGRR PGRPPK FIRLEITSENGEKSD DGTQDLLHFPTKE  
QFDEAEPATLNGLDQPEQTTIPIQLPQETQSSLEHEPETHTLHLQPQHEESV VPTQSTLTADDMRRAKRIRN  
AALQH L F IRKSF RPFKCLQCGKAFREKDKLDQHLRFHGREGN CPLTCDL CNKGFISST SLESHMKLHSDQKTY  
SCIFCPESFDRLDLLKDHVAIHINDGYFTCPTCKKRFPDFIQVKKHVRSFHSEKIYQCTE CDKAFCRPKLRL  
HMLRHSDRKDFLCSTCGKQFKRDKLREHMQRMHNP EREAKKADRISRSKTFKPRITSTDYDSFTFKRCLCMM  
GFRRRGMLVNHLSKRHPDMKIEEVPELTLPIIKPNR DYFCQYCDKVYKSASKRKAHILKNHPGAE LPPSIRKL  
RPAGPGE P DMLSTHTQLTG TIATPPVCCPHCSKQYSSKTKMVQH I RKKHPEFAQLSNTIHTPLTTAVISATP  
AVLTDSATGETVVTTDLLTQAMTELSQTLTTDYRTPQG DYQRIQYIPVSQSASGLQQPQH IQLQVQVASAT  
SPHQSQQSTVDVGQLHDPQPYPQHAIQVQHIQVSEPTASAPSSAQVSGQPLSPSAQQAQQGLSPSHIQSSST  
QGQALQQQQQQQNSVQHTYLP SAWNFRGYSSEIQM MTLPPGQFVITDSGVATPVTTGQVKAVTSGHYVLS  
ESQSELEEKQTSALSGGVQVEPPAHSDSLDPQ TNSQQQTQYIITTTTNGNGSSEVHITKP

# A.3 Anhang zum bioinformatischen Teil

## A.3.1 Workflow

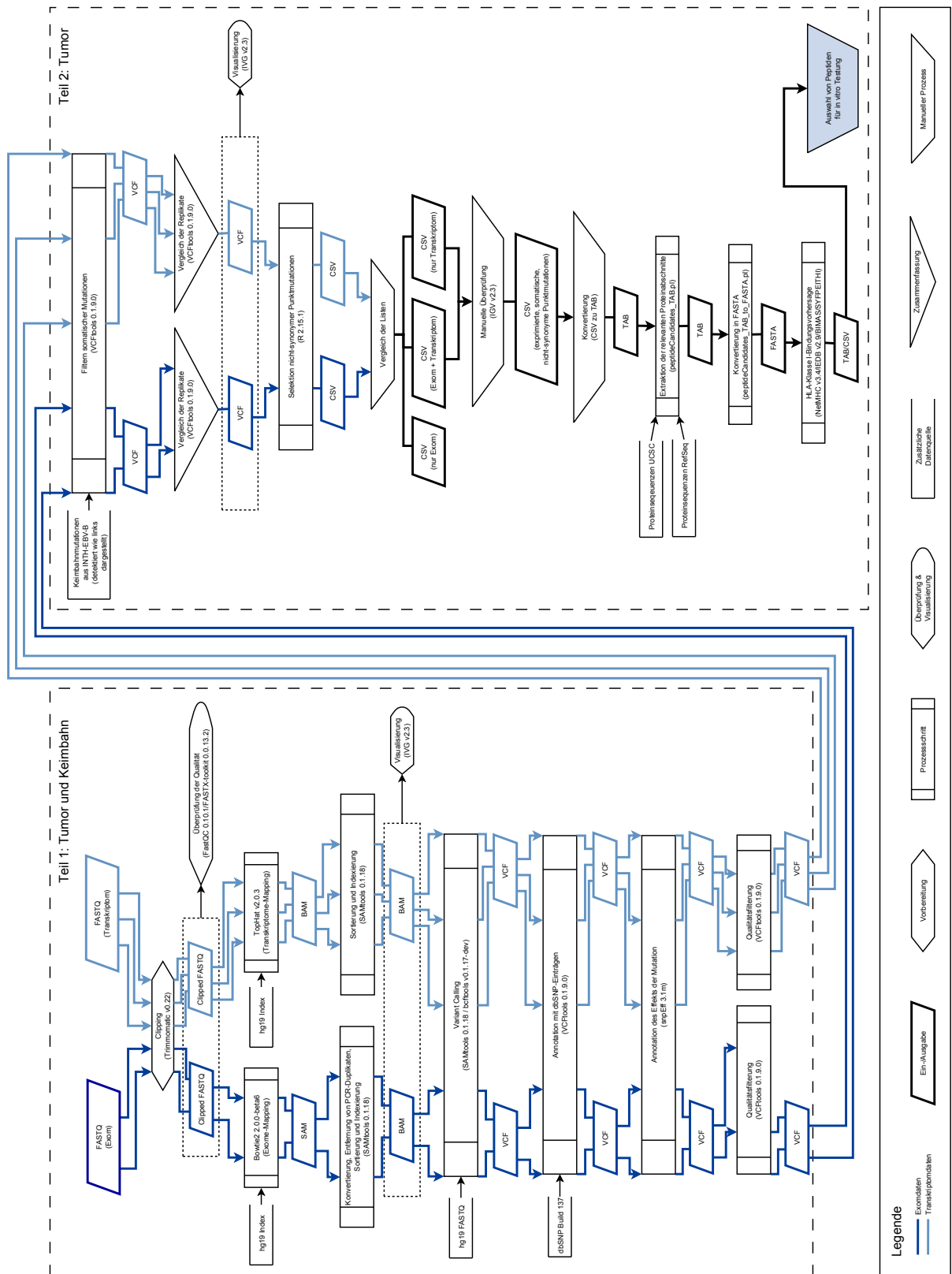


Abb. A.3: Rechnergestützter Workflow zur Lokalisierung von potentiell immunogenen, somatischen Punktmutationen in Ma-Mel-86. Der Teil 1 des Workflows wurde parallel sowohl für die Sequenzierungen der Tumor-DNA und -RNA durchgeführt als auch für die Sequenzierungen der Keimbahnkontrolle INTH-EBV-B. Im zweiten Teil wurden Keimbahnmutationen herausgefiltert und entsprechend nur noch die Daten der Tumorproben weiterprozessiert.

---

## A.3.2 Eigene *Perl*-Skripte

### A.3.2.1 mappedReads\_lengthDistribution.pl

```
#!/usr/bin/perl

# Author: Barbara Roeder (29.04.2014)
# Determination of mean mapped read length and number of mapped nucleotides

use strict;
use warnings;
use Getopt::Std;

#####
my $line_count = 0;
my $length_sum = 0;
my %opts;
#####

getopts('h', \%opts);
sub HELP_MESSAGE {
    die "\n$0:
Usage: samtools view <in.bam> | perl mappedRead_LengthDistribution.pl
-h print usage message and exit
\n";
}

foreach my $opt (keys(%opts)){
    if ($opt eq 'h') {HELP_MESSAGE();}
    else {die "ERROR: Unknown parameter.\n\n"; HELP_MESSAGE();}
}

while (<STDIN>) {
    my $line = $_;
    $line_count++;
    my @data = split /\t/, $line;
    my $length = length $data[10];
    $length_sum = $length_sum + $length;
}

my $length_mean = $length_sum / $line_count;
print STDOUT "Mean read length: $length_mean\n";
print STDOUT "Mapped NT: $length_sum\n";

exit;
```

### A.3.2.2 compareRefSeqUCSC.pl

```
#!/usr/bin/perl

# Author: Barbara Roeder (22.04.2013)
# Protein reference sequences provided by RefSeq (NCBI) and UCSC differ. Therefore,
# protein sequences of both reference databases were downloaded and a comparison and
# merge of sequences was performed.

use strict;
use warnings;
use Getopt::Std;

#####
my %opts;
my $ncbi_file;
my $ucsc_file;
my $output_file;
my @refseq;
my @ucsc;
my @trans_ucsc;
my $row_count = 0;
my %refseq_hash;
#####

getopts('n:u:o:h', \%opts);
foreach my $opt (keys(%opts)){
    if ($opt eq 'n') {$ncbi_file = $opts{'n'};}
```

```

elsif ($opt eq 'u') {$ucsc_file = $opts{'u'}};
elsif ($opt eq 'o') {$output_file = $opts{'o'}};
elsif ($opt eq 'h') {HELP_MESSAGE();}
else {die "ERROR: Unknown parameter.\n";}
}

sub HELP_MESSAGE {
    die "\n$0:
perl compareRefSeqUCSC.pl [-h] -n <refseq.fasta> -u <ucsc.tab> -o <output.tab>
-n FILE RefSeq sequences in FASTA format
-u FILE UCSC sequences with annotation in TAB format
-o FILE output file to create (TAB format)
-h print usage message and exit
\n";
};

open (REFSEQ, "<$ncbi_file") || die "ERROR: Could not open RefSeq file $ncbi_file.\n";
open (UCSC, "<$ucsc_file") || die "ERROR: Could not open UCSC file $ucsc_file.\n";
@refseq = <REFSEQ>;
@ucsc = <UCSC>;
close(REFSEQ);
close(UCSC);

# collect RefSeq sequences from FASTA file
foreach my $row (@refseq) {
    if ($row =~ /^>(NM_\d+)\s/) {
        chomp $refseq[($row_count + 1)];
        $refseq_hash{$1} = $refseq[($row_count + 1)];
        $row_count++;
    } else {
        $row_count++;
        next;
    }
}

open (OUT, ">$output_file");
print OUT "#Transcript_ID\tUCSC_ID\tadditional_ID\tisRefSeq\tSequence\n";

my $ucsc_hits = join ("\t", @ucsc);

# compare RefSeq sequences and remove redundancies
ID: foreach my $refseq_id (keys(%refseq_hash)) {
    my $refseq_printer = 0;

    # no RefSeq ID match in UCSC sequences
    if (index($ucsc_hits, $refseq_id) == -1) {
        print OUT "$refseq_id\t-\t-\t1\t$refseq_hash{$refseq_id}\n";
        next ID;
    }
    # RefSeq ID match => compare sequences
    } else {
        foreach my $line (@ucsc) {
            chomp $line;
            if ($line !~ /^#/) {
                @trans_ucsc = split /\t/, $line;
                if ($refseq_id eq $trans_ucsc[6] || index($trans_ucsc[3], $refseq_id.".") != ...
                    ... -1) {
                    # RefSeq is identical to UCSC => merge entries
                    if($refseq_hash{$refseq_id} eq $trans_ucsc[1]) {
                        print OUT "$refseq_id\t$trans_ucsc[0]\t$trans_ucsc[3]\t1\t$trans_ucsc[1]\n ...
                            ... ";
                        next;
                    }
                    # RefSeq is NOT identical to UCSC => keep separate entries
                } elsif ($refseq_hash{$refseq_id} ne $trans_ucsc[1]) {
                    print OUT "$refseq_id\t$trans_ucsc[0]\t$trans_ucsc[3]\t0\t$trans_ucsc[1]\n ...
                            ... ";
                    $refseq_printer++;
                    next;
                }
            }
        }
    }
}

# print RefSeq sequence if not identical to UCSC
if ($refseq_printer > 0) {
    print OUT "$refseq_id\t-\t-\t1\t$refseq_hash{$refseq_id}\n";
}

```

```

    }
  }
}

close(OUT);
exit;

```

### A.3.2.3 peptideCandidates\_TAB.pl

```

#!/usr/bin/perl

# Author: Barbara Roeder (22.04.2013)
# Perl program to extract relevant sequence section around mutation (option -a)

use strict;
use warnings;
use Getopt::Std;

#####
my %opts;
my $sequence_file;
my $mut_list;
my $output_file;
my $around;
my @seq;
my @mut;
my $row_count;
my @transcript;
my @sequence;
my @mut_sequence;
my @wt_sequence;
my $mut_sequence;
my @seq_mut;
#####

getopts('t:m:o:a:h', \%opts);
foreach my $opt (keys(%opts)){
  if ($opt eq 't') {$sequence_file = $opts{'t'};}
  elsif ($opt eq 'm') {$mut_list = $opts{'m'};}
  elsif ($opt eq 'o') {$output_file = $opts{'o'};}
  elsif ($opt eq 'a') {$around = $opts{'a'};}
  elsif ($opt eq 'h') {HELP_MESSAGE();}
  else {die "ERROR: Unknown parameter.\n";}
}

sub HELP_MESSAGE {
  die "\n$0:
perl peptideCandidates_TAB.pl [-h] -t <sequences.tab> -m <mutations.tab> -o <output. ...
... tab> -a INT
-t FILE wild-type sequences in TAB format as generated with compareRefSeqUCSC.pl
-m File mutations annotated with RefSeq transcript ID (NM_...), position of mutation ...
... (e.g. 600)
  and AA exchange (e.g. V600E) in TAB format
-o FILE output file to create (tab-separated)
-a INT print 2 x INT - 1 amino acids centered around mutation
  (INT indicates length and number of peptides containing mutation)
-h print usage message and exit
\n";
}

open(MUT, "<$mut_list") || die "ERROR: Could not open mutation list $mut_list.\n";
open(SEQ, "<$sequence_file") || die "ERROR: Could not open TAB file $sequence_file.\n" ...
... ;
@seq = <SEQ>;
@mut = <MUT>;
close (SEQ);
close (MUT);

my $seq_tmp = $sequence_file . "_mutated";
open(TMP, ">$seq_tmp") || die "ERROR: Could not create mutated TAB file $seq_tmp.\n";

# write tab-separated file with mutated protein sequences
ROW: foreach my $row (@seq) {
  chomp $row;
  if ($row =~ /~#/) {

```



```

print TMP "Transcript_ID\tUCSC_ID\tadditional_ID\tisRefSeq\twt_sequence\ ...
... tmut_sequence\n";
$row_count++;
next;
} else {
    @transcript = split /\t/, $row;
    @sequence = split //, $transcript[4];
LINE:   foreach my $line (@mut) {
        my @mutation = split /\t/, $line;
        my @AA_change = split /\d+/, $mutation[2];
        chomp $AA_change[1];
        if ($transcript[0] eq $mutation[0]) {
            if (scalar @sequence >= $mutation[1] && $sequence[($mutation[1]-1)] eq ...
                ... $AA_change[0]) {
                $sequence[($mutation[1]-1)] = $AA_change[1];
            } else {
                next LINE;
            }
        }
    }
}
$mut_sequence = join(" ", @sequence);
chomp $mut_sequence;
$transcript[5] = $mut_sequence;
$row = join ("\t", @transcript);
print TMP "$row\n";
}

close (TMP);

open (TAB_MUT, "<$seq_tmp");
@seq_mut = <TAB_MUT>;
close (TAB_MUT);

# now extract mutated sequences around mutation and print into $output_file
open (OUT, ">$output_file") || die "ERROR: Could not create output file $output_file.\ ...
... n";
print OUT "#Transcript_ID\tUCSC_ID\tadditional_ID\tisRefSeq\tAA_postion\tAA_change\ ...
... tmut_sequence\twt_sequence\n";

foreach my $line (@mut) {
    chomp $line;
    my @mutation = split /\t/, $line; # columns: Transkript, AA position, AA change
    my @AA_change = split /\d+/, $mutation[2];

    foreach my $row (@seq_mut) {
        if ($row =~ /^#/) {
            next; # skip header
        } else {
            chomp $row;
            @transcript = split /\t/, $row;
            if ($transcript[0] eq $mutation[0]) {
                @wt_sequence = split //, $transcript[4];
                @mut_sequence = split //, $transcript[5];

                # peptide sequence well within protein sequence
                if ($around <= ($mutation[1]-1) && ($mutation[1] + $around -1) <= scalar ...
                    ... @wt_sequence) {
                    if ($AA_change[0] eq $wt_sequence[($mutation[1]-1)] && $AA_change[1] eq ...
                        ... $mut_sequence[($mutation[1]-1)]) {
                        print OUT "$transcript[0]\t$transcript[1]\t$transcript[2]\t$transcript[3]\ ...
                            ... t$mutation[1]\t$mutation[2]\t";
                        for (my $i = $mutation[1]-$around; $i < ($mutation[1] + $around -1) ; $i ...
                            ... ++) {
                            print OUT "$mut_sequence[$i]";
                        }
                        print OUT "\t";
                        for (my $i = $mutation[1]-$around; $i < ($mutation[1] + $around -1) ; $i ...
                            ... ++) {
                            print OUT "$wt_sequence[$i]";
                        }
                        print OUT "\n";
                    }
                }

                # peptide sequence at the very beginning of protein sequence
            } elsif ($around > ($mutation[1] - 1)) {

```

```

if ($AA_change[0] eq $wt_sequence[($mutation[1]-1)] && $AA_change[1] eq ...
... $mut_sequence[($mutation[1]-1)]) {
# no restriction at end
if (($mutation[1] + $around - 1) <= scalar @wt_sequence) {
print OUT "$transcript[0]\t$transcript[1]\t$transcript[2]\t$transcript ...
... [3]\t$mutation[1]\t$mutation[2]\t";
for (my $i = 0; $i < ($mutation[1] + $around - 1) ; $i++) {
print OUT "$mut_sequence[$i]";
}
print OUT "\t";
for (my $i = 0; $i < ($mutation[1] + $around - 1) ; $i++) {
print OUT "$wt_sequence[$i]";
}
print OUT "\n";
# restricted at end
} elsif (($mutation[1] + $around - 1) > scalar $mut_sequence) {
for (my $i = 0; $i < scalar @mut_sequence ; $i++) {
print OUT "$mut_sequence[$i]";
}
print OUT "\t";
for (my $i = 0; $i < scalar @wt_sequence ; $i++) {
print OUT "$wt_sequence[$i]";
}
print OUT "\n";
}
}

# peptide sequence at the end of protein sequence
} elsif (($mutation[1] + $around - 1) > scalar @mut_sequence && $mutation [1] ...
... <= scalar @mut_sequence) {
if ($AA_change[0] eq $wt_sequence[($mutation[1]-1)] && $AA_change[1] eq ...
... $mut_sequence[($mutation[1]-1)]) {
# no restriction at beginning
if ($around <= ($mutation[1]-1)) {
print OUT "$transcript[0]\t$transcript[1]\t$transcript[2]\t$transcript ...
... [3]\t$mutation[1]\t$mutation[2]\t";
for (my $i = $mutation[1]-$around; $i < scalar @mut_sequence ; $i++) {
print OUT "$mut_sequence[$i]";
}
print OUT "\t";
for (my $i = $mutation[1]-$around; $i < scalar @wt_sequence ; $i++) {
print OUT "$wt_sequence[$i]";
}
}
print OUT "\n";
}
} else { next; }
}
}
}

unlink ($seq_tmp);
exit;

```

#### A.3.2.4 peptideCandidates\_TAB\_to\_FASTA.pl

```

#!/usr/bin/perl

# Author: Barbara Roeder (26.04.2013)
# Convert tab-separated output of "peptideCandidates_TAB.pl" into FASTA format
# and merge redundant protein fragments

use strict;
use warnings;
use Getopt::Std;

#####
my %opts;
my $input_file;
my $output_file;
my @pep_list;
my @row;
my %pep_hash;
my $name_mut;

```

```

my $name_wt;
my $names_combined;
#####

getopts('i:o:h', \%opts);
foreach my $opt (keys(%opts)){
  if ($opt eq 'i') {$input_file = $opts{'i'}};
  elsif ($opt eq 'o') {$output_file = $opts{'o'}};
  elsif ($opt eq 'h') {HELP_MESSAGE();}
  else {die "ERROR: Unknown parameter.\n";}
}

sub HELP_MESSAGE {
  die "\n$0:
  perl peptideCandidates_TAB_to_FASTA.pl [-h] -i <input.tab> -o <output.fasta>
  -i FILE tab-separated output file of peptideCandidates_TAB.pl
  -o FILE output to write in FASTA format
  -h print usage message and exit
  \n";
}

open (LIST, "<$input_file") || die "ERROR: Could not open input file $input_file.\n";
@pep_list = <LIST>;
close(LIST);

open(OUT, ">$output_file") || die "ERROR: Could not write output file $output_file.\n" ...
...;

foreach my $line (@pep_list) {
  next if $line =~ /^#/;

  chomp $line;
  @row = split /\t/, $line;
  $name_mut = join("_", $row[0], $row[1], $row[2], "mut");
  $name_wt = join("_", $row[0], $row[1], $row[2], "wt");

  push @{$pep_hash{$row[6]}}, $name_mut;
  push @{$pep_hash{$row[7]}}, $name_wt;
}

foreach my $key (sort keys %pep_hash) {
  $names_combined = join(":", @{$pep_hash{$key}});
  print OUT ">$names_combined\n$key\n\n";
}

close(OUT);
exit;

```

### A.3.2.5 adjust\_zygosity\_stringent.pl

```

#!/usr/bin/perl

# Author: Barbara Roeder (10.04.2013)
# Adjustment of zygosity in GT-Tag of vcf-file according to the the read counts ...
... reported in DP4
# User-defined threshold for homozygous or heterozygous variants (options -t and -s)

use strict;
use warnings;
use Getopt::Std;

#####
my %opts;
my $change_count = 0;
my $no_change_count = 0;
my $line_count = 0;
my $het_count = 0;
my $hom_count = 0;
my $remove_count = 0;
my $input_file = undef;
my $output_file = undef;
my $threshold = 0.75;
my $stringency = 0.1;
#####

```

```

getopts('i:o:t:s:h', \%opts);
foreach my $opt (keys(%opts)){
    if ($opt eq 'i') {$input_file = $opts{'i'}};
    elsif ($opt eq 'o') {$output_file = $opts{'o'}};
    elsif ($opt eq 't') {$threshold = $opts{t}};
    elsif ($opt eq 's') {$stringency = ($opts{s}/100)};
    elsif ($opt eq 'h') {HELP_MESSAGE()}
    else {die "ERROR: Unknown parameter.\n"}
}

sub HELP_MESSAGE {
    die "\n$0:
perl adjust_zygosity.pl [-h] -i <input.vcf> -o <output.vcf> [-t FLOAT -s INT]
-i FILE vcf input file to adjust
-o FILE output file to create with adjusted zygosity (in vcf format)
-t FLOAT threshold for homozygosity between 0 and 1 (i.e. only if
    more percent of the mapped reads differ from the reference
    the variant will be called homozygous) [0.75]
-s FLOAT threshold for stringency concerning heterozygosity (i.e. only
    if 50% +/- INT % of mapped reads differ from reference
    the variant will be called heterozygous) [10]
-h print usage message and exit
\n";
}

die "ERROR: Threshold (option -t) has to be between 0 and 1.\n" if ($threshold > 1 or ...
... $threshold < 0);

print "\nYour threshold: $threshold\n";

my $het_min=(0.5-$stringency)*100;
my $het_max=(0.5+$stringency)*100;
print "Heterozygous variants called between $het_min% and $het_max%.\n";

open(VCF,"<$input_file") || die "ERROR: Could not open input file $input_file.\n";
open(NEW_VCF, ">$output_file") || die "ERROR: Could not create output file ...
... $output_file.\n";

LINE: foreach my $line (<VCF>) {
    $line_count++;
    if ($line =~ /^#/){
        print NEW_VCF "$line";
        next LINE;
    }

    my @data = split /\t/, $line;
    my @info = split /;/, $data[7];
    my @DP = grep ( /DP4\=(\d+),(\d+),(\d+),(\d+)/ , @info);
    my @gt = split /:/, $data[9];

    # split DP4 and calculate percentage of reads supporting mutation
    (my $ref_for, my $ref_rev, my $alt_for, my $alt_rev) = ($DP[0] =~ /DP4\=(\d+),(\d+) ...
    ... ,(\d+),(\d+)/);
    my $ratio = ($alt_for + $alt_rev) / ($alt_for + $alt_rev + $ref_for + $ref_rev);

    # adjust zygosity, if necessary
    my $gt_before = $gt[0];

    if ($ratio >= $threshold && $gt[0] ne "1/1") {
        $gt[0] = "1/1";
        $change_count++;
        $hom_count++;
    } elsif ($ratio >= $threshold && $gt[0] eq "1/1") {
        $no_change_count++;
        $hom_count++;
    } elsif ($ratio < $threshold && $ratio <= 0.5 + $stringency && $ratio >= 0.5 - ...
    ... $stringency && $gt[0] eq "0/1") {
        $no_change_count++;
        $het_count++;
    } elsif ($ratio < $threshold && $ratio <= 0.5 + $stringency && $ratio >= 0.5 - ...
    ... $stringency && $gt[0] ne "0/1") {
        $gt[0] = "0/1";
        $change_count++;
        $het_count++;
    } elsif ($ratio < $threshold && ($ratio >= 0.5 + $stringency || $ratio <= 0.5 - ...
    ... $stringency)) {

```

```

    $remove_count++;
    next LINE;
} else { die "ERROR: Could not determine heterozygosity at line $line_count in ...
... $input_file:\n$line\n";
}

my $gt_now = join(":",@gt);
$data[9] = $gt_now;
chomp $data[9];
my $line_new = join("\t",@data);
print NEW_VCF "$line_new\n";
}

my $total = $change_count + $no_change_count;
print "Done!\nNumber of variants: $total\nNumber of changes: $change_count\n";
print "$output_file now contains $hom_count homozygous variants and $het_count ...
... heterozygous variants.\n";
print "$remove_count variants have been removed.\n";

close VCF;
close NEW_VCF;

exit;

```

### A.3.2.6 VCF\_tab\_to\_fasta\_alignment\_advanced.pl

```

#!/usr/bin/perl

# Program to convert output of VCFtools' vcf-to-tab to FASTA alignment.
# Modified by Barbara Roeder based on vcf_tab_to_fasta_alignment by C. Bergey

# Sample input file
# $ head input.vcf.tab
# chr10 94051 C ./ T/T
# chr10 94056 T ./ C/G
# chr10 94180 G A/A ./

# Sample output file
# >reference
# CTG
# >sample1
# CTA
# >sample2
# TSG

use strict;
use warnings;

my $exclude_het = 0;

my %iupac = (
    'G/G' => 'G',
    'C/C' => 'C',
    'T/T' => 'T',
    'A/A' => 'A',

    'G/T' => 'K',
    'T/G' => 'K',
    'A/C' => 'M',
    'C/A' => 'M',
    'C/G' => 'S',
    'G/C' => 'S',
    'A/G' => 'R',
    'G/A' => 'R',
    'A/T' => 'W',
    'T/A' => 'W',
    'C/T' => 'Y',
    'T/C' => 'Y',

    './.' => './',
);

my $input_tab = shift;
chomp $input_tab;

```

```

open (TAB, "<$input_tab")
  or die "ERROR: Could not open input file $input_tab.\n";

my $header = <TAB>;

my @col_names = split /\t/, $header;

# make temporary file without indels and multi nucleotide variants (MNVs)
my $temp_tab = $input_tab . "_clean";
open (TEMP, ">$temp_tab") || die "ERROR: Could not open temp file $temp_tab.\n";

my $num_cols = scalar @col_names;

LINE: foreach my $line (<TAB>) {
  my @data = split /\t/, $line;
  # indels
  if ((scalar @data) < $num_cols) {
    next LINE;
  }
  # MNVs
  for (my $i = 2; $i < $num_cols; $i++) {
    my $bp = $data[$i];
    chomp $bp;
    if ($bp =~ /\w{2,}/) {
      next LINE;
    }
  }

  # if necessary, exclude heterozygous mutations
  if ($exclude_het) {
    for (my $i = 2; $i < $num_cols; $i++) {
      my $bp = $data[$i];
      chomp $bp;
      if ($bp =~ /(\w)\s/(\w)/) {
        if ($1 ne $2) {
          next LINE;
        }
      }
    }
  }
  print TEMP $line;
}

close TAB;
close TEMP;

# Conversion to FASTA alignment
for (my $i = 2; $i < $num_cols; $i++) {
  my $ind = $col_names[$i];
  chomp $ind;
  print ">" . $ind . "\n";

  open (TEMP, "<$temp_tab") || die "ERROR: Could not open temp file $temp_tab.\n";

  my $count = 0;

  foreach my $line (<TEMP>) {
    my @data = split /\t/, $line;
    my $nuc = $data[$i];
    chomp $nuc;

    # for reference just print basepair
    if ($i == 2) {
      print $nuc;
      $count++;
    }

    # for samples
    ## symbol '.', i.e. no variant, therefore set to reference base
  } elsif ($nuc eq '.') {
    print $data[2];
    $count++;
  }

  ## substitutions
  } elsif ($nuc =~ /(\w)\s/(\w)/) {
    my $first = $1;
    my $second = $2;

```

```
# Homozygous
if ($first eq $second) {
    print $first;
    $count++;

# Heterozygous
} else {
    my $gt = $first . '/' . $second;
    if ( !exists($iupac{$gt}) ) { die "ERROR: Basepair is $nuc\n"; }
        $gt = $iupac{$gt};
    print $gt;
    $count++;
}
}
# newline every 100 bp
if ($count == 100) {
    print "\n";
    $count = 0;
}
}
close TEMP;
print "\n";
}
exit;
```

---

## **A.4 Inhaltsverzeichnis des elektronischen Anhangs**

**01 Eigene *Perl*-Skripte**

**02 Artifizielles *SNV-Alignment* für die phylogenetische Analyse**

**03 Tabellarische Übersicht der Vorhersageergebnisse**

**04 Abbildung 3.17 - Ergebnisse der Vorhersagealgorithmen**



# Literaturverzeichnis

- Alberts, B., Johnson, A., Walter, P., Lewis, J., Raff, M., & Roberts, K. 2008. *Molecular biology of the cell*. 5th edn. New York: Garland Science Taylor & Francis.
- Albino, A. P., Le Strange, R., Oliff, A. I., Furth, M. E., & Old, L. J. 1984. Transforming ras genes from human melanoma: a manifestation of tumour heterogeneity? *Nature*, **308**(5954), 69–72.
- Albino, A. P., Nanus, D. M., Mentle, I. R., Cordon-Cardo, C., McNutt, N. S., Bressler, J., & Andreeff, M. 1989. Analysis of ras oncogenes in malignant melanoma and precursor lesions: correlation of point mutations with differentiation phenotype. *Oncogene*, **4**(11), 1363–1374.
- Almand, B., Clark, J. I., Nikitina, E., van Beynen, J., English, N. R., Knight, S. C., Carbone, D. P., & Gabrilovich, D. I. 2001. Increased production of immature myeloid cells in cancer patients: a mechanism of immunosuppression in cancer. *Journal of Immunology (Baltimore, Md.: 1950)*, **166**(1), 678–689.
- Altman, J. D., Moss, P. A., Goulder, P. J., Barouch, D. H., McHeyzer-Williams, M. G., Bell, J. I., McMichael, A. J., & Davis, M. M. 1996. Phenotypic analysis of antigen-specific T lymphocytes. *Science (New York, N.Y.)*, **274**(5284), 94–96.
- Anders, S., & Huber, W. 2010. Differential expression analysis for sequence count data. *Genome Biology*, **11**(10), R106.
- Anders, S., & Huber, W. 2012. *Differential expression of RNA-Seq data at the gene level - the DESeq package. (Vignette)*.
- Andersen, R. S., Kvistborg, P., Frøsig, T. M., Pedersen, N. W., Lyngaa, R., Bakker, A. H., Shu, C. J., Straten, P. T., Schumacher, T. N., & Hadrup, S. R. 2012. Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. *Nature Protocols*, **7**(5), 891–902.
- Anichini, A., Vegetti, C., & Mortarini, R. 2004. The paradox of T-cell-mediated antitumor immunity in spite of poor clinical outcome in human melanoma. *Cancer immunology, immunotherapy: CII*, **53**(10), 855–864.
- Aptsiauri, N., Carretero, R., Garcia-Lora, A., Real, L. M., Cabrera, T., & Garrido, F. 2008. Regressing and progressing metastatic lesions: resistance to immunotherapy is predetermined by irreversible HLA class I antigen alterations. *Cancer immunology, immunotherapy: CII*, **57**(11), 1727–1733.
- Asemissen, A. M., Haase, D., Stevanovic, S., Bauer, S., Busse, A., Thiel, E., Rammensee, H.-G., Keilholz, U., & Scheibenbogen, C. 2009. Identification of an immunogenic HLA-A\*0201-binding T-cell epitope of the transcription factor PAX2. *Journal of Immunotherapy (Hagerstown, Md.: 1997)*, **32**(4), 370–375.
- Atak, Z. K., Gianfelici, V., Hulselmans, G., De Keersmaecker, K., Devasia, A. G., Geerdens, E., Mentens, N., Chiaretti, S., Durinck, K., Uyttebroeck, A., Vandenberghe, P., Wlodarska, I., Cloos, J., Foà, R., Speleman, F., Cools, J., & Aerts, S. 2013. Comprehensive analysis of transcriptome variation uncovers known and novel driver events in T-cell acute lymphoblastic leukemia. *PLoS genetics*, **9**(12), e1003997.

- 
- Atkins, M. B., Lotze, M. T., Dutcher, J. P., Fisher, R. I., Weiss, G., Margolin, K., Abrams, J., Sznol, M., Parkinson, D., Hawkins, M., Paradise, C., Kunkel, L., & Rosenberg, S. A. 1999. High-dose recombinant interleukin 2 therapy for patients with metastatic melanoma: analysis of 270 patients treated between 1985 and 1993. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **17**(7), 2105–2116.
- Avesson, L., & Barry, G. 2014. The emerging role of RNA and DNA editing in cancer. *Biochimica Et Biophysica Acta*, **1845**(2), 308–316.
- AWMF, Deutsche Krebsgesellschaft e. V., & Deutsche Krebshilfe e. V. 2013. *Malignes Melanom, S3-Leitlinie "Diagnostik, Therapie und Nachsorge des Melanoms", v1.1*. Berlin: Leitlinienprogramm Onkologie.
- Balch, C. M., Gershenwald, J. E., Soong, S.-J., Thompson, J. F., Atkins, M. B., Byrd, D. R., Buzaid, A. C., Cochran, A. J., Coit, D. G., Ding, S., Eggermont, A. M., Flaherty, K. T., Gimotty, P. A., Kirkwood, J. M., McMasters, K. M., Mihm, M. C., Morton, D. L., Ross, M. I., Sober, A. J., & Sondak, V. K. 2009. Final version of 2009 AJCC melanoma staging and classification. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **27**(36), 6199–6206.
- Baniyash, M. 2004. TCR zeta-chain downregulation: curtailing an excessive inflammatory immune response. *Nature Reviews. Immunology*, **4**(9), 675–687.
- Barnstable, C. J., Bodmer, W. F., Brown, G., Galfre, G., Milstein, C., Williams, A. F., & Ziegler, A. 1978. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens—new tools for genetic analysis. *Cell*, **14**(1), 9–20.
- Barrett, S. D., Bridges, A. J., Dudley, D. T., Saltiel, A. R., Fergus, J. H., Flamme, C. M., Delaney, A. M., Kaufman, M., LePage, S., Leopold, W. R., Przybranowski, S. A., Sebolt-Leopold, J., Van Becelaere, K., Doherty, A. M., Kennedy, R. M., Marston, D., Howard, W. A., Smith, Y., Warmus, J. S., & Teclé, H. 2008. The discovery of the benzhydroxamate MEK inhibitors CI-1040 and PD 0325901. *Bioorganic & Medicinal Chemistry Letters*, **18**(24), 6501–6504.
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., Irving, L. J., Karbelashvili, M. S., Kirk, S. M., Li, H., Liu, X., Maisinger, K. S., Murray, L. J., Obradovic, B., Ost, T., Parkinson, M. L., Pratt, M. R., Rasolonjatovo, I. M. J., Reed, M. T., Rigatti, R., Rodighiero, C., Ross, M. T., Sabot, A., Sankar, S. V., Scally, A., Schroth, G. P., Smith, M. E., Smith, V. P., Spiridou, A., Torrance, P. E., Tzonev, S. S., Vermaas, E. H., Walter, K., Wu, X., Zhang, L., Alam, M. D., Anastasi, C., Aniebo, I. C., Bailey, D. M. D., Bancarz, I. R., Banerjee, S., Barbour, S. G., Baybayan, P. A., Benoit, V. A., Benson, K. F., Bevis, C., Black, P. J., Boodhun, A., Brennan, J. S., Bridgham, J. A., Brown, R. C., Brown, A. A., Buermann, D. H., Bundu, A. A., Burrows, J. C., Carter, N. P., Castillo, N., Chiara E Catenazzi, M., Chang, S., Neil Cooley, R., Crake, N. R., Dada, O. O., Diakoumakos, K. D., Dominguez-Fernandez, B., Earnshaw, D. J., Egbujor, U. C., Elmore, D. W., Etschin, S. S., Ewan, M. R., Fedurco, M., Fraser, L. J., Fuentes Fajardo, K. V., Scott Furey, W., George, D., Gietzen, K. J., Goddard, C. P., Golda, G. S., Granieri, P. A., Green, D. E., Gustafson, D. L., Hansen, N. F., Harnish, K., Haudenschild, C. D., Heyer, N. I., Hims, M. M., Ho, J. T., Horgan, A. M., Hoschler, K., Hurwitz, S., Ivanov, D. V., Johnson, M. Q., James, T., Huw Jones, T. A., Kang, G.-D., Kerelska, T. H., Kersey, A. D., Khrebtukova, I., Kindwall, A. P., Kingsbury, Z., Kokko-Gonzales, P. I., Kumar, A., Laurent, M. A., Lawley, C. T., Lee, S. E., Lee, X., Liao, A. K., Loch, J. A., Lok, M., Luo, S., Mammen, R. M., Martin,

- J. W., McCauley, P. G., McNitt, P., Mehta, P., Moon, K. W., Mullens, J. W., Newington, T., Ning, Z., Ling Ng, B., Novo, S. M., O'Neill, M. J., Osborne, M. A., Osnowski, A., Ostadan, O., Paraschos, L. L., Pickering, L., Pike, A. C., Pike, A. C., Chris Pinkard, D., Pliskin, D. P., Podhasky, J., Quijano, V. J., Raczy, C., Rae, V. H., Rawlings, S. R., Chiva Rodriguez, A., Roe, P. M., Rogers, J., Rogert Bacigalupo, M. C., Romanov, N., Romieu, A., Roth, R. K., Rourke, N. J., Ruediger, S. T., Rusman, E., Sanches-Kuiper, R. M., Schenker, M. R., Seoane, J. M., Shaw, R. J., Shiver, M. K., Short, S. W., Sizto, N. L., Sluis, J. P., Smith, M. A., Ernest Sohna Sohna, J., Spence, E. J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C. L., Turcatti, G., Vandevondele, S., Verhovsky, Y., Virk, S. M., Wakelin, S., Walcott, G. C., Wang, J., Worsley, G. J., Yan, J., Yau, L., Zuerlein, M., Rogers, J., Mullikin, J. C., Hurles, M. E., McCooke, N. J., West, J. S., Oaks, F. L., Lundberg, P. L., Klenerman, D., Durbin, R., & Smith, A. J. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**(7218), 53–59.
- Berger, M. F., Levin, J. Z., Vijayendran, K., Sivachenko, A., Adiconis, X., Maguire, J., Johnson, L. A., Robinson, J., Verhaak, R. G., Sougnez, C., Onofrio, R. C., Ziaugra, L., Cibulskis, K., Laine, E., Barretina, J., Winckler, W., Fisher, D. E., Getz, G., Meyerson, M., Jaffe, D. B., Gabriel, S. B., Lander, E. S., Dummer, R., Gnirke, A., Nusbaum, C., & Garraway, L. A. 2010. Integrative analysis of the melanoma transcriptome. *Genome Research*, **20**(4), 413–427.
- Berger, M. F., Hodis, E., Heffernan, T. P., Deribe, Y. L., Lawrence, M. S., Protopopov, A., Ivanova, E., Watson, I. R., Nickerson, E., Ghosh, P., Zhang, H., Zeid, R., Ren, X., Cibulskis, K., Sivachenko, A. Y., Wagle, N., Sucker, A., Sougnez, C., Onofrio, R., Ambrogio, L., Auclair, D., Fennell, T., Carter, S. L., Drier, Y., Stojanov, P., Singer, M. A., Voet, D., Jing, R., Saksena, G., Barretina, J., Ramos, A. H., Pugh, T. J., Stransky, N., Parkin, M., Winckler, W., Mahan, S., Ardlie, K., Baldwin, J., Wargo, J., Schadendorf, D., Meyerson, M., Gabriel, S. B., Golub, T. R., Wagner, S. N., Lander, E. S., Getz, G., Chin, L., & Garraway, L. A. 2012. Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature*, **485**(7399), 502–506.
- Bernal, M., Ruiz-Cabello, F., Concha, A., Paschen, A., & Garrido, F. 2012. Implication of the  $\beta$ 2-microglobulin gene in the generation of tumor escape phenotypes. *Cancer immunology, immunotherapy: CII*, **61**(9), 1359–1371.
- Bertolotto, C. 2013. Melanoma: from melanocyte to genetic alterations and clinical options. *Scientifica*, **2013**, 635203.
- Blanchard, T., Srivastava, P. K., & Duan, F. 2013. Vaccines against advanced melanoma. *Clinics in Dermatology*, **31**(2), 179–190.
- Blom, B., & Spits, H. 2006. Development of human lymphoid cells. *Annual Review of Immunology*, **24**, 287–320.
- Bolger, A. M., Lohse, M., & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, **30**(15), 2114–2120.
- Brahmer, J. R., Drake, C. G., Wollner, I., Powderly, J. D., Picus, J., Sharfman, W. H., Stankevich, E., Pons, A., Salay, T. M., McMiller, T. L., Gilson, M. M., Wang, C., Selby, M., Taube, J. M., Anders, R., Chen, L., Korman, A. J., Pardoll, D. M., Lowy, I., & Topalian, S. L. 2010. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **28**(19), 3167–3175.

- 
- Brichard, V., Van Pel, A., Wölfel, T., Wölfel, C., De Plaen, E., Lethé, B., Coulie, P., & Boon, T. 1993. The tyrosinase gene codes for an antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas. *The Journal of Experimental Medicine*, **178**(2), 489–495.
- Buell, J. F., Gross, T. G., & Woodle, E. S. 2005. Malignancy after transplantation. *Transplantation*, **80**(2 Suppl), S254–264.
- Burnet, F. M. 1959. *The Clonal Selection Theory of Acquired Immunity*. London: Cambridge University Press.
- Burnet, F. M. 1970. The concept of immunological surveillance. *Progress in Experimental Tumor Research*, **13**, 1–27.
- Campbell, P. J., Pleasance, E. D., Stephens, P. J., Dicks, E., Rance, R., Goodhead, I., Follows, G. A., Green, A. R., Futreal, P. A., & Stratton, M. R. 2008. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(35), 13081–13086.
- Castelli, C., Storkus, W. J., Maeurer, M. J., Martin, D. M., Huang, E. C., Pramanik, B. N., Nagabhushan, T. L., Parmiani, G., & Lotze, M. T. 1995. Mass spectrometric identification of a naturally processed melanoma peptide recognized by CD8+ cytotoxic T lymphocytes. *The Journal of Experimental Medicine*, **181**(1), 363–368.
- Castle, J. C., Kreiter, S., Diekmann, J., Löwer, M., van de Roemer, N., de Graaf, J., Selmi, A., Diken, M., Boegel, S., Paret, C., Koslowski, M., Kuhn, A. N., Britten, C. M., Huber, C., Türeci, Ö., & Sahin, U. 2012. Exploiting the mutanome for tumor vaccination. *Cancer Research*, **72**(5), 1081–1091.
- Cha, E., Klinger, M., Hou, Y., Cummings, C., Ribas, A., Faham, M., & Fong, L. 2014. Improved survival with T cell clonotype stability after anti-CTLA-4 treatment in cancer patients. *Science Translational Medicine*, **6**(238), 238ra70.
- Chambers, B., Grufman, P., Fredriksson, V., Andersson, K., Roseboom, M., Laban, S., Camps, M., Wolpert, E. Z., Wiertz, E. J. H. J., Offringa, R., Ljunggren, H.-G., & van Hall, T. 2007. Induction of protective CTL immunity against peptide transporter TAP-deficient tumors through dendritic cell vaccination. *Cancer Research*, **67**(18), 8450–8455.
- Champiat, S., Ferté, C., Lebel-Binay, S., Eggermont, A., & Soria, J. C. 2014. Exomics and immunogenics: Bridging mutational load and immune checkpoints efficacy. *Oncoimmunology*, **3**(1), e27817.
- Chapman, P. B., Hauschild, A., Robert, C., Haanen, J. B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., Hogg, D., Lorigan, P., Lebbe, C., Jouary, T., Schadendorf, D., Ribas, A., O’Day, S. J., Sosman, J. A., Kirkwood, J. M., Eggermont, A. M. M., Dreno, B., Nolop, K., Li, J., Nelson, B., Hou, J., Lee, R. J., Flaherty, K. T., McArthur, G. A., & BRIM-3 Study Group. 2011. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *The New England Journal of Medicine*, **364**(26), 2507–2516.
- Chen, L., & Flies, D. B. 2013. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nature Reviews. Immunology*, **13**(4), 227–242.
- Chen, R.-L., Guo, W., Shi, Y., Wu, H., Wang, J., & Sun, G. 2013. Computational identification of specific splicing regulatory elements from RNA-seq in lung cancer. *European Review for Medical and Pharmacological Sciences*, **17**(13), 1716–1721.
-

- Chen, Y. T., Scanlan, M. J., Sahin, U., Türeci, O., Gure, A. O., Tsang, S., Williamson, B., Stockert, E., Pfreundschuh, M., & Old, L. J. 1997. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(5), 1914–1918.
- Cheng, P., Cheng, Y., Li, Y., Zhao, Z., Gao, H., Li, D., Li, H., & Zhang, T. 2012. Comparison of the gene expression profiles between smokers with and without lung cancer using RNA-Seq. *Asian Pacific journal of cancer prevention: APJCP*, **13**(8), 3605–3609.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, **6**(2), 80–92.
- Collins, F. S., Green, E. D., Guttmacher, A. E., Guyer, M. S., & US National Human Genome Research Institute. 2003. A vision for the future of genomics research. *Nature*, **422**(6934), 835–847.
- Conrad, C. T., Ernst, N. R., Dummer, W., Bröcker, E. B., & Becker, J. C. 1999. Differential expression of transforming growth factor beta 1 and interleukin 10 in progressing and regressing areas of primary melanoma. *Journal of experimental & clinical cancer research: CR*, **18**(2), 225–232.
- Cornish-Bowden, A. 1985. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Research*, **13**(9), 3021–3030.
- Cummins, D. L., Cummins, J. M., Pantle, H., Silverman, M. A., Leonard, A. L., & Chanmugam, A. 2006. Cutaneous malignant melanoma. *Mayo Clinic Proceedings*, **81**(4), 500–507.
- Danecek, P., Auton, A., Abecasis, G., Albers, C., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, **27**(15), 2156–2158.
- Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W. C., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R., & Futreal, P. A. 2002. Mutations of the BRAF gene in human cancer. *Nature*, **417**(6892), 949–954.
- del Campo, A. B., Aptsiauri, N., Méndez, R., Zinchenko, S., Vales, A., Paschen, A., Ward, S., Ruiz-Cabello, F., González-Aseguinolaza, G., & Garrido, F. 2009. Efficient recovery of HLA class I expression in human tumor cells after beta2-microglobulin gene transfer using adenoviral vector: implications for cancer immunotherapy. *Scandinavian Journal of Immunology*, **70**(2), 125–135.
- den Dunnen, J. T., & Antonarakis, S. E. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Human Mutation*, **15**(1), 7–12.
- Di Zazzo, E., De Rosa, C., Abbondanza, C., & Moncharmont, B. 2013. PRDM Proteins: Molecular Mechanisms in Signal Transduction and Transcriptional Regulation. *Biology*, **2**(1), 107–141.

- 
- Diaz-Montero, C. M., Salem, M. L., Nishimura, M. I., Garrett-Mayer, e., Cole, D. J., & Montero, A. J. 2009. Increased circulating myeloid-derived suppressor cells correlate with clinical cancer stage, metastatic tumor burden, and doxorubicin-cyclophosphamide chemotherapy. *Cancer immunology, immunotherapy: CII*, **58**(1), 49–59.
- Dissemond, J., Goette, P., Moers, J., Lindeke, A., Goos, M., Ferrone, S., & Wagner, S. N. 2003. Immunoproteasome subunits LMP2 and LMP7 downregulation in primary malignant melanoma lesions: association with lack of spontaneous regression. *Melanoma Research*, **13**(4), 371–377.
- Dissemond, J., Busch, M., Kothen, T., Mörs, J., Weimann, T. K., Lindeke, A., Goos, M., & Wagner, S. N. 2004. Differential downregulation of endoplasmic reticulum-residing chaperones calnexin and calreticulin in human metastatic melanoma. *Cancer Letters*, **203**(2), 225–231.
- Drew, S. I., Terasaki, P. I., Billing, R. J., Bergh, O. J., Minowada, J., & Klein, E. 1977. Group-specific human granulocyte antigens on a chronic myelogenous leukemia cell line with a Philadelphia chromosome marker. *Blood*, **49**(5), 715–718.
- Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J., & Schreiber, R. D. 2002. Cancer immunoediting: from immunosurveillance to tumor escape. *Nature Immunology*, **3**(11), 991–998.
- Dunn, G. P., Old, L. J., & Schreiber, R. D. 2004. The three Es of cancer immunoediting. *Annual Review of Immunology*, **22**, 329–360.
- Duro, E., Lundin, C., Ask, K., Sanchez-Pulido, L., MacArtney, T. J., Toth, R., Ponting, C. P., Groth, A., Helleday, T., & Rouse, J. 2010. Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Molecular Cell*, **40**(4), 632–644.
- Eklund, G., & Malec, E. 1978. Sunlight and incidence of cutaneous malignant melanoma. Effect of latitude and domicile in Sweden. *Scandinavian Journal of Plastic and Reconstructive Surgery*, **12**(3), 231–241.
- Elnemr, A., Ohta, T., Yachie, A., Kayahara, M., Kitagawa, H., Ninomiya, I., Fushida, S., Fujimura, T., Nishimura, G., Shimizu, K., & Miwa, K. 2001. Human pancreatic cancer cells express non-functional Fas receptors and counterattack lymphocytes by expressing Fas ligand; a potential mechanism for immune escape. *International Journal of Oncology*, **18**(1), 33–39.
- Falk, K., Röttschke, O., Stevanović, S., Jung, G., & Rammensee, H. G. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, **351**(6324), 290–296.
- Feramisco, J. D., Goldstein, J. L., & Brown, M. S. 2004. Membrane topology of human insig-1, a protein regulator of lipid synthesis. *The Journal of Biological Chemistry*, **279**(9), 8487–8496.
- Flaherty, K. T., Robert, C., Hersey, P., Nathan, P., Garbe, C., Milhem, M., Demidov, L. V., Hassel, J. C., Rutkowski, P., Mohr, P., Dummer, R., Trefzer, U., Larkin, J. M. G., Utikal, J., Dreno, B., Nyakas, M., Middleton, M. R., Becker, J. C., Casey, M., Sherman, L. J., Wu, F. S., Ouellet, D., Martin, A.-M., Patel, K., Schadendorf, D., & METRIC Study Group. 2012. Improved survival with MEK inhibition in BRAF-mutated melanoma. *The New England Journal of Medicine*, **367**(2), 107–114.
- Flockhart, R. J., Webster, D. E., Qu, K., Mascarenhas, N., Kovalski, J., Kretz, M., & Khavari, P. A. 2012. BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. *Genome Research*, **22**(6), 1006–1014.
-

- Florell, S. R., Boucher, K. M., Garibotti, G., Astle, J., Kerber, R., Mineau, G., Wiggins, C., Noyes, R. D., Tsodikov, A., Cannon-Albright, L. A., Zone, J. J., Samlowski, W. E., & Leachman, S. A. 2005. Population-based analysis of prognostic factors and survival in familial melanoma. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **23**(28), 7168–7177.
- Fog, C. K., Galli, G. G., & Lund, A. H. 2012. PRDM proteins: important players in differentiation and disease. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, **34**(1), 50–60.
- Fujiwara, K., Higashi, T., Nouse, K., Nakatsukasa, H., Kobayashi, Y., Uemura, M., Nakamura, S.-I., Sato, S., Hanafusa, T., Yumoto, Y., Naito, I., & Shiratori, Y. 2004. Decreased expression of B7 costimulatory molecules and major histocompatibility complex class-I in human hepatocellular carcinoma. *Journal of Gastroenterology and Hepatology*, **19**(10), 1121–1127.
- Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., Zinzindohoué, F., Bruneval, P., Cugnenc, P.-H., Trajanoski, Z., Fridman, W.-H., & Pagès, F. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science (New York, N.Y.)*, **313**(5795), 1960–1964.
- Garbe, C., Peris, K., Hauschild, A., Saiag, P., Middleton, M., Spatz, A., Grob, J.-J., Malvehy, J., Newton-Bishop, J., Stratigos, A., Pehamberger, H., & Eggermont, A. 2010. Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline. *European Journal of Cancer (Oxford, England: 1990)*, **46**(2), 270–283.
- Garcia-Lora, A., Algarra, I., & Garrido, F. 2003. MHC class I antigens, immune surveillance, and tumor immune escape. *Journal of Cellular Physiology*, **195**(3), 346–355.
- Gerlinger, M., Rowan, A. J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., Varela, I., Phillimore, B., Begum, S., McDonald, N. Q., Butler, A., Jones, D., Raine, K., Latimer, C., Santos, C. R., Nohadani, M., Eklund, A. C., Spencer-Dene, B., Clark, G., Pickering, L., Stamp, G., Gore, M., Szallasi, Z., Downward, J., Futreal, P. A., & Swanton, C. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *The New England Journal of Medicine*, **366**(10), 883–892.
- Gerlinger, M., Horswell, S., Larkin, J., Rowan, A. J., Salm, M. P., Varela, I., Fisher, R., McGranahan, N., Matthews, N., Santos, C. R., Martinez, P., Phillimore, B., Begum, S., Rabinowitz, A., Spencer-Dene, B., Gulati, S., Bates, P. A., Stamp, G., Pickering, L., Gore, M., Nicol, D. L., Hazell, S., Futreal, P. A., Stewart, A., & Swanton, C. 2014. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature Genetics*, **46**(3), 225–233.
- Gilmartin, A. G., Bleam, M. R., Groy, A., Moss, K. G., Minthorn, E. A., Kulkarni, S. G., Rominger, C. M., Erskine, S., Fisher, K. E., Yang, J., Zappacosta, F., Annan, F., Sutton, D., & Laquerre, S. G. 2011. GSK1120212 (JTP-74057) is an inhibitor of MEK activity and activation with favorable pharmacokinetic properties for sustained in vivo pathway inhibition. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, **17**(5), 989–1000.
- Gluzman, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*, **23**(1), 175–182.
- Graf, C., Heidel, F., Tenzer, S., Radsak, M. P., Solem, F. K., Britten, C. M., Huber, C., Fischer, T., & Wölfel, T. 2007. A neoepitope generated by an FLT3 internal tandem

- 
- duplication (FLT3-ITD) is recognized by leukemia-reactive autologous CD8<sup>+</sup> T cells. *Blood*, **109**(7), 2985–2988.
- Graham, F. L., Smiley, J., Russell, W. C., & Nairn, R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *The Journal of General Virology*, **36**(1), 59–74.
- Grauer, O. M., Nierkens, S., Bennink, E., Toonen, L. W. J., Boon, L., Wesseling, P., Suttmuller, R. P. M., & Adema, G. J. 2007. CD4<sup>+</sup>FoxP3<sup>+</sup> regulatory T cells gradually accumulate in gliomas during tumor growth and efficiently suppress antiglioma immune responses in vivo. *International Journal of Cancer. Journal International Du Cancer*, **121**(1), 95–105.
- Gray-Schopfer, V., Wellbrock, C., & Marais, R. 2007. Melanoma biology and new targeted therapy. *Nature*, **445**(7130), 851–857.
- Gros, A., Robbins, P. F., Yao, X., Li, Y. F., Turcotte, S., Tran, E., Wunderlich, J. R., Mixon, A., Farid, S., Dudley, M. E., Hanada, K.-I., Almeida, J. R., Darko, S., Douek, D. C., Yang, J. C., & Rosenberg, S. A. 2014. PD-1 identifies the patient-specific CD8<sup>+</sup> tumor-reactive repertoire infiltrating human tumors. *The Journal of Clinical Investigation*, **124**(5), 2246–2259.
- Hadrup, S. R., Bakker, A. H., Shu, C. J., Andersen, R. S., van Veluw, J., Hombrink, P., Castermans, E., Thor Straten, P., Blank, C., Haanen, J. B., Heemskerk, M. H., & Schumacher, T. N. 2009. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nature Methods*, **6**(7), 520–526.
- Hanahan, D., & Weinberg, R. A. 2011. Hallmarks of cancer: the next generation. *Cell*, **144**(5), 646–674.
- Hanson, H. L., Donermeyer, D. L., Ikeda, H., White, J. M., Shankaran, V., Old, L. J., Shiku, H., Schreiber, R. D., & Allen, P. M. 2000. Eradication of established tumors by CD8<sup>+</sup> T cell adoptive immunotherapy. *Immunity*, **13**(2), 265–276.
- Heemskerk, B., Kvistborg, P., & Schumacher, T. N. M. 2013. The cancer antigenome. *The EMBO journal*, **32**(2), 194–203.
- Hill, V. K., Gartner, J. J., Samuels, Y., & Goldstein, A. M. 2013. The genetics of melanoma: recent advances. *Annual Review of Genomics and Human Genetics*, **14**, 257–279.
- Hodis, E., Watson, I. R., Kryukov, G. V., Arold, S. T., Imielinski, M., Theurillat, J.-P., Nickerson, E., Auclair, D., Li, L., Place, C., Dicara, D., Ramos, A. H., Lawrence, M. S., Cibulskis, K., Sivachenko, A., Voet, D., Saksena, G., Stransky, N., Onofrio, R. C., Winckler, W., Ardlie, K., Wagle, N., Wargo, J., Chong, K., Morton, D. L., Stenke-Hale, K., Chen, G., Noble, M., Meyerson, M., Ladbury, J. E., Davies, M. A., Gershenwald, J. E., Wagner, S. N., Hoon, D. S. B., Schadendorf, D., Lander, E. S., Gabriel, S. B., Getz, G., Garraway, L. A., & Chin, L. 2012. A landscape of driver mutations in melanoma. *Cell*, **150**(2), 251–263.
- Huber, W., von Heydebreck, A., Suelmann, H., Poustka, A., & Vingron, M. 2003. Parameter estimation for the calibration and variance stabilization of microarray data. *Statistical Applications in Genetics and Molecular Biology*, **2**, Article3.
- Ikeda, H., Lethé, B., Lehmann, F., van Baren, N., Baurain, J. F., de Smet, C., Chambost, H., Vitale, M., Moretta, A., Boon, T., & Coulie, P. G. 1997. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity*, **6**(2), 199–208.
-



- Jackson, M. R., Cohen-Doyle, M. F., Peterson, P. A., & Williams, D. B. 1994. Regulation of MHC class I transport by the molecular chaperone, calnexin (p88, IP90). *Science (New York, N.Y.)*, **263**(5145), 384–387.
- Jäger, E., Chen, Y. T., Drijfhout, J. W., Karbach, J., Ringhoffer, M., Jäger, D., Arand, M., Wada, H., Noguchi, Y., Stockert, E., Old, L. J., & Knuth, A. 1998. Simultaneous humoral and cellular immune response against cancer-testis antigen NY-ESO-1: definition of human histocompatibility leukocyte antigen (HLA)-A2-binding peptide epitopes. *The Journal of Experimental Medicine*, **187**(2), 265–270.
- Jäger, E., Jäger, D., Karbach, J., Chen, Y. T., Ritter, G., Nagata, Y., Gnjatic, S., Stockert, E., Arand, M., Old, L. J., & Knuth, A. 2000. Identification of NY-ESO-1 epitopes presented by human histocompatibility antigen (HLA)-DRB4\*0101-0103 and recognized by CD4(+) T lymphocytes of patients with NY-ESO-1-expressing melanoma. *The Journal of Experimental Medicine*, **191**(4), 625–630.
- Jakhesara, S. J., Koringa, P. G., Bhatt, V. D., Shah, T. M., Vangipuram, S., Shah, S., & Joshi, C. G. 2013. RNA-Seq reveals differentially expressed isoforms and novel splice variants in buccal mucosal cancer. *Gene*, **516**(1), 24–32.
- Jang, S., & Atkins, M. B. 2014. Treatment of BRAF-mutant melanoma: the role of vemurafenib and other therapies. *Clinical Pharmacology and Therapeutics*, **95**(1), 24–31.
- Jenkins, M. K., Chu, H. H., McLachlan, J. B., & Moon, J. J. 2010. On the composition of the preimmune repertoire of T cells specific for Peptide-major histocompatibility complex ligands. *Annual Review of Immunology*, **28**, 275–294.
- Joseph, E. W., Pratilas, C. A., Poulikakos, P. I., Tadi, M., Wang, W., Taylor, B. S., Halilovic, E., Persaud, Y., Xing, F., Viale, A., Tsai, J., Chapman, P. B., Bollag, G., Solit, D. B., & Rosen, N. 2010. The RAF inhibitor PLX4032 inhibits ERK signaling and tumor cell proliferation in a V600E BRAF-selective manner. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(33), 14903–14908.
- Karanikas, V., Zamanakou, M., Kerenidi, T., Dahabreh, J., Hevas, A., Nakou, M., Gourgoulisanis, K. I., & Germeis, A. E. 2007. Indoleamine 2,3-dioxygenase (IDO) expression in lung cancer. *Cancer Biology & Therapy*, **6**(8), 1258–1262.
- Kaufman, H. L. 2012. Vaccines for melanoma and renal cell carcinoma. *Seminars in Oncology*, **39**(3), 263–275.
- Kershaw, M. H., Westwood, J. A., & Darcy, P. K. 2013. Gene-engineered T cells for cancer therapy. *Nature Reviews. Cancer*, **13**(8), 525–541.
- Kessler, J. H., Beekman, N. J., Bres-Vloemans, S. A., Verdijk, P., van Veelen, P. A., Kloosterman-Joosten, A. M., Vissers, D. C., ten Bosch, G. J., Kester, M. G., Sijts, A., Wouter Drijfhout, J., Ossendorp, F., Offringa, R., & Melief, C. J. 2001. Efficient identification of novel HLA-A(\*)0201-presented cytotoxic T lymphocyte epitopes in the widely expressed tumor antigen PRAME by proteasome-mediated digestion analysis. *The Journal of Experimental Medicine*, **193**(1), 73–88.
- Kisselev, A. F., Akopian, T. N., Woo, K. M., & Goldberg, A. L. 1999. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *The Journal of Biological Chemistry*, **274**(6), 3363–3371.
- Knudson, A. G. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, **68**(4), 820–823.

- 
- Knudson, A. G. 1974. Heredity and human cancer. *The American Journal of Pathology*, **77**(1), 77–84.
- Kondo, A., Sidney, J., Southwood, S., del Guercio, M. F., Appella, E., Sakamoto, H., Celis, E., Grey, H. M., Chesnut, R. W., Kubo, R. T., & Sette, A. 1995. Prominent roles of secondary anchor residues in peptide binding to HLA-A24 human class I molecules. *Journal of Immunology (Baltimore, Md.: 1950)*, **155**(9), 4307–4312.
- Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**(2), 283–292.
- Krapivner, S., Popov, S., Chernogubova, E., Hellénus, M.-L., Fisher, R. M., Hamsten, A., & van't Hooft, F. M. 2008. Insulin-induced gene 2 involvement in human adipocyte metabolism and body weight regulation. *The Journal of Clinical Endocrinology and Metabolism*, **93**(5), 1995–2001.
- Krauthammer, M., Kong, Y., Ha, H., Evans, P., Bacchiocchi, A., McCusker, J. P., Cheng, E., Davis, M. J., Goh, G., Choi, M., Ariyan, S., Narayan, D., Dutton-Regester, K., Capatana, A., Holman, E. C., Bosenberg, M., Sznol, M., Kluger, H. M., Brash, D. E., Stern, D. F., Materin, M. A., Lo, R. S., Mane, S., Ma, S., Kidd, K. K., Hayward, N. K., Lifton, R. P., Schlessinger, J., Boggon, T. J., & Halaban, R. 2012. Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nature Genetics*, **44**(9), 1006–1014.
- Krishnan, V. G., & Ng, P. C. 2012. Predicting cancer drivers: are we there yet? *Genome Medicine*, **4**(11), 88.
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., & Marra, M. A. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research*, **19**(9), 1639–1645.
- Kung, P., Goldstein, G., Reinherz, E. L., & Schlossman, S. F. 1979. Monoclonal antibodies defining distinctive human T cell surface antigens. *Science (New York, N.Y.)*, **206**(4416), 347–349.
- Kunz, M., Dannemann, M., & Kelso, J. 2013. High-throughput sequencing of the melanoma genome. *Experimental Dermatology*, **22**(1), 10–17.
- Lafuente, E. M., & Reche, P. A. 2009. Prediction of MHC-peptide binding: a systematic and comprehensive overview. *Current Pharmaceutical Design*, **15**(28), 3209–3220.
- Laing, M. E., Moloney, F. J., Kay, E. W., Conlon, P., & Murphy, G. M. 2006. Malignant melanoma in transplant patients: review of five cases. *Clinical and Experimental Dermatology*, **31**(5), 662–664.
- Landsberg, J., Kohlmeyer, J., Renn, M., Bald, T., Rogava, M., Cron, M., Fatho, M., Lennerz, V., Wölfel, T., Hölzel, M., & Tüting, T. 2012. Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. *Nature*, **490**(7420), 412–416.
- Langmead, B., & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**(4), 357–359.
- Lener, T., Moll, P. R., Rinnerthaler, M., Bauer, J., Aberger, F., & Richter, K. 2006. Expression profiling of aging in the human skin. *Experimental Gerontology*, **41**(4), 387–397.
- Lennerz, V., Fatho, M., Gentilini, C., Frye, R. A., Lifke, A., Ferel, D., Wölfel, C., Huber, C., & Wölfel, T. 2005. The response of autologous T cells to a human melanoma is dominated by mutated neoantigens. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(44), 16013–16018.
-

- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**(16), 2078–2079.
- Lindsten, T., Lee, K. P., Harris, E. S., Petryniak, B., Craighead, N., Reynolds, P. J., Lombard, D. B., Freeman, G. J., Nadler, L. M., & Gray, G. S. 1993. Characterization of CTLA-4 structure and expression on human T cells. *Journal of Immunology (Baltimore, Md.: 1950)*, **151**(7), 3489–3499.
- Linnemann, C., Heemskerk, B., Kvistborg, P., Kluin, R. J. C., Bolotin, D. A., Chen, X., Bresser, K., Nieuwland, M., Schotte, R., Michels, S., Gomez-Eerland, R., Jahn, L., Hombrink, P., Legrand, N., Shu, C. J., Mamedov, I. Z., Velds, A., Blank, C. U., Haanen, J. B. A. G., Turchaninova, M. A., Kerkhoven, R. M., Spits, H., Hadrup, S. R., Heemskerk, M. H. M., Blankenstein, T., Chudakov, D. M., Bendle, G. M., & Schumacher, T. N. M. 2013. High-throughput identification of antigen-specific TCRs by TCR gene capture. *Nature Medicine*, **19**(11), 1534–1541.
- Linnemann, c., Mezzadra, R., & Schumacher, T. N. M. 2014. TCR repertoires of intratumoral T-cell subsets. *Immunological Reviews*, **257**(1), 72–82.
- Livingstone, E., Zimmer, L., Vaubel, J., & Schadendorf, D. 2012. Current advances and perspectives in the treatment of advanced melanoma. *Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology: JDDG*, **10**(5), 319–325.
- Lodish, H. F., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (eds). 2000. The Three Roles of RNA in Protein Synthesis. *In: Lodish, H. F., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (eds), Molecular cell biology*, 4th ed edn. NCBI Bookshelf. New York: W.H. Freeman.
- Lohse, M., Bolger, A. M., Nagel, A., Fernie, A. R., Lunn, J. E., Stitt, M., & Usadel, B. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, **40**(Web Server issue), W622–627.
- Löwer, M., Renard, B. Y., de Graaf, J., Wagner, M., Paret, C., Kneip, C., Türeci, Ö., Diken, M., Britten, C., Kreiter, S., Koslowski, M., Castle, J. C., & Sahin, U. 2012. Confidence-based somatic mutation evaluation and prioritization. *PLoS computational biology*, **8**(9), e1002714.
- Lu, Y.-C., Yao, X., Li, Y. F., El-Gamil, M., Dudley, M. E., Yang, J. C., Almeida, J. R., Douek, D. C., Samuels, Y., Rosenberg, S. A., & Robbins, P. F. 2013. Mutated PPP1R3B is recognized by T cells used to treat a melanoma patient who experienced a durable complete tumor regression. *Journal of Immunology (Baltimore, Md.: 1950)*, **190**(12), 6034–6042.
- Lundegaard, C., Lamberth, K., Harndahl, M., Buus, S., Lund, O., & Nielsen, M. 2008. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Research*, **36**(Web Server issue), W509–512.
- Lundegaard, C., Lund, O., Buus, S., & Nielsen, M. 2010. Major histocompatibility complex class I binding predictions as a tool in epitope discovery. *Immunology*, **130**(3), 309–318.
- Ma, L., Nie, L., Liu, J., Zhang, B., Song, S., Sun, M., Yang, J., Yang, Y., Fang, X., Hu, S., Zhao, Y., & Yu, J. 2012. An RNA-seq-based gene expression profiling of radiation-induced tumorigenic mammary epithelial cells. *Genomics, Proteomics & Bioinformatics*, **10**(6), 326–335.

- 
- Maeurer, M. J., Gollin, S. M., Martin, D., Swaney, W., Bryant, J., Castelli, C., Robbins, P., Parmiani, G., Storkus, W. J., & Lotze, M. T. 1996. Tumor escape from immune recognition: lethal recurrent melanoma in a patient associated with downregulation of the peptide transporter protein TAP-1 and loss of expression of the immunodominant MART-1/Melan-A antigen. *The Journal of Clinical Investigation*, **98**(7), 1633–1641.
- Maher, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., & Chinnaiyan, A. M. 2009. Transcriptome sequencing to detect gene fusions in cancer. *Nature*, **458**(7234), 97–101.
- Mandelboim, O., Vadai, E., Fridkin, M., Katz-Hillel, A., Feldman, M., Berke, G., & Eisenbach, L. 1995. Regression of established murine carcinoma metastases following vaccination with tumour-associated antigen peptides. *Nature Medicine*, **1**(11), 1179–1183.
- Matsushita, H., Vesely, M. D., Koblodt, D. C., Rickert, C. G., Uppaluri, R., Magrini, V. J., Arthur, C. D., White, J. M., Chen, Y.-S., Shea, L. K., Hundal, J., Wendl, M. C., Demeter, R., Wylie, T., Allison, J. P., Smyth, M. J., Old, L. J., Mardis, E. R., & Schreiber, R. D. 2012. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature*, **482**(7385), 400–404.
- McCutcheon, M., Wehner, N., Wensky, A., Kushner, M., Doan, S., Hsiao, L., Calabresi, P., Ha, T., Tran, T. V., Tate, K. M., Winkelhake, J., & Spack, E. G. 1997. A sensitive ELISPOT assay to detect low-frequency human T lymphocytes. *Journal of Immunological Methods*, **210**(2), 149–166.
- McLeod, G. R., & Davis, N. C. 1992. Historical overview of melanoma. *Clinics in Dermatology*, **10**(1), 5–7.
- Méndez, R., Rodríguez, T., Del Campo, A., Monge, E., Maleno, I., Aptsiauri, N., Jiménez, P., Pedrinaci, S., Pawelec, G., Ruiz-Cabello, F., & Garrido, F. 2008. Characterization of HLA class I altered phenotypes in a panel of human melanoma cell lines. *Cancer immunology, immunotherapy: CII*, **57**(5), 719–729.
- Meric-Bernstam, F., Frampton, G. M., Ferrer-Lozano, J., Yelensky, R., Pérez-Fidalgo, J. A., Wang, Y., Palmer, G. A., Ross, J. S., Miller, V. A., Su, X., Eroles, P., Barrera, J. A., Burgues, O., Lluch, A. M., Zheng, X., Sahin, A., Stephens, P. J., Mills, G. B., Cronin, M. T., & Gonzalez-Angulo, A. M. 2014. Concordance of genomic alterations between primary and recurrent breast cancer. *Molecular Cancer Therapeutics*, **13**(5), 1382–1389.
- Metzker, M. L. 2005. Emerging technologies in DNA sequencing. *Genome Research*, **15**(12), 1767–1776.
- Metzker, M. L. 2010. Sequencing technologies - the next generation. *Nature Reviews. Genetics*, **11**(1), 31–46.
- Mittal, D., Gubin, M. M., Schreiber, R. D., & Smyth, M. J. 2014. New insights into cancer immunoediting and its three component phases—elimination, equilibrium and escape. *Current Opinion in Immunology*, **27**(Apr.), 16–25.
- Mocellin, S., Pasquali, S., Rossi, C., & Nitti, D. 2010. Interferon alpha adjuvant therapy in patients with high-risk melanoma: a systematic review and meta-analysis. *Journal of the National Cancer Institute*, **102**(7), 493–501.
- Morris, E. J., Jha, S., Restaino, C. R., Dayananth, P., Zhu, H., Cooper, A., Carr, D., Deng, Y., Jin, W., Black, S., Long, B., Liu, J., Dinunzio, E., Windsor, W., Zhang, R., Zhao, S., Angagaw, M. H., Pinheiro, E. M., Desai, J., Xiao, L., Shipps, G., Hruza, A., Wang, J., Kelly, J., Paliwal, S., Gao, X., Babu, B. S., Zhu, L., Daublain, P., Zhang, L., Lutterbach,

- B. A., Pelletier, M. R., Philippar, U., Siliphaivanh, P., Witter, D., Kirschmeier, P., Bishop, W. R., Hicklin, D., Gilliland, D. G., Jayaraman, L., Zawel, L., Fawell, S., & Samatar, A. A. 2013. Discovery of a novel ERK inhibitor with activity in models of acquired resistance to BRAF and MEK inhibitors. *Cancer Discovery*, **3**(7), 742–750.
- Murphy, K. 2012. *Janeway's immunobiology*. 8th edn. New York: Garland Science.
- Murray, R. J., Kurilla, M. G., Brooks, J. M., Thomas, W. A., Rowe, M., Kieff, E., & Rickinson, A. B. 1992. Identification of target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): implications for the immune control of EBV-positive malignancies. *The Journal of Experimental Medicine*, **176**(1), 157–168.
- Müschen, M., Moers, C., Warskulat, U., Even, J., Niederacher, D., & Beckmann, M. W. 2000. CD95 ligand expression as a mechanism of immune escape in breast cancer. *Immunology*, **99**(1), 69–77.
- Mutz, K.-O., Heilkenbrinker, A., Lönne, M., Walter, J.-G., & Stahl, F. 2013. Transcriptome analysis using next-generation sequencing. *Current Opinion in Biotechnology*, **24**(1), 22–30.
- Neefjes, J. J., Hämmerling, G. J., & Momburg, F. 1993. Folding and assembly of major histocompatibility complex class I heterodimers in the endoplasmic reticulum of intact cells precedes the binding of peptide. *The Journal of Experimental Medicine*, **178**(6), 1971–1980.
- Newell, E. W., Klein, L. O., Yu, W., & Davis, M. M. 2009. Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. *Nature Methods*, **6**(7), 497–499.
- Nguyen, M.-H., Ueda, K., Nakamura, Y., & Daigo, Y. 2012. Identification of a novel oncogene, MMS22L, involved in lung and esophageal carcinogenesis. *International Journal of Oncology*, **41**(4), 1285–1296.
- Nicholas, C., & Lesinski, G. B. 2011. Immunomodulatory cytokines as therapeutic agents for melanoma. *Immunotherapy*, **3**(5), 673–690.
- Nielsen, M., Lundegaard, C., Wornig, P., Lauemøller, S. L., Lamberth, K., Buus, S., Brunak, S., & Lund, O. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Science: A Publication of the Protein Society*, **12**(5), 1007–1017.
- Noppen, C., Lévy, F., Burri, L., Zajac, P., Remmel, E., Schaefer, C., Lüscher, U., Heberer, M., & Spagnoli, G. C. 2000. Naturally processed and concealed HLA-A2.1-restricted epitopes from tumor-associated antigen tyrosinase-related protein-2. *International Journal of Cancer. Journal International Du Cancer*, **87**(2), 241–246.
- Novellino, L., Renkvist, N., Rini, F., Mazzocchi, A., Rivoltini, L., Greco, A., Deho, P., Squarcina, P., Robbins, P., Parmiani, G., & Castelli, C. 2003. Identification of a mutated receptor-like protein tyrosine phosphatase kappa as a novel, class II HLA-restricted melanoma antigen. *Journal of Immunology (Baltimore, Md.: 1950)*, **170**(12), 6363–6370.
- Novellino, L., Castelli, C., & Parmiani, G. 2005. A listing of human tumor antigens recognized by T cells: March 2004 update. *Cancer immunology, immunotherapy: CII*, **54**(3), 187–207.
- O'Connell, B. C., Adamson, B., Lydeard, J. R., Sowa, M. E., Ciccia, A., Bredemeyer, A. L., Schlabach, M., Gygi, S. P., Elledge, S. J., & Harper, J. W. 2010. A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Molecular Cell*, **40**(4), 645–657.

- 
- O'Donnell, L., Panier, S., Wildenhain, J., Tkach, J. M., Al-Hakim, A., Landry, M.-C., Escibano-Diaz, C., Szilard, R. K., Young, J. T. F., Munro, M., Canny, M. D., Kolas, N. K., Zhang, W., Harding, S. M., Ylanko, J., Mendez, M., Mullin, M., Sun, T., Habermann, B., Datti, A., Bristow, R. G., Gingras, A.-C., Tyers, M. D., Brown, G. W., & Durocher, D. 2010. The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Molecular Cell*, **40**(4), 619–631.
- Okazaki, T., Chikuma, S., Iwai, Y., Fagarasan, S., & Honjo, T. 2013. A rheostat for immune responses: the unique properties of PD-1 and their advantages for clinical application. *Nature Immunology*, **14**(12), 1212–1218.
- Old, L. J., & Boyse, E. A. 1964. IMMUNOLOGY OF EXPERIMENTAL TUMORS. *Annual Review of Medicine*, **15**, 167–186.
- Park, J.-A., & Kim, K. C. 2010. Expression patterns of PRDM10 during mouse embryonic development. *BMB reports*, **43**(1), 29–33.
- Parker, K. C., Bednarek, M. A., & Coligan, J. E. 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *Journal of Immunology (Baltimore, Md.: 1950)*, **152**(1), 163–175.
- Paschen, A., Méndez, R. M., Jimenez, P., Sucker, A., Ruiz-Cabello, F., Song, M., Garrido, F., & Schadendorf, D. 2003. Complete loss of HLA class I antigen expression on melanoma cells: a result of successive mutational events. *International Journal of Cancer. Journal International Du Cancer*, **103**(6), 759–767.
- Paschen, A., Arens, N., Sucker, A., Greulich-Bode, K. M., Fonsatti, E., Gloghini, A., Striegel, S., Schwinn, N., Carbone, A., Hildenbrand, R., Cerwenka, A., Maio, M., & Schadendorf, D. 2006. The coincidence of chromosome 15 aberrations and beta2-microglobulin gene mutations is causative for the total loss of human leukocyte antigen class I expression in melanoma. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, **12**(11 Pt 1), 3297–3305.
- Paulsen, M. T., Veloso, A., Prasad, J., Bedi, K., Ljungman, E. A., Magnuson, B., Wilson, T. E., & Ljungman, M. 2014. Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods (San Diego, Calif.)*, **67**(1), 45–54.
- Peggs, K. S., Quezada, S. A., Chambers, C. A., Korman, A. J., & Allison, J. P. 2009. Blockade of CTLA-4 on both effector and regulatory T cell compartments contributes to the antitumor activity of anti-CTLA-4 antibodies. *The Journal of Experimental Medicine*, **206**(8), 1717–1725.
- Peng, Y., Schwarz, E. J., Lazar, M. A., Genin, A., Spinner, N. B., & Taub, R. 1997. Cloning, human chromosomal assignment, and adipose and hepatic expression of the CL-6/INSIG1 gene. *Genomics*, **43**(3), 278–284.
- Peters, B., & Sette, A. 2005. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC bioinformatics*, **6**, 132.
- Piwko, W., Olma, M. H., Held, M., Bianco, J. N., Pedrioli, P. G. A., Hofmann, K., Pasero, P., Gerlich, D. W., & Peter, M. 2010. RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. *The EMBO journal*, **29**(24), 4210–4222.
- Puntervoll, H. E., Yang, X. R., Vetti, H. H., Bachmann, I. M., Avril, M. F., Benfodda, M., Catricalà, C., Dalle, S., Duval-Modeste, A. B., Ghiorzo, P., Grammatico, P., Harland,

- M., Hayward, N. K., Hu, H.-H., Jouary, T., Martin-Denavit, T., Ozola, A., Palmer, J. M., Pastorino, L., Pjanova, D., Soufir, N., Steine, S. J., Stratigos, A. J., Thomas, L., Tinat, J., Tsao, H., Veinalde, R., Tucker, M. A., Bressac-de Paillerets, B., Newton-Bishop, J. A., Goldstein, A. M., Akslén, L. A., & Molven, A. 2013. Melanoma prone families with CDK4 germline mutation: phenotypic profile and associations with MC1R variants. *Journal of Medical Genetics*, **50**(4), 264–270.
- Quinlan, A. R., & Hall, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, **26**(6), 841–842.
- Rajasagi, M., Shukla, S. A., Fritsch, E. F., Keskin, D. B., DeLuca, D., Carmona, E., Zhang, W., Sougnez, C., Cibulskis, K., Sidney, J., Stevenson, K., Ritz, J., Neubergh, D., Brusica, V., Gabriel, S., Lander, E. S., Getz, G., Hachohen, N., & Wu, C. J. 2014. Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood*, **124**(3), 453–462.
- Rammensee, H.-G., & Singh-Jasuja, H. 2013. HLA ligandome tumor antigen discovery for personalized vaccine approach. *Expert Review of Vaccines*, **12**(10), 1211–1217.
- Rammensee, H.-G., Bachmann, J., Emmerich, N. P., Bachor, O. A., & Stevanović, S. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, **50**(3-4), 213–219.
- Rebaï, N., & Malissen, B. 1983. Structural and genetic analyses of HLA class I molecules using monoclonal xenobodies. *Tissue Antigens*, **22**(2), 107–117.
- Restifo, N. P., Esquivel, F., Kawakami, Y., Yewdell, J. W., Mulé, J. J., Rosenberg, S. A., & Bennink, J. R. 1993. Identification of human cancers deficient in antigen processing. *The Journal of Experimental Medicine*, **177**(2), 265–272.
- Rezvani, K., Yong, A., Tawab, A., Jafarpour, B., Eniafe, R., Mielke, S., Savani, B. N., Keyvanfar, K., Li, Y., Kurlander, R., & Barrett, A. J. 2009. Ex vivo characterization of polyclonal memory CD8+ T-cell responses to PRAME-specific peptides in patients with acute lymphoblastic leukemia and acute and chronic myeloid leukemia. *Blood*, **113**(10), 2245–2255.
- Robasky, K., Lewis, N. E., & Church, G. M. 2014. The role of replicates for error mitigation in next-generation sequencing. *Nature Reviews. Genetics*, **15**(1), 56–62.
- Robbins, P. F., Lu, Y.-C., El-Gamil, M., Li, Y. F., Gross, C., Gartner, J., Lin, J. C., Teer, J. K., Clifton, P., Tycksen, E., Samuels, Y., & Rosenberg, S. A. 2013. Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nature Medicine*, **19**(6), 747–752.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O. L., He, A., Marra, M., Snyder, M., & Jones, S. 2007. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods*, **4**(8), 651–657.
- Rodenko, B., Toebe, M., Hadrup, S. R., van Esch, W. J. E., Molenaar, A. M., Schumacher, T. N. M., & Ovaas, H. 2006. Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nature Protocols*, **1**(3), 1120–1132.
- Roelse, J., Grommé, M., Momburg, F., Hämmerling, G., & Neefjes, J. 1994. Trimming of TAP-translocated peptides in the endoplasmic reticulum and in the cytosol during recycling. *The Journal of Experimental Medicine*, **180**(5), 1591–1597.

- 
- Rosenberg, S. A., Packard, B. S., Aebersold, P. M., Solomon, D., Topalian, S. L., Toy, S. T., Simon, P., Lotze, M. T., Yang, J. C., & Seipp, C. A. 1988. Use of tumor-infiltrating lymphocytes and interleukin-2 in the immunotherapy of patients with metastatic melanoma. A preliminary report. *The New England Journal of Medicine*, **319**(25), 1676–1680.
- Rosenberg, S. A., Yang, J. C., Topalian, S. L., Schwartzentruber, D. J., Weber, J. S., Parkinson, D. R., Seipp, C. A., Einhorn, J. H., & White, D. E. 1994. Treatment of 283 consecutive patients with metastatic melanoma or renal cell cancer using high-dose bolus interleukin 2. *JAMA: the journal of the American Medical Association*, **271**(12), 907–913.
- Röttschke, O., Falk, K., Deres, K., Schild, H., Norda, M., Metzger, J., Jung, G., & Rammensee, H. G. 1990. Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. *Nature*, **348**(6298), 252–254.
- Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T., & Cresswell, P. 1996. Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity*, **5**(2), 103–114.
- Sahin, U., Türeci, O., Schmitt, H., Cochlovius, B., Johannes, T., Schmits, R., Stenner, F., Luo, G., Schobert, I., & Pfreundschuh, M. 1995. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proceedings of the National Academy of Sciences of the United States of America*, **92**(25), 11810–11813.
- Sanger, F., Nicklen, S., & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, **74**(12), 5463–5467.
- Schirle, M., Keilholz, W., Weber, B., Gouttefangeas, C., Dumrese, T., Becker, H. D., Stevanović, S., & Rammensee, H. G. 2000. Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *European Journal of Immunology*, **30**(8), 2216–2225.
- Schwartz, R. H., Mueller, D. L., Jenkins, M. K., & Quill, H. 1989. T-cell clonal anergy. *Cold Spring Harbor Symposia on Quantitative Biology*, **54 Pt 2**, 605–610.
- Schwinn, N., Vokhminova, D., Sucker, A., Textor, S., Striegel, S., Moll, I., Nausch, N., Tuettenberg, J., Steinle, A., Cerwenka, A., Schadendorf, D., & Paschen, A. 2009. Interferon-gamma down-regulates NKG2D ligand expression and impairs the NKG2D-mediated cytotoxicity of MHC class I-deficient melanoma by natural killer cells. *International Journal of Cancer. Journal International Du Cancer*, **124**(7), 1594–1604.
- Sekiya, T., Hirohashi, S., Nishimura, S., & Sugimura, T. 1983. Transforming activity of human melanoma DNA. *Gann = Gan*, **74**(6), 794–797.
- Seliger, B., Höhne, A., Knuth, A., Bernhard, H., Ehring, B., Tampé, R., & Huber, C. 1996. Reduced membrane major histocompatibility complex class I density and stability in a subset of human renal cell carcinomas with low TAP and LMP expression. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, **2**(8), 1427–1433.
- Senzer, N. N., Kaufman, H. L., Amatruda, T., Nemunaitis, M., Reid, T., Daniels, G., Gonzalez, R., Glaspy, J., Whitman, E., Harrington, K., Goldsweig, H., Marshall, T., Love, C., Coffin, R., & Nemunaitis, J. J. 2009. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. *Journal of Clinical*



- Oncology: Official Journal of the American Society of Clinical Oncology*, **27**(34), 5763–5771.
- Shankaran, V., Ikeda, H., Bruce, A. T., White, J. M., Swanson, P. E., Old, L. J., & Schreiber, R. D. 2001. IFN $\gamma$  and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature*, **410**(6832), 1107–1111.
- Shapira-Frommer, R., & Schachter, J. 2012. Adoptive immunotherapy of advanced melanoma. *Current Treatment Options in Oncology*, **13**(3), 340–353.
- Shinkai, Y., Rathbun, G., Lam, K. P., Oltz, E. M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., & Stall, A. M. 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell*, **68**(5), 855–867.
- Sidney, J., Southwood, S., & Sette, A. 2005. Classification of A1- and A24-supertype molecules by analysis of their MHC-peptide binding repertoires. *Immunogenetics*, **57**(6), 393–408.
- Sidney, J., Assarsson, E., Moore, C., Ngo, S., Pinilla, C., Sette, A., & Peters, B. 2008. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *ImmunoMe Research*, **4**, 2.
- Siegel, D. A., Huang, M. K., & Becker, S. F. 2002. Ectopic dendrite initiation: CNS pathogenesis as a model of CNS development. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, **20**(3-5), 373–389.
- Simpson, T. R., Li, F., Montalvo-Ortiz, W., Sepulveda, M. A., Bergerhoff, K., Arce, F., Roddie, C., Henry, J. Y., Yagita, H., Wolchok, J. D., Peggs, K. S., Ravetch, J. V., Allison, J. P., & Quezada, S. A. 2013. Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. *The Journal of Experimental Medicine*, **210**(9), 1695–1710.
- Skipper, J. C., Gulden, P. H., Hendrickson, R. C., Harthun, N., Caldwell, J. A., Shabanowitz, J., Engelhard, V. H., Hunt, D. F., & Slingluff, C. L. 1999. Mass-spectrometric evaluation of HLA-A\*0201-associated peptides identifies dominant naturally processed forms of CTL epitopes from MART-1 and gp100. *International Journal of Cancer. Journal International Du Cancer*, **82**(5), 669–677.
- Skolnick, M. H., Cannon-Albright, L. A., & Kamb, A. 1994. Genetic predisposition to melanoma. *European Journal of Cancer (Oxford, England: 1990)*, **30A**(13), 1991–1995.
- Slota, M., Lim, J.-B., Dang, Y., & Disis, M. L. 2011. ELISpot for measuring human immune responses to vaccines. *Expert Review of Vaccines*, **10**(3), 299–306.
- Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J. M., Desrichard, A., Walsh, L. A., Postow, M. A., Wong, P., Ho, T. S., Hollmann, T. J., Bruggeman, C., Kannan, K., Li, Y., Elipenahli, C., Liu, C., Harbison, C. T., Wang, L., Ribas, A., Wolchok, J. D., & Chan, T. A. 2014. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *The New England Journal of Medicine*, Nov.
- Sondak, V. K., Smalley, K. S. M., Kudchadkar, R., Grippon, S., & Kirkpatrick, P. 2011. Ipilimumab. *Nature Reviews. Drug Discovery*, **10**(6), 411–412.
- Sørensen, R. B., Junker, N., Kirkin, A., Voigt, H., Svane, I. M., Becker, J. C., Straten, P. T., & Andersen, M. H. 2009. The immunodominant HLA-A2-restricted MART-1 epitope is not presented on the surface of many melanoma cell lines. *Cancer immunology, immunotherapy: CII*, **58**(5), 665–675.

- 
- Spencer, D. H., Abel, H. J., Lockwood, C. M., Payton, J. E., Szankasi, P., Kelley, T. W., Kulkarni, S., Pfeifer, J. D., & Duncavage, E. J. 2013. Detection of FLT3 internal tandem duplication in targeted, short-read-length, next-generation sequencing data. *The Journal of molecular diagnostics: JMD*, **15**(1), 81–93.
- Spitler, L. E., Grossbard, M. L., Ernstoff, M. S., Silver, G., Jacobs, M., Hayes, F. A., & Soong, S. J. 2000. Adjuvant therapy of stage III and IV malignant melanoma using granulocyte-macrophage colony-stimulating factor. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **18**(8), 1614–1621.
- Stratton, M. R., Campbell, P. J., & Futreal, P. A. 2009. The cancer genome. *Nature*, **458**(7239), 719–724.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**(10), 2731–2739.
- Tanguay, S., & Killion, J. J. 1994. Direct comparison of ELISPOT and ELISA-based assays for detection of individual cytokine-secreting cells. *Lymphokine and Cytokine Research*, **13**(4), 259–263.
- Taswell, C., MacDonald, H. R., & Cerottini, J. C. 1980. Clonal analysis of cytolytic T lymphocyte specificity. I. Phenotypically distinct sets of clones as the cellular basis of cross-reactivity to alloantigens. *The Journal of Experimental Medicine*, **151**(6), 1372–1385.
- Taube, J. M., Klein, A., Brahmer, J. R., Xu, H., Pan, X., Kim, J. H., Chen, L., Pardoll, D. M., Topalian, S. L., & Anders, R. A. 2014. Association of PD-1, PD-1 Ligands, and Other Features of the Tumor Immune Microenvironment with Response to Anti-PD-1 Therapy. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, Apr.
- Taylor, K. H., Kramer, R. S., Davis, J. W., Guo, J., Duff, D. J., Xu, D., Caldwell, C. W., & Shi, H. 2007. Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Research*, **67**(18), 8511–8518.
- Teppo, L., Pakkanen, M., & Hakulinen, T. 1978. Sunlight as a risk factor of malignant melanoma of the skin. *Cancer*, **41**(5), 2018–2027.
- Thompson, J. A., Curti, B. D., Redman, B. G., Bhatia, S., Weber, J. S., Agarwala, S. S., Sievers, E. L., Hughes, S. D., DeVries, T. A., & Hausman, D. F. 2008. Phase I study of recombinant interleukin-21 in patients with metastatic melanoma and renal cell carcinoma. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **26**(12), 2034–2039.
- Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, **14**(2), 178–192.
- Topalian, S. L., Hodi, F. S., Brahmer, J. R., Gettinger, S. N., Smith, D. C., McDermott, D. F., Powderly, J. D., Carvajal, R. D., Sosman, J. A., Atkins, M. B., Leming, P. D., Spigel, D. R., Antonia, S. J., Horn, L., Drake, C. G., Pardoll, D. M., Chen, L., Sharfman, W. H., Anders, R. A., Taube, J. M., McMiller, T. L., Xu, H., Korman, A. J., Jure-Kunkel, M., Agrawal, S., McDonald, D., Kollia, G. D., Gupta, A., Wigginton, J. M., & Sznol, M. 2012. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *The New England Journal of Medicine*, **366**(26), 2443–2454.
-

- Tran, E., Turcotte, S., Gros, A., Robbins, P. F., Lu, Y.-C., Dudley, M. E., Wunderlich, J. R., Somerville, R. P., Hogan, K., Hinrichs, C. S., Parkhurst, M. R., Yang, J. C., & Rosenberg, S. A. 2014. Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. *Science (New York, N.Y.)*, **344**(6184), 641–645.
- Trapnell, C., Pachter, L., & Salzberg, S. L. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)*, **25**(9), 1105–1111.
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., & Pachter, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, **28**(5), 511–515.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., & Pachter, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, **7**(3), 562–578.
- Tschiedel, S., Gentilini, C., Lange, T., Wölfel, C., Wölfel, T., Lennerz, V., Stevanovic, S., Rammensee, H.-G., Huber, C., Cross, M., & Niederwieser, D. 2008. Identification of NM23-H2 as a tumour-associated antigen in chronic myeloid leukaemia. *Leukemia*, **22**(8), 1542–1550.
- Türeci, O., Sahin, U., Schobert, I., Koslowski, M., Scmitt, H., Schild, H. J., Stenner, F., Seitz, G., Rammensee, H. G., & Pfreundschuh, M. 1996. The SSX-2 gene, which is involved in the t(X;18) translocation of synovial sarcomas, codes for the human tumor antigen HOM-MEL-40. *Cancer Research*, **56**(20), 4766–4772.
- van Delft, J., Gaj, S., Lienhard, M., Albrecht, M. W., Kirpiy, A., Brauers, K., Claessen, S., Lizarraga, D., Lehrach, H., Herwig, R., & Kleinjans, J. 2012. RNA-Seq provides new insights in the transcriptome responses induced by the carcinogen benzo[a]pyrene. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, **130**(2), 427–439.
- van der Bruggen, P., Stroobant, V., Vigneron, N., & Van den Eynde, B. J. 2013. *Peptide database: T cell-defined tumor antigens*.
- van Rooij, N., van Buuren, M. M., Philips, D., Velds, A., Toebes, M., Heemskerk, B., van Dijk, L. J. A., Behjati, S., Hilkmann, H., El Atmioui, D., Nieuwland, M., Stratton, M. R., Kerkhoven, R. M., Kesmir, C., Haanen, J. B., Kvistborg, P., & Schumacher, T. N. M. 2013. Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **31**(32), e439–442.
- Veillette, A., Bookman, M. A., Horak, E. M., & Bolen, J. B. 1988. The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase p56lck. *Cell*, **55**(2), 301–308.
- Velculescu, V. E., Madden, S. L., Zhang, L., Lash, A. E., Yu, J., Rago, C., Lal, A., Wang, C. J., Beaudry, G. A., Ciriello, K. M., Cook, B. P., Dufault, M. R., Ferguson, A. T., Gao, Y., He, T. C., Hermeking, H., Hiraldo, S. K., Hwang, P. M., Lopez, M. A., Luderer, H. F., Mathews, B., Petroziello, J. M., Polyak, K., Zawel, L., & Kinzler, K. W. 1999. Analysis of human transcriptomes. *Nature Genetics*, **23**(4), 387–388.
- Viatte, S., Alves, P. M., & Romero, P. 2006. Reverse immunology approach for the identification of CD8 T-cell-defined antigens: advantages and hurdles. *Immunology and Cell Biology*, **84**(3), 318–330.
- Vigneron, N., Stroobant, V., Van den Eynde, B. J., & van der Bruggen, P. 2013. Database of T cell-defined human tumor antigens: the 2013 update. *Cancer Immunity*, **13**, 15.

- 
- Vignot, S., Frampton, G. M., Soria, J.-C., Yelensky, R., Commo, F., Brambilla, C., Palmer, G., Moro-Sibilot, D., Ross, J. S., Cronin, M. T., André, F., Stephens, P. J., Lazar, V., Miller, V. A., & Brambilla, E. 2013. Next-generation sequencing reveals high concordance of recurrent somatic alterations between primary tumor and metastases from patients with non-small-cell lung cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **31**(17), 2167–2172.
- Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., & Kinzler, K. W. 2013. Cancer genome landscapes. *Science (New York, N.Y.)*, **339**(6127), 1546–1558.
- Vroom, T. M., Scholte, G., Ossendorp, F., & Borst, J. 1991. Tissue distribution of human gamma delta T cells: no evidence for general epithelial tropism. *Journal of Clinical Pathology*, **44**(12), 1012–1017.
- Walshe, V. A., Hattotuwigama, C. K., Doytchinova, I. A., Wong, M., Macdonald, I. K., Mulder, A., Claas, F. H. J., Pellegrino, P., Turner, J., Williams, I., Turnbull, E. L., Borrow, P., & Flower, D. R. 2009. Integrating in silico and in vitro analysis of peptide binding affinity to HLA-Cw\*0102: a bioinformatic approach to the prediction of new epitopes. *PloS One*, **4**(11), e8095.
- Walter, S., Weinschenk, T., Stenzl, A., Zdrojowy, R., Pluzanska, A., Szczylik, C., Staehler, M., Brugger, W., Dietrich, P.-Y., Mendrzyk, R., Hilf, N., Schoor, O., Fritsche, J., Mahr, A., Maurer, D., Vass, V., Trautwein, C., Lewandrowski, P., Flohr, C., Pohla, H., Stanczak, J. J., Bronte, V., Mandruzzato, S., Biedermann, T., Pawelec, G., Derhovanessian, E., Yamagishi, H., Miki, T., Hongo, F., Takaha, N., Hirakawa, K., Tanaka, H., Stevanovic, S., Frisch, J., Mayer-Mokler, A., Kirner, A., Rammensee, H.-G., Reinhardt, C., & Singh-Jasuja, H. 2012. Multi-peptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. *Nature Medicine*, **18**(8), 1254–1261.
- Wang, L., Lawrence, M., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D. S., Zhang, L., Zhang, W., Vartanov, A. R., Fernandes, S. M., Goldstein, N. R., Folco, E. G., Cibulskis, K., Tesar, B., Sievers, Q. L., Shefler, E., Gabriel, S., Hacohen, N., Reed, R., Meyerson, M., Golub, T. R., Lander, E. S., Neubergh, D., Brown, J. R., Getz, G., & Wu, C. J. 2011. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *The New England Journal of Medicine*, **365**(26), 2497–2506.
- Wei, X., Walia, V., Lin, J. C., Teer, J. K., Prickett, T. D., Gartner, J., Davis, S., NISC Comparative Sequencing Program, Stemke-Hale, K., Davies, M. A., Gershenwald, J. E., Robinson, W., Robinson, S., Rosenberg, S. A., & Samuels, Y. 2011. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nature Genetics*, **43**(5), 442–446.
- Weinlich, G., Murr, C., Richardsen, L., Winkler, C., & Fuchs, D. 2007. Decreased serum tryptophan concentration predicts poor prognosis in malignant melanoma patients. *Dermatology (Basel, Switzerland)*, **214**(1), 8–14.
- Weinschenk, T., Gouttefangeas, C., Schirle, M., Obermayr, F., Walter, S., Schoor, O., Kurek, R., Loeser, W., Bichler, K.-H., Wernet, D., Stevanović, S., & Rammensee, H.-G. 2002. Integrated functional genomics approach for the design of patient-individual antitumor vaccines. *Cancer Research*, **62**(20), 5818–5827.
- Weiss, S. A., Chandra, S., & Pavlick, A. C. 2014. Update on vaccines for high-risk melanoma. *Current Treatment Options in Oncology*, **15**(2), 269–280.
- Whiteman, D. C., Pavan, W. J., & Bastian, B. C. 2011. The melanomas: a synthesis of epidemiological, clinical, histopathological, genetic, and biological aspects, supporting

- distinct subtypes, causal pathways, and cells of origin. *Pigment Cell & Melanoma Research*, **24**(5), 879–897.
- Wick, D. A., Webb, J. R., Nielsen, J. S., Martin, S. D., Kroeger, D. R., Milne, K., Castellarin, M., Twumasi-Boateng, K., Watson, P. H., Holt, R. A., & Nelson, B. H. 2014. Surveillance of the tumor mutanome by T cells during progression from primary to recurrent ovarian cancer. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, **20**(5), 1125–1134.
- Wolchok, J. D., Kluger, H., Callahan, M. K., Postow, M. A., Rizvi, N. A., Lesokhin, A. M., Segal, N. H., Ariyan, C. E., Gordon, R.-A., Reed, K., Burke, M. M., Caldwell, A., Kronenberg, S. A., Agunwamba, B. U., Zhang, X., Lowy, I., Inzunza, H. D., Feely, W., Horak, C. E., Hong, Q., Korman, A. J., Wigginton, J. M., Gupta, A., & Sznol, M. 2013. Nivolumab plus ipilimumab in advanced melanoma. *The New England Journal of Medicine*, **369**(2), 122–133.
- Wu, J., Feng, F., Yang, D., Yu, S., Liu, J., & Gao, Z. 2014. Investigation of key genes associated with prostate cancer using RNA-seq data. *The International Journal of Biological Markers*, **29**(1), e86–92.
- Xia, J., Jia, P., Hutchinson, K. E., Dahlman, K. B., Johnson, D., Sosman, J., Pao, W., & Zhao, Z. 2014. A meta-analysis of somatic mutations from next generation sequencing of 241 melanomas: a road map for the study of genes with potential clinical relevance. *Molecular Cancer Therapeutics*, **13**(7), 1918–1928.
- Xu, X., Zhu, K., Liu, F., Wang, Y., Shen, J., Jin, J., Wang, Z., Chen, L., Li, J., & Xu, M. 2013. Identification of somatic mutations in human prostate cancer by RNA-Seq. *Gene*, **519**(2), 343–347.
- Yang, T., Espenshade, P. J., Wright, M. E., Yabe, D., Gong, Y., Aebersold, R., Goldstein, J. L., & Brown, M. S. 2002. Crucial step in cholesterol homeostasis: sterols promote binding of SCAP to INSIG-1, a membrane protein that facilitates retention of SREBPs in ER. *Cell*, **110**(4), 489–500.
- Yao, L., Nakauchi, H., Honjo, T., & Kawakami, T. 1990. The cytoplasmic domain of the CD8 alpha-chain is required for its interaction with p56lck. *Immunology Letters*, **24**(4), 267–271.
- Yeung, J. T., Hamilton, R. L., Ohnishi, K., Ikeura, M., Potter, D. M., Nikiforova, M. N., Ferrone, S., Jakacki, R. I., Pollack, I. F., & Okada, H. 2013. LOH in the HLA class I region at 6p21 is associated with shorter survival in newly diagnosed adult glioblastoma. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, **19**(7), 1816–1826.
- Yewdell, J. W., Reits, E., & Neefjes, J. 2003. Making sense of mass destruction: quantitating MHC class I antigen presentation. *Nature Reviews. Immunology*, **3**(12), 952–961.
- Yu, H., Mashtalir, N., Daou, S., Hammond-Martel, I., Ross, J., Sui, G., Hart, G. W., Rauscher, F. J., Drobetsky, E., Milot, E., Shi, Y., & Affar, E. B. 2010. The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Molecular and Cellular Biology*, **30**(21), 5071–5085.
- Zhang, L., Conejo-Garcia, J. R., Katsaros, D., Gimotty, P. A., Massobrio, M., Regnani, G., Makrigiannakis, A., Gray, H., Schlienger, K., Liebman, M. N., Rubin, S. C., & Coukos, G. 2003. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *The New England Journal of Medicine*, **348**(3), 203–213.

- 
- Zhang, Q., Wang, P., Kim, Y., Haste-Andersen, P., Beaver, J., Bourne, P. E., Bui, H.-H., Buus, S., Frankild, S., Greenbaum, J., Lund, O., Lundegaard, C., Nielsen, M., Ponomarenko, J., Sette, A., Zhu, Z., & Peters, B. 2008. Immune epitope database analysis resource (IEDB-AR). *Nucleic Acids Research*, **36**(Web Server issue), W513–518.
- Zhou, J., Dudley, M. E., Rosenberg, S. A., & Robbins, P. F. 2005. Persistence of multiple tumor-specific T-cell clones is associated with complete tumor regression in a melanoma patient receiving adoptive cell transfer therapy. *Journal of Immunotherapy (Hagerstown, Md.: 1997)*, **28**(1), 53–62.

# Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst und nur mit den angegebenen Hilfsmitteln angefertigt habe. Die Dissertation wurde weder in gleicher noch in ähnlicher Form an einer anderen Universität oder in einem anderen Fachbereich vorgelegt.

---

Ort, Datum

---

Barbara Schrörs





# Danksagung



# Lebenslauf

