

**Molekulare Untersuchung des Wolf-
Hirschhorn-Syndroms**

-

**Transkriptidentifizierung und
Gencharakterisierung in der distalen
Chromosomenregion 4p16.3**

Dissertation

**Zur Erlangung des Grades
„Doktor der Naturwissenschaften“**

**am Fachbereich Biologie
der Johannes Gutenberg-Universität
in Mainz**

**Sabine Endeke
geb. in Laupheim (Württemberg)**

Mainz, 2000

Dekan: Prof. Dr. J. Markl

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung: 26.04.2001

1 EINLEITUNG	1
2 MATERIAL UND METHODEN	12
2.1 Herkunft Chromosom 4-spezifischer Cosmidklone	12
2.2 Isolierung von DNA	12
2.2.1 Isolierung von genomischer DNA aus Blut und Zellkulturen	12
2.2.2 Isolierung kleinerer Mengen Plasmid- und Cosmid-DNA	13
2.2.3 Isolierung großer Mengen Plasmid-, Cosmid-, P1- und PAC-DNA	14
2.3 Isolierung von RNA	14
2.4 DNA und RNA Standardmethoden	14
2.4.1 Verdauung von DNA mit Restriktionsendonukleasen	14
2.4.2 Agarose-Gelelektrophorese von DNA und RNA	15
2.4.3 Denaturierende Polyacrylamidgelelektrophorese	16
2.4.4 Wiedergewinnung von DNA aus Agarosegelen	16
2.4.5 Fällung von DNA und RNA	17
2.5 Subklonierung von DNA-Fragmenten in Plasmid-Vektoren	17
2.6 Polymerase-Kettenreaktion (PCR)	18
2.7 RT-PCR (Reverse Transkriptase-Polymerasekettenreaktion)	19
2.8 Radioaktive und nichtradioaktive Markierungstechniken	20
2.8.1 „Random primed oligo labelling“	20
2.8.2 „Nick-Translation“	20
2.8.3 Endmarkierung von Oligonukleotiden und PCR-Produkten	21
2.8.3.1 Endmarkierung von Oligonukleotiden	21
2.8.3.2 Endmarkierung von PCR-Produkten	21
2.8.4 „Run off“-Polymerisation (Lineare PCR)	22
2.9 Filterhybridisierung	22
2.9.1 DNA-DNA-Hybridisierung nach Southern	22
2.9.2 DNA-RNA-Hybridisierung	23
2.9.3 Plaquefilterhybridisierung	23
2.9.4 Koloniefilterhybridisierung	24

2.10 <i>in situ</i>-Hybridisierung	24
2.10.1 Fluoreszenz- <i>in situ</i> -Hybridisierung (FISH) von Metaphasechromosomen	24
2.10.2 RNA- <i>in situ</i> -Hybridisierung von Gewebeschnitten	26
2.10.2.1 Herstellung von Schnittpräparaten	26
2.10.2.2 Hybridisierung	27
2.11 Isolierung von PAC- bzw. P1-Klonen	28
2.12 Exonamplifikation	29
2.12.1 Subklonierung von Cosmid- und PAC-Klonen in den Vektor pSPL3	29
2.12.2 Kultur von COS7-Zellen	30
2.12.3 Elektroporation	31
2.12.4 RNA-Isolation aus COS7-Zellen	31
2.12.5 Synthese der cDNA	32
2.12.6 Amplifikation der DNA-Fragmente und Klonierung in pBluescript	32
2.12.7 Identifizierung von „echten“ Exonsequenzen	33
2.13 Herstellung und Screening einer hierarchisch angeordneten fetalen Hirn-cDNA-Bibliothek mittels PCR	35
2.14 <i>In vivo</i>-Excision von pBluescript aus λZapII mit Hilfe des ExAssist™/SOLR™ Systems	37
2.15 DNA-Sequenzierung	37
2.16 Computerauswertung von DNA- und Protein-Sequenzen	38
2.17 Reagenzien und Materialien	39
3 ERGEBNISSE	51
3.1 Exonamplifikation in der distalen Chromosomenregion 4p16.3	51
3.1.1 Auswahl von Cosmid-Klonen und Isolierung von PAC-Klonen zwischen den Markern D4S43 D4S142	52
3.1.2 Identifizierung transkribierter Bereiche durch Exonamplifikation	56
3.1.3 Homologievergleiche der Exonsequenzen	66
3.1.3.1 Exonsequenzen aus dem Contig I	66
3.1.3.2 Exonsequenzen aus dem Contig II	71
3.1.3.3 Exonsequenzen aus dem Contig III	75

3.2 Identifizierung und Charakterisierung korrespondierender cDNA-Klone	77
3.2.1 Der cDNA-Klon pFL12	77
3.2.2 Die cDNA-Klone λ CF16 und H7NIB	78
3.3 Molekulare Klonierung und chromosomale Lokalisation des humanen und murinen <i>LETMI (Letm1)</i>-Gens	79
3.3.1 Klonierung und Sequenzierung der kodierenden Region des menschlichen <i>LETMI</i> -Gens	79
3.3.2 Klonierung und Sequenzierung der kodierenden Region des murinen <i>Letm1</i> -Gens	80
3.3.3 Homologievergleiche der beiden Gene	82
3.3.4 Genomische Struktur des humanen und murinen <i>LETMI/Letm1</i> -Gens	88
3.3.5 Expressionsanalysen	91
3.3.5.1 Northern-Blot-Analyse	91
3.3.5.2 RNA- <i>in situ</i> -Hybridisierung an Gewebeschnitten	93
3.3.6 Identifizierung und Sequenzierung des prozessierten <i>LETMI</i> -Pseudogens	94
3.4 Detaillierte Analyse der WHS-kritischen Region zwischen D4S43 und dem <i>LETMI</i>-Gen	97
3.4.1 Identifizierung transkribierter Bereiche durch computergestützte Analysen	97
3.4.2 Charakterisierung der identifizierten Transkripte	98
3.4.2.1 Das <i>Gen 51 (WHSC1)</i>	105
3.4.2.2 Das <i>Gen 43 (WHSC2)</i>	108
3.4.2.3 Das <i>Gen 45</i>	114
3.4.2.4 Das <i>Gen 57</i>	123
3.4.2.5 Das <i>POLAP</i> -Gen	126
3.4.2.6 Weitere mögliche Transkripte des Menschen und der Maus	129
3.5 Untersuchung eines WHS-Patienten ohne cytogenetisch nachweisbare Deletion	130
3.6 Charakterisierung der WHSCR-orthologen Region der Maus	133
4. DISKUSSION	138
4.1 Identifizierung transkribierter Bereiche	139
4.1.1 Exonamplifikation	139
4.1.2 Detaillierte Untersuchung der WHSCR durch computergestützte Analysen	143
4.2 Das <i>LETMI</i>-Gen	145
4.3 Das <i>Gen 51 (WHSC1)</i>	151

4.4 Das Gen 43 (WHSC2)	152
4.5 Das Gen 45	154
4.6 Die Gene POLAP und 57	159
4.7 Mutationsscreening von Kandidatengenen bei einem WHS-Patienten ohne Deletion	160
4.8 Zusammenfassende Beurteilung der identifizierten Gene	161
4.9 Charakterisierung der WHSCR-orthologen Region der Maus	163
5. ZUSAMMENFASSUNG	166
6. LITERATURVERZEICHNIS	168
7. ANHANG	185
7.1 Danksagung	185
7.2 Lebenslauf	187
7.3 Veröffentlichungen	189

Verzeichnis der Abbildungen und Tabellen:

Abb. 1.1:	Übersicht über die distalen 3,5 Mb der ca. 5 Mb umfassenden Chromosomenregion 4p16.3 (verändert nach Riess et al., 1996).	5
Abb. 2.1:	Strategie der Subklonanalyse.	33
Abb. 2.2:	Vereinigungs- (Pooling) Strategie für die Erstellung der hierarchisch angeordneten cDNA-Bibliothek.	36
Abb. 3.1:	Gelelektrophoretische Auftrennung von Cosmid- und PAC-Klonen der distalen Region 4p16.3.	53
Abb. 3.2:	FISH-Analysen isolierter PAC-Klone.	54
Abb. 3.3:	Physikalische Karte der distalen Chromosomenregion 4p16.3 und Einteilung der isolierten PAC-Klone in die Contigs I-III.	55
Abb. 3.4:	BLASTX-Homologievergleich des Exons md.	66
Abb. 3.5:	Vergleich der Exonssequenzen 3A.B6 und 2A.A12 mit den cDNA-Sequenzen (5'-UTR) des <i>ZNF141</i> - und des <i>ZNF1</i> -Gens sowie der genomischen Sequenz HS4PTEL (Z95704).	68
Abb. 3.6.:	BLASTX-Homologievergleich des Exons 2B.G11.	68
Abb. 3.7.:	Nukleotidsequenz des <i>WEE1hu</i> -Pseudogens und der flankierenden genomischen Sequenzen.	70
Abb. 3.8:	Zeigt die cDNA-Sequenz des <i>DAGK4</i> -Gens (Schaap et al., unveröffentlicht).	73
Abb. 3.9:	BLASTX-Homologievergleich des Exons 6A.B12 mit der Proteinsequenz des murinen Troponin C (Acc.: P20801).	76
Abb. 3.10:	Nukleotidsequenz des cDNA-Klons pFL12.	78
Abb. 3.11:	Zeigt die Klonierungsstrategie des menschlichen <i>LETM1</i> -Gens.	79
Abb. 3.12:	Klonierungsstrategie des murinen <i>Letm1</i> -Gens.	81
Abb. 3.13:	Homologievergleich der cDNA-Konsensussequenzen und abgeleiteter Aminosäuresequenzen des humanen und murinen <i>LETM1/Letm1</i> -Gens (Accession-Nr. AF061026 und AF61025).	85
Abb. 3.14:	Vergleich der identifizierten EF-Hand-Motive des LETM1-Proteins mit der „PROSITE“-Konsensussequenz calciumbindender Proteine.	86
Abb. 3.15:	Zeigt den Homologievergleich der Proteinsequenzen des humanen und murinen LETM1s mit verschiedenen Spezies.	87
Abb. 3.16:	Übersicht über die Kartierung des humanen <i>LETM1</i> -Gens bezüglich des Contigs III und des Cosmid-Contigs nach Baxendale et al. (1993).	90

Abb. 3.17:	Northern-Blot-Analyse des humanen <i>LETMI</i> -Gens.	92
Abb. 3.18:	Northern-Blot-Analyse des murinen <i>Letm1</i> -Gens.	92
Abb. 3.19:	RNA- <i>in situ</i> -Hybridisierung an Gewebeschnitten embryonaler Mäuse (Tag 13,5 post coitus).	93
Abb. 3.20:	Konsensussequenz des prozessierten <i>LETMI</i> -Pseudogens und der flankierenden genomischen Bereiche.	96
Abb. 3.21:	Physikalische Karte der Chromosomenregion 4p16.3 (D4S166 bis <i>FGFR3</i>).	103
Abb. 3.22:	Homologievergleich des murinen cDNA-Klons AA062233 (<i>Whsc1</i>) mit der cDNA-Sequenz und abgeleiteter Aminosäuresequenz des <i>WHSCI</i> -Gens (Stec et al., 1998).	106
Abb. 3.23:	Northern-Blot-Analyse des murinen <i>Whsc1</i> -Gens.	107
Abb. 3.24:	Homologievergleich der cDNA-Konsensussequenzen des humanen und murinen <i>Gens 43</i> und abgeleiteter Aminosäuresequenzen.	111
Abb. 3.25:	Northern-Blot-Analyse des murinen <i>Gens 43</i> (<i>Whsc2h</i>).	113
Abb. 3.26:	Homologievergleich der cDNA-Konsensussequenzen des humanen und murinen <i>Gens 45</i> und abgeleiteter Aminosäuresequenzen.	117
Abb. 3.27:	Northern-Blot- und Master-Blot-Analyse des murinen <i>Gens 45</i> .	119
Abb. 3.28:	Northern-Blot-Analyse des humanen <i>Gens 45</i> .	120
Abb. 3.29:	RNA- <i>in situ</i> -Hybridisierung an Gehirnschnitten einer adulten Maus.	121
Abb. 3.30:	RNA- <i>in situ</i> -Hybridisierung an transversalen Gewebeschnitten durch adultes Hodengewebe der Maus.	122
Abb. 3.31:	Nukleotid- und Aminosäuresequenzvergleich des humanen und murinen <i>eRFS</i> -Gens sowie des cDNA-Klons AA118417 mit der genomischen Sequenz des Cosmids 58b6.	124
Abb. 3.32:	Northern-Blot-Analyse des cDNA-Klons AA118417.	125
Abb. 3.33:	Proteinvergleich (BLASTX) der partiellen cDNA-Sequenz des humanen <i>POLAP</i> -Gens mit der DNA-Polymerase I aus <i>E. coli</i> (Joyce et al., 1982).	127
Abb. 3.34:	Homologievergleich (BLASTX) des murinen cDNA-Klons AA475002 mit der Proteinsequenz des humanen <i>POLAP</i> -Gens nach Harris et al. (1999).	128
Abb. 3.35A-D:	Mini- und Mikrosatellitenanalysen.	132
Abb. 3.36:	Überlappung muriner P1- und PAC-Klone.	135
Abb. 3.37:	Die WHSCR-Region in der Maus (<i>Mmu</i> 5.20).	136
Abb. 4.1:	Genomische Lokalisation des humanen <i>LETMI</i> -Gens.	145

Abb. 4.2:	Lage des Translokationsbruchpunktes der Multiplen Myelom-Zelllinie LB1017 innerhalb des Introns 1 des <i>LETMI</i> -Gens.	150
Abb. 4.3:	Vergleichende Kartierung humaner und muriner orthologer Gene in konservierten Bereichen auf Chromosom 4p16.3 und Mmu 5 nach Kozak und Stephenson (1999).	165
Tab. 3.1A:	Übersicht über die isolierten Exonsequenzen des Contigs I.	63
Tab. 3.1B:	Übersicht über die Exonsequenzen des Contigs II.	64
Tab. 3.1C:	Übersicht über die Exonsequenzen des Contigs III.	65
Tab. 3.2:	Exonsequenzen aus dem <i>MYL5</i> - und dem <i>PDEB</i> -Gen.	72
Tab. 3.3:	Darstellung der Homologiesuche und Rückkartierungsdaten der <i>GAK</i> -Exons.	74
Tab. 3.4:	Übersicht über die Exonsequenzen mit schwachen Homologien zu bekannten Genen und/oder Sequenzidentität zu EST-cDNA-Klonen.	75
Tab. 3.5:	Vergleich der Exon/Intron-Struktur des humanen und murinen <i>LETMI</i> -Gens.	89
Tab. 3.6:	Übersicht über identifizierte cDNA-Klone des Menschen und der Maus mit Sequenzidentität zu den Cosmiden HSL58b6 bis 184d6.	102
Tab. 3.7:	Vergleich der Exon/Intron-Struktur des humanen und murinen <i>Gens 43 (WHSC2)</i> .	112
Tab. 3.8:	Vergleich der Exon/Intron-Struktur des humanen und murinen <i>Gens 45</i> .	118

Abkürzungsverzeichnis:

A	Adenin
Abb.	Abbildung
A. bidest	Aqua bidestillata
abs.	absolut
<i>A. thalianae</i>	<i>Arabidopsis thalianae</i>
AS	Aminosäure
Bp	Basenpaare
BSA	Rinderserumalbumin
C	Cytosin
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
ca.	circa
cDNA	komplementäre DNA
Ci	Curie
cM	centi-Morgan
C-terminal	carboxyterminal
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DTT	Dithiothreitol
DNA	„deoxyribonucleic acid“
dNTP	2' Deoxynukleosidtriphosphat
EBV	Epstein-Barr Virus
<i>E. coli</i>	<i>Escherichia coli</i>
EST	„expressed sequence tag“
EDTA	Ethylendiamintetraessigsäure
EtOH	Ethanol
FISH	Fluoreszenz in situ Hybridisierung
FITC	Fluoresceinisothiocyanat
G	Guanin
h	Stunde
HEPES	2-[4-(2-Hydroxyethyl)-1-piperazinyl]-ethansulfonsäure
Hs	<i>Homo sapiens</i>
IPTG	Isopropyl- β -D-thio-galactopyranosid
Kb	Kilobasenpaare
k.o.	„knockout“
kDa	Kilodalton
M	molar
Mb	Megabasenpaare
min	Minute
Mmu, Ms	<i>Mus musculus</i>
M-MLV	Moloney Murine Leukemia Virus
MOPS	3-(N-morpholin)Propansulfonsäure
μ	mikro
Nr.	Nummer
N-terminal	aminoterminal
OD	optische Dichte
OMIM	„Online Mendelian Inheritance in Men“
p	piko
PBS	„phosphate buffered saline“-Puffer
PCR	„polymerase chain reaction“
PEG	Polyethylenglykol
RT	Reverse Transkriptase

RT-PCR	„reverse transcriptase polymerase chain reaction“
Rn	Rattus norvegicus
RNA	„ribonucleic acid“
<i>S. cerevisiae</i>	Saccharomyces cerevisiae
SDS	Natriumdodecylsulfat
<i>S. pombe</i>	Schizosaccharomyces pombe
SSC	„standard saline citrate“-Puffer
T	Thymin
Tab.	Tabelle
TBE	Tris-Borat-EDTA-Puffer
Tris	Tris(hydroxymethyl)aminomethan
TRITC	Tetramethylrhodaminisothiocyanat
U	„Units“ (Enzymeinheiten)
Upm	Umdrehungen pro Minute
UV	Ultraviolett
V	Volt
vgl.	vergleiche
Vol.	Volumen
X-Gal	5-Bromo-4-Chloro-3-indolyl- β -D-galactosid
z. B.	zum Beispiel

Aminosäuren (Einbuchstabencode):

A	Alanin	I	Isoleucin	R	Arginin
C	Cystein	K	Lysin	S	Serin
D	Asparaginsäure	L	Leucin	T	Threonin
E	Glutaminsäure	M	Methionin	V	Valin
F	Phenylalanin	N	Asparagin	W	Tryptophan
G	Glycin	P	Prolin	Y	Tyrosin
H	Histidin	Q	Glutamin		

1 Einleitung

Das Wolf-Hirschhorn-Syndrom (WHS)

Das Wolf-Hirschhorn-Syndrom (4p⁻-Syndrom) wurde zuerst von Wolf et al. (1965) und Hirschhorn et al. (1965) beschrieben (OMIM-Nr.: 194190). Inzwischen sind über 265 Fälle dieses seltenen, komplexen Fehlbildungs-Retardierungs-Syndroms (mit einer Inzidenz von 1/50000 Lebendgeburten) bekannt oder beschrieben worden (Thomson, 1998; Battaglia et al., 1999). Neben pränatalen Wachstums- und Entwicklungsverzögerungen und schwerer mentaler Retardierung umfaßt das Syndrom eine Reihe weiterer Fehlbildungen (Johnson et al., 1976; Lurie et al., 1980; Dallapiccola et al., 1993). Charakteristisch sind die kraniofazialen Befunde wie u.a. Mikrozephalie, Mikrognathie und kraniale Asymmetrie, dabei können Hirnfehlbildungen wie Agenesie/Hypoplasie des Corpus callosum sowie Kleinhirnhypoplasie auftreten. Kennzeichnend sind insbesondere die typischen Gesichtsdysmorphien, die an einen griechischen Helm („greek helmet“) erinnern lassen. Zu diesen Dysmorphien gehören die prominente Glabella, Hämangiomen der hohen Stirn, Ptosis, Strabismus, Iriskolobomen, Hyperelorismus, eine breite, hakenförmig gebogene Nase, kurzes Philtrum, abfallende Mundwinkel (Karpfenmund) und kleine, einfach angelegte Ohren. Spaltbildungen wie (Lippen-Kiefer)-Gaumenspalten sind nicht selten. Die Fehlbildungen innerer Organe betreffen vor allem Herzfehler, Nieren- sowie Genitalfehlbildungen. Die letzteren umfassen meist Kryptorchismus, Hypospadie, Vagina- und Uterusfehlbildungen. Typisch sind auch Hernien, Rektusdiastase und Skalpdefekte. In etwa 50-100 % der bekannten Fälle werden Krampfanfälle mit stereotypem EEG beschrieben (Sgrò et al., 1995). Schließlich findet man ein Spektrum von Skelettfehlbildungen bzw. Anomalien einschließlich Wirbel- und Rippenfehlbildungen sowie eine verzögerte Ossifikation der Knochen. Kürzlich konnten auch Antikörperdefizienzen (Hanley-Lopez et al., 1998) und Fälle mit Hörbeeinträchtigungen bzw. bilateraler Taubheit (Lesperance et al., 1998; Battaglia et al., 1999) beschrieben werden.

Das Wolf-Hirschhorn-Syndrom (WHS) wird in über 90 % der Fälle durch Deletionen im terminalen Bereich des kurzen Arms von Chromosom 4 (4p16.3) hervorgerufen. Daneben spielen familiäre Translokationen (Martsolf et al., 1987; Stengel-Rutkowski et al., 1984) sowie *de novo* unbalancierte Translokationen, die aber immer mit dem Verlust von 4p16.3-Material

assoziiert sind, eine bedeutsame Rolle. In den letzten Jahren konnte durch cytogenetische und molekulargenetische Charakterisierung der Deletionen bei WHS-Patienten der eigentliche WHS-kritische Bereich (WHCSR) immer weiter eingegrenzt werden. Zu Beginn der vorliegenden Arbeit umfaßte er einen ca. 2 Mb großen Bereich der distalen 4p16.3-Region (Gandelmann et al., 1992; Estabrooks et al., 1994), konnte dann aber vom distalen Ende her immer weiter auf einen 165 Kb großen Bereich eingegrenzt werden (Johnson et al., 1994; Estabrooks et al., 1994; Reid et al., 1996; Altherr et al., 1997; Wright et al., 1997). Dabei spricht die Komplexität und Variabilität dieses Syndroms, der hohe Anteil an de novo Deletionen und die Beobachtung von Patienten mit kompletten und inkompletten Phänotyp (Petit et al., 1996) letztlich für eine Beteiligung mehrerer Gene an der Entstehung bzw. vollständigen Ausprägung dieses Syndroms. Bei der Beurteilung der WHS-kritischen Region (WHSCR) muß berücksichtigt werden, daß die Beeinträchtigung angrenzender Gene an der Entstehung dieses Syndroms nicht vollständig ausgeschlossen werden kann und daher bei der Charakterisierung der WHS-kritischen Region beachtet werden sollte. Entsprechende Positionseffekte, die durch Deletionen oder Translokationen auf angrenzende Bereiche (Gene) erfolgen können, haben oftmals im Abstand von bis zu einigen 100 Kb Einfluß auf die Expression eines Gens. Die Auswirkungen dieser Effekte konnten bereits für mehrere Gene wie beispielsweise für das *PAX6*- oder das *SOX9*-Gen gezeigt werden, bei denen die bis zu 125 Kb (*PAX6*) bzw. 850 Kb (*SOX9*) entfernt liegende Bruchpunkte in einer gestörten Gendosis resultieren und zu Erkrankungen führen (z. B. Aniridie bzw. Kampomelen Dysplasie; Übersichtsartikel: Kleinjahn und van Heyningen, 1998).

Genotyp-Phänotyp-Korrelationen

Für die Bestimmung krankheitsrelevanter Bereiche der distalen Region 4p16.3 wurden mehrere Patienten mit terminalen und interstitiellen Deletionen unterschiedlicher Größe untersucht. Durch Estabrooks et al. (1995a) konnte eine vorläufige Phänotypenkarte erstellt werden, die einige klinische Merkmale bestimmten chromosomalen Bereichen zuordnet. Solche Korrelationen sind jedoch schwierig da eine Reihe von Faktoren diese Untersuchung beeinflussen können, wie z. B. unterschiedliche Penetranz für bestimmte Merkmale, die Dichte der verwendeten Marker bei der Untersuchung der Patienten und Positionseffekte als Folge chromosomaler Veränderungen.

So wurden eine Reihe von Patienten beschrieben, die klinische Merkmale zeigen, die sich von denen des WHS unterscheiden und durch Deletionen proximal und distal der WHSCR verursacht werden. Dabei stellen die typischen kraniofazialen Befunde bei WHS-Patienten ein Hauptmerkmal zur Unterscheidung der meist milder betroffenen Patienten mit proximalen und distalen Deletionssyndromen dar (Estabrooks et al., 1995b; Anderson et al., 1997). So zeigen Patienten mit Deletionen distal der WHSCR ein Spektrum an meist milderem Dysmorphiezeichen und normaler bis moderater mentaler Retardierung, wobei die Entwicklungsverzögerungen leicht bis schwer sein können (Estabrooks et al., 1992; Estabrooks et al., 1995a; Sebastio et al., 1996; Anderson et al., 1997). Einige Merkmale dieser Patienten, wie schwere Entwicklungsverzögerungen, Mikrozephalie oder auch Genitaldefekte entsprechen auch denen des WHS (Anderson et al., 1997; Sebastio et al., 1996). Deletionen des distalen, telomernahen ca. 150 Kb-Bereichs (ca. D4S142 bis 4ptel) haben zumeist keine phänotypische Auswirkungen (Estabrooks et al., 1992). In seltenen Fällen lassen sich Wachstumsretardierung oder auch Entwicklungsverzögerungen feststellen, die durch Positionseffekte der durch Translokationen betroffenen Bereiche erklärt werden (Estabrooks et al., 1992).

Patienten mit Deletionen proximal der WHSCR weisen variable Phänotypen auf. Meist werden milde bis moderate geistige Retardierung mit entsprechenden Entwicklungsverzögerungen bei oftmals normalem Wachstum beschrieben (Francke et al., 1977; White et al., 1995). Insgesamt scheint hierbei eine Korrelation des Schweregrades der betroffenen Patienten mit dem distalen Ausmaß der Deletionen zu bestehen (White et al., 1995). Zudem liegen bei einigen dieser Patienten charakteristische klinische Merkmale vor, die auf ein definiertes, klinisch erkennbares Deletionssyndrom (proximales 4p⁻-Syndrom) schließen lassen (Chitayat et al., 1995; White et al., 1995; Kozma et al., 1999).

Duplikationen innerhalb dieser Chromosomenregion führen zumeist ebenfalls zu einem auffälligen, variablen Phänotyp, der meist eine mildere mentale Retardierung, Kleinwuchs, Entwicklungsverzögerungen, Mikrozephalie und eine typische Boxernase (Aplasie der Nasenknochen) beinhaltet (Dallapiccola et al., 1977; Gonzalez et al., 1977; Wyandt et al., 1993). Auch wurden Patienten mit Duplikationen um den Locus D4S96 beschrieben, bei denen Überwuchs und milde bis moderate geistige Retardierung beobachtet werden konnte (Partington et al., 1997).

Der klinische Phänotyp des Wolf-Hirschhorn-Syndroms überlappt mit dem des Pitt-Rogers-Danks-Syndroms (PRDS; Pitt et al., 1984; OMIM: 262350), das als milder eingestuft wird

(Clemens et al., 1996; Donnai, 1996). Cytogenetische Untersuchungen zeigten, daß die Deletionen beim PRDS häufig größer sind als beim WHS und den gesamten WHS-kritischen Bereich einschließen. Man geht deshalb davon aus, daß es sich hierbei um eine phänotypische Variation des gleichen Basisdefekts, d. h. der selben deletierten Gene oder regulatorischen Elemente handelt, wobei die phänotypischen Unterschiede durch die Demaskierung möglicher rezessiver Mutationen verursacht sein könnten (Clemens et al., 1996; Kant et al., 1997; Partington und Turner, 1999; Wright et al., 1998; Battaglia und Carey, 1998; Wright et al., 1999a). Imprintingeffekte können aufgrund der unterschiedlichen elterlichen Herkunft der Deletionen ausgeschlossen werden (Dallapiccola et al., 1993; Anvret et al., 1991).

Die Wolf-Hirschhorn-Syndrom-kritische Region (WHSCR)

Auf der Grundlage eines Patienten konnte die WHS-kritische Region (WHSCR) inzwischen auf einen Bereich von 260 Kb zwischen den Markern D4S43 und D4S3327 eingeschränkt werden (Wright et al., 1997). Über diesen milder betroffenen Patienten liegen jedoch nur wenige klinische Daten vor, die neben der Beschreibung der typischen kraniofazialen Befunde eine moderate psychomotorische Retardierung (Somer et al., 1995) beinhalten. Unter Berücksichtigung eines von White et al. (1995) publizierten Falls mit einer interstitiellen 4p-Deletion konnte der WHSCR-Bereich vom proximalen Ende her weiter auf eine 165 Kb große Region eingegrenzt werden, innerhalb derer bis dahin noch keine Gene bekannt waren (Wright et al., 1997). Dieser Bereich liegt in einem von Baxendale et al. (1993) veröffentlichten Cosmid-Contig, der inzwischen vom Sanger Center nahezu vollständig sequenziert wurde (siehe **Abb. 1.1**).

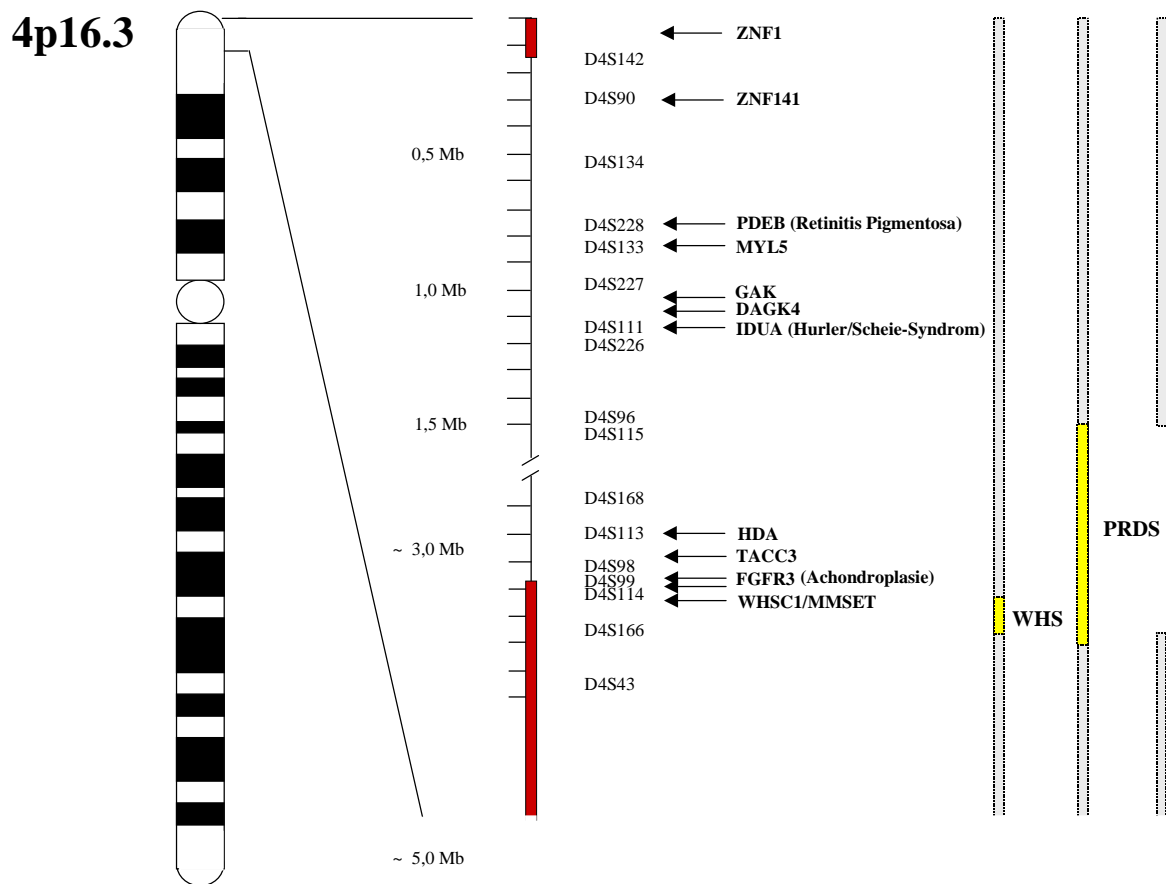


Abb. 1.1: Übersicht über die distalen 3,5 Mb der ca. 5 Mb umfassenden Chromosomenregion 4p16.3 (verändert nach Riess et al., 1996). Neben wesentlichen Markern und Genen sind Bereiche, die an der Ausbildung verschiedener Deletionssyndrome beteiligt sind, durch graue bzw. bei inzwischen weiter eingegrenzten minimalen kritischen Regionen durch gelbe Balken verzeichnet. Der bereits vom Sanger Center sequenzierte Cosmid-Contig von Baxendale et al. (1993) sowie der sequenzierte, telomernaher Bereich von 118 Kb (Flint et al., 1997) sind als rote Balken dargestellt.

Innerhalb dieser distalen Region 4p16.3 (D4S43 bis 4ptel) konnten bisher über 10 Gene beschrieben werden, von denen die meisten als Kandidatengene für die wesentlichen Merkmale des WHS ausgeschlossen werden können. Dies betrifft neben bereits im Text erwähnten Genen die Zinkfingergene *ZNF141* (Tommerup et al., 1993) und *ZNF1* (Flint et al., 1997), *GAK* („Cyclin G assoziierte Kinase“) (Kimura et al., 1997), *DAGK4* („Diacylglycerinkinase 4“) (Endele et al., 1996; Houssa et al., 1997); *MYL5* („myosin regulatory light chain gene 5“) (Collins et al., 1992), sowie das *HDA*- (McCombie et al., 1992) und das *TACC3*- („transforming acidic coiled-coil containing gene“) Gen (Still et al., 1999).

Durch „GRAIL“- und „BLAST“-Analysen der verfügbaren Nukleotidsequenzen konnten von Wright et al. (1997) 42 humane EST-cDNA-Klone innerhalb dieser 260 Kb großen Region identifiziert werden, die sich durch Homologien zu diesem Bereich auszeichnen und zu vermutlich insgesamt 9 unabhängigen cDNAs korrespondieren. 1998 wurde das erste Kandidatengen beschrieben (*WHSC1* = „Wolf-Hirschhorn syndrome candidate 1“), dem zwei der

vorher beschriebenen Transkripte zugeordnet werden konnten (Stec et al., 1998). Das Gen umspannt den distalen Bruchpunkt der WHSCR und zeigt ein komplexes alternatives Spleißmuster. Es stellt aufgrund seiner Expression in der frühen Entwicklung und der funktionellen Domänen des abgeleiteten Proteins ein Kandidatengen für verschiedene Merkmale des WHS dar. Das putative 136 kDa-Protein enthält neben einer PWWP-Domäne, einer HMG-Box und mehreren PHD („plant homeodomain“-Zinkfinger-ähnlichen Motiven eine SET-Domäne, die ihren Namen durch die *D. melanogaster* Proteine „Suppressor of variegation“ (suvar9), „Enhancer of zeste“ und Trithorax erhielt (Tschiersch et al., 1994). Sowohl die PHD-ähnlichen Motive als auch die SET-Domäne kommen bei einer Reihe evolutionär konservierter Proteine vor, die einen Chromatin-vermittelten Einfluß auf die Transkription bestimmter Gene ausüben, wie beispielsweise Mitglieder der Trithorax-Gruppe bei *D. melanogaster* (trx und ash1) und trx-homologer Proteine des Menschen (ALL-1/HRX/TRX) und der Maus (mll-1) (Mazo et al., 1990; Tripoulas et al., 1996; Aasland et al., 1995; Gu et al., 1992; Hanson et al., 1999). Aufgrund der signifikanten Homologien zu diesen Proteinen und der DNA-bindenden Eigenschaften des putativen WHSC1-Proteins wird eine entsprechende Funktion als Transkriptionsfaktor oder Koregulator vermutet (Stec et al., 1998). Interessanterweise konnte für dieses Gen gezeigt werden, daß es einige Translokationsbruchpunkte (t(4;14)) bei Patienten mit Multiplen Myelomen enthält und hierbei u.a. Fusionsproteine mit dem Translokationspartner *IgH* entstehen. Zeitgleich wurde *WHSC1* über die Suche nach Genen, die durch Translokationen t(4;14) bei Patienten mit Multiplen Myelomen betroffen werden, isoliert und dementsprechend auch als *MMSET*-Gen („Multiple Myeloma SET domain“) bezeichnet (Chesi et al., 1998).

„Contiguous gene“-Syndrome (CGS)

Das Wolf-Hirschhorn-Syndrom wird aufgrund der Beteiligung mehrerer benachbarter Gene an dem Phänotyp zu den „contiguous gene“-Syndromen gerechnet. Schmickel (1986) hat als erster den Begriff „contiguous gene syndrome“ geprägt um zu verdeutlichen, daß mehrere unabhängige, benachbarte Gene innerhalb eines umschriebenen chromosomalen Abschnitts an der Ausbildung komplexer, aber spezifischer Phänotypen beteiligt sein können. Diese Phänotypen können aus chromosomalen Aberrationen (Deletionen oder Duplikationen) resultieren, die eine Störung der normalen Gendosis der in dem Bereich lokalisierten Gene verursachen.

Die Größe der Deletionen bzw. Duplikationen kann zwischen wenigen hundert und einigen Megabasen liegen, so daß üblicherweise zwischen 10 und 100 Gene beeinflusst werden (Ledbetter und Ballabio, 1995). Obwohl es keine detaillierten Untersuchungen gibt, kann man davon ausgehen, daß nur einige wenige dieser Gene Dosis-sensitiv sind und somit zum Phänotyp beitragen (Ledbetter und Ballabio, 1995). Die Identifizierung dieser Gene stellt hierbei einen Hauptaspekt der Untersuchung dieser Syndrome dar. Durch Budarf und Emanuel wurde 1997 der Begriff der sogenannten „segmental aneusomy syndromes“ (SAS) geprägt um zu verdeutlichen, daß die krankheitsverursachenden Gene bei diesen Syndromen nicht unmittelbar benachbart („contiguous“) sein müssen. Für das Verständnis der Ätiologie und Pathogenese dieser komplexen Krankheitsbilder ist die Analyse der entsprechenden chromosomalen Regionen von besonderem Interesse.

Entsprechende Projekte zur Untersuchung anderer „contiguous gene“-Syndrome, die durch die Haploinsuffizienz mehrerer benachbarter Gene verursacht werden, beschäftigen sich seit einiger Zeit z. B. mit dem Cri-du-Chat-Syndrom, dem DiGeorge-Syndrom, dem Langer-Giedion-Syndrom, dem Miller-Dieker-Syndrom und dem Williams-Beuren-Syndrom (Simmons et al., 1997; Gong et al., 1996; Lüdecke et al., 1995; Review: Ledbetter und Ballabio, 1995; Budarf und Emanuel, 1997; Francke, 1999). Obwohl inzwischen detaillierte Untersuchungen vorgenommen wurden, konnten die molekulargenetischen Grundlagen dieser Erkrankungen bisher nicht vollständig geklärt werden. So konnte beispielsweise für das Miller-Dieker-Syndrom, das durch eine (sub)mikroskopische Deletion (ca. 400 Kb) in der Chromosomenregion 17p13.3 hervorgerufen wird gezeigt werden, daß die Lissencephalie - ein Bestandteil des Krankheitsbildes - auf die Haploinsuffizienz des *LIS1*-Gens zurückgeführt werden kann (Chong et al., 1996). Die Ursachen weiterer Merkmale dieser Erkrankung, wie faziale Veränderungen und zahlreiche Fehlbildungen bleiben jedoch bis heute ungeklärt. Ähnlich sind auch die Verhältnisse beim Langer-Giedion- und dem Williams-Beuren-Syndrom (Momeni et al., 2000; Ewart et al., 1993). Das Williams-Beuren-Syndrom wird durch Deletion mehrerer Gene in der Region 7q11.23 verursacht wobei isolierte Merkmale dieser Erkrankungen auch einzeln auftreten. So konnte gezeigt werden, daß die auch isoliert vererbte supraaortale Aortenstenose (SVAS) auf die Haploinsuffizienz des *Elastin*-Gens zurückzuführen ist (Ewart et al., 1993). Jedoch können auch hier nicht alle Merkmale des Syndroms durch ein einzelnes Gen erklärt werden. Obwohl inzwischen mehr als 16 Gene in der deletierten Region identifiziert wurden, konnte für keines dieser Gene eine direkte Beteiligung

an den weiteren Merkmalen dieses Syndroms nachgewiesen werden (Osborne et al., 1996; Francke, 1999; Peoples et al., 2000).

Dies zeigt, daß die Aufklärung von „contiguous gene“-Syndromen nicht nur die Identifizierung sondern auch die Charakterisierung von Genen voraussetzt. Da für einzelne Gene zu meist keine Mutationen nachgewiesen werden können, ist die Kenntnis der Funktion der einzelnen Gene wichtig um beurteilen zu können, ob es sich letztendlich um Kandidatengene handelt oder nicht. Um die vollständigen Ursachen dieser Erkrankungen zu klären haben zahlreiche Gruppen ihre Forschungsbemühungen auf die orthologen Bereiche der Maus ausgedehnt, wobei das Ziel die Etablierung entsprechender Mausmodelle für diese komplexen Syndrome ist (Hirotsume et al., 1997; Galili et al., 1997; Kimber et al., 1997; Sutherland et al., 1998; Valero et al., 1998; DeSilva et al., 1999).

Besonders gut untersucht sind hier die Verhältnisse beim DiGeorge- bzw. velo-kardio-fazialen Syndrom. Es handelt sich um zwei sehr ähnliche kongenitale Krankheitsbilder, die durch verschiedene kraniofaziale und kardiale Fehlbildungen gekennzeichnet sind und mit Mikrodeletionen in der Chromosomenregion 22q11.2 assoziiert sind. Die Deletionen umfassen einen ca. 2 Mb großen Bereich, wobei der minimale deletierte Bereich auf einen ca. 250 Kb großen Abschnittes eingegrenzt wurde (Budarf et al., 1995; Gong et al., 1996). Alle bisherigen Untersuchungen sprechen dafür, daß mehrere Gene in diesem Bereich gemeinsam für die komplexen Veränderungen verantwortlich sind (Galili et al., 1997). Obwohl die genomische Sequenz der deletierten Region bekannt ist und inzwischen eine Reihe möglicher Kandidatengene isoliert wurden, konnte dennoch zuerst für keines dieser Gene eine direkte Beteiligung am Krankheitsbild nachgewiesen werden (Augousseau et al., 1986; Wilming et al., 1997; Wadey et al., 1999; Yamagishi et al., 1999). Daher bestand das Ziel einiger Arbeitsgruppen darin ein Mausmodell zu etablieren, bei dem die genetischen Veränderungen entsprechend denen des Menschen induziert werden und somit der Effekt der einzelnen Gene besser untersucht werden kann (Galili et al., 1997; Kimber et al., 1997). Dementsprechend wurde die Analyse dieser Chromosomenregion auf den syntänen Bereich der Maus auf dem Chromosom 16 ausgedehnt. Obwohl die orthologe genomische Region der Maus einigen größeren Rearrangements unterlag ist die deletierte Region mit den darin enthaltenden Genen zwischen beiden Spezies weitgehend konserviert, so daß prinzipiell ein Mausmodell für diese Erkrankung generiert werden kann (Galili et al., 1997; Puech et al., 1997; Sutherland et al., 1998). Inzwischen konnte durch Lindsay et al. (1999) gezeigt werden, daß durch das Einfügen einer gezielten, ca. 1,2 Mb umfassenden Deletion in das Mausgenom, die den distalen Teil der

minimalen DiGeorge-kritischen Region (MDGCR) umfaßt, ein DiGeorge-ähnlicher Phänotyp bei der Maus erzeugt werden kann. Der Phänotyp kann durch Kreuzung mit Tieren, die eine Duplikation derselben Region aufweisen, komplementiert werden. Obwohl nicht alle der Merkmale dieser Erkrankung in den hemizygoten knockout-Mäusen nachgewiesen werden konnten, stellt dieser Ansatz ein gutes Beispiel für die Bedeutung eines Mausmodelles zum Verständnis solch komplexer Erkrankungen dar. Dies wird durch Arbeiten von Kimber et al. (1999) unterstrichen, die zeigen konnten, daß die Deletion eines ca. 150 Kb großen Bereichs innerhalb der proximalen MDGCR im Mausgenom scheinbar keinen wesentlichen Einfluß auf die Entstehung dieses Syndroms ausübt, wodurch der Bereich als Kandidatenregion ausgeschlossen werden konnte.

Fortschritte konnten auch bei der Aufklärung der mechanistischen Ursachen von Mikrodeletionssyndromen gemacht werden. Inzwischen weiß man, daß die Deletionsbruchpunkte bei einer Reihe von Syndromen wie beispielsweise dem Williams-Beuren-Syndrom, dem Smith-Magenis-Syndrom und dem DiGeorge/velo-kardio-fazialen-Syndrom, innerhalb größerer chromosomaler Sequenzwiederholungen zu finden sind. Diese als „Duplikons“ bezeichneten Elemente, die nicht in den entsprechenden orthologen Regionen der Maus nachgewiesen werden konnten, scheinen maßgeblich an der Entstehung der jeweiligen chromosomalen Rearrangements beteiligt zu sein (Francke 1999; Chen et al., 1997; Edelmann et al., 1999; Shaikh et al., 2000). Sie flankieren bei all diesen Syndromen eine sogenannte „gemeinsame deletierte Region“, die bei der Mehrzahl der jeweiligen Patienten aufzufinden ist. Jedoch gibt es bisher aufgrund der zumeist unterschiedlich großen interstitiellen und auch terminalen und darüber hinaus oftmals wenig charakterisierten Deletionen innerhalb der distalen Chromosomenregion 4p16.3 keine Hinweise, die auf einen ähnlichen Mechanismus bei der Entstehung des WHS schließen lassen.

Die distale Region 4p16.3

Die distale Region 4p16.3 ist eine sehr gendichte Region (Riess et al., 1996), in der aufgrund cytogenetischer und molekularer Analysen eine Reihe von krankheitsrelevanten Genen kartiert wurden. Hierzu gehören beispielsweise das *HD*- („Huntington Disease“) und das *FGFR3*- („Fibroblast growth factor receptor 3“) Gen, die durch Mutationen zu der neurodegenerativen Erkrankung Chorea Huntington (Gusella et al., 1983; The Huntington's Disease

Collaborative Research Group, 1993; OMIM: 143100) bzw. der häufigsten Form des erblichen Kleinwuchses, der Achondroplasie (ACH, OMIM: 100800), führen (Thompson et al., 1991; Francomano et al., 1994; Velinov et al., 1994; Shiang et al., 1994; Rousseau et al., 1994). Dies führte dazu, daß der Bereich genetisch und physikalisch detailliert kartiert wurde (Bucan et al., 1990; Whaley et al., 1991; Weber et al., 1991a; Bates et al., 1990; Bates et al., 1992; Baxendale et al., 1993).

Mutationen in dem *FGFR3*-Gen führen zudem zur Hypochondroplasie (HCH, Bellus et al., 1995, OMIM: 146000) und zur Thanatophoren Dysplasie Typ I und II (TD I und II, Tavormina et al., 1995; OMIM: 187600 und 187601), einer milden bzw. letalen Variante der ACH sowie zu Kraniosynostosen mit assoziierten Hauterkrankungen (Meyers et al., 1995; Hollway et al., 1995; Tavormina et al., 1999) und einer Form der autosomal dominanten angeborenen, bilateralen Taubheit (Hollway et al., 1998) (siehe OMIM: 134934). Dieses Gen wird derzeit auch als mögliches Kandidatengen für eine als Cherubinismus bezeichnete Erkrankung, die sich primär durch Veränderungen der Kieferknochen auszeichnet und in der chromosomalen Region D4S127 bis 4pter lokalisiert werden konnte, diskutiert (Mangion et al., 1999; Tiziani et al., 1999; OMIM: 118400).

Darüber hinaus beinhaltet die Chromosomenregion 4p16.3 die Gene für die α -L-Iduronidase (*IDUA*) und Phosphodiesterase- β (*PDEB*) (Scott et al., 1990; Weber et al., 1991b). Mutationen im *IDUA*-Gen führen zu dem autosomal rezessiven Hurler-Scheie-Syndrom (Bunge et al., 1994, Scott et al., 1993; OMIM: 252800) und Mutationen innerhalb des *PDEB*-Gens verursachen ca. 5% der autosomal rezessiv vererbten, genetisch heterogenen Retinitis pigmentosa sowie die autosomal dominant vererbte, stationäre Nachtblindheit (McLaughlin et al., 1995; Bayes et al., 1995; Gal et al., 1994).

Dem humanen Chromosom 4 konnten anhand von Zellfusionsstudien bereits 1988 durch Pereira-Smith und Smith und 1991 durch Ning et al. zellproliferationssupprimierende Eigenschaften zugeordnet werden. Neuere Untersuchungen bestätigen diese Ergebnisse und zeigten, daß Heterozygotieverluste („loss of heterozygosity“; LOH) u.a. die Chromosomenregion 4p betreffen und bei einer Reihe an unterschiedlichen Tumoren auftreten (Elder et al., 1994; Bell et al., 1996; Caron et al., 1996; Tamura et al., 1996; Hu et al., 2000). So wurden bei ca. 50 % der beschriebenen Blasenkarzinome Heterozygotieverluste innerhalb 4 verschiedener Regionen auf Chromosom 4 beschrieben (Elder et al., 1994; Polascik et al., 1995), wobei eine dieser Regionen inzwischen auf einen 30 Kb großen Bereich proximal des Markers D4S43 eingegrenzt werden konnte (Bell et al., 1996). Das *SH3BP2*-Gen, das in

dieser Region identifiziert wurde, scheint jedoch nicht an der Entstehung dieser Tumore beteiligt zu sein (Bell et al., 1997). Heterozygotieverluste, die die distale Chromosomenregion 4p16.3 betreffen, konnten auch bei 19,5 % von 86 untersuchten Neuroblastomen nachgewiesen werden (Caron et al., 1996). Darüber hinaus konnten Translokationsbruchpunkte zwischen den Chromosomen 4 und 14 t(4;14)(p16.3;q32) beschrieben werden, die mit einem malignen Neoplasma der B-Zellen (Multiple Myelome) assoziiert sind (Bergsagel et al., 1996). Hierbei stellen Translokationen des *IgH*-Lokus auf Chromosom 14 mit verschiedenen Translokationspartnern, die häufig mehrere Regionen innerhalb des kurzen Arms von Chromosom 4 betreffen (Nishida et al., 1997; Avet-Loiseau et al., 1998), einen Mechanismus in der Pathogenese dieser B-Zelltumoren dar (Rabbitts, 1994). Die Bruchpunkte liegen hierbei bis zu 120 Kb proximal des *FGFR3*-Gens, für das bereits eine Überexpression in diesen Tumoren und in einigen Fällen auch aktivierende Mutationen, wie sie auch bei der Thanatophoren Dysplasie Typ II auftreten, nachgewiesen werden konnten (Richelda et al., 1997; Chesi et al., 1997), allerdings treten diese Mutationen sehr selten auf (Fracchiolla et al., 1998; Avet-Loiseau et al., 1998). Inzwischen konnte eine direkte Beteiligung des *FGFR3*-Gens an der Zellproliferation und der Apoptose von B-Zelllinien nachgewiesen werden (Plowright et al., 2000).

Zielsetzung

Die Zielsetzung der vorliegenden Arbeit war die Identifizierung und vorläufige Charakterisierung neuer Gene, die an der Entstehung des Wolf-Hirschhorn-Syndroms beteiligt sein könnten. Zunächst sollten hierfür nicht nur neue Transkriptionseinheiten identifiziert, sondern auch die genaue physikalische Kartierung der Region 4p16.3 durchgeführt werden. Ausgewählte neue cDNAs sollten sequenziert und durch Homologievergleiche bzw. Expressionsstudien sowie Untersuchungen an Patienten weiter charakterisiert werden. Zur Vorbereitung eines Mausmodells für das WHS sollten darüber hinaus die orthologen murinen Transkripte isoliert und eine vergleichende Kartierung der syntänen Mausregion durchgeführt werden.

2 Material und Methoden

2.1 Herkunft Chromosom 4-spezifischer Cosmidklone

Ein Großteil der in der vorliegenden Arbeit verwendeten Cosmidklone wurde im Rahmen der Identifizierung des Gens für die Chorea Huntington isoliert und wurde freundlicherweise von S. Baxendale (MIT, Cambridge) und M. MacDonald (MIT, Cambridge) bereitgestellt. Diese Cosmide wurden mittels einer regionalen Hybridzelllinie in die Region 4p16.3 kartiert und auf zwei Not I/Mlu I/Nru I-Restriktionsfragmentkarten verankert (Bucan et al., 1990; Whaley et al., 1991). Die Cosmide IS92, CD1, CD2, 5J80, 2R164, B31 und 2R88 sind auf einem telomernahen 350 Kb Nru I-Fragment verankert und beinhalten die D-Marker D4S90 (CD1, CD2) und D4S142 (B31, 2R88) (Whaley et al., 1991). Das Cosmid 8C101I14 hingegen konnte nicht eindeutig lokalisiert werden und kartiert in einem Bereich distal von D4S111. Die Cosmide 175.1, 157.2, 157.4, 157.5, 157.6 und 417.1, 417.5, 417.6 kartieren in eine ca. 250 Kb große Region distal des Markers D4S111, in der eine Häufung unmethylierter CpG-Inseln gefunden wurde (Bucan et al., 1990; Whaley et al., 1991). Zwei weitere Cosmide, pC385.1 und pC385.12 beinhalten genomische Bereiche um das *FGFR3*-Gen und wurden freundlicherweise von J. Wasmuth (Irvine, Californien) zur Verfügung gestellt.

Als Vektorsystem dieser Cosmide diente mit Ausnahme des Cosmids 8C101I14 der pCOS2 Vektor (Ehrich et al., 1987), der über eine Klonierungskapazität von 32-40 Kb verfügt. Als Vektorsystem für das Cosmid 8C101I14 wurde pWE15 verwendet (Wahl et al., 1987).

2.2 Isolierung von DNA

2.2.1 Isolierung von genomischer DNA aus Blut und Zellkulturen

Die Isolierung genomischer DNA aus Blut erfolgte mit Hilfe von Quiagen-tip 100 Säulen (Quiagen, Hilden) entsprechend den Angaben des Herstellers.

5 ml Vollblut wurden mit 20 ml eiskaltem Puffer G1 versetzt und für 10 min auf Eis inkubiert. Nach 15minütiger Zentrifugation (4°C, 1300 g) wurde das Pellet in 5 ml eiskaltem PBS resuspendiert und erneut abzentrifugiert. Nach erneuter Resuspension des Pellets in Puffer G2 wurden 150 µl Poteinase K Lösung (20 mg/ml) zugesetzt und für 1 h bei 50°C inkubiert. Die nun klare Lösung wurde auf eine mit Puffer QBT äquilibrierte Quiagen-tip 100 Säule aufgetragen. Die Säule wurde anschließend mit 10 ml Puffer QC gewaschen und die DNA mit 5 ml Puffer QF eluiert. Nach Isopropanolfällung und Waschen der DNA in 70 % Ethanol konnten zwischen 80 und 100 µg DNA gewonnen werden.

Für DNA aus EBV-transformierten, lymphoblastoiden Zelllinien (Neitzel, 1986; Elliott et al., 1991), wurden die in (RPMI-Medium, Sigma, USA; 20 % fetales Kälberserum) Suspensionskultur wachsenden Zelllinien abzentrifugiert (10 min 4°C, 3200 Upm), 2 mal in 1 x PBS gewaschen und in Lysepuffer resuspendiert. Die Zelllinie des WHS-Patienten O. S. (Marczinek et al., 1995) wurde freundlicherweise durch P. Nürnberg (Charité, Berlin) zur Verfügung gestellt.

2.2.2 Isolierung kleinerer Mengen Plasmid- und Cosmid-DNA

Für die Plasmidsequenzierung (vgl. **2.15**) wurde die Plasmid-DNA mit Hilfe des „RPM Sample Kits“ (Dianova, Hamburg) entsprechend den Angaben des Herstellers isoliert. Die Präparation kleinerer Mengen Plasmid- bzw. Cosmid- DNA für Restriktionsanalysen bzw. für das Erstellen einer hierarchisch angeordneten cDNA-Bibliothek (**2.13**) erfolgte in abgewandelter Form nach Birnboim & Doly (1979). Hierzu wurden 15 ml LB Medium mit geeignetem Antibiotikumzusatz mit dem gewünschten Cosmid-/Plasmidklon angeimpft und bei 37°C im Schüttelinkubator für 12-16 Stunden inkubiert, abzentrifugiert und in 200 µl Puffer 1 resuspendiert. Die Bakterien wurden durch Zugabe von 400 µl Puffer 2 lysiert und der Ansatz für 5 min auf Eis inkubiert. Zur Fällung bakterieller Proteine und chromosomaler DNA wurde 300 µl eiskaltes 3 M NaAcetat (pH 4,8) zugegeben. Nach Zentrifugation wurde der plasmidhaltige Überstand vorsichtig abgehoben und durch eine Phenol/Chloroform/Isoamylalkohol-Extraktion (25:24:1) und eine Chloroformextraktion gereinigt, mit Ethanol gefällt und nach einem Waschschrift in 70 %igem Ethanol vakuumgetrocknet. Die DNA wurde schließlich in einem geeigneten Volumen ¼ x TE-Puffer aufgenommen.

2.2.3 Isolierung großer Mengen Plasmid-, Cosmid-, P1- und PAC-DNA

Die Isolierung großer Mengen Plasmid-, Cosmid-DNA sowie die Isolierung von P1- oder PAC-DNA wurden mit Hilfe von Quiagen-tip 100 bzw. 500 entsprechend den Angaben des Herstellers (Quiagen, Hilden) durchgeführt. Die Bakterienkultur wurde je nach dem verwendeten Vektor in 100-500 ml LB-Medium angesetzt und über Nacht bei 37°C im Schüttelinkubator angezogen. Die nach Birnboim und Doly (1979) mit der Methode der alkalischen Lyse erhaltene Suspension wurde mit zwei Zentrifugationsschritten von festen Bestandteilen befreit und im Anschluß über Quiagen-tip 100/500-Säulen nach Vorschrift des Herstellers aufgereinigt (vgl. 2.2.1). Die luftgetrocknete DNA wurde in einem geeigneten Volumen A. bidest. (für die direkte Sequenzierung, vgl. 2.15) oder ¼ x TE-Puffer zur Langzeitaufbewahrung gelöst.

2.3 Isolierung von RNA

Die Isolierung von Gesamt-RNA aus menschlichen bzw. murinen Geweben erfolgte je nach Menge des verwendeten Gewebes bzw. der gewünschten RNA-Ausbeute mit dem QUIAGEN „RNeasy Total RNA Kit“ (Quiagen, Hilden) oder mit dem QUIAGEN „Total RNA Midi-Kit“ (Quiagen, Hilden). Die Präparationen wurden entsprechend den Angaben des Herstellers durchgeführt.

2.4 DNA und RNA Standardmethoden

2.4.1 Verdauung von DNA mit Restriktionsendonukleasen

Die Verdauung von DNA mit Restriktionsendonukleasen erfolgte entsprechend den Angaben des Herstellers und unter Verwendung der mitgelieferten Restriktionspuffer. Die verwendete Enzymmenge wurde je nach Aktivität des entsprechenden Enzyms, der Qualität und Menge der zu restringierenden DNA variiert.

2.4.2 Agarose-Gelelektrophorese von DNA und RNA

Die elektrophoretische Auftrennung von DNA-Fragmenten erfolgte auf horizontalen 0,8-3 %igen Agarose- oder LMP- („low melting point“) Agarose-Gelen. Als Puffersystem wurde 1 x TBE verwendet. Die aufzutrennende DNA wurde vor der Elektrophorese mit 10-15 % DNA-„Blue-Marker“, bzw. mit DNA-„Orange-Marker“ versetzt. Als Molekulargewichtsstandard wurden Hind III-restringierte λ -DNA (Pharmacia, Freiburg), sowie die 100 Bp-„Leiter“ (Gibco BRL, Eggenstein) verwendet. Nach Ethidiumbromidfärbung (5 μ g/ml) wurden die Gele unter UV-Licht ausgewertet und mit Hilfe eines Gel-Imaging-System Transilluminators der Firma Fröbel dokumentiert.

RNA-Moleküle wurden auf horizontalen 1 %igen Agarose-Gelen aufgetrennt (6,5 % Formaldehyd). Als Laufpuffer wurde 1x MOPS verwendet. Vor der Elektrophorese wurde die RNA durch Zusatz von 1/10 Vol. Ethidiumbromid (5 μ g/ μ l) und 3 Vol. „RNA Loading“-Puffer zur Auflösung von Sekundärstrukturen für 5 min bei 70°C inkubiert. Als Molekulargewichtsstandard wurde die „RNA-Leiter“ der Firma Gibco BRL (Eggenstein) verwendet.

Für die Aufreinigung von DNA-Fragmenten mittels Phenol/Chloroform-Extraktionen sowie für die Verwendung von DNA-Fragmenten für die radioaktive Hybridisierung wurden 1 %ige LMP- („low melting point“) Agarosegele der Fa. FMC (USA) mit niedrigem Schmelzpunkt verwendet (vgl. 2.4.4).

Zur Größenbestimmung von Not I-restringierter PAC- bzw. P1-DNA wurde die „Contour-clamped-homogenous-electric field“ (CHEF)-Elektrophorese verwendet, die aufgrund wechselnder elektrischer Felder die Auftrennung großer DNA-Fragmente erlaubt. Die DNA-Proben wurden hierfür in Agaroseblöckchen eingegossen und auf einem 1 %igen LMP-Agarosegel (GTG-Agarose, Seakem) unter Verwendung von 0,5 x TBE als Laufpuffer aufgetrennt (PFGE-Steuereinheit, BioRad; Fragmentgrößenbereich von 20 bis 150 Kb, Laufzeit 27 h). Als Molekulargewichtsstandard wurde ein λ -DNA-Concatemer (Pharmacia) von jeweils 49,5 Kb-Länge verwendet.

2.4.3 Denaturierende Polyacrylamidgelelektrophorese

Zur Darstellung polymorpher Mikrosatellitenmarker wurden 8 %ige, denaturierende Polyacrylamidgele (8 M Harnstoff) verwendet. Die Auftrennung der durch 5 minütiges Aufkochen denaturierten, radioaktiv markierten PCR-Produkte (vgl. **2.8.3.2**) auf den Gelen (210 x 580 x 2 mm) erfolgte für 4 ½ h bei 75 W oder für ca. 16 h bei 17 W unter Verwendung von 1 x TBE als Laufpuffer. Die Gele wurden danach mit Haushaltsfolie überzogen und durch Autoradiographie dokumentiert (ca. 16 h bei 4°C).

Die Auftrennung von einzelsträngig fluoreszenzmarkierten PCR-Produkten (3,5 µl des denaturierten Ansatzes: 1,5 µl des 1:8 verdünnten PCR-Produktes, 1,5 µl Formamid, 0,5 µl Bluemarker und 0,5 µl TAMRA 350-Standard (Applied Biosystem GmbH, Weiterstadt)) erfolgte, ähnlich der unter **2.15** beschriebenen automatischen Sequenzierung, mit Hilfe eines automatischen Sequenziergerätes (ABI Prism TM 373). Die Auswertung erfolgte mittels der GeneScan TM Analysis Software 2.1 unter Verwendung eines internen Längenstandards.

2.4.4 Wiedergewinnung von DNA aus Agarosegelen

Nach gelelektrophoretischer Trennung von DNA-Fragmenten auf 1-2 %igen Agarosegelen wurden die Ethidiumbromid-gefärbten Banden unter UV-Licht aus dem Gel ausgeschnitten. Die Isolation der Fragmente erfolgte nach Angaben des Herstellers mit Hilfe des „QIAquick Gel Extraktion Kits“ der Firma Quiagen (Hilden).

Bei Verwendung von LMP-Agarosegelen mit niedrigem Schmelzpunkt wurden die ausgeschnittenen Gelstücke mit einem Gewichtsvolumen 0,2 M NaCl/1 x TBE versehen und im Heizblock bei 65°C für 15-30 min aufgeschmolzen. Die DNA wurde durch Zusatz von 0,75 Vol. Phenol (mit 0,2 M NaCl/1 x TBE gesättigt) aus der Agarosesuspension extrahiert. Nach Reextraktion der phenolischen Phase mit 300 µl 0,2 M NaCl/1 x TBE wurden die DNA-haltigen Überstände vereinigt und mit dem 4 fachen Volumen 1-Butanol weiter aufgereinigt und extrahiert. Die DNA wurde danach aus der unteren, wäßrigen Phase wie unter **2.4.5** beschrieben gefällt.

Für die Wiedergewinnung von PCR-Produkten, die ausschließlich als „Template“ für weitere PCR-Reamplifikationen eingesetzt wurden, wurden die ausgeschnittenen Banden mit ca. 1

Vol. A. bidest. überschichtet und für ca. 1-2 h bei 37°C oder über Nacht bei 4°C inkubiert. Ein Aliquot dieser Suspension wurde als Template für die PCR-Reaktion verwendet.

2.4.5 Fällung von DNA und RNA

DNA wurde nach Zusatz von 1/10 Vol. 3 M NaAcetat (pH 5,2) und 2,5 Vol. EtOH abs. für 15 min bei -70°C, für 2-3 h bei -20°C oder durch kurzes Eintauchen in flüssigem Stickstoff gefällt. Nach 30 minütiger Zentrifugation bei 13000 Upm wurde das DNA-Pellet mit 70 %igem EtOH gewaschen, für 15 min zentrifugiert und vakuumgetrocknet.

PCR-Produkte wurden zur Entfernung kleinerer Oligonukleotide durch Zugabe von 1 Vol. 4 M Ammoniumacetat und zwei Vol. Isopropanol (100 %) bei RT gefällt. Die Ansätze wurden hierbei für mind. 30 min bei RT (13000 Upm) pelletiert, mit 70 %igem EtOH gewaschen, vakuumgetrocknet und in A. bidest. gelöst.

2.5 Subklonierung von DNA-Fragmenten in Plasmid-Vektoren

Für die Subklonierung von DNA-Fragmenten wurden der Spleißvektor pSPL3 (vgl. 2.12) und der Plasmidvektor pBluescript II KS+ (Stratagene, Heidelberg) verwendet.

Der Plasmidvektor pBluescript ermöglicht die durch die α -Komplementation des lacZ-Gens auf X-Gal/IPTG-haltigen Agarplatten hervorgerufene „Blau-Weiß“-Selektion rekombinanter Klone. Hierfür wurden zur Transformation *E. coli*-Zellen mit der LacZ Deletion Δ M15 (hier: DH5 α , Hanahan, 1985) verwendet.

Zur Vermeidung von Religationen wurden die Enden der linearisierten Plasmid-DNA in der Regel mit alkalischer Phosphatase aus Kälberdarm (Boehringer, Mannheim) nach Anweisung des Enzymherstellers dephosphoryliert. Neben einer Hitzeinaktivierung (30 min, 70°C) wurde zur Inaktivierung des Enzyms stets eine Phenol/Chloroform-Extraktion mit anschließender Fällung und/oder Gelaufreinigung durchgeführt.

Die Ligationsreaktion erfolgte in der Regel nach einem Standardprotokoll unter Verwendung von T4-Ligase (New England Biolabs) nach Angaben des Herstellers. Die Transformation kompetenter Bakterien des Stammes *E. coli* DH5 α wurde nach einer modifizierten Methode

von Hanahan (1985) durchgeführt. Für die Transformation wurden 200 µl kompetenter Zellen mit jeweils der Hälfte des Ligationsansatzes vermischt, für 30 min auf Eis inkubiert und nach einem Hitzeschock (1 min 42°C) mit 1 ml SOC oder LB-Medium versetzt und für 1 h bei 37°C geschüttelt. Für die Subklonierung genomischer Fragmente in den Spleißvektor pSPL3 wurde in der Regel der gesamte Transformationsansatz ausplattiert. Rekombinante weiße Kolonien wurden in 96er Mikrotiterplatten mit LB-Medium (100 µg/ml Ampicillin) gepickt und nach über Nacht Inkubation mit 1 Vol. 2 x FM versetzt und bei -70°C eingefroren. Zur Ermittlung der Integratgrößen wurde eine PCR mit den pBluescript-spezifischen Primern T₃A und T₇A und den spezifischen Primern Exon S und Exon B (vgl. 2.17), wie unter 2.6 beschrieben, durchgeführt. Die Selektion definitiv positiver Klone erfolgte durch die Sequenzierung der T₃A/T₇A Amplifikationsprodukte und durch Koloniefilterhybridisierung (vgl. 2.9.4).

2.6 Polymerase-Kettenreaktion (PCR)

Die PCR (Saiki et al., 1988) wurde in 25-100 µl-Ansätzen mit jeweils 40 mM KCl, 10 mM Tris/HCl (pH 8,3), 0,1 mg/ml Gelatine, 1,5 mM MgCl₂, je 10 bis 20 pmol der beiden Primer, je 200 µM der vier dNTPs sowie 0,5-1 U Taq-Polymerase (Perkin-Elmer, Weiterstadt; Gibco BRL, USA) durchgeführt.

Die Amplifikationsbedingungen wurden entsprechend der verwendeten Primerpaare sowie deren DNA-Matrizen variiert und wurden ausgehend von einem Standardprogramm (erster Denaturierungsschritt: 3 min 94°C, 35 Zyklen: 1 min 94°C, 1 min 58°C, 1 min 72°C, 35 Zyklen, Extensionsschritt 7 min 72°C) entsprechend verändert. Als Templat wurden 50-100 ng genomische, 10-20 ng Plasmid- oder Phagen-DNA, 2 µl cDNA (2.7), 1-2 µl von Bakterien-dauer-Glycerinkulturen sowie 5 µl fragmenthaltige Suspension (vgl. 2.4.4) pro PCR-Reaktion eingesetzt.

Die Amplifikation von DNA-Fragmenten > 3 Kb erfolgte mit dem „Expand™ High Fidelity PCR System“ (Boehringer, Mannheim) nach einem leicht modifizierten Protokoll des Herstellers (200 µM dNTPs). Neben dem vom Hersteller empfohlenen PCR-Protokoll wurde das Standard-PCR-Protokoll mit Elongationszeiten von 3-8 min verwendet.

Bei sehr GC-reichen DNA-Matrizen wurde zusätzlich 10 % DMSO bei der PCR zugegeben. Bei Reaktionsansätzen mit Phagen-DNA wurde das nicht-ionische Detergenz NP40 (0,1 %) zur Solubilisierung kleinerer Phagenproteine verwendet. Als weitere Zusätze wurden bei Auftreten unspezifischer PCR-Produkte Glycerol (5-10 % Endkonzentration) oder Ammoniumsulfat (60 mM Endkonzentration) verwendet.

Die Amplifikation erfolgte entweder in einem DNA Thermal Cycler (Perkin Elmer, Weiterstadt) oder in einem PTC 200TM bzw. PTC 100TM Gerät (MJ Research, Inc, USA).

Zur Überprüfung der PCR-Reaktion wurde in der Regel 1/10 des Reaktionsansatzes auf einem entsprechend konzentrierten Agarose- oder Polyacrylamidgel aufgetrennt.

2.7 RT-PCR (Reverse Transkriptase-Polymerasekettenreaktion)

Für die Erststrangsynthese der isolierten RNA (reverse Transkription) und der anschließenden Amplifikation spezifischer cDNA-Fragmente wurde der „RT-PCR-Kit“ (Perkin Elmer, Weiterstadt) nach Angaben des Herstellers verwendet. Pro Reaktion wurden zwischen 1 und 2 µg Gesamt-RNA aus unterschiedlichen humanen und murinen Geweben eingesetzt. Zur reversen Transkription wurden entweder spezifische Gegenstrang-Primer oder Random Hexamere bzw. oligo d(T)-Primer verwendet.

Für die Herstellung größerer Mengen cDNA wurden je Ansatz 4 µg Gesamt-RNA mit DEPC-H₂O auf ein Volumen von 15,9 µl aufgefüllt, 10 min bei 70°C denaturiert und für 2-3 min auf Eis inkubiert. Nach Zugabe eines Reaktionsgemisches (6,25 µM Random Hexamers und/oder Oligo d(T)₁₆, (Perkin Elmer, Weiterstadt), 1 mM der 4 dNTPs, 1 U/µl RNase-Inhibitor (Perkin Elmer, Weiterstadt), 1 x 1st strand buffer (Gibco BRL, USA), 10 mM Dithiothreitol (DTT, Gibco BRL, USA), 2,5 U/µl MMLV-Reverse Transkriptase (Gibco BRL, USA)) wurde der Ansatz für 90 min bei 37°C inkubiert. Nach 10minütiger Inkubation bei 94°C (Hitzeinaktivierung des Enzyms) wurde der Ansatz bis zur weiteren Verwendung bei -20°C gelagert.

2.8 Radioaktive und nichtradioaktive Markierungstechniken

2.8.1 „Random primed oligo labelling“

Für die radioaktive Markierung von DNA-Fragmenten nach Feinberg und Vogelstein (1983) wurden 50-100 ng denaturierte Sonden-DNA, 30 μCi [$\alpha\text{-P}^{32}$] dCTP, 6 μl „Oligolabelling“-Puffer (OLB), 60 μg BSA, 3 U Klenow-DNA-Polymerase (Boehringer Mannheim) in einem 30 μl Ansatz für 2-3 h bei 37°C inkubiert. In den meisten Fällen wurde für die spezifische Markierung von PCR-Produkten OLB-Puffer ohne Zusatz von Hexameren verwendet. Hierbei wurden jeweils die beiden PCR-Primer in einer Konzentration von 0,6 μM eingesetzt.

Zur Abtrennung von nicht eingebaute Nukleotiden wurden die markierten DNA-Fragmente über TE-äquilibrierte Sephadex® G-50 Säulen („Nick Columns“, Pharmacia, Freiburg) aufgereinigt.

2.8.2 „Nick-Translation“

Für Digoxigenin- (Digoxigenin-11-dUTP, Boehringer Mannheim) oder Biotin- (Biotin-16-dUTP, Boehringer Mannheim) markierte Sonden für die Fluoreszenz-*in situ*-Hybridisierung (FISH) an Metaphasechromosomen wurde die „Nick-Translation“ eingesetzt (MacGregor und Mizuno, 1976). In der vorliegenden Arbeit wurde der „Nick-Translations-Kit“ der Fa. Gibco BRL (Eggenstein) entsprechend der Angaben des Herstellers verwendet.

Für die Reaktion wurde ca. 500-800 ng Pst I-restringierte, Phenol/Chloroform-gereinigte Cosmid- oder PAC-DNA verwendet. Die Reaktion wurde durch Einfrieren bei -20°C beendet. Nicht eingebaute Nukleotide wurden durch eine anschließende Fällung des Ansatzes entfernt (vgl. 2.4.5).

2.8.3 Endmarkierung von Oligonukleotiden und PCR-Produkten

2.8.3.1 Endmarkierung von Oligonukleotiden

Die 3'-Endmarkierung von Oligonukleotiden für die RNA-*in situ*-Hybridisierung an Gewebeschnitten (2.10.2) erfolgte entweder mit Hilfe des „DIG Oligonucleotide Tailing Kits“ (Boehringer, Mannheim) nach Angaben des Herstellers oder mit Hilfe der terminalen Deoxynukleotidyltransferase (TdT). Hierfür wurden ca. 80 ng des jeweiligen Oligonukleotids mit 12 µl [α -³⁵S] dATP (Du Pont NEG 034H), 5 µl Reaktionspuffer („cobalt reaction buffer“, Gibco, BRL) auf 18,2 µl mit H₂O aufgefüllt und nach Zugabe von 1,8 µl TdT (Takara, Amersham, Braunschweig) vermischt und für 60 min bei 37°C inkubiert. Die Aufreinigung des markierten Oligonukleotids erfolgte über „Nensorb 20“-Säulen (NEN, Du Pont, USA). Nach Zugabe von 3 ml Methanol wurde die Säule mit 3 ml 0,1 M Tris-HCl pH 8,0 äquilibriert. Im Anschluß wurde die mit 0,5 ml 0,1 M Tris-HCl vermischte Sonde aufgetragen und die uneingebauten Nukleotide mit einer weiteren Zugabe von 1,5 ml 0,1 M Tris-HCl pH 8,0 entfernt. Das markierte Oligonukleotid wurde anschließend mit 0,5 ml 20 %igem Ethanol eluiert, wobei nach Verwerfen des ersten austretenden Tropfens die nachfolgenden 11 in einem Eppendorfgefäß aufgefangen und bis zur weiteren Verwendung bei -20°C gelagert wurden.

2.8.3.2 Endmarkierung von PCR-Produkten

Die radioaktive Endmarkierung von PCR-Produkten für die Auftrennung auf denaturierenden Polyacrylamidgelen (2.4.3) erfolgte über die Übertragung des γ -Phosphatrestes von [γ -³²P] ATP mit Hilfe der Polynukleotidkinase (PNK) an die freie 5'-Hydroxygruppe der PCR-Produkte. Der Reaktionsansatz (10-20 ng des entsprechenden, isopropanolgefällten PCR-Produkts, 10 µCi [γ -³²P] ATP, 1 x PNK-Puffer (NEB), 10 U T4-PNK (NEB)) wurden bei 37°C für 30 min inkubiert. Die Reaktion wurde durch Zugabe von 3 µl Stopplösung (95% Formamid, 0,05 % Bromphenolblau, 0,5 % Xylenxyanol und 20 mM EDTA) abgebrochen und bis zur weiteren Verwendung bei -20°C gelagert.

Die nichtradioaktive Markierung von PCR-Produkten für die GeneScan Fragmentanalysen auf ABI-Sequenzierautomaten (PE Applied Biosystems, Weiterstadt) erfolgte mit Hilfe eines

Standard- (50 µl) PCR-Ansatzes (vgl. **2.6**) unter Verwendung eines nichtmarkierten und eines fluoreszenzmarkierten (Phosphoramidit [TET], PE Applied Biosystems, Weiterstadt) Primers.

2.8.4 „Run off“-Polymerisation (Lineare PCR)

Für die Herstellung strangspezifischer Sonden für Northern-Blot-Analysen wurden ca. 200 ng linearisierter Plasmid-DNA bzw. 10-50 ng eines PCR-Produktes als Templat eingesetzt. Der Reaktionsansatz setzte sich wie folgt zusammen: 0,2 µM eines Vektor-, Integrat- oder PCR-Produkt-spezifischen Primers, 5 µl 10 x PCR-Puffer, je 1 µl dATP, dGTP, dTTP (jeweils 1,25 mM), 1 µl dCTP (6,25 µM), 30 µCi [α -³²P] dCTP sowie 1 U Taq-Polymerase (Gibco BRL, USA). Nach einem initialen Denaturierungsschritt (2 min bei 94°C) erfolgte die Amplifikation bei: 1 min 94°C, 2 min 60°C, 3 min 72°C, 40 Zyklen und einem terminalen Elongationsschritt (7 min 72°C). Die Aufreinigung der markierten Einzelstränge erfolgte wie unter **2.8.1** beschrieben.

2.9 Filterhybridisierung

2.9.1 DNA-DNA-Hybridisierung nach Southern

DNA-Fragmente aus Agarosegelen wurden mit der von Southern (1975) beschriebenen Methode auf Nylonmembranen (Hybond-N⁺, Amersham) transferiert, wobei entweder ein uni- oder bidirektionaler Aufbau („Sandwich“-Blot, Sambrook et al., 1989) verwendet wurde. Als Transferpuffer wurden für beide Verfahren entweder 10 x SSC oder 0,4 M NaOH verwendet. Bei ersterem Puffer wurden die Agarosegele zunächst für 30 min in Denaturierungspuffer und anschließend für 45 min in Neutralisierungspuffer inkubiert. Der DNA-Transfer erfolgte je nach Größe der zu transferierenden DNA Fragmente für 16-48 h. Die Nylonmembran wurde nach Abschluß des Transfers kurz in 2 x SSC gewaschen und für 2 h bei 80°C im Ofen zur Fixierung der DNA „gebacken“ oder durch „UV-Crosslinking“ (UV StratalinkTM 1800, Stratagene, USA) behandelt.

Die Filter wurden anschließend bei 65°C für zwei bis vier Stunden im Hybridisierungssofen prähybridisiert (50 mM Na₂HPO₄, 1 mM EDTA, 1% SDS). Die denaturierte, radioaktiv markierte Sonde (vgl. **2.8.1**) wurde in einem geeigneten Volumen Hybridisierungslösung zugegeben (2×10^5 bis 1×10^6 Cpm/ml „Church“-Hybridisierungs-Puffer, Church und Gilbert, 1984) und für 14-18 h bei 65°C inkubiert. Das Waschen der Filter erfolgte in Lösungen mit abnehmender Na₂HPO₄-Konzentration (100 mM, 50 mM, 20 mM) mit mehrmaligem Wechsel der Waschpuffer für zwei bis vier Stunden bei 65°C.

Die noch feuchten Filter wurden im Anschluß mit Frischhaltefolie umwickelt und mit einem Röntgenfilm (Hyperfilm MP RNP8 oder RNP6, Amersham) bedeckt. Die Exposition der Autoradiographie erfolgte in einer mit Verstärkerfolie ausgekleideten Kassette (Quanta III/Cronex-Kassette, DuPont) bei -70°C für 2 Stunden bis zu 7 Tagen.

2.9.2 DNA-RNA-Hybridisierung

Für die DNA-RNA-Hybridisierungen wurden kommerziell erhältliche humane und murine „Multiple Tissue Northern Blots™“ sowie ein „Mouse RNA Master Blot™“ der Firma Clontech (USA) verwendet.

Die Hybridisierung der Filter erfolgte weitgehend nach den Angaben des Herstellers unter Verwendung der mitgelieferten „ExpressHyb™“-Hybridisierungslösung. Bei bereits mehrfach verwendeten Filtern wurde die Hybridisierungszeit auf 12 h ausgedehnt. Des weiteren wurden die letzten Waschschrte (0,1 x SSC, 0,1 % SDS für 2 x 20 min) zumeist bei 68°C statt 50°C durchgeführt. Als Hybridisierungssonden wurden radioaktiv markierte (RT)-PCR-Produkte oder cDNA-Klon-Integrate verwendet. Die Dokumentation durch Autoradiographie erfolgte wie unter **2.9.1** beschrieben.

2.9.3 Plaquefilterhybridisierung

Die Isolierung von Klonen aus einer humanen kraniofazialen (J. Murray, Iowa, USA) bzw. aus einer Amniozyten-Phagenbibliothek (H. Leffers, Aarhus, Dänemark, unveröffentlicht) erfolgte nach der von Benton und Davis (1977) beschriebenen Methode der Plaquefilterhybridisierung. Hierbei wurde rekombinante virale DNA aus den Plaques der auf einem Bakterien-

rasen ausplattierten Phagen auf zwei Nylonfilter (Hybond-N⁺) im Duplikat durch kurzes Auflegen der Membranen übertragen. Die Filter wurden zur Fixierung der DNA für 2 h bei 80°C gebacken und dann wie unter **2.9.1** geschildert mit einer radioaktiv markierten Sonde (**2.8.1**) hybridisiert, gewaschen und anschließend autoradiographiert. Plaque-Regionen, deren Signale in Duplikat vorhanden waren wurden aus dem Weichagar ausgestochen und in 500 µl SM-Medium mit 50 µl CHCl₃ überführt und bei 4°C aufbewahrt. Nach wiederholtem Ausplattieren und Hybridisieren wurden Einzelplaques isoliert. Von diesen Einzelplaque-Phagensuspensionen wurde jeweils ein Aliquot mit DMSO in einer Endkonzentration von 7% versetzt und als Dauerkultur bei -70°C gelagert.

2.9.4 Koloniefilterhybridisierung

Zur Identifizierung rekombinanter Bakterienklone wurden die ausplattierten Bakterienkolonien oder die in 96er-Loch-Mikrotiterplatten als Dauerglycerinkultur aufbewahrten Bakterien (vgl. **2.12.6**) unter Verwendung eines (96er)-Stempels (Sigma, USA) auf Hybond-N⁺-Filter (Amersham, Braunschweig) übertragen und auf Ampicillin-LB-Agarplatten über Nacht bei 37°C inkubiert. Die Koloniefilter wurden im Anschluß nach der von Sambrook et al. (1989) geschilderten Methode prozessiert. Alternativ wurden die Filter ohne weitere Vorbehandlung einem einfachen 1minütigen Autoklavieren (120°C, 1 Bar, Aufbrechen der Bakterien und gleichzeitiges Fixieren der DNA) unterzogen. Die Hybridisierung und Detektion wurde wie unter **2.9.1** beschrieben durchgeführt.

2.10 *in situ*-Hybridisierung

2.10.1 Fluoreszenz-*in situ*-Hybridisierung (FISH) von Metaphasechromosomen

Metaphasen von menschlichen Lymphozyten wurden wie von Lemieux et al. (1992) und Klever et al. (1991) beschrieben präpariert und denaturiert.

Ca. 1 µg aufgereinigte Cosmid- oder PAC-DNA wurde mit Pst I geschnitten und nach einer Phenol/Chloroform-Aufreinigung mit DIG-11-dUTP bzw. Biotin-16-dUTP (Boehringer,

Mannheim) mit Hilfe der Nick-Translation (vgl. **2.8.2**) markiert. Die markierte DNA wurde in Gegenwart von Cot-1-DNA (GibcoBRL, Eggenstein) im Verhältnis 1:60 sowie einzelsträngiger Forellen-Sperma-DNA (Pharmacia, Schweden, Konz.: 1 µg/mg Cot-1-DNA) mit Ethanol gefällt (vgl. **2.4.5**) und in 50 % Formamid, 2 x SSC, 10 % Dextransulfat resuspendiert. Zur Absättigung repetitiver Sequenzen zur Unterdrückung unspezifischer Hybridisierungen (CISS-Hybridisierung, „chromosomal in situ suppression hybridization“) wurde die Probe für 10 min bei 75°C denaturiert und anschließend für 15-30 min bei 37°C inkubiert.

Ca. 0,25-1 µg der Probe wurde auf die Objektträger überführt (ca. 20-30 µl), der anschließend vorsichtig eingedeckelt und mit Fixogum abgedichtet wurde. Als Positivkontrolle der Hybridisierung und zur leichteren Identifizierung des humanen Chromosoms 4 wurde ein digoxigenin- bzw. biotinmarkierter Chromosom 4-spezifischer Alpha-Satellit (Oncor, USA) mithybridisiert (Konzentration: 10 ng/Objektträger). Die Hybridisierung erfolgte für 16 h bei 37°C. Die Objektträger wurden danach 3 x 5 min bei 42°C in 50 % Formamid, 1 x SSC und 3 x 5 min in 0,3 x SSC gewaschen und für 15 min in BSA, 4 x SSC, 0,1% Tween 20 (Sigma) inkubiert. Digoxigeninsignale wurden dann mit anti-DIG-monoklonalen Antikörpern (Maus-IgG, Sigma, USA, Verdünnung 1:1000), TRITC-markierten anti-Maus IgG Antikörpern aus Kaninchen (Sigma, USA, Verdünnung 1:100) und zur Verstärkung des Signals mit TRITC-markierten anti-Kaninchen IgG Antikörpern aus Ziege (Sigma, USA, Verdünnung 1:64) nach der von Boehringer (Mannheim) beschriebenen Methode nachgewiesen. Der Nachweis von Biotinsignalen erfolgte unter Verwendung von 20 µl einer 1 %igen Fluorescein-Avidin-Lösung (Vector Laboratories, USA, in 4 x SSC/0,1 % Tween 20) (45 min 37°C), anschließendem Waschen (37°C, 4 x SSC/0,1 % Tween 20 für 3 x 10 min) und nachfolgender 30 minütiger Inkubation mit einem biotinylierten Anti-Avidin-Antikörper (Vector Laboratories, USA, Verdünnung 1:100). Nach dreimaligem Waschen (wie oben) erfolgte erneut eine Inkubation mit 1 %iger Fluorescein-Avidin-Lösung. Die Hybridisierungen wurden in der Regel mit zwei unterschiedlichen Sonden und nachfolgender gleichzeitiger Detektion der biotinylierten und DIG-markierten Sonden durchgeführt (Zwei-Farben-FISH, Lichter et al., 1990).

Die Chromosomen wurden mit einer 0,4 %igen DAPI-Lösung (4,6-Diamidine-2-Phenylindol-Dihydrochlorid, Boehringer, Mannheim, in 4 x SSC/0,1 % Tween 20) gegengefärbt und die fertigen Präparate in VectashieldTM-Mounting Medium (Vector Laboratories, USA) eingedeckelt und mit Nagellack versiegelt. Die Bildanalyse wurde mit einem Leica (Bensheim) DMRBE Mikroskop mit einem 100x/1.30 PC Fluotar Oelimmersions-Objektiv durchgeführt

und mit einer High-Performance-CCD-Kamera (Applied Imaging, Großbritannien) und einem assoziierten Computer-Software Paket (Cytovision 2.21, Applied Imaging, Großbritannien) dokumentiert.

2.10.2 RNA-*in situ*-Hybridisierung von Gewebeschnitten

Die Untersuchung der Genexpression erfolgte an fetalen Mäusen bzw. an Hirn- und Hodengewebe adulter Mäuse. Zur Präparation embryonaler Mäuse wurden weibliche Mäuse (H/H, C57/BL6) nach erfolgter Kopulation (weißlicher Vaginalpropf am darauffolgenden Morgen) isoliert und nach 13,5 oder 14,5 Tagen getötet.

2.10.2.1 Herstellung von Schnittpräparaten

Für die RNA-*in situ*-Hybridisierung an Gewebeschnitten mittels radioaktiv markierter Oligonukleotide wurden die Embryonen über Nacht (16-24h) bei Raumtemperatur in 4% PFA in PBS fixiert. Nach einer weiteren über Nacht-Inkubation in 0,5 M Sucrose, 1 x PBS bei 4°C wurden die Embryonen als Blöckchen in Tissue Tek (Miles, USA) eingebettet, in flüssigem Stickstoff eingefroren und anschließend bei -70°C gelagert.

Von diesen Blöckchen wurden mit einem Kryostat Schnitte von 7-10 µm Dicke angefertigt, auf Objektträger (ProbeOn, Fisher Biotech, USA) aufgebracht und für 20 min bei 40°C auf einem Heizblock getrocknet. Die Schnitte wurden anschließend in 4% PFA/1x PBS präfixiert und zum Abbau zellulärer Proteine für 8 min mit Pronase E (0,125 mg/ml in 1x P-Puffer) behandelt. Im Anschluß erfolgte für 10 min eine Acetylierung mit frischer Lösung AC. Die einzelnen Schritte wurden jeweils durch 3 Waschschrirte (jeweils 5 min bei Raumtemperatur in 1 x PBS) beendet. Anschließend wurden die Schnitte durch jeweils 3 minütige Inkubation in einer aufsteigenden Ethanolreihe bei Raumtemperatur (70%, 90%, 100%) dehydriert. Die luftgetrockneten Schnitte (ca. 1-2 h) wurden bis zur weiteren Verwendung bei -70°C trocken gelagert.

Für die nichtradioaktive RNA-*in situ*-Hybridisierung an Schnitten adulter Hoden- und Hirngewebe wurden die freipräparierten Gewebe direkt auf Trockeneis eingefroren, in

Blöckchen eingebettet und für die Erstellung von Gewebeschnitten verwendet. Sie wurden ohne weitere Behandlung bis zur weiteren Verwendung bei -70°C gelagert.

Die histologische Hämatoxylin-Eosin-Färbung der Schnittpräparate erfolgte nach den Angaben des Herstellers (Sigma).

2.10.2.2 Hybridisierung

Ca. 10^7 Cpm der ^{35}S -markierten Oligonukleotide wurden mit 0,99 ml Hybridisierungslösung (900 μl Hybridisierungsmischung, 50 μl Heringssperma DNA, 40 μl 5 M DTT) vermischt und für 30 min bei 42°C inkubiert. Jeweils 300 μl dieser Lösung wurden luftblasenfrei auf einen Objektträger pipettiert und anschließend in einer feuchten Kammer (4 x SSC, 50 % Formamid) für 14-16 h bei 42°C inkubiert. Nach anschließendem Waschen (1x 3 min 1 x SSC, RT; 4 x 15 min 1 x SSC, 55°C ; 1 x 3 min 1 x SSC, RT; 1 x 3 min H_2O , RT) und Dehydrierung (1 min 60 %iger Ethanol, 1 min 95 %iger Ethanol) wurden die Schnitte unter Zuhilfenahme eines handelsüblichen Föns getrocknet. Die Objektträger wurden in einer Autoradiographiekassette mittels eines doppelseitigen Klebebandes fixiert, mit einem Röntgenfilm bedeckt (Hyperfilm β -max, Amersham, Braunschweig) und für ca. 4-5 Tage bei RT exponiert. Nach der Entwicklung des Röntgenfilmes wurden die Schnitte im Dunkeln für ca. 10 sec mit photographischer Lösung bedeckt (Kodak NTB 2), über Nacht bei RT getrocknet und für 3-4 Wochen bei 4°C exponiert.

Für die RNA-*in situ*-Hybridisierung an Gewebeschnitten mittels DIG-markierter Oligonukleotidsonden (vgl. 2.8.3.1) wurden 30 μl der 1:50 (2 x SSC/4% PFA) verdünnten Sonden luftblasenfrei auf jeweils einen zuvor für 15 min mit 4% Paraformaldehyd/1x PBS behandelten Gewebeschnitt pipettiert und für 3 h bei 33°C in einer feuchten Kammer inkubiert. Die Objektträger wurden anschließend für 3 x 5 min in 2 x SSC (33°C), 2 x 10 min in 0,1 x SSC (33°C) gewaschen und für 5 min bei RT in Puffer PBT inkubiert. Nach 15 minütiger Inkubation in Blockierungslösung (1% (w/v) „Blocking Reagenz“, Boehringer Mannheim in Puffer 1*) wurden pro Objektträger ca. 50 μl 1:100 in Blockierungslösung verdünnte AP-konjugierte anti-Digoxigenin Fab-Fragmente (Boehringer, Mannheim) zugegeben und für 1 h bei 37°C inkubiert. Die Substratreaktion erfolgte nach anschließendem Waschen der Schnitte für 2 x 5 min in PBT, 1 x 5 min in NBT nach Angaben des Herstellers mit einer NBT/BCIP-haltigen Lösung (Boehringer, Mannheim). Die Reaktion wurde nach ca. 30 bis 60 min bis über

Nacht-Inkubation durch 5minütige Inkubation in PBS abgestoppt und die Objektträger nachfolgend mit einem Eindeckmittel eingedeckelt.

2.11 Isolierung von PAC- bzw. P1-Klonen

Für die Isolierung von PAC- („P1 derived artificial chromosomes“) Klonen (Ioannou et al., 1994) wurde eine gesamtgenomische, hierarchisch angeordnete pCYPAC2-Bibliothek verwendet, die freundlicherweise von Pieter de Jong (Buffalo, USA) zur Verfügung gestellt wurde. Diese Bibliothek umfasst ca. 123 000, in 384-Loch-Mikrotiterplatten angeordnete Einzelklone. Durch die weitere Anordnung dieser Bibliothek nach einem von Munroe et al. (1995) beschriebenen Vereinigungsschema in 20 Spalten- und 16 Reihen-„Poole“ (Prawitt, 1999) konnten nach PCR-Screening über positive Signale meist mehrere mögliche Plattenpositionen bestimmt werden. Die Bestimmung der Einzelpositionen erfolgte nachfolgend über Koloniefilterhybridisierung (vgl. 2.9.4).

Alternativ wurden für die Isolierung von P1-/PAC-Klonen kommerziell erhältliche, automatisiert hergestellte „high density“-Filter vom Humanen Ressourcenzentrum im Berlin (RZPD) verwendet. Für das Screening wurden Filter der murinen genomischen Bibliotheken 703 (C57/BL6-Bibliothek, F. Francis unveröffentlicht, mit dem Vektorsystem pAd10SacBII (Pierce et al., 1992; Präfix ICRFP); 711 (129/SvevTACfBr-Bibliothek, P. de Jong und K. Osoegawa, unveröffentlicht, mit dem Vektorsystem pPAC4, Frengen et al., unveröffentlicht; Präfix RPCIP), sowie der humanen genomischen Blutzell-Bibliothek 704 (P. de Jong, Vektorsystem pCYPAC2, Präfix RPCIP) benutzt. Die Auswertung der über die Hybridisierung ermittelten Einzelklon-Positionen wurde anhand der vom RZPD vorgegebenen Kodierungsschlüssel vorgenommen. Positive Klone wurden vom RZPD angefordert, auf LB-Agaroseplatten (50 µg/ml Kanamycin) durch Ausstreichen vereinzelt und nochmals mittels PCR oder Koloniefilterhybridisierung überprüft.

2.12 Exonamplifikation

Die Methode der Exonamplifikation wurde von Buckler et al. (1991) entwickelt und ermöglicht die direkte Isolierung von Exons aus klonierter genomischer DNA. Die Grundlagen dieser Methode stellen intakte Spleißstellen und ein speziell entwickelter Vektor (pSPL1) dar, der zu einem 6031 Bp Vektor (pSPL3) weiterentwickelt wurde (Church et al., 1994). Dieser Vektor kann als Shuttle-Vektor sowohl in bakteriellen (*E. coli*) als auch in eukaryotischen Zellen (COS7) propagiert werden, wobei die Selektion positiver, bakterieller Klone durch ein Ampicillinresistenzgen vereinfacht wird. Für die Expression klonierter DNA in Eukaryotenzellen besitzt der Vektor einen SV40 Promotor (Buckler et al., 1991).

Die multiple Klonierungsstelle liegt innerhalb von HIV1-tat-Intronsequenzen, an die sich beidseitig HIV1-tat-Exonsequenzen und β -Globinsequenzen aus Kaninchen anschließen und somit intakte Spleißdonor- bzw. akzeptorstellen bilden. Zur besseren Identifikation positiver Klone wurden zusätzlich zu einer Bst XI-Schnittstelle innerhalb der multiplen Klonierungsstelle halbe Bst XI-Schnittstellen nahe den Spleißstellen eingefügt (Church et al., 1994).

Nach Transfektion in COS7-Zellen wird klonierte genomische DNA transkribiert und *in vivo* prozessiert. Wird kein funktionelles Exon an die vektoreigenen Spleißstellen gespleißt, wird die Bst XI-Schnittstelle regeneriert und kann nach RT-PCR mit den Vektorprimern SA2 und SD6 zur Eliminierung dieser Hintergrundprodukte verwendet werden. Erfolgt ein Spleißvorgang durch die Verwendung kryptischer Spleißstellen innerhalb der einklonierten bzw. der vektoreigenen DNA, bleibt die Bst XI-Schnittstelle, die die multiple Klonierungsstelle (MCS) flankiert, erhalten und kann ebenfalls zur Eliminierung der entsprechenden Hintergrundprodukte verwendet werden. Erfolgreich gespleißte Exons werden mit den weiter intern liegenden Primern SA4 und SD2 amplifiziert und enthalten somit die sich unmittelbar an die Spleißstellen anschließenden Sal I- (5') und Bgl II- (3') Schnittstellen, die eine gerichtete Klonierung in den Vektor pBlueskript erlauben.

2.12.1 Subklonierung von Cosmid- und PAC-Klonen in den Vektor pSPL3

Um eine Vielzahl möglichst überlappender Fragmente zu erhalten, die innerhalb eines als sehr effizient beschriebenen Größenbereichs von 3,5-4,5 Kb liegen (Burfoot und Campbell, 1994),

wurden die Cosmid- bzw. PAC-Klone (**2.1** und **2.11**) mit Bam HI/Bgl II und in getrennten Ansätzen mit Pst I oder auch mit Eco RI restringiert. Die Fragmente wurden in die jeweiligen Pst I-, Bam HI/Bgl II- oder Eco RI-Schnittstellen des pSPL3 ligiert, der zur Vermeidung einer Religation zuvor mit alkalischer Phosphatase dephosphoryliert wurde (vgl. **2.5**).

Auf eine Isolation der Integrate vor der Subklonierung konnte mit Ausnahme des Cosmids 8C101I14 verzichtet werden, da die verwendeten Vektoranteile der Cosmid- bzw. PAC-Klone keine Ampicillin-Resistenzgene enthalten und somit nicht zur Bildung künstlicher Vektorprodukte beitragen können. Diese künstlichen Konstrukte, die sich aus einem Replikationsorigin und einem funktionstüchtigen Ampicillinresistenzgen zusammensetzen, können bei der nachfolgende Ampicillin-Selektion zu einem sehr hohen Anteil an Hintergrundklonen führen (Buckler et al., 1991; Löbber, 1996). Das Integrat des Cosmids 8C101I14 wurde deshalb vor der Subklonierung mittels Not I-Verdau und nachfolgender Gelaufreinigung (**2.4.4**) isoliert.

Die Ligations- und Transformationsbedingungen wurden so gewählt, daß je nach dem Ursprungsklon eine möglichst effiziente Subklonierung aller Fragmente gewährleistet werden konnte. Nach Ligation und Transformation kompetenter *E. coli* DH5 α -Zellen (vgl. **2.5**) wurde der gesamte Ansatz auf ampicillinhaltigen LB-Agarplatten ausgestrichen und über Nacht bei 37°C inkubiert. Eine durchschnittliche Kolonienzahl von 200-500 Kolonien (10fache Redundanz) wurde hierbei als ausreichend erachtet. Die entstandenen Kolonien wurden mit LB-Ampicillin-Medium abgeschwemmt und für 1-2 h bei 37°C geschüttelt. Plasmid-DNA wurde wie unter **2.2.3** beschrieben präpariert.

2.12.2 Kultur von COS7-Zellen

COS7-Zellen (Gluzman, 1981) wurden in Dulbecco's modified Eagle Medium (DMEM, MacPherson und Stoker, 1962), das mit 10% hitzeinaktiviertem, fetalen Kälberserum, Glutamin (2mM) und Antibiotika (Penicillin-Streptomycin; 100 μ g/ml) komplettiert worden war, bei 37°C und 0,5 % CO₂ kultiviert. Die Zellen wurden in T-175 Kulturflaschen (Costar, Bodenheim) zur 75 bis 80 %igen Konfluenz angezüchtet und dann für die Elektroporation verwendet.

2.12.3 Elektroporation

Die COS7-Zellmenge einer T-175-Zellkulturschale wurde für ca. 3 Transfektionen verwendet. Hierfür wurde das Medium abpipettiert und die adhärennten COS7-Zellen mit 10 ml 0,05 % Trypsin/0,02 % EDTA in PBS-Puffer gewaschen, abgelöst, in ein 50 ml Falcon-Röhrchen überführt und abzentrifugiert (1200 Upm, 4°C, 8 min in einer Sorvall RT 6000D, Dupont, USA). Der Überstand wurde verworfen, das Zellpellet in 30 ml eiskaltem PBS resuspendiert und erneut abzentrifugiert. Das Zellpellet wurde dann in 2,1 ml eiskaltem PBS-Puffer aufgenommen.

Ca. 10-20 µg DNA in 100 µl PBS wurden in 0,4 cm Elektroporationsküvetten (BioRad, München) mit 0,7 ml der vorbereiteten COS-Zellen aufgefüllt, gemischt und 10 min auf Eis inkubiert. Nach erneutem Resuspendieren wurde die Küvette in den GENE Pulser (BioRad, München) eingesetzt und bei 1,2 kV und 25 µF ein Puls gesetzt. Zur Regeneration der Zellen wurde die Küvette 10 min auf Eis inkubiert. Anschließend wurde die Zellsuspension in 10 ml präequilibriertem DMEM-Medium in 100 mm x 100 x 15 Nunclon-Zellkulturpetrischalen (Intermed. Nunc, Dänemark) überführt und für 48-72 h im Brutschrank inkubiert.

2.12.4 RNA-Isolation aus COS7-Zellen

Ca. 48-72 h nach der Transfektion wurde die RNA isoliert. Hierbei erfolgten alle weiteren Schritte auf Eis.

Nach Abziehen des Mediums wurden die Zellen 3 x mit 10 ml PBS gewaschen, mit einem Zellschaber gesammelt und in 10 ml PBS aufgenommen. Nach einer kurzen Zentrifugation wurde das Zellpellet dann in TKM-Lösung resuspendiert, die Zellen mit Triton-X-100 aufgebrochen und zur Trennung der RNA von den restlichen Zellbestandteilen erneut zentrifugiert. Der RNA-haltige Überstand wurde einer Phenol und einer Phenol/Chloroform-Extraktion (1:1) unterzogen und anschließend mit 5 M NaCl und EtOH abs. bei -70°C über Nacht gefällt. Das RNA-Pellet wurde mit 70 %igem EtOH gewaschen, getrocknet, in 20 µl ¼ TE-Puffer gelöst. 1 µl hiervon wurde zur Konzentrationsbestimmung auf einem 1 %igen Agarosegel aufgetrennt.

2.12.5 Synthese der cDNA

Die reverse Transkription der RNA erfolgt in 25 µl Ansätzen mit 40 mM KCl, 10 mM Tris/HCl (pH 8,3), 1 mg/ml Gelatine, 1,5 mM MgCl₂, 4 mM Dithiothreitol, 1 µM des Vektor-Primers SA2 und je 200 µM der vier dNTPs sowie etwa 2,5-5 µg RNA. Dem Gemisch wurde nach 5 minütiger Inkubation bei 65°C und nachfolgender Inkubation auf Eis RNase-Inhibitor (3,5 U RNasin, Promega) sowie die Reverse Transkriptase (MMLV R, 200 U, BRL) zugegeben. Der Ansatz wurde für 90 min bei 42°C inkubiert und im Anschluß für die Zweitstrangsynthese verwendet.

2.12.6 Amplifikation der DNA-Fragmente und Klonierung in pBluescript

Die Zweitstrangsynthese und die ersten nachfolgenden Amplifikationsschritte wurden unter Verwendung des gesamten RT-Ansatzes (2.12.5) in sechs PCR-Zyklen (jeweils 1 min 94°C, 1 min 60°C, 5 min 72°C) in 100 µl Ansätzen durchgeführt (1/10 Vol. PCR-Puffer, 1 µM Primer SD6, je 200 µM der 4 dNTPs sowie 2,5 U Taq-Polymerase (AmpliTaq®, Perkin-Elmer, Weiterstadt). Nach Bst XI-Restriktion mit 50 U (über Nacht, 50°C) wurde am darauffolgenden Tag für 1 h mit 10 U Bst XI nachgeschnitten. Von diesem Ansatz wurden 10 µl für die zweite PCR verwendet (jeweils 1 µM der Primer SA4 und SD2, Extension von nunmehr 3 min, 35 Zyklen, 100 µl Ansätzen, sonst wie oben angegeben). Nach Aufreinigung mit einer Phenol/Chloroform/Isoamylalkohol (25:24:1) und einer Chloroform-Extraktion wurden die PCR-Produkte gefällt und mit Sal I und Bgl II geschnitten.

Die nachfolgende Ligation erfolgte in den mit Sal I und Bam HI geschnittenen Vektor pBluescript II SK+ (Stratagene, Heidelberg). Anschließend wurden die Konstrukte durch Transformation in *E. coli* DH5α überführt, selektiert (vgl. 2.5) und zur weiteren Charakterisierung in 96-Loch-Mikrotiterplatten gepickt.

2.12.7 Identifizierung von „echten“ Exonsequenzen

Für die weitere, systematische Analyse der großen Zahl der entstehenden Exonklone wurden folgende Strategien alternativ verwendet:

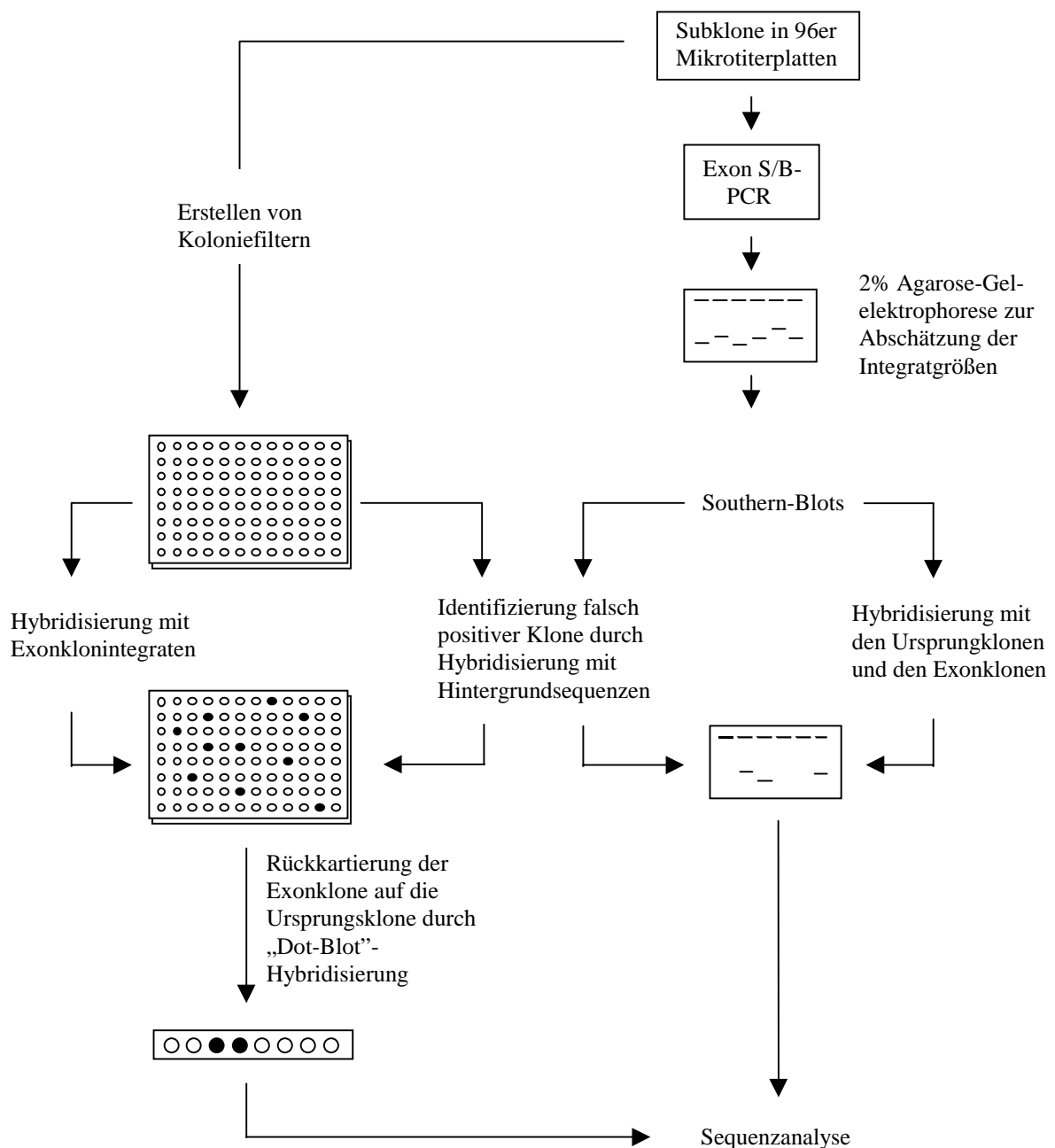


Abb. 2.1: Strategie der Subklonanalyse. Erläuterung siehe Text.

Die Integratgrößen der Subklone wurden durch PCR mit den speziell entwickelten Primern Exon S und B oder den Bluescript-spezifischen Primern T₃A und T₇A bestimmt (vgl. **2.17**). Die Sequenzierung einer Reihe willkürlich ausgewählter PCR-Produkte ergab eine Reihe an Hintergrundsequenzen, deren Zahl im Verlaufe der Exonamplifikationsexperimente erweitert werden konnte. Diese Sequenzen setzen sich vor allem aus repetitiven Elementen (Alu-, LINE-Elemente), HIV- und anderen Anteilen des Spleißvektors sowie der Cosmid- und PAC-Vektoren (wie beispielsweise Replikationsoriginanteilen) zusammen. Diese Hintergrundsequenzen wurden für die Hybridisierung der oben aufgeführten Kolonie- und Southern-Blot-Filter (**2.9.4**, **2.9.1**) zur Identifizierung falsch positiver Klone verwendet. Aus den Exonklonen, die keine Hybridisierungssignale mit den Hintergrundsequenzen aufwiesen, wurden weitere Klone mit unterschiedlichen Integratgrößen ausgewählt und nach PCR-Amplifikation, Gelaufreinigung (**2.4.4**) und radioaktiver Markierung (**2.8.1**) für die Hybridisierung der gleichen Filter verwendet. Die wiederholte Durchführung dieses Verfahrens ermöglichte letztlich die vollständige, effiziente Erfassung aller putativ „echten“ Exonklone und darüber hinaus die Vermeidung von Mehrfachsequenzierungen identischer Klone.

Für die Rückkartierung der Exonklone auf die Ursprungsklone wurden entweder Southern-Blot-Filter der gelelektrophoretisch aufgetrennten Subklonintegrate bzw. Dot-Blot-Filter mit DNA der Ursprungsklone oder den Subklonintegraten (Exons) wie unter **2.9.1** beschrieben, hybridisiert.

2.13 Herstellung und Screening einer hierarchisch angeordneten fetalen Hirn-cDNA-Bibliothek mittels PCR

Eine normalisierte fetale Hirn cDNA-Bibliothek (NIB, „normalised infant brain“) im Plasmid-Vektorsystem LafmidBA (Soares et al., 1994) wurde freundlicherweise von Marcelo Bento Soares (NIH, Bethesda) zur Verfügung gestellt.

Die Anordnung dieser Plasmidbibliothek erfolgte nach einem von Munroe et al. (1995) beschriebenen Vereinigungsschema („Pooling“-Strategie) und wurde so gewählt, daß mindestens 150 000 Einzelklone der ursprünglich ca. 47 000 Klone umfassenden Bibliothek repräsentiert sind (3-4 fache Redundanz).

Etwa 1 ng Plasmid-DNA wurde über Elektroporation mit einem BioRad (München) „GenePulser“ unter Verwendung von 0,1 cm Elektroporationsküvetten (BioRad) und *E.coli* DH10B-Zellen nach den Angaben des Herstellers transformiert. Der Bakterientiter wurde nach 10-12 stündiger Inkubation bei 37°C bestimmt. Aliquots des Transformationsansatzes wurden auf 576 LB-Agarplatten ausplattiert und für 12-16 h bei 37°C inkubiert. Die entstandenen Bakterienkolonien wurden mit ca. 5 ml LB-Medium (Ampicillin) abgeschwemmt. Jeweils 500 µl dieser Suspension wurden in 96-Loch-Mikrotiterplatten (Costar) überführt und mit 500 µl 2x FM vermischt und bei -70°C eingefroren. Die restlichen Ansätze wurden zur Erhöhung der DNA-Ausbeute vor der Plasmid-Präparation (vgl. 2.2.2) für ca 2 h bei 37°C geschüttelt. Die Plasmid-DNAs wurden in 20 µl TE gelöst und in 6 Mikrotiterplatten entsprechend der Anordnung der Dauerglycerinkulturen überführt. Durch OD-Messungen und 1 %ige Agarosegele mit Referenzproben bekannter Konzentration konnte die Konzentration der Plasmid-DNA-Präparationen abgeschätzt werden. Jeweils 8 µl der Einzelwells wurden zu 6 „Pools“ der 6 einzelnen Mikrotiterplatten mit Hilfe von Mehrkanalpipetten vereinigt und jeweils 5 µl wurden zu 24 Reihen- und 24 Spalten-Pools nach dem nachfolgenden Schema vereinigt:

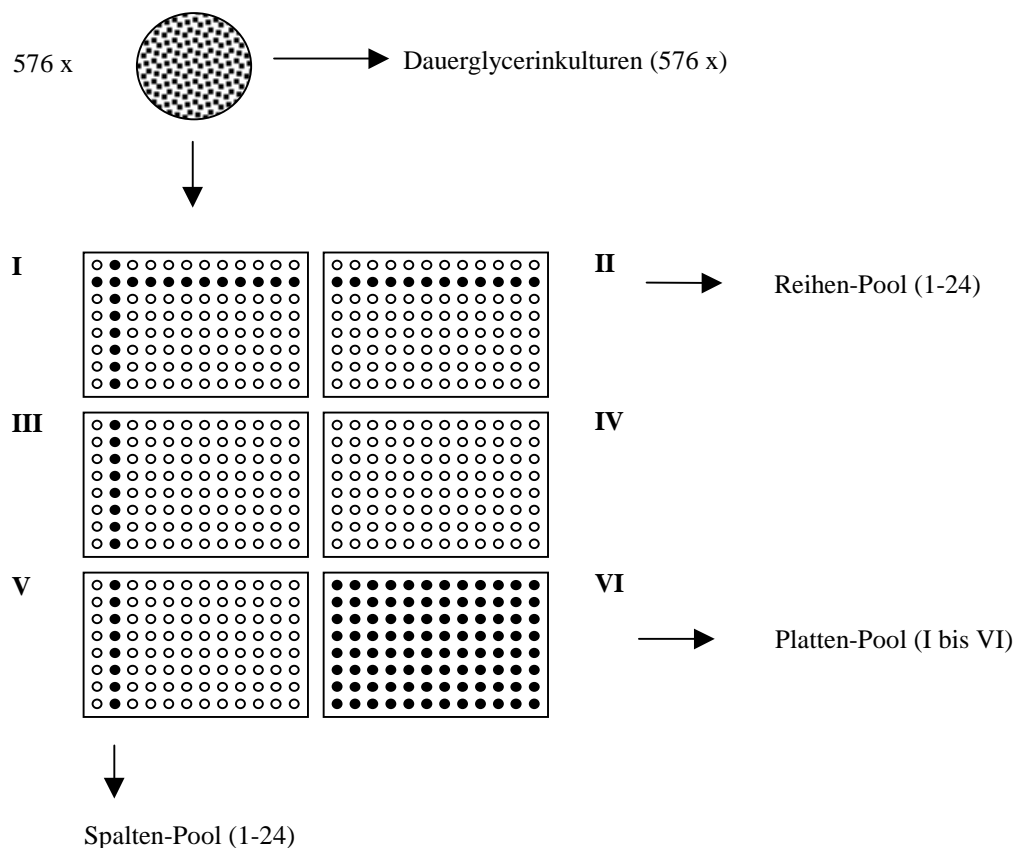


Abb. 2.2: Vereinigungs- (Pooling) Strategie für die Erstellung der hierarchisch angeordneten cDNA-Bibliothek. Die Plasmid-DNA der 576 Präparationen wurde in 6 96er Loch-Mikrotiterplatten gesammelt und vereinigt. Plasmid-DNA jeweils einer Mikrotiterplatte wurde für die Platten-Pools vereinigt. DNA der einzelnen Längs- und Querreihen dreier bzw. zweier Platten wurde für die Herstellung der jeweils 24 Reihen- und 24 Spalten-Pools vereinigt.

Für ein nachfolgendes PCR-Screening wurden die Pools weiter mit $\frac{1}{4}$ TE auf eine Konzentration von ca. 50-100 ng (Plattenpools) und 10-20 ng (Spalten- und Reihen-Pools) verdünnt. Somit konnte das PCR-Screening mit den 6 Platten-Pools begonnen werden, bei positiven Signalen wurden Reihen- und Spalten-Pool-PCRs angeschlossen, die Rückschlüsse auf die Einzelposition der Klone liefern, die ebenfalls durch PCR überprüft wurden. Die einer positiven Einzelposition entsprechenden Bakterienglycerinkultur, die ca. 80 unterschiedliche Einzelklone enthält, wurde in verschiedenen Verdünnungen ausplattiert. Durch Koloniefilterhybridisierung (2.9.4) mit dem radioaktiv markierten PCR-Produkt konnten so Einzelklone identifiziert werden.

2.14 *In vivo*-Excision von pBluescript aus λ ZapII mit Hilfe des ExAssist™/SOLR™ Systems

Die aus der kraniofazialen- (J. Murray) sowie der Amniocyten- (H. Leffers, unveröffentlicht) cDNA-Bibliotheken durch Plauefilterhybridisierung identifizierten cDNA-Klone λ cf16, AMA3 und AMA11 wurden durch *in vivo*-Excision in pBluescript enthaltene Bakterienklone überführt. Hierzu wurden jeweils 100 μ l der Phagensuspensionen mit 200 μ l XL1-Blue Zellen ($OD_{600}=1,0$) und 1 μ l ExAssist Helferphagen ($>1 \times 10^6$ Pfu/ml) für 15 min bei 37°C inkubiert. Nach Zusatz von 3 ml 2x YT-Medium wurde der Ansatz bis zur Ausbildung eines klaren Lysates (ca. 3-4 h) bei 37°C geschüttelt. Nach 20 minütiger Inkubation des Ansatzes bei 70°C folgte eine 15minütige Zentrifugation bei 4000g. Jeweils 1-50 μ l des nun pBluescript enthaltenden Überstandes wurden mit 200 μ l SOLR-Zellen ($OD_{600}=1,0$) für 15 min bei 37°C inkubiert. 100 μ l des Ansatzes wurden auf LB-Ampicillin-Agarplatten ausplattiert und bei 37°C für 12-16 h bebrütet.

2.15 DNA-Sequenzierung

Zur DNA-Sequenzierung wurde die auf Sanger et al. (1977) beruhende Kettenabbruchmethode verwendet, die von Lee et al. (1992) für die Markierung von DNA mit fluoreszierenden Didesoxynukleotiden modifiziert wurde.

Die automatische Sequenzierung wurde auf Sequenziergeräten der Typen 373A, 377 oder 377XL der Firma ABI (PE Biosystems, Weiterstadt) durchgeführt. Neben ca. 500 ng bis 1 μ g aufgereinigter Plasmid-, ca. 4 μ g Cosmid- oder 5 bis 10 μ g PAC-DNA (vgl. 2.2.3) wurden 50 bis 300 ng isopropanolgefällte PCR-Produkte (2.4.5) als Matrize verwendet. Die DNA wurde mit Hilfe des „Prism Ready Reaction Dye Deoxy Terminator Cycle Sequencing Kits“ (Applied Biosystems/PE Biosystems, Weiterstadt) in einer linearen PCR-Reaktion sequenziert (nach einer einleitenden Denaturierung 1 min bei 95°C folgten 25 Zyklen mit jeweils 95°C für 18 sec und 59-60°C für 4 min, für PAC- und Cosmid-Direktsequenzierungen wurde ein leicht

modifiziertes Programm verwendet: nach einer einleitenden Denaturierung von 4 min für 95°C wurden 35 Zyklen wie oben angeschlossen).

Die Aufreinigung der Sequenzierungsansätze sowie die Polyacrylamidgelelektrophorese und die automatisierte Auswertung der Sequenzsignale erfolgte gemäß den Angaben des Herstellers (PE Biosystems, Weiterstadt).

2.16 Computerauswertung von DNA- und Protein-Sequenzen

Die Auswertung der Sequenzdaten erfolgte über eine NEXT-Workstation auf dem Zentralrechner (Genius) des Deutschen Krebsforschungszentrums in Heidelberg unter Verwendung des Programmpakets HUSAR sowie über verschiedene Programmpakete, die über das World Wide Web zur Verfügung stehen (ExpASy Molecular Biology Server, <http://www.expasy.ch/>; Pedro's BioMolecular Research Tools, http://www.biophys.uni-duesseldorf.de/bionet/research_tools.html). Für Sequenzvergleiche auf Nukleotidebene wurde die Programme „BLASTN“, „BLAST 2“ und „Gapped BLAST“ und auf Proteinebene die Programme „BLASTX“ am NCBI (National Center for Biotechnology Information) verwendet (Pearson und Lipman, 1988; Altschul et al., 1990; Altschul et al., 1994; Altschul et al., 1997). Als Exon- und Genvorhersageprogramme wurden „GRAIL II“-Analysen („gene recognition and analysis internet link“, Uberbacher und Mural, 1991; Xu et al., 1994) und „GENSCAN“-Analysen verwendet (Burge und Karlin, 1997; Burge und Karlin, 1998).

Die genomischen Sequenzen der hierfür verwendeten Cosmidklone und des PAC-Klons entstammen den Datenbanken des Sanger Centers (Accession Nr.: Z49235 (HSL108f12), Z68226 (HSL141a8), Z67997 (HSL206d7), Z49862 (HSL58b6), Z69638 (HSL174g8A), Z69374 (HSL174g8B), Z69375 (HSL174g8C), Z79579 (HSL158b1), Z49237 (HSL27h9), Z68165 (HSL96a2), Z68164 (HSL19h1), Z68276 (HSL190b4), Z49236 (HSL184d6), Z69651 (HSL75b9A), Z69652 (HSL75b9B), Z69653 (HSL75b9C), Z69376 (HSL75b9D).

2.17 Reagenzien und Materialien

Puffer und Lösungen

Agar Platten	15 g Agar-Agar ad 1000 ml LB- bzw. NZYDT-Medium
Church-Hybridisierungs-Puffer	1 mM EDTA 0,5 M Na ₂ HPO ₄ , pH 7,2 7 % SDS
Denaturierungslösung	1,5 M NaCl 0,5 M NaOH
Denhardts-Lösung, 100 x	10 g Ficoll 10 g Polyvinylpyrrolidon 10 g BSA ad 500 ml
DNA-„BlueMarker“	0,25 % Bromphenolblau 0,25 % Xylencyanol 40 % Sucrose
DNA-„OrangeMarker“	15 g Sucrose 0,175 g Orange G ad 50 ml H ₂ O
DTM	je 100 µM dATP, dGTP, dTTP in 250 mM TrisHCl, 25 mM MgCl ₂ , 50 mM β-Mercaptoethanol; (pH 7,0)
2x FM	65 % Glycerin 0,1 M MgSO ₄ 0,025 M Tris Cl (pH 8,0)
Hybridisierungslösung	100 ml Formamid 40 ml 20 x SSC 2 ml 100 x Denhardts-Lösung 10 ml 20 % Sarkosyl 20 ml 0,2 M NaPO ₄ (pH 7,0) 20 g Dextranulphat ad 200 ml mit H ₂ O
LB-Medium	10 g Trypton 5 g Hefeextrakt 10 g NaCl ad 1000 ml Aqua bidest., (pH 7,5)
Lösung AC	295 ml H ₂ O 4 ml Triethanolamin 0,525 ml HCl konz. 0,75 ml Essigsäure-Anhydrid

Lösung I	50 mM Glucose 25 mM Tris 10 mM EDTA, (pH 8,0)
MOPS (10x)	0,2 M 3(N-morpholin)Propansulfonsäure 0,05 M Na-Acetat 0,01 M EDTA (pH 7,0)
Neutralisierungslösung	0,5 M Tris-HCl, pH 7,0 1,5 M NaCl
NZYDT-Medium	5 g NaCl 5 g Hefeextrakt 10 g NZ-Amine, (pH 7,5)
OL	90 OD U 5'-pd(N ₆)/ml TE
OLB	1M HEPES/ DTM/ OL (25/25/7)
PBS-Puffer	120 mM NaCl 2,7 mM KCl 10 mM Kalium-Phosphatpuffer; (pH 7,4)
PBS-Puffer*	137 mM NaCl 2,7 mM KCl 1,15 mM KH ₂ PO ₄ 6,5 mM Na ₂ HPO ₄ 2H ₂ O
PBT	PBS* 2 mg/ml BSA 0,1% Tween-20
PTW	PBS* + 0,1 % Tween-20
Puffer 1*	100 mM Tris/HCl 150 mM NaCl (pH 7,5)
Puffer 1	50 mg Tris-HCl 10 mM EDTA 100 µg/ml RNase A; (pH 8,0)
Puffer 2	200 mM NaOH 1 % SDS
Puffer 3	2,55 M Kaliumacetat; (pH 4,8)
Puffer QBT	750 mM NaCl 50 mM MOPS 15 % Isopropanol 0.15 % Triton X-100; (pH 7,0)
Puffer QC	1,0 mM NaCl 50 mM MOPS 15 % Isopropanol; (pH 7,0)
Puffer QF	1,24 M NaCl sonst wie Puffer QC
RNA-Loading buffer	500 µl Formamid (deionisiert) 100 µl 10 x MOPS-Puffer 150 µl Formaldehyd (filtriert, 37 %)

SM-Medium	0,1 M NaCl 50 mM Tris-HCl, pH 7,5 8 mM MgSO ₄ 0,1 % Gelatine
SOB-Medium	20 g Trypton 5 g Hefeextrakt 0,5 g NaCl ad 1000ml A.bidest mit 2,5 mM KCl und 10 mM MgCl ₂ (pH 7,0)
SOC-Medium	SOB-Medium 20 mM Glucose
Standard-Saline-Citrat-Puffer (1x SSC)	0,15 M NaCl 0,015 M Natriumcitrat; (pH 7,0)
Tris-Borat-Elektrophoresepuffer (1xTBE)	90mM Tris 90 mM Borsäure 1,25 mM Na ₂ EDTA; (pH 8,3)
Tbf I	30 mM K Acetat 50 mM MnCl ₂ 100 mM KCl 10 mM CaCl ₂ 15 % Glycerin; (pH 5,8)
Tbf II	10 mM NaMOPS, pH 7,0 75 mM CaCl ₂ 10 mM KCl 15 % Glycerin
TE-Puffer	10 mM Tris-HCl; (pH 7,8) 1 mM Na ₂ EDTA
TKM	10 mM Tris, pH 7,5 10 mM KCl 1 mM MgCl ₂
2x YT-Medium	10 g NaCl 10 g Hefeextrakt 16 g Trypton; ad 1000 ml A. bidest.

Radioisotope, Enzyme, Antikörper und Nukleinsäuren

Anti-Mouse-TRITC (Rabbit IgG-Konjugat)	Boehringer (Mannheim)
Anti-Rabbit-TRITC (Goat IgG-Konjugat)	Boehringer (Mannheim)
[α - ³² P] dCTP, [γ - ³² P] ATP	Amersham (Braunschweig)
Alkalische Phosphatase	Boehringer (Mannheim)
Cot-1-DNA	Gibco BRL (USA)
DNase I	Boehringer (Mannheim)
[α - ³⁵ S] dATP	Du Pont (USA)

Klenow-DNA-Polymerase	New England Biolabs (USA)
Lysozym	Boehringer (Mannheim)
Proteinase K	Boehringer (Mannheim)
Reverse Transkriptase	Perkin Elmer (USA)
	Gibco BRL (USA)
RNAse A	Boehringer (Mannheim)
Salmon sperm DNA	Pharmacia (Schweden)
Taq-DNA-Polymerase	Gibco BRL (USA)
	Boehringer (Mannheim)

Bakterienstämme

DH5 α	F ⁻ , end A1, hsdR17, (r _k ⁻ , m _k ⁺), supE44, thi-1, Lambda-, recA1, gyrA96, relA1, 80dlacZ Δ M15
DH10B	F ⁻ , mcrA Δ (mrr-hsdRMS-mcrBC), 80dlacZ Δ M15, Δ lacX74, deoR, recA1, endA1, araD139, Δ (ara, leu)7696, galU, galK, λ -, rpsL, nupG
SOLR TM	e14(mcrA), Δ (mcrCB-hsdSMR-mrr)171, bcC, recB, recJ, umuC: Tn5(kan ^r), uvrC, lac, RyA96, rel A1, thi-1, endA1, lambda ^R , [F ⁺ proAB, lacI ^q Z Δ M15] Su ⁻ (nonsuppressing)
XL1-Blue	recA1, endA1, gyrA96, thi-1, hsdR17, supE44, elA1, lac, [F ⁺ proAB, lacI ^q Z Δ M15], Tn10(tet ^r)

Molekulargewichtsstandards

100 Bp-Leiter	Gibco BRL (USA)
123 Bp-Leiter	Gibco BRL (USA)
λ x Hind III-Marker	Boehringer (Mannheim)
λ DNA PFGE-Marker	Pharmacia (Schweden)
RNA-Leiter	Gibco BRL (Eggenstein)

Chemikalien und Materialien

Chemikalien:

Agar-Agar	Difco (USA)
Agarose	Eurogentec (Belgien)
Ampicillin	Ratiopharm (Ulm)
Antifade-Lösung (Vecta- Shield Mounting Medium)	Vector Laboratories (USA)
BSA	NEB (USA)
Blocking Reagent	Boehringer (Mannheim)
Colcemid	Gibco BRL (USA)
DAPI	Boehringer (Mannheim)
EDTA	Boehringer (Mannheim)
Ethanol	Riedel-de-Haen (Seelze)
Ethidiumbromid	Oncor (USA)
Fetales Kälberserum	Sigma (USA)
Fluorescein-Avidin DN	Vector Laboratories (USA)
Formamid (Ultrapure)	Gibco BRL (USA)
Hefeextrakt	Difco (USA)
IPTG	Boehringer (Mannheim)
Kanamycin	Sigma (USA)
Orange G	Sigma (USA)
Phenol	Roth (Karlsruhe)
RPMI-Medium	Sigma (USA)
SDS	INC (USA)
Szintillationsflüssigkeit „Rotszint“	Roth (Karlsruhe)
Trypton	Difco (USA)
Trypsin-EDTA	Sigma (USA)
Tris	Boehringer (Mannheim)
Tris-HCl	Gerbu (Gaiberg)
X-Gal	Eurogentec (Belgien)

Nichtaufgeführte Chemikalien (Reinheitsgrad „p.a.“ oder reinst wurden von folgenden Firmen bezogen: Gibco BRL (USA), Merck (Darmstadt), Roth (Karlsruhe), Serva Feinbiochemica (Heidelberg) und Sigma (USA).

Materialien:

Nylonmenbran (Hybond TM -N ⁺)	Amersham Life Science (Braunschweig)
Röntgenfilme (Hyperfilm TM -MP)	Amersham Life Science (Braunschweig)
Röntgenkassetten (Quanta III, Cronex Kassette)	Du Pont (USA)

Geräte:

Bakterienschüttler	Lab-Therm, Labor-Shaker A. Kühner AG (Basel)
Brutschränke	SpecialHyb Bachofer (Reutlingen) CO ₂ -Auto-Zero Heraeus (Hanau) IR 1500 Automatic CO ₂ Inkubator Flow Laboratories (USA)
DNA-Sequenziergeräte	ABI PRISM TM 377, 373A und 377XL Applied Biosystems GmbH, bzw. PE Biosystems (Weiterstadt) Software: Analysis 2.1.2 und Genescan 672 Power Macintosh 7200/90
Elektroporationsanlage	BioRad Laboratories GmbH (München)
Gel-Imaging-System (Software: Bio-1DV.96)	Fröbel Labortechnik Vilbert Lourmet (Frankreich)
Hybridisierungsöfen	Techne Hybridiser HB-2 Bachofer Laborgeräte (Reutlingen)
Mikroskope	Axiophot Zeiss (Frankfurt) DNRBE Leica
PCR-Geräte	PTC-100 TM , PTC-200 TM

	MJ-Research Inc. (USA) DNA Thermal Cycler Perkin Elmer Cetus (USA)
Photometer	GeneQuant RNA7DNA Calculator Pharmacia (Freiburg)
Pulsfeldgelelektrophoreseanlage	CHEF Mapper Einheit BioRad (München)
Szintillationszähler	LS 1801-Scintillator Beckmann (USA)
UV-Crosslinker	UV Stratalink™ 1800 Stratagene (USA)
Vakuumtrockner	Univapo 100H, Refrigerated Aspirator Uni-Equip (Martinsried)
Wasserbäder	M20S Lobotec (Wiesbaden) HB Uni-Equip Uni-Equip (Martinsried) LKB2219 Multitemp II Thermostatic Circ. LKB-Produkte AB (Schweden)
Zentrifugen	Biofuge A Heraeus Christ (Osterode am Harz) Omnifuge 2.0 RS Heraeus Sepatech (Hanau) Sorvall RC5C Du Pont (Bad Homburg) Sorvall RT 600D Du Pont (USA) Hettich Mikro Rapid/K Hettich (Tuttlingen)

Primersequenzen:***Vektorspezifische Primer und Primer für die Exonamplifikation:***

T ₃ A:	5'-ATTAACCCTCACTAAAGGG-3'
T ₇ A:	5'-AATACGACTCACTATAGGG-3'
ab1:	5'-GAATTGTGAGCGGATAAC-3'
ab2:	5'-GTTTTCCCAGTCACGACG-3'
SP6:	5'-GATTTAGGTGACACTATAG-3'
M13uni:	5'-GTAAAACGACGGCCAGT-3'
M13rev:	5'-GGAAACAGCTATGACCATG-3'
PAC R:	5'-CGACGATAGTCATGCCCC-3'
PAC S:	5'-TCTGCCGTTTCGATCCTC-3'
Ex S:	5'-CCTCGAGGTCGACCCAGCA-3'
Ex B:	5'-ACTAGTGGACTCCAGGTC-3'
SA2:	5'-ATCTCAGTGGTATTTGTGAGC-3'
SD6:	5'-TCTGAGTCACCTGGACAACC-3'
SA4:	5'-CACCTGAGGAGTGAATTGGTCG-3'
SD2:	5'-GTGAACTGCACTGTGACAAGC-3'

Exonprimer:

1-5':	5'-GGTAGTGTGGACGTTGCCT-3'
1-3':	5'-GAGCACCAGGGCAGTGAG-3'
2-5':	5'-GAGGAACCCGTGAGGACCAG-3'
2-3':	5'-TGTTTTAGCATCTTCCTCAGGC-3'
3-5':	5'-GGAGGCCTCAGTGATTCG-3'
3-3':	5'-GGCTTCAGGGATGTCGTAGA-3'
4-5':	5'-GCATGTATGTGTGTGAATGTGC-3'
4-3':	5'-GCAGGGGACTGAGGACTTAA-3'
5-5':	5'-GCGGGAAAAGATACCTTCAT-3'
5-3':	5'-TCAGTCACTGGCGGATGTTA-3'
6-5':	5'-AGAGAGTGTGAGAAGGACTGGG-3'
6-3':	5'-AGACTCCTTGACAGGAGAGGTG-3'
11-5':	5'-ACTTACTCAGACGTTTCCACCA-3'
11-3':	5'-ACTTTTCATCTGAAGACCACCA-3'
12-5':	5'-CACACCAGAAGCCTCTCACA-3'
12-3':	5'-GGCATTTTCTGCATCATGAA-3'
14-5':	5'-GATGAAGATGACTGGGCATC-3'
15-5':	5'-CATGGCGTCTTCAACAAAAA-3'
15-3':	5'-TTCTGCGCCAGTCCTAGGT-3'
16-5':	5'-ACGTCATCAGTGTGCTGAG-3'
16-3':	5'-GTTGACCTTGCCATCCTTGT-3'
17-5':	5'-GCTCCTGTAACCGCCAGG-3'
17-3':	5'-CAGGGCGTCACTGAAGCT-3'
D1A6-5':	5'-GGTGGGCGTGGCCTACGTGT-3'
D1A6-3':	5'-CTGTAGCCGGTCCAGCAGGG-3'
D1D4-5':	5'-CACGCGAAGCTGCCCCCAGA-3'
D1D4-3':	5'-GACCTCCGTTGGTCAGGTCG-3'
z66-5':	5'-GAATTCTGGAGACTTCGAGC-3'

D-Marker, PAC-Klon- und Gen-spezifische Primer:

D4S43-2A:	5'-GACTGGTTGTTTGAGGGCGTTG-3'	Gusella et al. (1992)
D4S43-2B:	5'-TCCTTGACTCTGC TTCAGC-3'	
D4S114-A102.2F:	5'-CTGTCATAAGAGGGCCAGTACTC-3'	Tagle et al. (1993)
D4S114-A102.2R:	5'-TATGGCTCACCTCTCATCTCTGTG-3'	
D4S115A:	5'-ATCAAGTCGAGGGACCTGGGCT-3'	Gusella et al. (1992)
D4S115B:	5'-CAGACAGCAGAGTCCACGGACAG-3'	
D4S115.PCR.2.2:	5'-GGTGATGACTTCTCATGAGC-3'	Allitto et al. (1992)
D4S115.PCR.2.1:	5'-GGTGATGTAGAGGAGGCACCT-3'	
D4S166A:	5'-GGCCTGATCTGAGGTTGTGGAAACG-3'	Gusella et al. (1992)
D4S166B:	5'-CTGTACCCAGGAGCATTCTCTCCC-3'	
D4S168A:	5'-TCCCAGGACTCTTCTGGGCACAGGG-3'	Gusella et al. (1992)
D4S168B:	5'-CACCAGGGCTCTCAGGACCT-3'	
D4S227A:	5'-GATCAGGTATCTACTTTCTTATGTAC-3'	Weber et al. (1993)
D4S227B:	5'-GTCCACCCACGGAACTGATGTTTC-3'	
ZNF141-A:	5'-CATTTACCCAAGACCATTGGCC-3'	
ZNF141-B:	5'-GAGTTAGGTGTGAAAAC TTC-3'	
HDA-A:	5'-CATGTGCC TTGCGGCTTCAG-3'	
HDA-B:	5'-GCCACGTGCACCTCGGTAG-3'	
112M14A:	5'-GCTAATGTCCACTTTATCGC-3'	
112M12B:	5'-ACTTTGACATACAGCTCCTC-3'	
121C2A:	5'-CTGACTTGGACCAACTGTCCC-3'	
121C2B:	5'-GTAGAGGCGTTTCCAGTGTCC-3'	
178E7A:	5'-GCATGGACCTGCCCTCTCTG-3'	
178E7B:	5'-GGATGCTGGTCTCAGGTCTAG-3'	

FGFR3:

8F:	5'-AGGGCGGTGCTGGCGCTCGC-3'	Wüchner et al. (1997)
9R:	5'-AGACAGTGCGGAGCAGCAGC-3'	
13F:	5'-CTGGCAGCCCGTCTGAGGAGC-3'	
14R:	5'-CTGCTCCCAGCATCTCAGGGCA-3'	
16F:	5'-CAGCGCAGCCCTGGCCTATTC-3'	
17R:	5'-CCTGAAGGGCTGCCAGTCCCT-3'	

LETM1:

16-1F:	5'-CATGGCGTCCATCTTACT-3'
16-1R:	5'-GCTCTTCACCTCTGCGACC-3'
16-6:	5'-GAGGTGAAGAGCTAGAACCAC-3'
16-10:	5'-GAAGGCAATCGGGCCTCAG-3'
16-25:	5'-CTTGAGGAGAGCGCCTCCC-3'
16-33:	5'-GACGTCTGTGCCGGGACAAG-3'
16-34:	5'-GAACCGGTAGTCCGCGCTG-3'
16-34R:	5'-CCACCGCGGACTACCGG TTC-3'
16-SPL1:	5'-CGTTGAGGATGCGCCAGAG-3'
AMA11-T7-1:	5'-CTGCTTGCTGAAGGAACC-3'
LETM1-1F:	5'-GACGAAAAGACACTCCCCGC-3'
LETM1-1R:	5'-GTGACAACCAGCCCCAACC-3'
LETM1-2F:	5'-CCCCGTTCCTGGTACAG-3'
LETM1-2R:	5'-GTCACTCGTGACTCTCCACC-3'
LETM1-3F:	5'-GTGACATGGCATGGCACGGG-3'
LETM1-3R:	5'-CAATTAGCCAGGTGTGGTGG-3'
LETM1-4F:	5'-CTGAGGTTGTGCCTATAAACC-3'
LETM1-4R:	5'-GTCACAAACAAGGAAGCCATC-3'

LETM1-5F:	5'-CACTTAGATGAGAGTCAGTG-3'
LETM1-5R:	5'-CTCCACTCTGCTGAGCACAG-3'
LETM1-6F:	5'-CTGAGGATGAGAGGCAGCAC-3'
LETM1-6R:	5'-CGATCGTTCAGCATCTTCCAG-3'
LETM1-7F:	5'-CTGTTAGCCACTTATTCTGGAG-3'
LETM1-7R:	5'-GTGGCCTTTTCGAGGCTGATG-3'
LETM1-8F:	5'-CTGAGAGCTATTCTGGTGATG-3'
LETM1-8R:	5'-GAGGAATGAGTGCGCTTGCAC-3'
LETM1-9F:	5'-CACATGTCAGGAGTCATCGG-3'
LETM1-9R:	5'-CAGAGCTGCAGGGCAAGCATG-3'
LETM1-10F:	5'-CAGTGTGGCCTCTGCCATATG-3'
LETM1-10R:	5'-CCAGTGGCCTCATCCTCAAC-3'
LETM1-11F:	5'-CCAATGACTGTCCAGTGAGAA-3'
LETM1-11R:	5'-CTCTTTGGCCAGCTTATCAAG-3'
LETM1-12F:	5'-CGAACTCCTAACCTCAGGTG-3'
LETM1-12R:	5'-GGACTGCTGGTCAAAGGTGG-3'
LETM1-13F:	5'-GAGGTTGCCTGGCTCAAGC-3'
LETM1-13R:	5'-CCAGAGCAGGTGGCGGATG-3'
LETM1-14F:	5'-CAGTGATTGCGTGCATGCTG-3'
LETM1-14R:	5'-CACAGCAGCCAGCTCTGTG-3'

Letm1:

m16-3:	5'-CGATGTTGATATTGCCGTCC-3'
m16-5:	5'-GCATCACTGAGCTCATCAG-3'
m16A1:	5'-CCATAGAGGAAATGGCCCTG-3'
m16A2:	5'-CTGCGGCCTCGTGAGGTTATC-3'
m16-21:	5'-GATAGACTGGGTCTCAAATGTAG-3'
m16-6:	5'-GGTAGCTGAAATTGTGGCCAC-3'
m16-6R:	5'-CTTCACTTCTGCGGCCTC-3'
m16-18:	5'-GCAGCTCATGAGCAAGATGG-3'
m16-19:	5'-GCAGCTCATGAGCAAGATGG-3'
m16-21:	5'-GATAGACTGGGTCTCAAATGTAG-3'
m16-25:	5'-CTTGAGCTGAGCGCTTGCTG-3'
m16-29:	5'-GATTTGTGCGGACCTCTTCC-3'
m16-39:	5'-CTGTATCCTGTATGGATTCCG-3'
m16-38:	5'-GCCTATGAGGACACATAGCAG-3'
m16-40R:	5'-GGTGCCAGTGAACGTGTTGGC-3'
m16-H7:	5'-GCAGGCAGCATGTCGAGCAC-3'
m16-5A:	5'-CCTCACCCGTCCTTGAAGTTC-3'

murines Gen 51 (Whsc1):

m51-3:	5'-GAAAACCAGAAGGCGCAGAG-3'
m51-4:	5'-CTTCAGTAATTCGGAAAGGG-3'
AA139269-5':	5'-CATGTCCTCAGTAACTCATAAC-3'
AA139269-3':	5'-GAAGCAGCATAGAAATGCGTTCAC-3'
mWhsc1-9A:	5'-GTCTGATGCTTGCAAACCAC-3'
mWhsc1-9B:	5'-CATGCTCAGTTAAGGATGCAG-3'
mWhsc1-20A:	5'-CAGAGTGTCTGAACAGGATGC-3'
mWhsc1-20B:	5'-CTTGATGATCTTGGTCTCAGG-3'
mWhsc1-25A:	5'-GGGAATGTCCTTGGCATCAC-3'
mWhsc1-25B:	5'-GCTCACAGCAGTAGGACTGC-3'

Gen 43 (WHSC2):

ProTra2:	5'-CGGCGCGCGCTTTCCCCCAAG-3'
96a2-4:	5'-GGGATTCTGCTCCTCCAGC-3'
43-1F:	5'-GTTGGGGTCCCTAGCGAGG-3'
43-1R:	5'-GAGGAGGGTCCGGAGTTG-3'
43-2F:	5'-GTGCCCATCCGGAAGCTC-3'
43-2R:	5'-GATCCATTTCCCTCCACC-3'
43-3F:	5'-CCTCTCAGCCGGGAGGTC-3'
43-3R:	5'-CTTGAAGGGGCAGTCTTG-3'
43-4F:	5'-CAAGGGCCAAGGGCTCTC-3'
43-4R:	5'-GAAGCCCTTAGCTCTGCC-3'
43-5F:	5'-GACGTTTGCCGGTCACTG-3'
43-5R:	5'-CCTTAGTGACGGCACCAG-3'
43-6F:	5'-CTGGTGCCGTCACCTAAGG-3'
43-6R:	5'-CAACCCACCCCAACTGGG-3'
43-7F:	5'-CCTGAGCTGACACTGCAG-3'
43-7R:	5'-GCAGTGGTGCCAGACATG-3'
43-8F:	5'-GAAGCATGGCCTTCTTG-3'
43-8R:	5'-GTCCTGGCAGGCAGCTTG-3'
43-9F:	5'-GCACTAGGGTTTGGGAAC-3'
43-9R:	5'-CCTACCTGGTGGCCTGTG-3'
43-10F:	5'-GGATGCAGGAGGTGGTCC-3'
43-10R:	5'-GGAGCTGCTGGGTGAGTG-3'
43-11F:	5'-CACTCACCCAGCAGCTCC-3'
43-11R:	5'-GGAGCAGGGAGGACAAGC-3'

murines Gen 43 (Whsc2h):

mProTra2	5'-CGCGTAAGATGCCGGCCG-3'
m43-3:	5'-GTGGCTGCACAACAAGCTGG-3'
m43-4:	5'-GCAAGGGTGGTCAGGGCATT-3'
m43-12A:	5'-CGCGGGAGATGCAGCGTC-3'
m43-8C:	5'-GTTGGCTCACTGTTTCAGAG-3'
m43-10:	5'-CTGACTATGCTGCTGGCC-3'
m43-7A:	5'-GTGCCCTGTTTCGACTTTAACC-3'
m43-11:	5'-GGGCAGCAAACATCTGTTTC-3'
m43-E2:	5'-CACATTGGTCATGGGCTTG-3'
m43-13:	5'-GAACCCAGACACAGGTTTG-3'

Gen 45:

45-1:	5'-CAGGCTGCAGGCTGTCTTC-3'
45-2:	5'-CTCCCTCTGCTGGCCATGG-3'
45-3:	5'-GTCCCGGGTCCCTGAAC-3'
45-4:	5'-GAGCAGCGTGGTCCCTCAGTGC-3'
45-6:	5'-GCATCGAGCAGCTCCCTCAGG-3'
45A-F:	5'-GTCACTGAGGTTCGGACGTG-3'
45A-R:	5'-CAGGAGCCGGTGGAGATAG-3'
45B-F:	5'-GATGAGGGACACTCCACTC-3'
45B-R:	5'-CGCTCTTTGGCACTGGGTG-3'
45C-F:	5'-CGGGAACAAAGGCGAGGCG-3'
45C-R:	5'-CTGGGGCTTGGACAAGGGTC-3'
45D-F:	5'-GCCCAGCGAGCCTCTAGTC-3'
45D-R:	5'-CTTGCGCTCTGGAGGTCTC-3'

murines Gen 45:

m45-1: 5'-GCTCAGGTCCTTGCTGTCG-3'
 m45-2: 5'-GCGGTTTTCCGTAGGTCCTG-3'
 m45-3: 5'-CACTGGCCAAAGATCTTCAG-3'
 m45-4: 5'-CTTGAACCGCGGTAAGGGTG-3'

Pol4p:

mpol-1: 5'-GGGAAAGAGCGGCTTGCT-3'
 mpol-2: 5'-GTGACACTGGCCAATAAC-3'
 mpol-3: 5'-CACAGATTGAATTGCGCATTC-3'
 mpol-4: 5'-CCACTGTGAAGTCAGGGTAG-3'

Weitere, EST-cDNA-Klon spezifische Primer:

m57-1: 5'-GGGGATATAACTACGATGAAG-3'
 m57-2: 5'-CAAGGCTGCACACCATCTTG-3'
 h131-C: 5'-GAGCTCCAGCCACAGATTTTC-3'
 m139-A: 5'-GTATTCACCTTGTGAAGGGATTTG-3'
 m139-B: 5'-GACAGCTTACTTGGAGATGCC-3'
 m21-1F: 5'-CCTCAGCTGAGGGTAGGGTAG-3'
 m21-1R: 5'-GGGAACTCGGCTCGTCACCTG-3'

Primer für die Mikrosatellitenanalysen:

58b6-B_R: 5'-CACTAAGTACTCCATACCTCTG-3'
 58b6-A_R: 5'-GCACTCGGAGAAAAGGGAG-3'
 27h9-A_R: 5'-GAGACCAGCTTAGGTAACG-3'
 27h9-B_R: 5'-GGCTGTGGTGAGTTGTTTG-3'
 19h1-A_R: 5'-CTGTCCAGTGTGCATGTCATC-3'
 19h1-B_R: 5'-CTGACTCTCAGTGGAAATCCTG-3'
 75b9a-A_R: 5'-GGCTTGAAAGTCTGTGAAGC-3'
 75b9a-B_R: 5'-CTTTCTGTGCCACGATGG-3'
 75b9b-A_R: 5'-GGTACTCTAGAGGCTAAGGTG-3'
 75b9b-B_R: 5'-CCCAAAGTGCTGGGATTATAG-3'

Oligonukleotide für die in RNA-situ-Hybridisierung:

m1601s: 5'-GAGCTGGCCAAGTTTCTCCAGGACCCATATGAGGAAATG-3'
 m1601as: 5'-CATTTCCCTCATATGGTGTCTGGAGAAACTTGGCCAGCTC-3'
 m1602s: 5'-GTCCACCTTGACAGCCCTTCCAGAAATTGTGGCAAAGGAG-3'
 m1602as: 5'-CTCCTTTGCCACAATTTCTGGAAGGGTCTGCAAGGTGGAC-3'
 m4501s: 5'-CCATGGCCCCCGCGCTCAGGTCCTTGCTGTCCGCCACGGAC-3'
 m4501as: 5'-GTCCGTGGCGACAGCAAGGACCTGAGCGCGGGGCCATGG-3'
 m4502s: 5'-GCCTGCTGGCCAGCTGGACGCTGAAGATCTTTGGCCAGTG-3'
 m4502as: 5'-CACTGGCCAAAGATCTTCAGCGTCCAGCTGGCCAGCAGGC-3'
 m4503s: 5'-CTACAGCCTGGACGCACGGACGGAGACCCCTCTGCTGCAGG-3'
 m4503as: 5'-CCTGCAGCAGGAGGGTCTCCGTCCGTGCGTCCAGGCTGTAG-3'

3 Ergebnisse

3.1 Exonamplifikation in der distalen Chromosomenregion 4p16.3

Zu Beginn der vorliegenden Arbeit konnte die für das Wolf-Hirschhorn-Syndrom kritische Region (WHSCR) auf einen, zwischen den Markern D4S43 und D4S142 liegenden, ca. 2 Mb umfassenden Bereich eingegrenzt werden (Gandelmann et al., 1992; Johnson et al., 1994). Der proximale Teil dieser Region (D4S43 bis ca. D4S114) ist Bestandteil eines Cosmid-Contigs, der im Zuge der Isolation des Gens für die Huntington'sche Erkrankung durch Baxendale et al. (1993) erstellt und 1996 fast vollständig durch das Sanger Center sequenziert wurde. Innerhalb des sich anschließenden distalen Bereichs (D4S114 bis D4S142) konnten u. a. durch Bucan et al. (1990), Whaley et al. (1991) und Weber et al. (1991a) mehrere kleinere Cosmid-Contigs erstellt werden, von denen einige im Rahmen dieser Arbeit für die weitere Klonierung dieser Region verwendet und erweitert wurden (**3.1.1**). Während der Erstellung dieser Arbeit konnte die WHSCR auf einen 165 Kb großen Bereich innerhalb des Cosmid-Contigs von Baxendale et al. (1993) eingegrenzt werden (Wright et al., 1997), so daß letztlich im Verlauf der vorliegenden Arbeit die Erstellung und Analyse eines durchgehenden Contigs abgebrochen wurde.

Für die systematische Suche nach Genen, die eine Rolle bei der Ätiologie und Pathogenese des Wolf-Hirschhorn-Syndroms spielen, standen zu Beginn dieser Arbeit für die gesamte Region zwischen D4S43 und D4S142 mit Ausnahme kleinerer Bereiche (Weber et al., 1991b; McCombie et al., 1992) zunächst noch keine genomischen Sequenzdaten zur Verfügung. Um eine rasche und effiziente Isolierung transkribierter Bereiche zu ermöglichen, wurde die Methode der Exonamplifikation (Buckler et al., 1991) angewendet. Diese Methode ermöglicht mit Hilfe eines speziell entwickelten Spleißvektors (Buckler et al., 1991; Church et al., 1994) die direkte Isolierung kurzer DNA-Fragmente (Exons) aus klonierter genomischer DNA. Diese Fragmente sind Ausgangspunkt für die gezielte Isolierung korrespondierender cDNA-Klone und können zudem als physikalische Marker zur weiteren Kartierung dieser Region verwendet werden.

3.1.1 Auswahl von Cosmid-Klonen und Isolierung von PAC-Klonen zwischen den Markern D4S43 und D4S142

Als Ausgangspunkt für die Erstellung größerer, überlappender Klon-Contigs zwischen D4S43 und D4S142 dienten 16 Cosmidklone, die von Bucan et al. (1990) und Whaley et al. (1991) isoliert und kartiert wurden sowie zwei von J. Wasmuth (Irvine, Californien) zur Verfügung gestellte Cosmidklone. Diese Klone bilden mehrere kleinere Contigs in der Chromosomenregion 4p16.3 (vgl. **2.1** und **Abb. 3.3**). Ein Teil dieser Cosmide liegt in der Region zwischen D4S90 und D4S142 und deckt damit die telomernahe Region der WHSCR nach Gandelmann et al. (1992) ab. Ein weiterer Teil dieser Cosmide stammt aus einem Bereich zwischen D4S111 und D4S228, der sich durch das hohe Vorkommen unmethylierter CpG-Inseln und damit einer potentiell erhöhten Gendichte auszeichnet (Bucan et al., 1990). Die Cosmide pC385.1 und pC385.12 enthalten das *FGFR3*-Gen und stellen damit die nächsten Ausgangspunkte für die Isolation weiterer Anschlußklone des Cosmid-Contigs von Baxendale et al. (1993) dar.

Im Rahmen dieser Arbeit wurden zur Erweiterung kleinerer Contigs zusätzliche PAC-Klone isoliert und angeordnet. Die Isolierung dieser Klone erfolgte entweder über PCR-Screening der von P. de Jong entwickelten, hierarchisch angeordneten, genomischen PAC-Bibliothek (siehe **2.11**) oder über Hybridisierung kommerziell erhältlicher Koloniefilter der humanen genomischen Bibliothek 704 (RZPD, Humanes Ressourcenzentrum, Berlin). Als Sonden wurden bereits bekannte und kartierte D-Marker und Gene verwendet (318L12: HDA-A/HDA-B; 184O23: D4S114-A102.2F/D4S114A-102.2R; 121C2: D4S227A/D4S227B; 288H13: D4S115A/D4S115B; 111E4: D4S166A/D4S166B; 300F20: D4S43A/D4S43B; 112M14: ZNF141A/ZNF141B; **Abb. 3.3**). Des Weiteren wurden Primerpaare aus den identifizierten Exonklonen der zeitgleich durchgeführten Exonamplifikationsexperimente der Cosmidklone (siehe **3.1.2**) abgeleitet und für die weitere Isolierung überlappender Klone verwendet (167F14: 16-5'/16-3'; 255K2: 4-5'/4-3'; 129P5: 5-5'/5-3'; 36P21: 12-5'/12-3'). Durch die gelelektrophoretische Auftrennung dieser vollständig restringierten Klone wurde eine genauere Anordnung einzelner PAC-Klone bezüglich der bereits vorliegenden Cosmide ermöglicht (**Abb. 3.1**).

Zusätzlich wurden unter Verwendung der vektorspezifischen Primer PAC R und PAC S die Randsequenzen der isolierten PAC-Klone über Direktsequenzierung bestimmt. Die erhaltenen Sequenzdaten ermöglichten nach Homologievergleichen mit den Sequenzdatenbanken die Generierung weiterer Einzelkopie-Sonden bzw. STS's für die Isolierung weiterer Klone (24H21: 112M14A/112M14B; 178E7: 121C2A/121C2B; RPCIP704E01186Q14 (186E1): 178E7A/178E7B; vgl. **Abb. 3.3**).

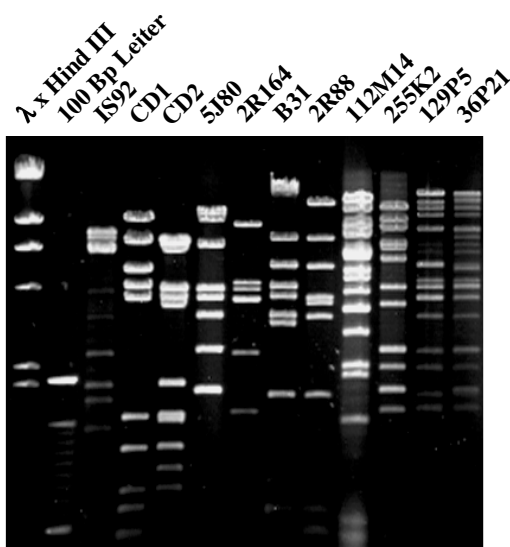


Abb. 3.1: Gelelektrophoretische Auftrennung von Cosmid- und PAC-Klonen der distalen Region 4p16.3. Als Beispiel für die Anordnung der Klone ist die Auftrennung Eco RI-restringierter Cosmid- und PAC-Klone im Bereich der Marker D4S90 bis D4S142 dargestellt.

Um eine Isolierung möglicher chimärer und/oder falsch positiver Klone ausschließen zu können, wurden PAC-Klone, die nicht eindeutig mittels obiger Methoden zugeordnet werden konnten, weiterhin durch FISH auf ihre genomische Lokalisation hin überprüft (**Abb. 3.2**). Nur in einem Fall konnte ein PAC-Klon isoliert werden (167F14), der keine Hybridisierungssignale mit der Chromosomenregion 4p16.3 aufwies und dessen weitere Charakterisierung im Zusammenhang mit der molekularen Klonierung des humanen *LETMI*-Gens in Kapitel 3.3 dargestellt ist.

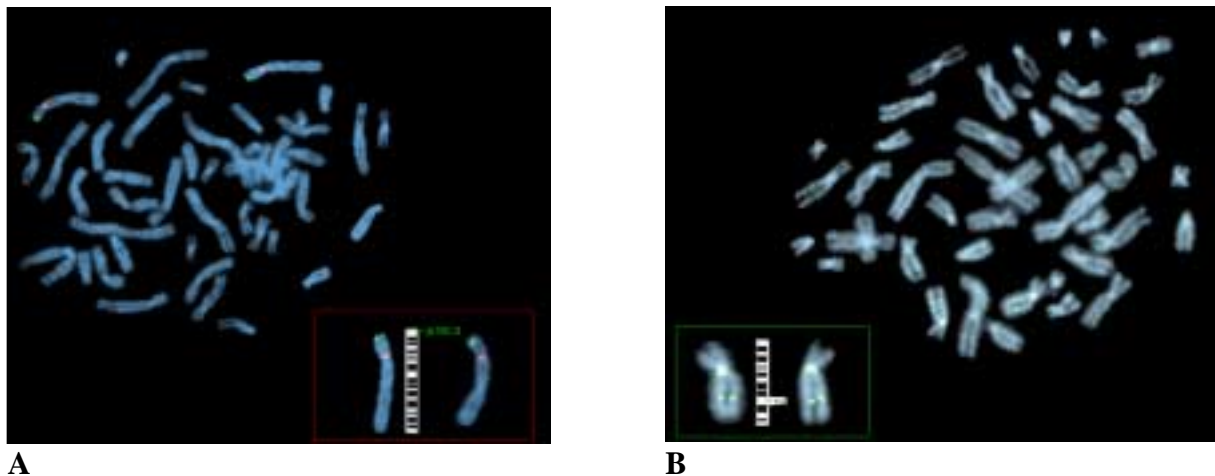


Abb. 3.2: FISH-Analysen isolierter PAC-Klone. **A:** Fluoreszenz-*in situ*-Hybridisierung des PAC-Klons 184O23. Die Hybridisierung der Sonde an die distale Chromosomenregion 4p16.3 ist durch grüne Fluoreszenzsignale erkennbar. Der Chromosom-4-spezifische α -Satellit innerhalb dieser Simultanhybridisierung ist rot dargestellt. **B:** Chromosomale Lokalisation des PAC-Klons 167F14. Hybridisierungssignale (grün) sind auf Chromosom 8q21-22 erkennbar. Zur eindeutigen Identifizierung des humanen Chromosoms 8 wurde zeitgleich eine Chromosom-8-spezifische α -Satellit-Sonde (in grün dargestellt) mithybridisiert.

Die Cosmid- und PAC-Klone können in drei größere Contigs (I bis III) eingeteilt werden (**Abb. 3.3**). Contig I liegt in dem telomernahen Bereich um die Marker D4S90 und D4S142. Er erstreckt sich über ca. 300-350 Kb und stellt den Anschluß an den inzwischen sequenzierten subtelomerischen Bereich (HS4PTEL, Acc.-Nr.: Z95704) der distalen Region 4p16.3 dar (Flint et al., 1997). Contig II erstreckt sich über einen Bereich von ca. 250-300 Kb und beinhaltet die D-Marker D4S111 und D4S228, die im α -L-Iduronidase-Gen (*IDUA*, Scott et al., 1990) bzw. im Phosphodiesterase β -Gen (*PDEB*; Weber et al., 1991b) liegen. Contig III grenzt an den Cosmid-Contig von Baxendale et al. (1993) an. Er umfaßt den Bereich zwischen dem D-Marker D4S114, dem *FGFR3*- und dem *HDA*-Gen (ca. 150 Kb). Einen Überblick über die Lage und Anordnung der einzelnen Cosmid- und PAC-Klone, die insgesamt einen Bereich von ca. 1 Mb der distalen Region 4p16.3 umfassen, sowie über die verwendeten Marker bzw. Gene ist in **Abb. 3.3** dargestellt.

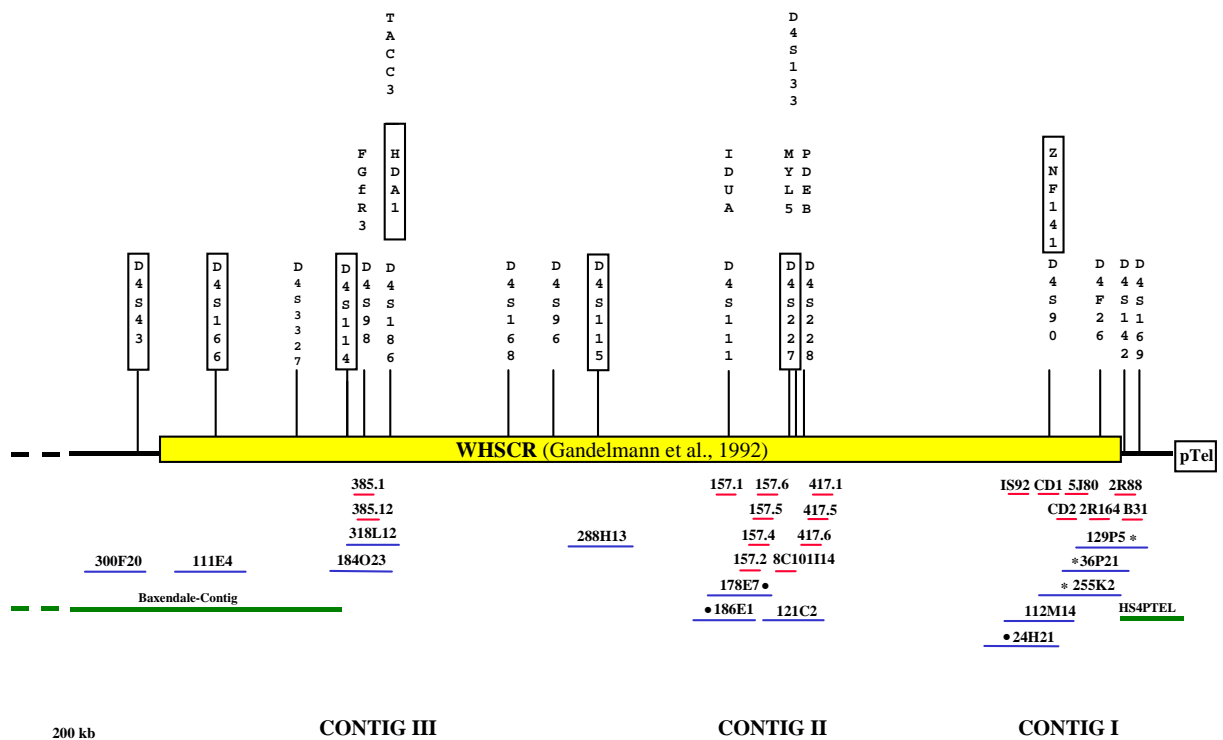


Abb. 3.3: Physikalische Karte der distalen Chromosomenregion 4p16.3 und Einteilung der isolierten PAC-Klone in die Contigs I-III. Zur Orientierung sind einige D-Marker und zu Beginn der Arbeit bekannte Gene sowie die WHSCR nach Gandelmann et al. (1992) eingezeichnet. Rote Linien kennzeichnen die zur Verfügung gestellten Cosmid-Klone (vgl. 2.1). Blaue Linien bezeichnen die in der vorliegenden Arbeit zusätzlich isolierten PAC-Klone. Sonden, die aus bekannten STS-Markern, Genen oder Exonklonen bzw. den Randsequenzdaten der PAC-Klone abgeleitet und zur Isolierung von PAC-Klonen verwendet wurden (2.11) sind eingerahmt oder durch „*“ (Exonklone) bzw. „•“ (Randsequenzen) gekennzeichnet. Durch die Randsequenzierungsdaten bzw. anhand von PCR-Analysen konnte die Orientierung mehrerer PAC-Klone relativ zu den bekannten sequenzierten Bereichen (grüne Linien) sowie zu anderen Klonen bestimmt werden. Nachfolgend sind die Ergebnisse dieser Sequenzvergleiche aufgeführt:

318L12S: HDAD/8C10E14 (McCombie et al., 1992): Bp 25065-24872; 184O23R: 184d6: Bp 23462-23702, 300F20R: 124h12: Bp 16483-16011; 300F20S: 161c2: Bp 28123-28568; 111E4R: 108f12: 13247-13484; 111E4S: 27h9: Bp 16310-15978; 36P21S: HS4PTEL: Bp 55613-55877; 129P5R: HS4PTEL: Bp 82208-82334.

Durch die Randsequenzierung des PAC-Klons 184O23 (PAC S) konnten humane und murine EST-cDNA-Klone des Menschen und der Maus mit einer 88-100 %igen Sequenzidentität in den EST-Datenbanken gefunden werden (Accession-Nr. AA098912, Hillier et al., 1996; Accession-Nr. AA617284, Marra et al., 1996; WashU-Merck EST Projekt, unveröffentlicht). Diese EST-Klone ließen sich in weiteren Homologievergleichen dem humanen und murinen *TACC3*-Gen zuordnen (Still et al., 1999).

3.1.2 Identifizierung transkribierter Bereiche durch Exonamplifikation

Mit Hilfe dieser, unter **2.12** ausführlich beschriebenen Methode der Exonamplifikation wurden zunächst die 18 vorliegenden Cosmid- sowie 7 PAC-Klone aus den Contigs I-III untersucht.

Um eine Vielzahl möglichst überlappender Fragmente zu erhalten, wurden die Cosmidklone mit Bam HI/Bgl II sowie Pst I, die PAC-Klone mit Bam HI/Bgl II sowie Eco RI in jeweils getrennten Reaktionsansätzen restringiert und nachfolgend in den Spleißvektor pSPL3 subkloniert. Aufgrund der bekannten schlechteren Effizienz der Exonamplifikation bei Verwendung großer gepoolter genomischer Bereiche (Church et al., 1994; Trofatter et al., 1995) wurde die isolierte DNA der Cosmid-Subklone in „Poolen“ von maximal 2-3 Cosmiden und die DNA der Subklone jeweils nur eines PAC-Klons für die Transfektion verwendet. Die entsprechenden Poolen sind im folgenden aufgeführt (¹: Bam HI/Bgl II- und Pst I-Verdau; ²: Bam HI/Bgl II- und Eco RI-Verdau):

¹ Pool 0	157.2, 157.4
¹ Pool 1	CD2, 157.5, 417.6
¹ Pool 2	B31, 2R164, CD1
¹ Pool 3	5J80, 157.1, 2R88
¹ Pool 4	8C101, IS92
¹ Pool 5	157.6, 417.1, 417.5
¹ Pool 6	pC385.1, pC385.12
² Pool 7	IS92, 2R164

Nach Isolierung der RNA, RT-PCR und Bst XI-Verdau wurden die erhaltenen Fragmente gerichtet in den Vektor pBluescript kloniert. Von diesen potentiellen Exonklonen wurden 4224 Subklone (je 192 bis 288 Subklone für die einzelnen vereinigten Cosmid- oder PAC-Klone) in Mikrotiterplatten überführt und analysiert. Aufgrund dieser großen Anzahl wurden eine Reihe von Untersuchungen angewendet, um falsch positive Exons zu identifizieren und die weiteren Untersuchungen auf „echte“ Exonklone zu beschränken. Mit diesen ausführlich unter **2.12.7** dargestellten Methoden konnten letztlich Mehrfachsequenzierungen weitestgehend vermieden werden (**Tab. 3.1A bis C**).

„Positive“ Exonklone wurden mit „pBluescript“ spezifischen Primern T₃A und T₇A sequenziert. Von diesen untersuchten Klonen konnten letztlich 67 unterschiedliche Sequenzen

mit einer durchschnittlichen Größe von 145,5 Bp erhalten werden. Die Anzahl der aus den einzelnen Ursprungsklonen isolierten Exonklone variierte zwischen 0 (Cosmid 417.6, CD1) und 10 (157.6) pro Cosmid (ca. 3,3 unterschiedliche Exonklone/Cosmid) und zwischen 2 (112M14) und 5 (255K2) je PAC-Klon (ca. 3,5 unterschiedliche Exons pro PAC-Klon).

Die Exonsequenzen sind auf den folgenden Seiten entsprechend den verschiedenen Ursprungsklonen (Cosmid-Poole 1-7, PAC-Klone 129P5-178E7) aufgeführt, wobei Exonsequenzen, die repetitiven Elementen (beispielsweise Alu- oder LINE-Elemente) bzw. falsch positiven Sequenzen (beispielsweise *E. coli*-, Vektor-Anteile, endogene Retroviren) entsprechen, nicht berücksichtigt sind. Bereiche von Exonsequenzen, die Sequenzhomologien zu anderen Exons aufweisen und zumeist zusammengespleißten Anteilen desselben Transkripts (vgl. 3.1.3) entsprechen, sind zur leichteren Orientierung farbig oder anderweitig markiert dargestellt.

Pool O

0A.A9 305 Bp (vgl. 1A.C10 (Bp 1-151), 0A.D1 115 Bp (Bp 191-305))

ATGTCGTGTC GGCAGGTGAG CCAGACGCGG TTCTACGTGG CAGAGAGCAG GGATGTAGCC CCGCACGTCT CCCTGTTTGT
TGGCGGCCTG CCTCCCGGCC TGTCTCCCGA GGAGTACAGC AGCCTGCTGC ATGAGGCCGG GGCTACCAAA GCCACCGTGG
TGTCGGTGAG TCACATCTAC TCCTCCCAAG GCGCGGTAGT GTTGGACGTT GCCTGCTTTG CCGAGGCCGA GCGGCTGTAC
ATGCTGCTGA AGGACATGGC TGTGCGGGGC CCGCTGCTCA CTGCCCTGGT GCTCCCCGAC CTGCT

0A.B5, 1A.E3, 1a 207 Bp (vgl. 1A.E4)

GGTGGGCGTG GCCTACGTGT CCGTGCAGAT GACCCCGAAG AGCACGGCCC GCTCTGTGGT GCTGGAGGTC CTGCCGCTGC
TCGGCCGCCA GGCCGAGAGT CCCGAGAGCT TCCAGCTGGT GGAGGTGGCG ATGGGCTGCA GGCACGTCCA GCGGACGATG
CTGATGGACG AACAGCCCCCT GCTGGACCGG CTACAGGACA TCCGGCA

0A.D1/1A.D6 115 Bp

GCGCGGTAGT GTTGGACGTT GCCTGCTTTG CCGAGGCCGA GCGGCTGTAC ATGCTGCTGA AGGACATGGC TGTGCGGGGC
CGGCTGCTCA CTGCCCTGGT GCTCCCCGAC CTGCT

0A.D4, 5B.H8 151 Bp

CACGCGAAGC TGCCCCCAGA CAGCTGTCCC CTCCTTGTGT TCGTGAACCC CAAGAGTGGA GGCCTCAAGG GCCGAGACCT
GCTCTGCAGC TTCCGGAAGC TACTGAACCC TCATCAGGTC TTCCACTGTA CCAACGGAGG TCCTCTTCCC G

0A.E3 175 Bp

GTCGGCCACG GGACCTGACG CAACAGGATG GACGAGTCCC CTGAGCCTCT GCAGGTGGT GCCACCACCC CCACTCTCCC
CGCTGCTCGC GGGAGCCAGA GGGCCCTGCG GTCCTCGGTG GTCTTACCAG CCCCTCGTCA TCCCAGGGCC CTCCGCGCCT
GTGAGGACTC CCTCA

Pool 1

1A.B6, 5A.C2 183 Bp

AAAATCCCCT GAGGAACCCT GAGGACCAGC CGAGGAGGTT CAGTCCGTGC CAGATCCAGT GGGTTTCTCT ACTCCACAAT
AACCCATCGG AGCATGAGCC ACAACCCCTCC ACCTCCATAT GAAAGAATCA CAAAAGCTGG AAGAGAAGTG TGTAAGGACC
GCCTGAGGAA GATGCTAAAA GAG

1A.B9 79 Bp

TCTGCAATTT CATGTCTCAT GAGAAGTGCC TGAAGCACGT GAGGATCCCG TGCACGAGTG TGGCACCCAG CCTGGTCCG

1A.C7 124 Bp

GTCCTCTCCA GGGGTCGCAG GGGACTGAGG ACTTAAAAGG AAGGAGTTCT GAGGGCAGAC TCTTCTCTCC CCCTGCTGCT
ACAATGTGAT TTCTGCACAT TCACACACAT ACATGCTAAG CAAT

1A.C10 266 Bp (vgl. 0A.A9, 0A.D1)

ATGTCTGTGC GGCAGGTGAG CCAGACGCGG TTCTACGTGG CAGAGAGCAG GGATGTAGCC CCGCACGTCT CCCTGTTTGT
TGGCGGCCTG CCTCCCGGCC TGTCTCCCGA GGAGTACAGC AGCCTGTGTC ATGAGGCCGG GGCTACCAAA GGC CGGTAG
TGTTGGACGT TGCCTGCTTT GCGGAGCCG AGCGGCTGTA CATGCTGCTG AAGGACATGG CTGTGCGGGG CCGGCTGCTC
ACTGCCCTGG TGCTCCCGA CCTGCT

1A.D12 165 Bp

AGAAACCCTA CAAATGTAAA GAATGTGGCA AAGCCTTTAG ATGGATCCAG AAGCCTGAAT GAACATACGA ATATTCATAT
TGGAGAGAAA CCCTACACAT GTGAAGAATG TGGCAAAGAT TTTACTTGGT CCTCAACCCT TACTGTACAC CAGAGAATTC
AGACA

1A.E4 91 Bp

GGTGGGCGTG GCCTACGTGT CCGTGCAGGT GACCCCGAAG AGCACGGCCC GCTCTGTGGT GCTGGAGGTC CTGCCGCTGC
TCGGCCGCCA G

1A.E6 46 Bp

CAGACTGGCA ATCAGAAGCA AACAAACAAC AGATGGCACC ATTTCA

1A.E12, 5A.A5 236 Bp (vgl. 5A.A9)

GCCGCATCAT GGCTGTGATG CCTTCCCTCC CTCCAGGCCT GCATGATGAC CATCACCTTC CTGCCTTACA CGGCCTGAT
TGTGGGTAC GCATTCCTACT TCCCGCACCT GCTGAGCCCG CAGATCCAGC GCTCTGCCCA CAGGGCTCTG TACCGACGAC
ACGTCTGGG CATCGTCTC CAAGGCCCGG CCTGTGCTT TGCAGCGGCC ATCTTCTCTC TCTTCTTGT CCCCTT

1B.E10 105 Bp

GAAGACTGAA CTTTCTTCAG GATCTTGATC AGTAAGACTG GTGCTGAGAA AAAAAAGTCT CCAGTTGTAT ATGCAGATTA
CAGAGCTGAT ATAATCCATG TGACA

1B.F9 250 Bp

CTGGGGCTCG GGGGCCGACC TGTGGGGCTC CGACAGCGAC ACCAGGTTTG AGAAGCCACG CATGGACGAC GGGCTGCTGG
AGGTGTGTGG CGTGACGGGC GTCGTGCACA TGGCCAGGT CCAGGGTGGG CTGCGCTCCG GAATCCGGAT TGCCAGGGT
TCCTACTTCC GAGTCACGCT CCTCAAGGCC ACCCCGGTGC AGGTGGACGG GGAGCCCTGG GTCCAGGCC CCGGGCACAT
GATCATCTCA

1B.H10 157 Bp

CTTACATGGC GTCTTCAACA AAAAGTAATT GTTTAGAGAA GTGACAAGAT GAAGGAGGAC TCTGAGTCCT AGTCTCCCTC
CTCGAGTCCC TGTGTGAGCA GAGGTAAACA ATGAAAATC ACCAACAGG AGAAATTCAC CTAGGACTGG CGCAGAA

Pool 2**2A.A5, 3A.B3, de 130 Bp (vgl. ce Bp 1- Bp 130)**

GCCACCGCGG GAAAAGATAC CTTCAATTACA GGACAACAGC ACACTGAGAG CAACACACCA CCACAGCGGT TAATTTTGCA
AAGACTGCGT TAACATCCGC CAGTGACTGA CATTTTACAG AAGACAGACT

2A.A8 192 Bp

GTGCAGGAAT AAGAGAGAGT GTCAGAAGGA CTGGGAGATG GAGCAATGGC CTCTGTGGAG GAGTAAGCCC TCCTCACCAG
AGGCCGTGAG GATCCTTTGG ACTCTTCAA GCCCATAGGC TGCAATGGCT GCATTATCCA AAGAGCAAAA ATGGTGACTC
ACCTCTCCTG TCAAGGAGTC TCACCCATCT CA

2A.A12 107 Bp (vgl. 3A.B6, cd)

GGAA GCCTCG GTGATTCTGC CACAGCCTCA GCCTCTGTGG CTCTGTGACC TGCCGGTATT GGATGATTGG TATCTAAGAC
TCTGGGACAC TCCTGAAGTC GGGAAAT

2A.B1, ke 67 Bp

GGTAAGGACA GCTGCCTTCA GCTCATTCTT CCCCAGAGT TTGCAAAAACC TCCCAGCCTT CCAAGAA

2B.G11 106 Bp

GGATAAGGAA CCTTGTCTCA AACCTCAATC CCCAGTGTG CCTCTGAAGA AGGAGATGAA GATGACTGGG CATCCAACAA
AGTTACCTGT AAAATTCGTG ATTGTG

Pool 3**3A.A3, cc 134 Bp**

ATTCTGCCAC AGGATTCTGT TCACACCAGA AGCCTCTCAC ACAGGTGACA AGACCTGGAA AACTCAAAGA CTCCACTCTG
AAAGTGGGGT TTAAGAAGAC TGTGTTGACT TTTCATGATG CAGAAAATGC CTTC

3A.A12, ca 77 Bp

ACCTGATCAC CTGTTTGGAG CAAAGAAAAG AGCCCTTTAA TGGGAAGACA CATGAGACAG TAGCCAAAGC CCCA

3A.B5 117 Bp

CCCATCCTGC CTTCCGTGGG GGTCTCTGCA GGTCCCAGA GTCCCTTATCG GGTCCACTGT GGGCACAAC CGCCCACTCC
CCAAGAATCC ACAGGTGATT AGCTGCCCTG GACTGCA

3A.B6, fa 107 Bp (vgl. 2A.A12, cd)

GGAGGCCTCA GTGATTCGGC CACAGCCTCA GCCTCCGTGC CTCTGTGACC TCGGGGTATT GGATGATTCG TAGCTAAGAC
TCTACGACAT CCCTGAAGCC GGGAAAT

3A.H9, lf 68 Bp

CATCAACTGA CTTTCCACCA TCACAATATC CATTCTGTGT AAGCATATGA GCATCTGCC AAGCAGAC

3B.E3 115 Bp (vgl. cd)

ACTTACTCAG ACGTTTCCAC CATCTCCAGC CTCTGTAACC TTTGAAGGGC CTGCAGAGAA TGCCATGTTT CAGGATGACG
ATTGGTGGTC TTCAGATGAA AAGTATTTGT GTTGG

3B.F11 102 Bp

GTCCGGACAG GCCGAGATGA CGCCGAGCCC CCTGTTGCTG CTCTGTGCTG CGCCGCTGCT GCTGGGGGCC TTCCACCGG
CCGCCGCCG CCGAGACCTG GA

3B.G5 81 Bp (gp, d)

CCCATAGGCT GCAATGGCTG CATTATCCAA AGAGCAAAA TGGTGACTC ACCTCTCTG TCAAGGAGTC TCACCCATCTC
A.

Pool 4**4A.B3 153 Bp (4A.D11)**

GCAAGACCAA CGTCAAGGAC GACGAGCTGG ACGCCATGCT CAAAGAGGCC TCGGGGCCA TCAACTTCAC CATGTTTCTG
AACCTGTTT GGGAGAAGCT GAGCGCATCA AGCGTCTGCT GATGTCCAG GCTGACAAGA TGACGGCGGA AGA

4A.B11, 5B.E12 72 Bp

GCCTCTCCAT CCTAAACAAA CAGAAACCTG TGCCTACTCC TCTGAGAGCC GTGTCCTCCA TTGTAAAATG CA

4A.B12 82 Bp

CTGCAATGTT CCTAGAGAGA AGCCAGCACT CGCCAGATCC CACGGCACTG ACCAGCTTGA GCCCATCCAG GAACTTCTCT
GA

4A.C11, ic 138 Bp (vgl. if)

ATTCTGCCAC AGGATTCCTC TGTTCACC ACCAAGCCTCC CACACAATA TATCAGGTCA CTGGACAAGA TCCACAACT
GATGTGCTC AGCTCCTGG ATTCAGCCCC CTTCCTAGGG GTATGCACCA ATCTCCACC TTGCCTGAGT TGCAGTCATC
TTTGTGGAT GTCCAGGGC TGA

4A.D11 48 Bp

CATCAAGCGT CTGCTGATGT CCCAGGCTGA CAAGATGACG GCGGAAGA

4B.E6 96 Bp

GAAATTGGGA GCACCAGCTT TGGAGAGGCA GTAGGAGTAG TCACTCCAAA CTCTGATCTC CTCTAAGCAG TTCCGAGAGA
TTTCAGTGTG GGGTCA

Pool 5**5A.A9 164 Bp**

GCACTGATTG TGGGGTACGC ATTCCACTTC CCGCACCTGC TGAGCCCAGCA GATCCAGCGC TCTGCCACA GGGCTCTGTA
CCGACGACAC GTCCTGGGCA TCGTCTCCA AGCCCCGCC CTGTGCTTTG CAGCGGCCAT CTTCTCTCTC TTCTTTGTCC
CCTT

5A.C9 129 Bp

TCGGCGCTCG ACTTCTTGGC GGGTCCAGGC TCCTGGGAG GTGCTTCCGG CCGCGACCAG AGTGACTTCG TGGGGCAGAC
GTTGGAAGTG GGCAGAGCTG GGTGCGGGT GCGGCGGGT
CTGGCCGAA

5A.C11 33 Bp

GGCAGGACTC AAATCTTCCC CTGACTTCCT GTT

5A.D3 290 Bp

AATCCCGGTC GTCGACTCTA CTAGTACAG GGGCCCCGGA CCTGGGTCCC GAGGGGCAGG TGGACGTGGC CCCACCGGAA
CTCCTCGCAC TGAGCCTTCA TCCTTGGGAC CCGTTAGGCC TAAGGCCTCG CGTCGGGTGG GACCTGGACC GGGTACACGT
GCTGCGGGCA GTGCGGGTGT TGGAGTCTGT CGGGCAGCAG GTACGCACCG AAGAGTTTGG ACCACAGCGA CACGACCGGG
ATTCTGGACC TCTAGGTGAT CAAGATCTCG CTGGCGGTG CGCCACCTCG

5B.C8 120 Bp

GCCCCATCAC ACCCTGTCTA CCACACCTGC CATACTGAGC AACGGGAAGA GGTGGTTTGC AGCCCCATGA GCAGCCCTGA
CAGTCCCGGA ATCGGGCACC CTGGTTCCCC AGGTGGCACA

5B.F3 44 Bp

ACTGGGAATC AGAAGCAAAC AAACAACAGA TGGCACCATT TCA

5B.G8 183 Bp

GCTCCTGTAA CCGCCAGGCA GTGGCCCCGC CATGTCCAG CCCCAGGACC CAGAGCAGGC ACTGGATACA CCGGGGGACT
GCCCCCAGG CAGGAGAGAC GAGGACGCTG GGGAGGGGAT CCAGTGCTCC CAACGCATGC TCAGCTTCAG TGACGCCCTG
CTGTCCATCA TCGCCACCGT CAT

5B.H1 112 Bp

CGTGCTCCGT GCATCCTACA CGCATCCTGG TGAGAGGTGT CCTCAGCCCT TTGCATGGAC GCTGGTGGGC TGTGTGGGG
AAGGTTGCC AGCTCTGCTC ATGGGGCCTG CA

Pool 6**6A.B10 85 Bp**

TCTCAGAGTC CAGGGAGCAA ATCCACCCA CCTCGCTGTA GTCTGCACA TCCTCCTTCA GCAGCTCCAG CTCCTCCTTC
TCCTT

6A.B12 138 Bp

GGAGAACGTC ATCAGTGTG CTGAGCTCAT CAACGCCATG AAGCAAGTCA AGCACATTCC CGAAAGCAAG CTCACCAGCC
TGGCCGACG ACTGGATGAA AACAAGGATG GCAAGGTCAA CATCGACGAC CTCGTCAA

6A.C7 187 Bp (6A.C11)

ACGCTCCATC CTCGGGAGAT GACGAAGACG GGGAGGACGA GGCTGAGGAC ACAGGTGTGG ACACA GGTAA GGAACTGAG
GTTCGAAGAG GGCACCCGAG CCTGAGCTGG CAGCCTAGAG GTCCAGGCTG ACCCTACCC CACCAAAGGT GGGTCCACTC
CTGGGTCATT CACTCCTGAA TCACTCA

6A.C11 65 Bp

ACGCTCCATC CTCGGGAGAT GACGAAGACG GGTAGGACGA GGCTGAGGAC ACAGGTGTGG
ACACA

6A.G1 84 Bp

GAGCGAGTGG AAGCCGGCAG AGCTGTCCAC GTGGCTGCAG TCAGAGACCT TGAAGGACAC TGCCCCGGTG CTGGAGGGCT
TGAA

Pool 7**kb 172 Bp**

AGACCTGGCC AAGAGGAGGA ATTCTCCAGA AAGGAGCCAG GGATCTCAGT ATAAGAAAA CTCTCGGTCG GGTTCAGTGA
CAATCACAAG CAATTTCACT CTGACCAGAT TAGACTCTGG CTGGCAGTGT GGAAAAAGCC CAACTTTGGG AGAATCAAAC
CAGGACCTTC CA

if 143 Bp (vgl. 4A.C11 Bp 1-74)

ATTCTGCCAC AGGATTCTTC TGTGCCACC AGAAGCCTCC CACACAATA TATCAGGTCA CTGGACAAGA TCCAGACGTG
ATAGCCCCTG TCCCTGATCA GCTAACTCTG AGGCAAGGGG AAGGACAAAA CTACTCTACT CCA

PAC-Klon 129P5**gn 227 Bp**

GTCATGGGAC GAGAGCCTGA ACCTGGCTGA ACTGAGGAGA CCACCACACC AGCACCAAAG GAGCAACTAG GAGACCGTGA
TATTCACGTG GACTCTCTCA GCTTGCAGCC TCTGCTCTCT CTAGGGGCAC TGTCCCACCC ACCCTTGAGC CTGAGAAGAA
AAGTTGGTAG AGAGGTTGTT CCTCACTTC CAACAGAAAA GAGGTGTCCA GGCCCTCAAT TCCCTGA

gi 79 Bp

CTTAATTTGT ATTTGAATGA ACCCTCTTTA CTACCGGATT GGTGGGGTGT GAGCTGAGTT ACAAGCCCCG ATGTTTAAA

PAC-Klon 255K2**cd 232 Bp (3B.E3, 3A.B6)**

ACTTACTCAG GACGTTTCCA CCATCTCCAG CCTCTGTAAC CTTTGAAGGG CCTGCAGAGA ATGCCATGTT TCAGGATGAC
 GATTGGTGGT CTTCAGATGA AAAGTATTTG TGTGGGACC CAGCAGGAGG CCTCAGTGAT TCGGCCACAG CCTCAGCCTC
 CGTCGCTCTG TGACCTGCGG GTATTGGATG ATTCGTAGCT AAGACTCTAC GACATCCCTG AAGCCGGGAA AT

ce 245 Bp (2A.A5)

AATCTGTCAC TGTAAAATGT CAGTCACTAG TGGAGCTTAC AGTGGCCTTT CCAAAAATTAA CTGCTGTGGT GGTGTGTTCG
 TCTCAGTGTG CTGTTGCACT TTAATGAAAA TATCTTTTCC TGTAGTGACC TAGTCCTCCA TGCCACACACA GCAATATGGA
 GAGAAAGGGA ATTCAACCAT TTATTCATGT GCTCTGGACT GCAAGTTCTT TGATGGGAGG GACCACAACCT GCTTCGTCTC
 TTTTT

cb 278 Bp

GAAAAATGAT GGCACGCGAG CAATTACTGT TGCTGCAATG GAAATATGGG CTACACTGAC CTGCTGCTAC CAAAACCAAG
 CAGAGTAGAC CCTGTGACTT CCTGGAACAA AGGATCCATT ATAGTACCAC TCAGCTAGAG TAAGCAGGCA CAGCACACAG
 GGTCCCTTAC CCCAACACAA ATACTTTTCA TCTGAAGACC ACCAATCGTC ATCCTGAAAC ATGGCATTCT CTGCAGGCC
 TTCAAAGGTT ACAGAGGCTG GAGATGGTGG AAACGTCT

PAC-Klon 112M14**md 141 Bp**

CTATGTGTTT TCATTTACC CAAGACCATT GGCCAGTGCA GGGCATAGAA GATTCATTCC ACAAACCTTAT ACTGAGAAGA
 TATGAGAAAT GTGGACATGA TAATTTACAA TTAAGAAAAG GCTGTAAAAG TTTGAATGAG T

da 61 Bp

GTCTCTCCA AGGGTCAAAG GAGACTGAGG ACTGCTTCTA GAGTTTTCCA CTAACAGGCC A

PAC-Klon 24H21**x 82 Bp**

GTCTCTTGAT GGGAAAGACC ACGTCTGGAC TGTAGCTGAG ATTGAATTTG AGAGAGGTTA CAGAAGTGCT TCAGAATCCT
 CA

ae 170 Bp

TCTCCAACCC AAGATCCAGA TGTCATATGC ACTGTCCCCT GATCAGCTAA CTCTGAGACA AGAGGAAAGA TAAAAGTACT
 CCAGTACCAC AGTATACAGG ATGGCAGACT AGTGATTATG GTTGTGATTA TGGCTCTGAA TATTTTGTGG ACCTGACCTC
 CCACTGCTAA

PAC-Klon 121C2**hd 242 Bp**

TGCCCTCACT TCAGCTCATT TGCTGACGAG CTCACTGACT ACAAGACAAA GAATATGCTG GCCACACCCA TCATGAATGG
 CAAAGACGTC GTGGCGGTGA TCATGGCAGT GAACAAGCTC AACGGCCCAT TCTTCACCAG CGAAGACGAA GATGTGTTCT
 TGAAGTACCT GAATTTTGCC ACGTTGTACC TGAAGATCTA TCACCTGAGC TACCTCCACA ACTGCGAGAC GCGCCGCGGC
 CA

hm 215 Bp (vgl. hc, hh)

GTGCTGCTGT GGTTCGGCCAA CAAGGTGTTT GAGGAGCTGA CGGACATCGA GAGGCAGTTC CACAAGGCCT TTTACACGGT
 GCGGGCCTAC CTCAACTGCG AGCGGTACTC CGTGGGCCTC CTGGACATGA CCAAGGAGAA GGAATTTTTT GACGTGTGGT
 CTGTGCTGAT GGGAGAGTCC CAGCCGTACT CCGGCCACG CACGCCTGAT GGCCG

hh 281 (vgl. hc, hm)

GTGCTGCTGT GGTTCGGCCAA CAAGGTGTTT GAGGAGCTGA CGGACATCGA GAGGCAGTTC CACAAGGCCT TTTACACGGT
 GCGGGCCTAC CTCAACTGCG AGCGGTACTC CGTGGGCCTC CTGGACATGA CCAAGGAGAA GGAATTTTTT GACGTGTGGT
 CTGTGCTGAT GGGAGAGTCC CAGCCGTACT CCGGCCACG CACGCCTGAT GGCCGCGGAA TTGTCTTCTA CAAAGTGATC
 GACTACATCC TCCACGGCAA GGAGGAGATC AAGGTCACTC C

hc 141 Bp (vgl. hc, hm)

GTGCTGCTGT GGTTCGGCCAA CAAGGTGTTT GAGGAGCTGA CGGACATCGA GAGGCAGTTC CACAAGGCCT TTTACACGGT
 GCGGGCCTAC CTCAACTGCG AGCGGTACTC CCGTGGGCCT CTGGACATG ACCAAGGAGA A

PAC-Klon 178E7

bf 117 Bp

GGACTTTAAG ATTGAAGATG GCAAAGCGGT GATTCCCCTG GCGGTCACGG TGCAAGGAGA CGTGCTCATC GTCATCTATC
ACGCCGGTGC CACTCTGGGC GGCCGGCTGC AGGCCAA

lc 174 Bp

AGGCTATTAT CCAATGAAGA GAAAAGAAC AGAGCCATCA TTCAAGAAGT TTGCTTCATG AAAAAGCTTT CCGGCCACCC
GAACATTGTC CAGTTTGT CTGCAGCGTC TATAGGAAAA GAGGAGTCAG ACACGGGGCA GGCTGAGTTC CTC'TTGCTCA
CAGAGCTCTG TAAA

le 275 Bp

GTGTTGACCC AAGCAGAAAT GAAAGAGCAA GGCCACACAG ACGTGTTCAG GAACAACCCA TGATCCACGC ACAGCTGGAC
ATACGAAGGA GCTGTGTGCG TCCACGCCAC GGGTCCCTGC TCAGCATGGA GAGGGGTGAG CTGCTGGTGT GACGGCGCCA
CGGGTCCCTG CTCAGCATGG AGAGGGGTGA GCTGCTGGTG TGACGGCGCC ACGGGTCCCT GCTCAGCATG GAGAGGGGTG
AGCTGCTGGT GTGACTGCAG CACAGGGCAT GCCCA

bl 248 Bp

GCCCTGGGCT GCATCTTGTA CCTGCTGTGC TTCCGGCAGC ACCCTTTTGA GGATGGAGCG AAAC'TTCGAA TAGTCAATGG
GAAGTACTCG ATCCCCCGC ACGACACGCA GTACACGGTC TTCCACAGCC TCATCCGCGC CATGCTGCAG GTGAACCCGG
AGGAGCGGCT GTCCATCGCC GAGGTGGTGC ACCAGCTGCA GGAGATCGCG GCCGCCCGCA ACGTGAACCC CAAGTCTCCC
ATCACAGA

In den anschließenden Tabellen (**Tab. 3.1A bis 3.1C**) sind die wesentlichen Informationen zu den jeweiligen Exonsequenzen, wie vorhandene offene Leserahmen und Sequenzlängen gemäß ihrer Herkunft aus den Contigs I-III angegeben. Exonklone, die keinen offenen Leserahmen aufweisen, entstammen zum Teil untranslatierten Bereichen bekannter oder unbekannter Gene, wobei Sequenzfehler durch die mehrfachen Amplifikationsschritte allerdings nicht völlig auszuschließen sind.

Exon	L. (Bp)	ORF	IS92	CD1	CD2	2R164	5J80	B31	2R88	24H21	112M14	255K2	36P21	129P5	Primer	PCR-Pr. (Bp)
md	141	+3	•	•	-	-	-	-	-	•	• U	-	-	-	-	-
if	143	-	• U	-	-	-	-	-	-	-	-	-	-	-	-	-
4A.C11, ic	138	+1, 3	+• U	-	-	-	-	-	-	-	-	-	-	-	-	-
4B.E6	96	-	+• U	n	n	n	n	n	n	n	n	n	n	n	4-5/3'	98
1A.C7	124	-	-	-	+ U	-	-	-	-	+	+	-	-	-	-	-
da	61	-	-	-	•	-	-	-	-	•	• U	-	-	-	-	-
1A.D12	165	+3	n	n	+ U	n	n	n	n	n	n	n	n	n	-	-
1B.E10	105	-1	-	-	+ U	-	-	-	-	n	n	n	n	n	-	-
gi	79	-	-	-	-	•	-	-	-	-	-	-	-	• U	-	-
2A.B1, ke	67	+2	-	-	+ U	-	-	-	-	-	-	-	-	-	-	-
kb	172	+2	-	-	• U	-	-	-	-	-	-	-	•	•	-	-
cb	278	-	-	-	-	•	•	-	-	-	-	• U	•	-	-	-
cd	232	-	-	-	-	•	•	-	-	•	•	• U	•	•	-	-
ce	245	-	-	-	-	•	•	-	-	-	-	• U	•	•	-	-
gn	227	-	-	-	-	-	-	-	-	•	•	•	•	• U	-	-
3B.E3	115	+2, 3	-	-	-	+ U	-	-	-	-	-	+	+	-	11-5/3'	104
3A.H9,if	68/70	-1, -2	-	-	-	+• U	-	-	-	-	-	•	• U	•	-	-
3A.B6, fa	107	-	-	-	-	+• U	(+)-	-	-	-	-	•	• U	• U	3-5/3'	100
3A.A3, cc	134	+1	-	-	-	+• U	-	-	-	-	-	+• U	(+)	-	12-5/3'	108
3A.A12, ca	75	+3	-	-	-	+• U	-	-	-	-	-	+• U	+•	(+)	-	-
2A.A12	107	+2	n	-	(+)	+	+ U	-	-	n	n	+	+	+	-	-
2B.G11	106	+2	-	-	-	(+)	+ U	+	n	n	n	n	n	n	14-5'	-
2A.A8	192	+1, 2	-	-	-	-	+• U	+•	+•	-	-	-	+	+	6-5/3'	168
3B.G5, d, gp	81	+1, 3	-	-	-	-	+• U	+•	+• U	-	-	-	• U	• U	-	-
2A.A5, 3A.B3, de	130	-	-	-	-	-	+• U	+• U	+• U	-	-	-	+	+	5-5/3'	104
ae	170	-2	-	-	-	-	-	-	-	• U	•	-	-	-	-	-
x	82	+2, -2	-	-	-	-	-	-	-	• U	•	-	-	-	-	-

Tab. 3.1A: Übersicht über die isolierten Exonsequenzen des Contigs I. Legende siehe Seite 65.

Exon	L. (Bp)	ORF	157.1	157.2	157.4	157.5	157.6	8C1011I 4	417.1	417.5	417.6	121C2	178E7	Primer	PCR-Pr. (Bp)
3A.B5	117	+2	+ U	n	n	n	n	n	n	n	n	n	n		
3B.F11	102	+2, +3	+ U	-	-	-	-	-	-	-	-	n	n		
0A.B5, 1A.E3, 1a	207	+1, 2	-	+ -	+ •	+ •	+ U	-	-	-	-	-	+ U		
0A.A9	305	+1	-	-	+ U	+ U	n	n	-	-	-	-	+	D1A6-5'/3'	569
0A.D1, 1A.D6	115	-	-	+ U	+ U	+ U	n	-	-	-	(+)	-	+	1-5'/3'	100
0A.D4, 5B.H8	151	+1, 3	-	+ U	+ U	+ U	+ U	-	-	-	-	-	+	D1D4-5'/3'	132
0A.E3	175	+2	-	+ U	+ U	+ U	n	n	n	n	n	n	n		
1A.B9	79	+3	n	n	n	+ U	n	n	n	n	n	n	n		
1A.E4	91	+1, 2	n	+ U	+ U	+ U	+ U	-	-	-	-	n	n		
1A.C10	266	+1	n	n	n	+ U	n	n	n	n	+	n	n		
1B.F9	250	+2	-	+ U	+ U	+ U	+	-	-	-	-	n	n		
4A.B11, 5B.E12	72	+2	n	n	n	n	n	+ U	+ U	+ U	n	n	n		
4A.B3	105	+2	-	-	-	-	-	+ U	n	n	n	n	n		
4A.B12	82	+1	n	n	n	n	n	+ U	n	n	n	n	n		
4A.D11	48	+2	n	n	n	n	n	+ U	n	n	n	n	n		
5A.D3	290	+3	-	+ U	+ U	+ U	+ U	-	-	n	n	n	n		
5B.G8	183	+3	-	-	+ U	+ U	+ U	-	-	-	-	-	+	17-5'/3'	160
5A.C11	33	+2, 3	n	n	n	n	+ U	n	n	n	n	n	n		
5A.A9	164	+1, 3	-	-	-	+ U	+ U	-	-	-	-	n	n		
1A.E12, 5A.A5	236	+1	-	-	(+)	+ U	+ U	-	-	-	-	n	n		
5B.F3	44	+1, 2, 3	-	-	-	-	+ U	-	-	-	-	n	n		
5B.H1	112	+2, 3	-	-	-	-	+ U	-	-	-	-	n	n		
5A.C9	129	+1, 2	-	-	-	-	+ U	-	-	-	-	n	n		
5B.C8	120	+1, -3	-	-	-	-	-	-	-	+ U	-	n	n		
1A.E6	46	+1, 2, 3	-	-	-	-	-	-	+ U	+	+	n	n		
1A.B6, 5A.C2	183	+1	-	-	-	+ U	+ U	-	-	-	-	n	n	2-5'/3'	172
1B.H10	157	-1	-	-	-	-	-	-	+ U	+	+	n	n	15-5'/3'	153
1c	174	+1	-	-	-	-	-	-	-	-	-	-	+ U		
bf	117	+2, -1	-	-	-	-	-	-	-	-	-	-	+ U		
bl	248	+1, -2, -3	-	-	-	-	-	-	-	-	-	-	+ U		
1e	275	-2	-	•	-	-	-	-	•	-	•	-	+ U		
hd	242	-	-	-	-	-	-	-	-	-	•	-	+ U		
hm	215	+1	-	-	-	-	-	-	-	-	-	•	+ U		
hh	281	-3	-	-	-	-	-	-	-	-	-	•	+ U		
hc	141	+1, -3	-	-	-	-	-	-	-	-	-	•	+ U		

Tab. 3.1B: Übersicht über die Exonsequenzen des Contigs II. Legende siehe Seite 65.

Exon	L. (Bp)	ORF	pC385.1	pC385.12	318L12	184023	Primer	PCR-Pr. (Bp)
6A.B12	138	+2	+	+	+	+	16-5'/3'	116
6A.B10	85	+1, 2	+	+	+	+		
6A.C7	187	-	+	+	+	+		
6A.C11	65	+1	+	+	+	+		
6A.G1	84	+1, 2	+	+	+	+		

Tab. 3.1C: Übersicht über die Exonsequenzen des Contigs III. Neben den Exonsequenzen, die entsprechend ihrer Herkunft aus den Cosmid- oder PAC-Klonen (gelb hinterlegt, eindeutige Ursprungsklone durch U gekennzeichnet) dargestellt sind und teilweise unabhängig aus den PAC- und Cosmid-Klonen isoliert werden konnten, sind mögliche offene Leserahmen (ORF), die Exonlängen (L.) sowie die wesentlichen abgeleiteten Primer und die Größen (in Bp) der an genomischer DNA erhaltenen PCR-Produkten angegeben. Die meisten Exonklone wurden mittels PCR und/oder Hybridisierungstechniken zur weiteren Anordnung der einzelnen Contigs verwendet. Hybridisierungssignale mit den unter **2.12.7** beschriebenen Verfahren sind bei den Exonklonen durch „+“, bzw. bei den im Rahmen einer Diplomarbeit erhaltenen Ergebnissen (Grimm, 1998) durch „●“ dargestellt („n“ = nicht bestimmt; „(+)= schwache oder unklare Hybridisierungssignale) und grau hinterlegt.

3.1.3 Homologievergleiche der Exonsequenzen

Die Exonsequenzen wurden auf Homologien zu bereits bekannten Nukleotid- bzw. Proteinsequenzen untersucht. Die Ergebnisse sind in den nachfolgenden Kapiteln aufgrund der Übersichtlichkeit entsprechend den einzelnen Contigs (I-III) dargestellt. Innerhalb der Contigs können die Sequenzen nach den Ergebnissen der Homologievergleiche in zwei oder drei verschiedene Kategorien (**A**, **B** und/oder **C**) eingeteilt werden. Gruppe **A** umfaßt Exonsequenzen, die aus bekannten Genen isoliert werden konnten. Die Gruppe **B** enthält Exonsequenzen, die signifikante Homologien zu bekannten Genen aufweisen (E oder P-Werte < 0,001), aber mit großer Wahrscheinlichkeit nicht identisch sind. Die Gruppe **C** schließlich beinhaltet Exonsequenzen, die keine bzw. nur schwache Homologien (E-Werte > 0,001) zu bekannten Genen oder Proteinen aufweisen.

3.1.3.1 Exonsequenzen aus dem Contig I

Exon aus dem ZNF141-Gen

A

Ein Exon (md), das aus dem PAC-Klon 112M14 isoliert werden konnte, zeigt sowohl auf DNA- als auch auf Proteinebene eine 100 %ige Identität zu dem menschlichen Zinkfingeren *ZNF141* (vgl. **Abb. 3.4**). Dieses Gen konnte 1993 von Tommerup et al. isoliert und in die distale Chromosomenregion 4p16.3 auf den Cosmid-Klon CD1 und damit in die Nähe des Markers D4S90 kartiert werden.

```

sp|Q15928|Z141_HUMAN ZINC FINGER PROTEIN 141 (L15309) Length = 474
Score = 263 (120.5 bits), Expect = 9.7e-29, P = 9.7e-29
Identities = 46/46 (100%), Positives = 46/46 (100%), Frame = +3

Query (md):          3 MCSHFTQDHWPVQGIEDSFHKLILRRYEKCGHDNLQLRKGCKSLNE 140
                    MCSHFTQDHWPVQGIEDSFHKLILRRYEKCGHDNLQLRKGCKSLNE
Sbjct (ZNF141):     77 MCSHFTQDHWPVQGIEDSFHKLILRRYEKCGHDNLQLRKGCKSLNE 122

```

Abb. 3.4: BLASTX-Homologievergleich des Exons md. Die 100 %ige Homologie zu dem *ZNF141*-Gen befindet sich in dem unter **3.1.2** ermittelten offenen Leserahmen +3 (**Tab. 3.1A**).

Exonsequenzen mit Homologien zu weiteren Zinkfingerengen**B**

Insgesamt 5 verschiedene Exonsequenzen aus den Klonen B31 (2A.A12), 5J80 (3A.A12, 3A.B6), CD2 (1A.D12) und 255K2 (cd) zeigen auf Nukleinsäureebene und teilweise auch auf Proteinebene eine signifikante Sequenzidentität zu verschiedenen Mitgliedern der Krüppel-Familie (C₂-H₂-Typ) von Zinkfingerengen. Die Exonsequenzen weisen dabei Homologien zu der KRAB-Domäne (3A.A12), den Zinkfingermotiven (1A.D12) oder den 5'-Bereichen (3A.B6 und 2A.A12) dieser Genfamilie auf (vgl. **Abb. 3.5**). Die höchsten Homologien (auf Nukleinsäureebene 85-96 % und auf Proteinebene ca. 70-83 %) betreffen dabei die Zinkfingerogene *ZNF1* und *ZNF141*, die in der distalen Region 4p16.3 auf das Cosmid CD1 (*ZNF141*) und in eine Region distal des Markers D4S142 (*ZNF1*; HS4PTEL: ca. Bp 43245-Bp 76560) kartiert wurden (Tommerup et al., 1993; Flint et al., 1997). Aufgrund der genauen Zuordnung aller 4 Exonklone zu ihren Ursprungsklonen (vgl. **Tab. 3.1A**), die zwischen diesen kartierten Zinkfingerengen liegen, können weitere Zinkfingerogene oder auch Pseudogene zwischen D4S90 (CD1) und dem *ZNF1*-Gen vermutet werden. So weist beispielsweise eines dieser Exons aus dem Cosmid B31 (2A.A12) eine 85 %ige Nukleotidsequenzidentität zu dem *ZNF1*- bzw. dem *ZNF141*-Gen und eine zu erwartende 100 %ige Identität mit der genomischen Sequenz der distalen Region HS4PTEL (Z95704: Bp 114428-Bp 114555; Flint et al., 1997), die das Cosmid B31 beinhaltet, auf. Das Exon wird darüber hinaus von putativen funktionellen Spleißstellen flankiert (**Abb. 3.5**). Daneben zeigt dieses Exon eine 97 %ige Sequenzidentität zu einem EST-cDNA-Klon des WashU-Merck EST Projektes (Acc.: R73035; Hillier et al., 1995, unveröffentlicht) und repräsentiert somit vermutlich ein neues, bisher nicht näher charakterisiertes Transkript, das einem putativen, neuen *ZNF*-Gen der distalen Chromosomenregion 4p16.3 zugeordnet werden könnte. Das Exon 3A.B6, welches aus dem weiter centromerwärts gelegeneren Cosmid 5J80 isoliert werden konnte, zeigt eine 88 %ige Identität zu derselben genomischen Region und könnte auf ein weiteres Zinkfingerengen oder Pseudogen in der Region des Cosmids 5J80 hindeuten (**Abb. 3.5**). Sequenzierfehler können dabei aufgrund der unabhängigen und zumeist mehrfachen Isolation derselben Exonsequenzen zumeist ausgeschlossen werden.

Diese Ergebnisse bestätigen die Annahme weiterer *ZNF*-Gene oder Pseudogene zwischen den Markern D4S90 und D4S142 (Tommerup et al., 1993; Flint et al., 1997), wobei die einzelnen Mitglieder dieses Clusters aufgrund der hohen Homologien untereinander möglicherweise durch Genduplikation auseinander hervorgegangen sein könnten.

```

HS4PTEL 114445 agGGAAGCCTCGGTGATTCTGCCACAGCCTCAGCCTCT-GTGGCTCTGTGACCTGCCGGTATTGGATGATTTCGTATCTAAGACTCTGGGACA
2A.A12 1 GGAAGCCTCGGTGATTCTGCCACAGCCTCAGCCTCT-GTGGCTCTGTGACCTGCCGGTATTGGATGATTTCGTATCTAAGACTCTGGGACA
3A.B6 1 GGAAGCCTCAAGTATTCGCCACAGCCTCAGCCTCC-GTCGCTCTGTGACCTGCCGGTATTGGATGATTTCGTAGCTAAGACTCTACGACA
ZNF141 52 GGAAGCCTTGGTATTCGCCACAGCCTCAGCCTCC-GTCGCTCTGTGACCTGCCGGTATTGGATGATTGGTAGCTAAGACTCCCGAATA
ZNF1 34 GGAAGCCTCGGTGATTCAAGCAGCCTCTGCCTCCCGTGTCTGTGACCTGAGGGTATTGGACAATTTGTAGCTAAGACTCCCGGATA

HS4PTEL CTCTGAAGTCGGGAAATgt 114556
2A.A12 CTCCTGAAGTCGGGAAAT 107
3A.B6 TCCCTGAAGCCGGGAAAT 107
ZNF141 CTTCAAGTCGGGAAAT 158
ZNF1 CCTT-GAAGTCGGGAAAT 150

```

Abb. 3.5: Vergleich der Exonsequenzen 3A.B6 und 2A.A12 mit den cDNA-Sequenzen (5'-UTR) des ZNF141- und des ZNF1-Gens sowie der genomischen Sequenz HS4PTEL (Z95704). Sequenzunterschiede der einzelnen Exons zu der genomischen Sequenz HS4PTEL sind rot und blau markiert. Spleißdonor- und akzeptorstellen, die vermutlich im Zuge der Exonamplifikation verwendet wurden, sind in kursiv dargestellt.

Exonsequenzen mit Homologien zum WEE1hu-Gen

Vier Exonsequenzen (3A.H9, 2B.G11, 2A.A5 und ce) zeigen entweder in dem Plus- oder dem Minusstrang auf Nukleinsäureebene eine 73-92 %ige Identität und auf Proteinebene eine 72-80 %ige Identität (3A.H9 und 2B.G11) zu dem humanen *WEE1hu*-Gen (Igarashi et al., 1991). Dieses Gen konnte 1993 von Taviaux und Demaille in die Chromosomenregion 11p15.5 kartiert werden.

Das Exon 2B.G11 verfügt über einen durchgehenden offenen Leserahmen (+2, vgl. **Tab. 3.1A**), der aber nicht dem des WEE1-Proteins entspricht (vgl. **Abb. 3.6**).

```

emb|CAA43979| (X62048) Wee1 Hu [Homo sapiens], Length = 646
Score = 52.3 bits (123), Expect = 5e-07 Identities = 24/30 (80%)
Positives = 27/30 (90%), Frame = +1

Query: 13 LSQTSIPSAASEEGDEDDWASNKVTKIRD 102
          +S+TSIP+AASEEGDEDDWASNKV KI D
Sbjct: 434 ISRTSIPNAASEEGDEDDWASNKVMFKIGD 463

```

Abb. 3.6: BLASTX-Homologievergleich des Exons 2B.G11.

Während die Exons 2B.G11, 2A.A5 und ce aus den Cosmiden B31, 2R88 und dem PAC-Klon 36P21 isoliert werden konnten, stammt das Exon 3A.H9 aus dem centromerwärts gelegeneren Cosmid 5J80 (siehe **Tab. 3.1A**).

Die Ergebnisse legen die Vermutung nahe, daß es sich bei diesen Exons um Bestandteile zweier verschiedener, *WEE1hu*-ähnlicher Gene oder Pseudogene handeln könnte. Zur Klärung dieser Fragestellung wurden weiterführende Untersuchungen im Rahmen einer medizinischen Doktorarbeit durchgeführt (Pak, 1999).

Sequenzanalysen der entsprechenden chromosomalen Regionen der PAC-Klone 129P5, 36P21 sowie dem Cosmid 2R88 ergaben über 2978 Bp eine durchgehende, über 80 %ige Sequenzidentität zu der cDNA-Sequenz des *WEE1hu*-Gens und entspricht damit einem prozessierten *WEE1hu*-Pseudogen (Pak, 1999; siehe **Abb. 3.7**). Diese Ergebnisse konnten von Flint et al. (1997) im Zuge der Sequenzierung und Charakterisierung der distalen Region 4p16.3 bestätigt werden. Das *WEE1hu*-Pseudogen liegt dabei in einer Region von Bp 98702 bis Bp 94663 der genomischen Sequenz HS4PTEL (Z95704).

Durch Direktsequenzierung an dem Cosmid 5J80 wurde gezeigt, daß sich die an das Exon 3A.H9 anschließende Sequenz von der des Pseudogens unterscheidet und keine Homologien zu bekannten Genen oder Proteinen aufweist (Pak, 1999). Weiterhin konnte aus dieser Region auch das Exon ce isoliert werden, das ebenfalls deutliche Sequenzunterschiede zu dem *WEE1hu*-Pseudogen aufweist (vgl. **Abb. 3.7** und **Tab. 3.1A**). Unklar ist, ob es sich dabei um ein nicht prozessiertes *WEE1hu*-Pseudogen oder um ein funktionelles Gen handelt.

GGTGTGAGAGAGGAGGTGAGGACAGTGGCCTGGATCACTAACTGGAAGGATGGGTTTGGGGAGGAGCAGGTCTCCGGAGGAGGATGCA 90
ATGTGGGCAGCCTGCACACATCACTTAGAGCTTCTCTCCAGCTGTGTCTAGTGGGTACTTGCCTGTGCTTTCACATGTGGCCT 180
CTGTGCTGGAAATCTTTCCTTCCCTGGAAGAAGAGCTTAACTGCTCTCAAGGGTACCTTCTGTGAATCTGTCTGTCCGGCCCTCA 270
CCCCCTAGCCCTCTCTTTCCTCATGCACTTGTGGACTCAGCACGCTACACAGCCTTGACCATATTGTGCCACAGAAGGTGCCCC 360
AGGATCAGGACAGGTCGCATCATCGCGAAGCCTGTGCTCCTCAGAGTGTCCGCCACACAGCAGGGCTCAGCGAGAGACTGACAGGG 450
GACCTGGCTGGACTCTAGCTCTCTTCTTACAGGTGTGCCCTTGGGGCGCTCACCTTGCCTCCGTGGGCTCCGTTTCTGCATCTGT 540
CAAAATCGGTGTGCGCTTTCCTTCCCTTCCCGCGAGGGTCTGTGGGATGGTCCACTCAAGGATCCTGAAAACACTGGTTAGCACGCC 630
TGTGCTGGGAGACAGAGCTCTTAAACCTCTTCTCGCTGTCCCTGTCTTGGCTGGAGGAGACGAAAGAGCCGGGCAGAGGAACA 720
GGAGGAAACTGCCGCAATCCCTATTCCAAATGCACTTTACACACATTTCCCATTAATCCTCCGAATCTCAAGTGTAGCCGGGAC 810
CGAGCTGAGGGTGGCGATGGCGCTGCAGTGAAGGAGACCCTCGGGTGAACAGGAGGGTGAACCCGGCCGCCCGCCCTGTGGAC 900
CGTCCCGCCTGCCCCGATCCCCGCGCCCGGAGCCAGCGCAGCACAGAGGACCCAGGGCCCGCACGCGCGGACCCGCCCCCTG 990
GGCTCAGGGTCCCGACCTCGCAGCCGGGCTCTCCTGTCCTGCGCGCTCCAGGGCCATGCGTTTCCGAGCCGCGCGCCGGGCC 1080
GCCTGCACCTTTCGCGCAAAAGTCGTTCTCGCCCTCGCGCACCCGCGGGGGAGGAGCAGGAGGAGGAAAGGAGGACTCGGCCCTCAGG 1170
GGACCGCTGGCCGCTGCCGCCCGCTGGAGCCCATGGAGCCCGGGCCGAGCCCGGGAGCGGGAGGAGGACAAGCTGTGGGCCCTGCC 1260
CAGCGCGGAGGGGGGGCGAGGGCCCGGGCGACTCACGGGACACGCGGGTCCGCTCCTCGTCGAGTTCGAGGCCCGCGGCTGCCT 1350
GATTCCTGGCGCTTTCCTCGCGGGGGCGCCGGGGCCCTGGGAACGCTCCTGTGCGGGGTGCGGGGGCGCGCGGGCACGGAGGGGG 1440
TCTGCTGGTCCCTGTGGGACCCCGGGTGCCTCCCGCAGTCTCCCGCAAGCGGGCTCTGACACCCGCGCAGCCGAGAGTCTGCGT 1530
CCAAAGCTCAGGGATGGATCGCAGCGCTTCACTCTGCTTAGTTCCTCTTACAGAGACACAGAAATACAGGAAAAGGGATTTGATC 1620
GGCCACAGATTCTCCTGTAATTCACCCCTTTACTCCGGATTCTGCTTGCATCTGCTAGGACAGCGCTGTTGAGTACAGAGGAG 1710
GTATTGGGATGACTCCTGCGGTGAAGACATGGAAGCCAGTGTATGATGAGTTTGAAGATGAACTACACCTGTTAGAGAATTACAGTTA 1800
CTGAAAGCGATGTGAATCTCGCCCTACAACAGAAATTCATGAGCCAGAGAAAATGGGCTTCGAGAATTTGATTCGTGTTTTAACAGC 1890
GTGGAGAGGCTGGATACATGATTTATGCCATTAAGCGATCAGAAAAGTCACTGCCCCTCCGTTGATGAGCAGAAACCTTTGAGAGAA 1980
GTGATGCTCAGCAGTGTCCGAGCGGCTTTCGCTACTTTGACACTTGTCTGCTGGGAGATGCTCATATGCTTACACAGAAATGGATA 2070
-----A-----
CCGTGATGGTGAAGTTTGTGATGCTATAAGTAAAAATCCGAATCACAGGATTACTTTAAAGAACAGAGTTGAAGGATCTCTCTTCT 2160
TT-----A-----CA-----
GCAAGTTGGCCGGGCTCGAGGTATATTCAATCAATGTCTTAAATTCACAGGATAAGAACCTTGTCTCAAACCTCAATCCCCAGTGTCT 2250
GCCTCTGAAGAAGAGATGAGATGACTGGGCATCCACAAAGTTACTCTGTAATAATCGTGATTGTGCGTATGTAACAAGGATCTCTAAC 2340
CCACAAGTGAAGAGACTGATAGTCTTTTTTTTTTTTGAACAAGTGAAGTTTACAGGAGAAATATACCATCTACCAAAAGCAGATATCT 2430
TGCACCTCCCTCACAGTGTGCTGTGCTGTGCTGTCTTCCCAGGAATGGAGACCAATGGCATGGAATCAGAGAGAATAGATTAC 2520
CTCAGATACCAAGTGTCTTCCCAAGAAATTTTCCAGAGTTGCTAAAAGTTATGATTCATTCATATCCAGAGAGAAGACCTCAGCAATGG 2610
CCCTCTAAAGCATTCACTGCTGCTGTCTGCTCTAGGAAGAGTAGAGAACCCTGTGTATAGGATTGAATTTGAAAACCTTCAAAAATAC 2700
ACTTTTACAAAAGAACTCAGGAGGCATAGATGGTAAAGGCTGCAGCTGAGGAAGGAGCACTTCTACTGATCAGAGGCCACTAGTCTC 2790
ACCAGCTATGGAACACAACTCTGGCTTATTAGAAAAGAAATGAAGTGTCTGTCAGCCTCTAGTAAAACACTCATTTCCCACTTAC 2880
TCATCCCACAAAACACTATGAAAAAAGCAAGCTAGGTGCAAACTCACTGCTAGAATCCAATTTGCAATTACTCTCTATTGGTGTGAG 2970
TGGTTTTACTGATTAGGACTTTTATTGTGAATTACAATGAAAACGTATTTTGACCATAGCTTTTTTCCAGGCTTTTATATAGTCTTACC 3060
AGTCTGTCTCTGTAATAATGTCAGTCACTGGCGGATGTAAACGAGTCTTTGCAAAAATTAACCGCTGTGGTGGTGTGCTCTCAGTGT 3150
A-----A-----A-----A-----T-----GC-----CA-----TG-----C-----
GCTGTGTGCTGTAATAAGAGTATCTTTTCCCGGCTGGCCTGAAGAAAGGATTAAGGGCTTGTATCACAACATATTTCTCTCAATAAAG 3240
-----CA-----T-----AA-----T-----T-----A-----
CAAAATGCTAAAACCGCCATTACTACTCAGCCTTCAATGCACCTGTGTATGAAATTTCTCTCTCTCTCTATATATATATGTTTT 3330
AAGTGAATTATACTGTATATCCAAGTGGGAGCACTTTGTAGCCCTTAGCATGAGACATGGGATGATAAATCTGTGGAAGTATTGCCTTAT 3420
AAATTTGCTATTTTCAAGTTTGTCTGCTGTAACCTTTTAGCATTAACAATTAAGTGTGTAATAGGCTTTCTTTTTGAAACAAATATG 3510
TAAAGTGTACAGCTGGCTTGTATGAAAAGCAGCTATTGGCCCTTTTTTTTTTTTCAAGTCTTAAAGCTTTGGAATTTGTATTAATCCAG 3600
TATAATAATTAAGTGTGTTGTAATTTTGTCTGTAATGAACTTCCCTGACTCTTAAAAAATAACCTCCTCCCAATTTGTATTTGAT 3690
CATATTGTTAGTCTGTGTTCAATGTTTGTCTGTTTCTCTCAAGTGTATGTATATAAACTTATCTTGGTATCCTCTACACAGCTG 3780
TTTACCGTCCCGCCCTCAAGTTTTTTTACAAAATAGATTTTTTCTGTGATTAGCCGTTTGTACTATAACTGGCTAATGATTTTGCA 3870
AAAAAATAA 3960
AATATTGACTGGCAATATAAAGTCAACCTGAGTTTAACTTAATTGTCTAGAAAAAATAAAGAAAGGGTTTATGGTGGACTACCA 4050
GGCAGAAGCCCTGGTGTGAGCAGCAGGTGTAGTTTGGGAAACAGCACAGCCATGCTGAGGAGGAGAGAGAGGTGATGACCAAAACACAG 4140
GAAAATCCAAAACCCAGAAAGTCAAGGAGCCAAATGAAAGCCTTTATCCTGCATCAACCATTTAACTCATAAACTGCAAGTGAAC 4230
ATGGCTTTTAGAGATTTTATAGTTCTTTATTTCTGTATCCTGTAAAGGAAATGGATTTTGTGATCAGACTTCCCGAGGAAATTT 4320
TTTACCAGCTGTCGCCAGGAATTAATAACAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 4410
GAAGCATCCTCTGATAGGCTCTAGGAAATGAGTTAGTTGCTGTCAAGTATTACACCCCTGCACAGACATTAATTTACCACAGCC 4500
TTGGTATGGAATGGGAGGATTTATTAAGCAGGACTGGCTGGTGGCC 4550

3A.H9

2B.G11

2A.A5

ce

Abb. 3.7: Nukleotidsequenz des WEE1hu-Pseudogens und der flankierenden genomischen Sequenzen. Eine Sequenzidentität von über 80 % zum WEE1hu-Gen (Acc.-Nr.: X62048) ist zwischen den Basenpaaren 1019-3997 vorhanden und beinhaltet die vollständige Sequenz der WEE1hu-cDNA. Bereiche ohne Sequenzübereinstimmungen sind in kursiv dargestellt. Die vorliegende genomische Sequenz zeigt keine signifikanten Übereinstimmungen zu der von Shiozawa et al. ermittelten Sequenz der 5'-untranslatierten Region des WEE1hu-Gens (Acc.-Nr.: AB019580, unveröffentlicht). Eine Folge von Adeninresten im Anschluß an den Bereich mit Sequenzidentität zu dem WEE1hu-Gen ist unterstrichen. Kurze, flankierende Sequenzwiederholungen, wie sie bei einer Vielzahl anderer, beschriebener prozessierter Pseudogene gefunden werden können (Buiting et al., 1996) sind nicht vorhanden. Die Exons 3A.H9, 2B.G11, 2A.A5 und ce sind eingerahmt. Die Exons 2A.A5 und 2B.G11 haben eine 100 %ige Identität zu der genomischen Sequenz HS4PTEL (Z95704) und werden von putativen, kryptischen Spleißstellen flankiert (kursiv dargestellt). Die Exons 3A.H9 und ce hingegen zeigen deutliche Sequenzunterschiede und stammen aus dem centromerwärts gelegeneren Cosmid 5J80 bzw. dem PAC-Klon 255K2.

Exons mit schwachen Homologien

C

In diese Gruppe gehören eine Reihe von Exonsequenzen, die keine Homologien zu bekannten Genen oder Proteinen aufweisen. 8 dieser Exons (1A.C7, 1B.E10, 2A.B1, da, gi, kb, 3B.G5 und 2A.A8) stellen vermutlich Bestandteile neuer Gene oder auch Pseudogene dar. Sechs dieser Exons konnten aus dem Cosmid CD2 (1A.C7, 1B.E10), dem PAC-Klon 112M14 (da) oder dem Cosmid 2R164 (2A.B1, kb und gi) isoliert werden, die anderen Exons stammen aus den Cosmiden B31 und 2R88 und weisen erwartungsgemäß eine 100 %ige Übereinstimmung zu der sequenzierten Region HS4PTEL auf (Flint et al., 1997), die diese Cosmide beinhaltet. 9 weitere Exons derselben Gruppe, die keine Homologien zu bekannten Genen aufweisen (3A.A3, 3B.E3, 4A.C11, 4B.E6, if, gn, cb, x und ae) und die aus den Cosmiden IS92, 5J80 sowie den PAC-Klonen 255K2 und 24H21 isoliert werden konnten, zeigen eine 83-91 %ige Sequenzidentität zu verschiedenen chromosomalen Regionen, darunter der genomischen Sequenz der distalen Region 4p16.3 (HS4PTEL; Flint et al., 1997). Die weiteren Übereinstimmungen betreffen zwei weitere chromosomale Regionen auf Chromosom 21 (Acc.-Nr.: AC006557, Birren et al., 1999, Whitehead Institute/MIT Center for Genome Research, Cambridge, unveröffentlicht) und Chromosom 18 (Acc.-Nr.: AC004527, Weiss et al., University of Utah, Utah, unveröffentlicht). Interessanterweise liegen diese Sequenzübereinstimmungen innerhalb eines jeweils ca. 7-9 Kb großen Bereichs (HS4PTEL: ca. Bp 42.000-49.000; Chr. 18: ca Bp 26.000-35.000; Chr. 21: 122.200-131.000), in denen bereits bekannte (*ZNFI*) oder aufgrund von BLASTN-Homologievergleichen ebenfalls putative Zinkfingerogene oder Pseudogene liegen.

3.1.3.2 Exonsequenzen aus dem Contig II***Exonsequenzen des PDEB- und MYL5-Gens***

A

Es konnten zwei Exonsequenzen mit 100 %iger Homologie zum *MYL5*-Gen (Collins et al., 1992) und 4 Exonsequenzen mit 99-100 %iger Homologie zu dem *PDEB*-Gen (Weber et al., 1991b) isoliert werden (**Tab. 3.2**). Sowohl das *MYL5*-Gen („Myosin regulatory light chain gene 5“) als auch das *PDEB*-Gen (3'5'-cGMP-Phosphodiesterase β -Untereinheit) wurden im Zuge ihrer Isolierung sehr genau innerhalb der distalen Chromosomenregion 4p16.3 zwischen

den D-Markern D4S227 und D4S228 kartiert, aus der die verwendeten Cosmide stammen (vgl. **Abb. 3.3**).

Exon-Sequenz	Länge in Bp	Homologievergleich (BLASTN) Gen	% Identität	Position im Gen/Exon	Herkunft aus dem Cosmid-/PAC-Klon
4A.D11	48	<i>MYL5</i>	100	6	8C101I14
4A.B3	153	<i>MYL5</i>	100	4,6	8C101I14
hd	242	<i>PDEB</i>	100	2, 3	121C2
hc	141	<i>PDEB</i>	98	4	121C2
hm	215	<i>PDEB</i>	99	4,5	121C2
hh	281	<i>PDEB</i>	99	4,5,6	121C2

Tab. 3.2: Exonsequenzen aus dem *MYL5*- und dem *PDEB*-Gen. Dargestellt sind neben der Exonlänge die über Homologievergleiche ermittelte Sequenzidentität (BLASTN) der isolierten Exonsequenzen, die Position im Gen bezüglich der bekannten Exonstruktur sowie die Herkunft aus dem jeweiligen Cosmid- oder PAC-Klon. Die Angaben der Exonbereiche zeigen, daß durch Exonamplifikation auch korrekt oder fehlerhaft aneinandergefügte Exons eines Gens isoliert werden können. Durch die Isolierung der *MYL5*-Exons konnte das Cosmid 8C101I14 (vgl. **2.1**) in der Nähe des Markers D4S133 plaziert werden.

Exonsequenzen des DAGK4-Gens

Insgesamt konnten 8, zum Teil überlappende Exons (**Abb. 3.8**) isoliert werden, die mit 99-100 %iger Sequenzidentität dem Diacylglycerinkinase 4-Gen (*DAGK4*, Schaap et al., unveröffentlicht, GenBank Accession-Nr.: L38707) entsprechen. Die Sequenzierung des mit den Primern D1A6-5' und D1A6-3' (vgl. **Tab. 3.1B**) an genomischer DNA erhaltenen PCR-Produkts zeigte, daß sich die Exonsequenz 0A.B5 aus 3 zusammengespleißten Exons des *DAGK4*-Gens zusammensetzt. Durch die nachgewiesene Exon-Intron-Struktur und die meist 100 %ige Übereinstimmung der isolierten Exons zu der cDNA-Sequenz des *DAGK4*-Gens konnte die Isolierung eines Pseudogens weitgehend ausgeschlossen werden. Die aus den kartierten Cosmiden 157.2-6 isolierten Sequenzen (**2.1, Tab. 3.1B**) ermöglichten daher die erstmalige genomische Lokalisation dieses Gens (Endele et al., 1996). Es kartiert zwischen den Markern D4S111 und D4S227/D4S228 und damit zwischen *IDUA* und *MYL5* bzw. *PDEB*. Die Transkriptionsorientierung innerhalb des Contigs II konnte anhand der Rückkartierungsdaten nicht bestimmt werden.

Das homologe murine *Dagk4*-Gen wurde von Pilz et al. (1995) ca. 0,00-10,97 cM von dem *Pdeb*-Gen entfernt auf dem murinen Chromosom 5 (Mmu 5) kartiert. Es handelt sich hierbei

Exonsequenzen des GAK-Gens

Aus dem Cosmid 157.6 und dem PAC-Klon 178E7 konnten insgesamt 4 Exonsequenzen mit einer 100 %igen Homologie zu dem humanen Cyclin G assoziierten Kinase-Gen (*GAK*, Kimura et al., 1997) isoliert werden (**Tab. 3.3**). Dieses Gen wurde durch Kimura et al. (1997) mittels FISH an humanen R-bandierte Metaphase-Chromosomen in den distalen Bereich von 4p16 kartiert.

Durch die vorliegenden Untersuchungen konnte somit die Lokalisation dieses Gens zwischen die Marker D4S111 und D4S228 eingegrenzt werden. Es befindet sich direkt distal des *DAGK4*-Gens und weist vermutlich eine telomerwärts gerichtete Transkriptionsorientierung auf, die sich aus der Lage der isolierten Exons innerhalb des Contigs II ergibt.

Exon-Sequenz	Position im <i>GAK</i> -Gen	Länge (Bp)	Rückkartierung/Herkunft
5A.C9	61- 144	129	157.6
lc	208- 381	174	178.E7*
bl	742- 989	248	178.E7*
bf	1857-1972	117	178.E7*

Tab. 3.3: Darstellung der Homologiesuche und Rückkartierungsdaten der *GAK*-Exons. * Die Exons lc, bl und bf konnten ausschließlich auf den PAC-Klon 178.E7 rückkartiert werden (vgl. **Tab. 3.1B**). Zu dem Exon 5A.C9 liegen keine genaueren Rückkartierungsdaten vor.

Exonsequenzen mit schwacher Homologie

C

Die restlichen 17 Exonsequenzen, die aus den Cosmiden 157.1 (3A.B5, 3B.F11), 157.2 (0A.E3), 157.6 (5A.C2, 5A.D3, 5B.G8, 5A.C11, 5A.A9, 5A.A5, 5B.F3, 5B.H1), 8C101I14 (4A.B11, 4A.B12), 417.1 (1A.E6, 1B.H10) 417.5 (5B.C8) und dem PAC-Klon 178E7 (le) isoliert werden konnten (**Tab. 3.1B**) zeigen keine signifikante Sequenzidentität zu bekannten Genen. Diese Sequenzen stellen möglicherweise Exons bisher unbekannter Gene dar, was durch signifikante Übereinstimmungen einzelner Exons zu verschiedenen EST-cDNA-Klonen deutlich wird (**Tab. 3.4**). Trotzdem kann die Isolierung falsch positiver Exons über die Nutzung kryptischer Spleißstellen nicht ausgeschlossen werden. In der **Tabelle 3.4** sind die Ergebnisse der Homologievergleiche auf DNA- bzw. Proteinebene von den Exons dargestellt,

die über einen putativen offenen Leserahmen verfügen. Weiterhin sind Exons aufgelistet, die Sequenzähnlichkeiten zu EST-cDNA-Klonen der „dbEST“-Datenbanken aufweisen.

Exon	Homologievergleich (BLASTX)	%	E-Wert	Homologievergleich (BLASTN)	%	E-Wert		
0A.E3	Kollagen, Typ X, alpha 1 (X60382)	(14/27)	51	0,19	-			
1A.B6	-			AI086274	(145/146)	99	2e-74	
3B.F11	-			AI245701	(94/95)	98	8e-45	
5A.D3	Homo sapiens sco-spondin-mucin-like (AC004877)	(17/34)	49	10,0	-			
5B.C8	-			AA226727	(119/120)	99	4e-59	
5B.G8	Treptomycetes coelicolor hypothetical protein (AL079308)	(20/34)	58	3,6	R14407	(139/141)	98	3e-70
				AA050746	(63/72)	87	4e-11	
5B.H1	Mus musculus Ese2L protein (AF132479)	(19/29)	64	8,9	-			

Tab. 3.4: Übersicht über die Exonsequenzen mit schwachen Homologien zu bekannten Genen und/oder Sequenzidentität zu EST-cDNA-Klonen. Neben den Homologievergleichen (BLASTX) sind die Homologievergleiche (BLASTN) mit den EST-Datenbanken sowie die Accession-Nummern der identifizierten cDNA-Klone des Menschen und der Maus (in kursiv) angegeben.

3.1.3.3 Exonsequenzen aus dem Contig III

Exonsequenzen des FGFR3-Gens

A

Erwartungsgemäß konnten aus den Cosmidklonen pC385.1 und 1.2 zwei Exonsequenzen (6A.C7 und 6A.C11) isoliert werden, die mit einer 100 %igen Identität dem humanen *FGFR3*-Gen entsprechen (Keegan et al., 1991). Das Exon 6A.C11 (65 Bp) konnte anhand der durch Wüchner et al. (1997) aufgeklärten genomischen Struktur dem Exon 4 des *FGFR3*-Gens zugeordnet werden. Das Exon 6A.C11 hingegen setzt sich aus zwei verschiedenen Sequenzanteilen zusammen (vgl. 3.1.2). Während der Bereich von Bp 1-66 dem korrekt gespleißten Exon 4 entspricht, zeigt die restliche Nukleotidsequenz (Bp 67-Bp 187) eine 98 %ige Sequenzidentität zu einem Bereich innerhalb des Intronbereichs zwischen Exon 2 und 3 (Acc.:Y09852, Wüchner et al., unpubliziert), was vermutlich auf eine fehlerhafte Klonierung während der Exonamplifikation zurückzuführen ist.

Exonsequenzen mit schwachen Homologien

C

3 weitere Exonsequenzen, die aus den Cosmiden pC385.1 und 1.2 isoliert werden konnten, zeigten keine signifikanten Homologien zu bereits bekannten Genen. Eines dieser Exons, 6A.B12, wies auf Aminosäureebene schwache Homologien (E-Werte von 0,006 bis 3,8) zu dem Ca^{2+} -bindenden EF-Hand-Motiv verschiedener Ca^{2+} -bindender Proteine aus diversen Spezies auf, wie beispielsweise zu dem humanen und murinen Troponin C (Parmacek et al., 1990; Gahlmann et al., 1988) oder einer calciumabhängigen Proteinkinase aus *A. thalianae* (Hong et al., 1996). In **Abbildung 3.9** ist der Homologievergleich mit dem murinen Troponin C (Parmacek et al., 1990) dargestellt.

```
Swissprot: TPCS-mouse P20801 mus musculus troponin C, sceletal muscle (stnc). 2/91
Plus strand HSPs:
Score: 58 (25.5 bits), Expect 0 0.083, P 0 0.080
Identities 0 9/29 (31%), Positives 0 19/29 (65%), Frame +2
```

```
50 KHIPESKLTSLAALDENKDGKVNIDDLV 136
    +H E ++ SL D+N DG+++ D+ +
124 EHVTEEEIESLMKDGKNNDRIDFDFEFL 152
```

Abb. 3.9: BLASTX-Homologievergleich des Exons 6A.B12 mit der Proteinsequenz des murinen Troponin C (Acc.: P20801). Die Homologie bezieht sich auf den Leserahmen +2 (vgl. **Tab. 3.1C**).

3.2 Identifizierung und Charakterisierung korrespondierender cDNA-Klone

Für eine schnelle und effiziente Isolierung korrespondierender cDNA-Klone zu den putativen Exonsequenzen wurden neben computergestützten Analysen in den EST-Daten-Banken des NCBI drei cDNA-Bibliotheken durchsucht. Neben einer kraniofazialen cDNA-Bibliothek (J. Murray, unveröffentlicht) auf λ ZAP II-Basis wurde eine fetale, hierarchisch angeordnete Leber- und Milz-cDNA-Bibliothek verwendet, die ein systematisches PCR-Screening erlaubt (Löbber, 1996) sowie eine normalisierte Gehirn-cDNA-Bibliothek (NIB: „normalized infant brain“; Soares et al., 1994), die freundlicherweise von Herrn Marcelo Bento Soares zur Verfügung gestellt wurde. Die letztgenannte cDNA-Bibliothek wurde im Rahmen der vorliegenden Arbeit nach dem von Munroe et al. (1995) beschriebenen Vereinigungsschema („Pooling“-Strategie) angeordnet (siehe 2.13), so daß ein Screening mittels PCR ermöglicht wurde.

Für das Screening dieser cDNA-Bibliotheken wurden die in **Tab. 3.1A-C** aufgelisteten Primer sowie die radioaktiv markierten Integrate eines Großteils der Exonklone der Gruppen B und C verwendet. Insgesamt konnten 3 cDNA-Klone isoliert werden, die nachfolgend dargestellt sind.

3.2.1 Der cDNA-Klon pFL12

Für das Exon 3A.A3 konnte mittels der abgeleiteten Primer 12-5' und 12-3' ein cDNA-Klon (pFL12) aus der hierarchisch angeordneten fetalen Leber- und Milz-cDNA-Bibliothek (Löbber, 1996) isoliert werden.

Die Nukleotidsequenz des 1123 Bp langen cDNA-Klon-Integrates pFL12 wurde in beiden Orientierungen mittels einer „Primer-Walking“-Strategie bestimmt und ist in **Abb. 3.10** dargestellt. Das Exon 3A.A3 zeigte mehrere Sequenzabweichungen zu der ermittelten Sequenz und entspricht daher nicht dieser cDNA, die deshalb in eine andere chromosomale Region kartiert. Durch Homologievergleiche in EST-Datenbanken konnten mehrere überlappende EST's identifiziert werden, die aus verschiedenen fetalen Geweben (AA628705, Hillier et al., 1997; T89587, T89859, T79074, Hillier et al., 1995, WashU-NCI humanes EST

Projekt, unveröffentlicht) und Tumor-cDNA-Bibliotheken stammen (AI263148, AI221100 National Cancer Institute, Cancer Genome Anatomy Project (CGAP); AA496479, Hillier et al., 1997, WashU-NCI humanes EST Projekt). Die Nukleotidsequenz dieser Klone zeigen eine 86-90 %ige Identität zu den genomischen Sequenzen derselben Regionen auf Chromosom 4, 18 und 21, die sich bereits durch Sequenzübereinstimmungen zu Exonsequenzen der Gruppe C des Contigs I auszeichneten.

```

GGCACGAGGCTTACTCTAACCCCTAAGAAGTAGAGGCACATTGATCCACCTAGAATCTGCATACAAAAGTTGGCCTCTGCCTGGGATTC      90
CCAACTGGACCAGACTTTGGGTTGAAGACATACGAAAACTCACAGAAGACATTTCTGCATCATGAGAGGTCAATGTAGACATCTTA      180
AGCCCCATTTTGTAGAGTGGGGCCCTTTGAGTTTTCCAGATCTTGTTCAGTGACCTGTCTACAGCTGTGTGAGAGGCTTCTGATGGGAACAG      270
AATCCGTGGCAGAATCTTTAAAGTGAACAAGCATCTTAGCAGTGGGAGGTAAGGCCACAAAATATCCAGAGCCTTAATGACAACCATA      360
ATCACAAGTGTACCTTCCCTCTATACTGTGTTATGGGAGTAGAGTACTTTTGTCTCTTCTTACCTCAGAGTTAGCTGATCAGGGAC      450
AGGGGATATCACATCTGGAGCCAAGATGTGCAGGACCACCAGGACAGCGCCATTTTCTGTTTGCACCACAAAAGTCAGCCTGAGTGTCT      540
CCTCCGTGGATCACTATGAGGCCATCATTACTCAGCATCTTTGAGTAATGAGACATTTGAGGATGTCATGTGCAGACTGGGGTCAGTCTG      630
ACAAGGTCTAATTTCTGCTCCCATTTTCAGAGGAAGATAAGTGAATCATGCAGGTTTGTCTCCCTCCCTCACAGAAAGAAATCTCCTTGGTTT      720
GTACCCAGATAAGAATGCTCCAGTTTTCTGGTTCTTTGGTGAAAAATGAGGAATCTGGAGACTCAAATGATCAATAAGCTAATTTGCTT      810
CTATTTCAATGATCATTTAGAAAAAGACATAAAGCAGTCATGGTCCCTACCATCAAGGAACTTGAGTCTAGAGCACATGAATAATGAT      900
TGAATTCAGCATCGTGTTCAGTACAAAGACAGAACTGTATGGAGACTTAAGCTTAATTTGTGTTACTTCCCTTTATTGTTTTGTGAG      990
TTTTAATGCCACCATCTGATCAGCTGTTTATTGACAGAGAGAAATTGTTTATTATTGTTGTTTACCTTGCTAAGAATAATA      1080
ATTAGCTTCTAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA      1123

```

Abb. 3.10: Nukleotidsequenz des cDNA-Klons pFL12. Der cDNA-Klon enthält keinen offenen Leserahmen und endet in einer Folge von von poly(A)-Resten. Zwei putative Polyadenylierungssignale sind unterstrichen.

3.2.2 Die cDNA-Klone λ CF16 und H7NIB

Mit dem Exon 6A.B12 als radioaktive Sonde sowie der Primer 16-5' und 16-3' konnte ein cDNA-Klon (λ CF16) aus einer menschlichen kraniofazialen cDNA-Bibliothek (J. Murray) und ein weiterer cDNA-Klon (H7NIB) aus der hierarchisch angeordneten NIB-cDNA-Bibliothek (Soares et al., 1994) identifiziert werden. Beide cDNA-Klone entsprechen einem neuen Gen, das in den folgenden Kapiteln dargestellt ist.

3.3 Molekulare Klonierung und chromosomale Lokalisation des humanen und murinen *LETM1* (*Letm1*)-Gens

3.3.1 Klonierung und Sequenzierung der kodierenden Region des menschlichen *LETM1*-Gens

Die Grundlage für die Klonierung des *LETM1*-Gens lieferten die unter 3.2.2 erwähnten cDNA-Klone λ CF16 und H7NIB. Die Klonierungsstrategie ist in **Abb. 3.11** dargestellt. Alle Klone wurden doppelsträngig mittels einer „Primer-Walking“-Strategie sequenziert. Die Isolierung und Sequenzierung der Klone H7NIB, AMA3 und AMA11 wurde im Rahmen einer medizinischen Doktorarbeit durchgeführt (Pak, 1999).

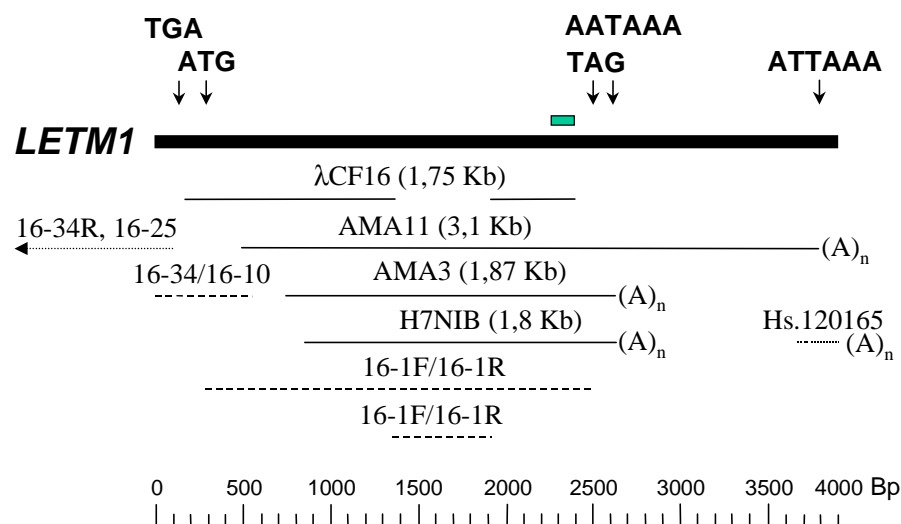


Abb. 3.11: Zeigt die Klonierungsstrategie des menschlichen *LETM1*-Gens. Das Exon 6A.B12, welches als Ausgangspunkt für die Isolation von cDNA-Klonen verwendet wurde, ist als grünes Kästchen oberhalb der als Balken dargestellten cDNA-Konsensussequenz eingezeichnet. Die Nukleotidsequenzen der cDNA-Klone, die aus einer kraniofazialen- (λ CF16), einer hierarchisch angeordneten Hirn-cDNA-Bibliothek (H7NIB) sowie einer Amniozyten-cDNA-Bibliothek (AMA3 und AMA11) isoliert wurden, sind als durchgehende oder (bei fehlenden Nukleotidsequenzen siehe λ CF16) als unterbrochene Linien eingezeichnet. Wesentliche RT-PCR-Produkte sowie genomische Sequenzierungen am PAC-Klon 184023, die zur Überprüfung und Verlängerung der bestehenden cDNA-Sequenz verwendet wurden, sind als gestrichelte oder punktierte Linien dargestellt. Zur Orientierung sind neben den möglichen Polyadenylierungssignalen das erste mögliche Startcodon sowie die, den offenen Leserahmen begrenzenden Stopcodons TGA und TAG eingezeichnet. Durch Homologievergleiche identifizierte EST-cDNA-Klone, mit deren Hilfe die cDNA im 3'-Bereich erweitert werden konnte (siehe UniGene-Cluster Hs.120165), sind als punktierte Linie verzeichnet. Die unterschiedlich langen, 3'-untranslatierten Bereiche der hier gezeigten cDNA-Klone werden durch eine Vielzahl weiterer EST-cDNA-Klone (UniGene-Cluster Hs.120165) bestätigt.

Die vollständige, durch die überlappenden cDNA-Klone erhaltene Nukleotidsequenz wurde durch RT-PCR an verschiedenen fetalen und adulten Geweben verifiziert und ist in **Abb. 3.13** dargestellt. Die Sequenzdaten sämtlicher RT-PCR-Produkte entsprechen der abgeleiteten Konsensussequenz. Eine Ausnahme bildet der um 528 Bp kürzere cDNA-Klon λ CF16 (**Abb. 3.11**). RT-PCR-Produkte, die dem cDNA-Klon λ CF16 entsprechen, konnten unter Verwendung unterschiedlichster Primerkombinationen nicht erhalten werden. Vermutlich entstand dieser Klon durch ein Rekombinationsereignis während der Propagierung der Bakteriophagen (vgl. **Abb. 3.13**).

Die abgeleitete cDNA-Konsensus-Sequenz enthält einen durchgehenden offenen Leserahmen von 2364 Bp. Der vollständige, durch zwei Stopcodons flankierte offene Leserahmen erstreckt sich, beginnend am ersten putativen Methionin, über 2217 Bp bis zum 3'-gelegenen Stopcodon und kodiert für ein putatives, aus 739 Aminosäuren bestehendes Protein (**Abb. 3.13**). Die Nukleotidsequenz um das erste putative Startcodon (GAGCACATGG) zeigt Abweichungen von der Konsensussequenz eukaryotischer Translationsstartpunkte GCC(A/G)CCATGG (Position -6 bis +4) (Übersichtsartikel: Kozak, 1996). Während die Position +4 durch den hochkonservierte Purinrest G besetzt ist, wird die zweite konservierte Position -3 durch einen Pyrimidinrest anstatt des konservierten Purinrestes (meist A) eingenommen. Das Vorhandensein zumindest einer dieser beiden Positionen erfüllt jedoch die Minimalbedingungen eukaryotischer Translationsstartpunkte (Kozak, 1996).

3.3.2 Klonierung und Sequenzierung der kodierenden Region des murinen *Letm1*-Gens

Um weiterführende Untersuchungen dieses Gens im Maus-Modell zu ermöglichen, wurde das murine *Letm1*-Gen isoliert und sequenziert. Die Klonierungs- und Sequenzierungsstrategie ist in **Abb. 3.12** dargestellt.

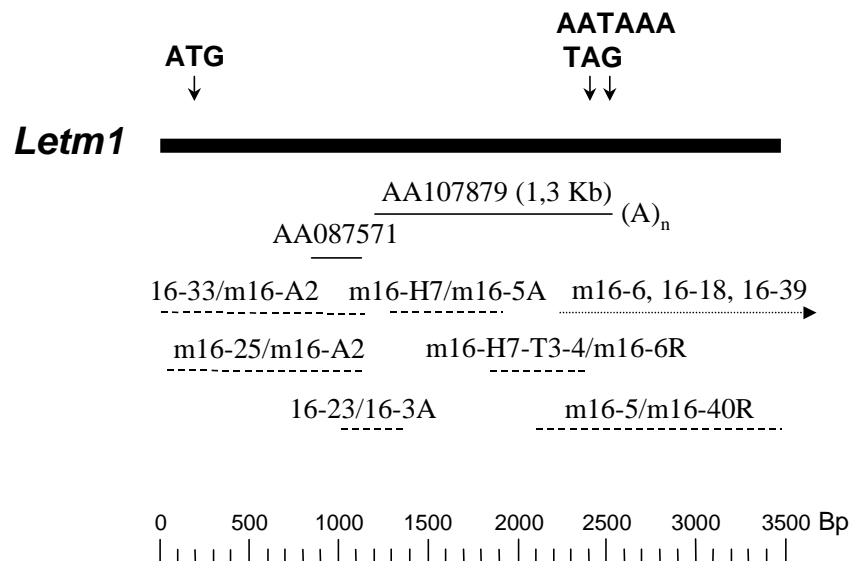


Abb. 3.12: Klonierungsstrategie des murinen *Letm1*-Gens. Als Grundlage für die vollständige Klonierung der murinen *Letm1*-cDNA diente der doppelsträngig sequenzierte cDNA-Klon AA107879. Weitere, durch Homologievergleiche identifizierte cDNA-Klone sind dargestellt, sofern sie durch die Ableitung von Primersequenzen für die weitere Klonierung der cDNA verwendet wurden (AA087571). RT-PCR-Produkte und genomische Sequenzierungen an den murinen genomischen P1-Klonen, die zur Überprüfung und Verlängerung der cDNA-Sequenz dienten, sind durch gestrichelte oder gepunktete Linien dargestellt. Die Mehrzahl der durch Homologievergleiche identifizierten 3'-EST-cDNA-Klone (UniGene-Cluster Mm. 26947) korrespondieren zumeist mit der Nutzung des in fett dargestellten Polyadenylierungssignales AATAAA.

Über erste Homologievergleiche der *LETMI*-cDNA mit EST-Datenbanken konnten mehrere murine cDNA-Klone aus verschiedenen embryonalen cDNA-Bibliotheken (Tag 8,5 bis 14,5 post coitus) mit Sequenzübereinstimmungen zu dem *LETMI*-Gen identifiziert werden (Acc.-Nr.: AA107879, AA087571 und AA014714). Die vorliegenden, ca. 300-400 Bp langen EST's zeigten eine mehr als 80 %ige Sequenzidentität zu der humanen *LETMI*-cDNA. Die Klone wurden vom Humanen Ressourcenzentrum in Berlin angefordert, vereinzelt und über Sequenzierung mit vektorspezifischen Primern überprüft. Nur einer dieser Klone (AA107879) enthielt die ursprüngliche, über Homologievergleiche ermittelte Sequenz und wurde nachfolgend auf beiden Strängen sequenziert. Aus dieser Sequenz sowie der Sequenz des ESTs (AA087571) wurden Primer abgeleitet, die in Kombination mit Primern aus dem humanen *LETMI*-Gen für RT-PCR-Analysen an verschiedenen fetalen und adulten murinen Geweben verwendet wurden. Durch Sequenzierung der erhaltenen PCR-Produkte konnte die cDNA-Sequenz in 5'-Richtung verlängert werden. Mittels einer kombinierten Strategie aus genomischer Sequenzierung an den murinen genomischen P1-Klonen ICRFP703D0215 und

ICRFP703K232 (vgl. **3.3.4** und **3.6**) sowie RT-PCR konnte die cDNA-Sequenz auch in Richtung 3' verlängert werden. Die Konsensussequenz von 3480 Bp sowie die abgeleitete Aminosäuresequenz (738 AS) ist in **Abb. 3.13** dargestellt.

3.3.3 Homologievergleiche der beiden Gene

Die Sequenzidentität der beiden cDNA-Sequenzen beschränkt sich hauptsächlich auf einen Bereich um den offenen Leserahmen und beträgt 81 % (**Abb. 3.13**). Die untranslatierte Region weist mit Ausnahme eines Bereichs um das putative Polyadenylierungssignal keine signifikanten Sequenzhomologien auf (23 %). Beide Gene kodieren für ein putatives Protein, das sich aus 739 Aminosäuren (bzw. 738 AS bei der Maus) zusammensetzt und ein kalkuliertes Molekulargewicht von 83,4 (82,9) kDa aufweist. Die abgeleiteten Aminosäuresequenzen sind zu 84 % identisch. Mittels verschiedener Softwareprogramme wurden die strukturellen Merkmale der putativen Proteine identifiziert, die aufgrund der nachfolgend dargestellten möglichen Proteindomänen zur Benennung des Gens als *LETMI* bzw. *Letml* (Leucin-Zipper und EF-Hand enthaltendes Transmembranprotein 1) führten (Endele et al., 1999) (**Abb. 3.13**).

Durch die Programme „TMPred“ und „SOSUI“ (Hofman und Stoffel, 1993; Hirokawa et al., 1998) konnte bei beiden Proteinen eine putative Transmembrandomäne identifiziert werden, die aufgrund der Ladungsverteilungen für Typ II-Transmembranproteine charakteristisch ist (Hartmann et al., 1989). C-terminal zu dieser Domäne weisen die Proteine mehrere α -helikale Bereiche auf, von denen einige zur Ausbildung von „coiled coils“ befähigt sein könnten („COILS“, Version 2.2, Lupas et al., 1991). Sie könnten damit an möglichen Protein-Protein-Interaktionen beteiligt sein (Kohn et al., 1997). Unter Verwendung des Programms „PROSITE“ (Bairoch et al., 1997) konnten mehrere mögliche Phosphorylierungsstellen für Proteinkinase C (PKC) und Caseinkinase II (CK2) sowie eine mögliche Glykosylierungsstelle identifiziert werden (**Abb. 3.13**). Im C-terminalen Bereich beider Proteine befinden sich neben einem möglichen Leucin-Zipper-Motiv zwei Bereiche mit Homologien zu dem EF-Hand-Motiv Ca^{2+} -bindender Proteine (EF-Hand I und II). Durch BLASTP-Homologievergleiche des Bereichs zwischen den beiden EF-Hand-Motiven (AS 297-317) konnten darüber hinaus Homologien (57%) zu der Linkerregion zwischen den EF-Hand-Motiven II und III der Caltractin (Centrin)-Protein-Familie in Pflanzen (*Dunaliella salina*) identifiziert werden.

EF-Hand-Motive stellen calciumbindende Motive dar, die aus einer α -Helix (Aminosäurereste 1-10), einer Schleife um das Ca^{2+} -Ion (10-21) und einer zweiten α -Helix (19-29) gebildet werden (Übersichtsartikel: Kawasaki und Kretsinger, 1994; Celio, 1996). Mehrere Aminosäuren innerhalb dieser Schleife (Position 1, 3, 5, 7, 9 und 12) sind an der koordinierten Bindung eines Calciumions beteiligt und sind daher besonders hoch konserviert. Die als (X, Y, Z sowie -X und -Z) bezeichneten Aminosäurereste sind über Sauerstoffatome der Seitenketten oder der Peptidbindung (-Y) an der Bindung des Calciumions beteiligt.

Ein Vergleich der beiden möglichen EF-Hand-Motive (EF-Hand I und II) des humanen und murinen LETM1-Proteins mit der „PROSITE“-Konsensussequenz EF-Hand-calciumbindender Proteine (PS00018) ist in **Abb. 3.14** dargestellt. Die EF-Hand II des humanen und murinen LETM1-Proteins enthält an den Positionen X, Y, Z, -X und -Z sowie an den Positionen 6, 8 und 13 konservierte Aminosäurereste und entspricht daher der Konsensussequenz. Eine Chou-Fasman-Analyse zeigt, daß diese Aminosäuren durch α -helikale Bereiche flankiert werden (Chou und Fasman, 1974). Die EF-Hand I weist hingegen mehrere Abweichungen zu dieser Konsensussequenz und keine flankierenden α -helikalen Bereiche auf.

	1	2	3	4	5	6	7	8	9	10	12	13
	X	Y			Z		-Y		-X		-Z	
	[D]-x-[DNS]-{ILVFW}-[DENSTG]-[DNQGRHK]-{GP}-[ILVMC]-[DENQSTAGC]-x(2)-[DE]-[LIVMFYW]											
LETM1-I	D	Y	S	E	D	L	Q	E	I	KK	E	L
Letm1-I	D	Y	S	E	D	L	Q	E	I	KK	E	L
LETM1-II	D	E	N	K	D	G	K	V	N	ID	D	L
Letm1-II	D	D	N	K	D	G	N	I	N	ID	D	L

Abb. 3.14: Vergleich der identifizierten EF-Hand-Motive des LETM1-Proteins mit der „PROSITE“-Konsensussequenz calciumbindender Proteine.

Ein Hinweis auf die Calciumbindungsfähigkeit von EF-Hand-Motiven kann durch den Tufty-Kretsinger-Test (Tufty und Kretsinger, 1975; Babitch, 1990) erhalten werden. Hierbei werden 16 konservierte Aminosäurereste des aus 29 AS bestehenden EF-Hand-Motivs auf ihr Vorhandensein überprüft. Die Anwendung dieses Tests ergibt für die EF-Hand II des LETM1-Proteins einen Score-Wert von 13 (von insgesamt 16). Sie könnte damit in der Lage

sein, ein Calciumion zu binden. Die EF-Hand I ergibt hingegen einen Score-Wert von 5 und scheint damit tatsächlich keine calciumbindenden Fähigkeiten aufzuweisen.

BLASTP-Homologievergleiche des humanen und murinen LETM1-Proteins mit den Proteinsequenz-Datenbanken ergaben signifikante Homologien (25-43 %ige Identität) zu verschiedenen putativen Transmembranproteinen aus *S. cerevisiae* (U40828, U48418, S66710), *S. pombe* (D98117, Z98559/CAB11168), *A. thalianae* (AC001229/AAB60908) und *C. elegans* (Z81094/CAB03156.1). Die Homologien beschränken sich dabei hauptsächlich auf einen zentralen Bereich des LETM1-Proteins, der die Transmembrandomäne beinhaltet. Die weiteren Proteindomänen sind, mit Ausnahme des putativen Proteins aus *C. elegans* (Z81094), nicht zwischen diesen Spezies konserviert. Das *C. elegans*-Protein (784 AS) enthält ebenfalls C-terminal der Typ II-Transmembrandomäne (AS 148 bis AS 164) ein mögliches EF-Hand-Motiv (DESDGIVDKQLV, AS 609 bis AS 621), das nach dem Tufty-Kretsinger Test eine mögliche Calciumbindungsfähigkeit aufweist (Score-Wert von 11).

Interessanterweise ist zwischen all diesen Proteinen, einschließlich des humanen und murinen LETM1-Proteins, eine spezifische Phosphorylierungsstelle für PKC und CK2 in einem bestimmten Abstand N-terminal der Transmembrandomäne konserviert. Eine Gegenüberstellung der Proteinsequenzen um diesen hochkonservierten Bereich ist in **Abb. 3.15** dargestellt.

N		PKC/CK2		Transmembrandomäne	C
LETM1	ELKHYHGFRLWIDTKIAARMLWRILNGHSL	TRRE	RRQFLRICADLFR	LVPFLVFVVVPFMEFLLPVA	VKLFNMLPST
Letm1	ELRHYYHGFRLWIDTKIAARMLWRILNGHTL	TRRE	RRQFLRICADLFR	LVPFLVFVVVPFMEFLLPV	VKLFNMLPST
U48418	ALRHYWDGSKLLGLEIKISSKLLMKSAAGYPL	TRRE	NLQLKRTTQDIVR	LVPFAAFLIIPFAELLLPFA	LKLFNLLPST
U40828	ALRHYWDGSKLLGLEIKISSKLLMKSAAGYPL	TRRE	NLQLKRTTQDIVR	LVPFAAFLIIPFAELLLPFA	LKLFNLLPST
D89117	GVLHFWDGTILLGVEIKISSKLVYKMAVGYEL	TRRE	SRQLTRTLKDIGR	LVPFSVFVVVPFAELLLPIA	VKLFNMLPST
Z98559	GVLHFWDGTKLLGVEIKISSKLVYKMAVGYEL	TRRE	SRQLTRTLKDIGR	LVPFSVFVVVPFAELLLPIA	VKLFNMLPST
AC001229	TLQHYWLGTKLLWADVRI SVRLLVKLANGKGL	SRRE	RQQLTRTTADIFR	LVPVAVFIIIVPFMEFLLPVF	LKLFNMLPST
Z81094	ELKHYHGFRLLALET RVSAKYLWTVLRGATL	SRRE	RQQLVRTVSDLFR	LVPFSFFIIVPFMELALPIF	IKLFPGLPST
S66710	ALKHYANGTKLLGYEIKVSTKLLIKFAQGYEL	SRRE	RNQLRRTMGDVFR	LIPFSAFLIIPFAELFLPFA	LKLFNMLPST
Konsensus	H G LL - + +	G L RRE	Q R D R	LVPF F PF E LP	KLFPN LPST

Abb. 3.15: Zeigt den Homologievergleich der Proteinsequenzen des humanen und murinen LETM1s mit verschiedenen Spezies. Die Homologie (25-43 %ige Identität) zwischen diesen Proteinen ist hauptsächlich auf den hier dargestellten Bereich um die Transmembrandomäne (blau) beschränkt. All diese Proteine enthalten eine mögliche PKC- und CK2-Phosphorylierungsstelle (in rot) in einem spezifischen Abstand N-terminal der Transmembrandomäne. Die abgeleitete Konsensussequenz zeigt identische oder konservierte positiv (+) und negativ (-) geladene Aminosäurereste.

Während die meisten dieser Proteine bisher nicht weiter charakterisierte putative Transmembranproteine darstellen, wurde für ein Protein (Mrs7p) aus *S. cerevisiae* (U48418) eine mitochondriale Lokalisation gezeigt (Waldherr et al., 1993). Analysen mit den Programmen „PSORT“ und „PSORT II“ (Nakai und Kanehisa, 1992) ergaben für die meisten dieser Proteine, einschließlich des putativen humanen und murinen LETM1-Proteins (zu 48 bzw. 17 % bzw. Scorewerten von 0,772 und 0,832) ebenfalls die Vorhersage einer möglichen mitochondrialen Lokalisation in der Zelle (innere mitochondriale Membran) (zwischen 17 und 56 %), wobei mögliche Schnittstellen für mitochondriale Präsequenzen vorhergesagt wurden (Gavel und van Heijne, 1990) (vgl. **Abb. 3.13**).

3.3.4 Genomische Struktur des humanen und murinen *LETM1/Letm1*-Gens

Die Lage, Anzahl und Größe der Intronsequenzen des menschlichen *LETM1*-Gens und des orthologen murinen Gens wurden mit Hilfe der PCR unter Verwendung der genomischen PAC-Klone 184O23 und 318L12 beim Menschen (vgl. **3.1**) bzw. der P1-Klone ICRFP703-D0215 und ICRFP703K232 (vgl. **3.6**) der Maus bestimmt. Hierbei wurden zunächst willkürlich verschiedene cDNA-Primer-Kombinationen für eine PCR an genomischer DNA verwendet, die Hinweise auf die Lage und schließlich die Größe von Intronbereichen ergaben. Zur Bestimmung großer Intronbereiche (>3 Kb) wurde das „Expand High-Fidelity PCR System“ (Boehringer Mannheim) verwendet. Exon-Intron-Grenzen wurden über die Sequenzierung der PCR-Produkte bzw. über Direktsequenzierung der PAC- und P1-DNA (184O23, 318L12 bzw. ICRFP703D0215 und ICRFP703K232) bestimmt.

Beide Gene bestehen aus 14 Exons und 13 Introns. Die Größen der Exons sind mit Ausnahme der Exons 9 und 10, die sich um jeweils 3 zusätzliche (Exon 9) bzw. fehlende Basenpaare (Exon 10) bei der Maus unterscheiden, weitgehend identisch. Die Lage der Spleißstellen ist ebenfalls zwischen beiden Genen konserviert. Damit ist die Genstruktur der beiden Gene gleich, wobei Unterschiede in den Introngrößen vorliegen. Beide Gene erstrecken sich genomisch über einen Bereich von jeweils ca. 40 Kb. Die Größe der Exons und Introns sowie die Exon-Intron-Grenzen sind in **Tab. 3.5** dargestellt.

Exon Nr.	Intron (3')	Exon (5')	Exon (3')	Intron (5')	Exon Größe (Kb)	Intron
hu As			Y T V P R			
hu Exon 1			ACACCGTCCC	<i>GGGgtaagcggcg</i>	≥ 0,379	> 7,5
mu Exon 1			CCGCCTCCC	<i>GGGgtaagctggt</i>	≥ 0,231	n.d.
hu As		G S P G D		L R N C L		
hu Exon 2	<i>cctgggtacag</i>	GTAGTCCAGGGGATC	GTTGAGGAACTGCCT	<i>gtgagtaaat</i>	0,061	7,5
mu Exon 2	<i>tttgtgtag</i>	GTAGTCTGAGGGATC	GTTGACGAGCCGTG	<i>gtaagtaagc</i>	0,061	8,0
hu As		N V P F G		R E R R Q		
hu Exon 3	<i>ccttttcaag</i>	GAATGTTCCATTTGG	CGGGAGCGCAGGCAG	<i>gtaatgggga</i>	0,451	5,0
mu Exon 3	<i>ccttttcaag</i>	GAGTGTCTGTCTCG	CGGGAGCGCAGGCAG	<i>gtaatgggaca</i>	0,451	7,0
hu As		F L R I C		T Q S L K		
hu Exon 4	<i>cccctttccag</i>	TTCCTCCGGATCTGC	ACTCAGTCACTCAAG	<i>gtaagggcag</i>	0,144	1,5
mu Exon 4	<i>atctttgcaag</i>	TTCCTCCGGATTTGT	ACCCAGTCTATCAAG	<i>gtaaggaagg</i>	0,144	0,74
hu As		E E R L K		V F F Q K		
hu Exon 5	<i>gccctttgcaag</i>	GAGGAGAGGCTGAAG	GTGTTTTTCCAGAAG	<i>gtgctgtgat</i>	0,138	1,95
mu Exon 5	<i>gctctcacag</i>	GAGGAAAGGCTGAAG	GCATTTTTCCAGAAG	<i>gtgcagtact</i>	0,138	0,65
hu As		I R E T G		K A D D K		
hu Exon 6	<i>actttccgcaag</i>	ATCCGGGAAACAGGG	AAGGCAGACGACAAG	<i>gtgagcctgg</i>	0,204	n.d.
mu Exon 6	<i>actttctgcaag</i>	ATCCGGGAGACAGGG	AAGGCTGATGATAAG	<i>gtgaggtggg</i>	0,204	8,2
hu As		L I A E E		G Q L K Q		
hu Exon 7	<i>tgttttctag</i>	CTGATTGCTGAGGAA	GGTCAGCTGAAAGCAG	<i>gtgctggaat</i>	0,120	1,9
mu Exon 7	<i>tctcttccag</i>	CTGATTCTGAGGAA	GGCCAGCTGAAACAG	<i>gtatgtggcc</i>	0,120	3,6
hu As		W L D L H		L P E I V		
hu Exon 8	<i>ttggttgcaag</i>	TGGCTGGACCTGCAC	CTCCAGAGATTGTG	<i>gtactgtgtg</i>	0,132	0,51
mu Exon 8	<i>ttcctttccag</i>	TGGCTGGACTTGCAC	CTTCCAGAAATTGTG	<i>gtatgtattg</i>	0,132	0,44
hu As		A K E A Q		R S E V A		
hu Exon 9	<i>tcttctgaag</i>	GCAAAGGAAGCACAG	CGCTCGGAGGTGGCG	<i>gtaagcggtg</i>	0,144	0,76
mu Exon 9	<i>tcctctgaag</i>	GCAAAGGAGGCCAG	GCCTCTGAGGCTGTG	<i>gtaagcagcc</i>	0,141	0,65
hu As		K D F E P		L E G L K		
hu Exon 10	<i>tgcctttccag</i>	AAGGATTTTGGAGCCC	CTGGAGGGCTTGAAG	<i>gtaatgagggc</i>	0,132	3,0
mu Exon 10	<i>tgcatttccag</i>	AAGGACATCCAGCCA	CTGGAAGGTCTGAAG	<i>gtgagctgca</i>	0,135	1,0
hu As		E E E I T		Q D Y S E		
hu Exon 11	<i>tttatcccag</i>	GAGGAAGAGATCAGC	CAGGACTACAGCGAG	<i>gtgggtggga</i>	0,135	2,5
mu Exon 11	<i>ttctctccag</i>	GGTGAAGAGATTACC	CAGGACTACAGTGA	<i>gtgggtggat</i>	0,135	1,0
hu As		D L Q E I		G M P T G		
hu Exon 12	<i>atthttattag</i>	GACTTGCAGGAGATC	CGGCATGCCACGGG	<i>gtgagtgggg</i>	0,188	1,0
mu Exon 12	<i>tttaaatag</i>	GACTTACAGGAGATC	GAGCACTCCAACAGG	<i>gtgagtgggc</i>	0,188	1,0
hu As		E N V I S		D D L V K		
hu Exon 13	<i>gcttccacag</i>	GGAGAACGTCATCAG	GACGACCTCGTCAAG	<i>gtgggcctcg</i>	0,139	1,2
mu Exon 13	<i>cactccccag</i>	GGAGAGTGTATCAG	GATGACCTTGTCAAG	<i>gtgagtcagt</i>	0,139	1,1
hu As		V I E L V				
hu Exon 14	<i>acgtgtgtag</i>	GTGATTGAGCTGGTG			≥ 1,540	
mu Exon 14	<i>ctgtgtgtag</i>	GTAATCGACTTGGTG			> 1,261	

Tab. 3.5: Vergleich der Exon/Intron-Struktur des humanen und murinen *LETMI*-Gens. Neben der Anzahl und Größe der jeweiligen humanen und murinen Exons und Introns sind die Nukleotidsequenzen der Exon-Intron-Grenzen dargestellt. Intronsequenzen sind durch kursive Kleinbuchstaben, Spleißdonor- und Akzeptorstellen durch Fettdruck gekennzeichnet. N.d.: nicht bestimmt.

Die Aufklärung der genomischen Organisation des *LETMI*-Gens erlaubte auch die genaue Lokalisation und Orientierung dieses Gens in Bezug auf den von Wright et al. (1997) weiter eingegrenzten WHS-kritischen Bereich. Durch die ursprüngliche Isolierung des Exons

34R erhaltenen PCR-Produkts von 894 Bp zwischen dem Cosmid-Fragment L75b9B und C (siehe **Abb. 4.1**). Der sich 5' an das Exon 1 angrenzende Bereich zeichnet sich durch einen hohen GC-Gehalt von ca. 70% (ca. über 1 Kb) aus, der typisch für die oftmals mit Haushaltsgenen assoziierten CpG-Inseln ist (Bird, 1987; Antequera und Bird, 1993). Durch die genaue Lokalisierung des Exons 1 ergibt sich, daß die regulatorischen Elemente dieses Gens weniger als 80 Kb von dem distalen Deletionsbruchpunkt der 165 Kb großen WHSCR entfernt sind (Wright et al., 1997) (siehe **Abb. 3.16**).

3.3.5 Expressionsanalysen

3.3.5.1 Northern-Blot-Analyse

Northern-Blot-Analyse des humanen und murinen LETMI-Gens

Für die Northern-Blot-Analyse des humanen *LETMI*-Gens wurden kommerziell erhältliche „Multiple Tissue Northern Blots“ mit poly(A)⁺-RNA aus verschiedenen fetalen und adulten menschlichen Geweben verwendet. Die Hybridisierung mit radioaktiv markierten Sonden aus unterschiedlichen Bereichen der *LETMI*-cDNA ergab eine ubiquitäre Expression in allen untersuchten fetalen und adulten Geweben (**Abb. 3.17**). Um die unter **3.3.1** beschriebene, mögliche Nutzung alternativer Polyadenylierungsstellen zu überprüfen, wurden über PCR zwei verschiedene Sonden aus der putativen kodierenden Region und der 3'-untranslatierten Region abgeleitet. Mehrere Hybridisierungssignale (5,5; 5,1; 3,8 und 2,8 Kb) konnten mittels der aus den Exons 1-3 abgeleiteten Sonde (16-1F/16-SPL1) erhalten werden (vgl. **Abb. 3.17A**), wobei die stärksten Hybridisierungssignale von 5,5 und 2,8 Kb vermutlich mit den unterschiedlichen cDNA-Längen von 3,907 und 2,646 Kb korrespondieren (vgl. **3.3.1** und **Abb. 3.13**). Die Hybridisierungssonde aus der 3'-untranslatierten Region ergab hingegen zwei deutlich sichtbare 5,5 Kb- und 2,0 Kb-Banden (**Abb. 3.17B** und **C**). Aus RT-PCR-Analysen an verschiedenen fetalen und adulten Geweben mit unterschiedlichsten Primern ergaben sich keine Hinweise auf alternative Spleißprodukte.

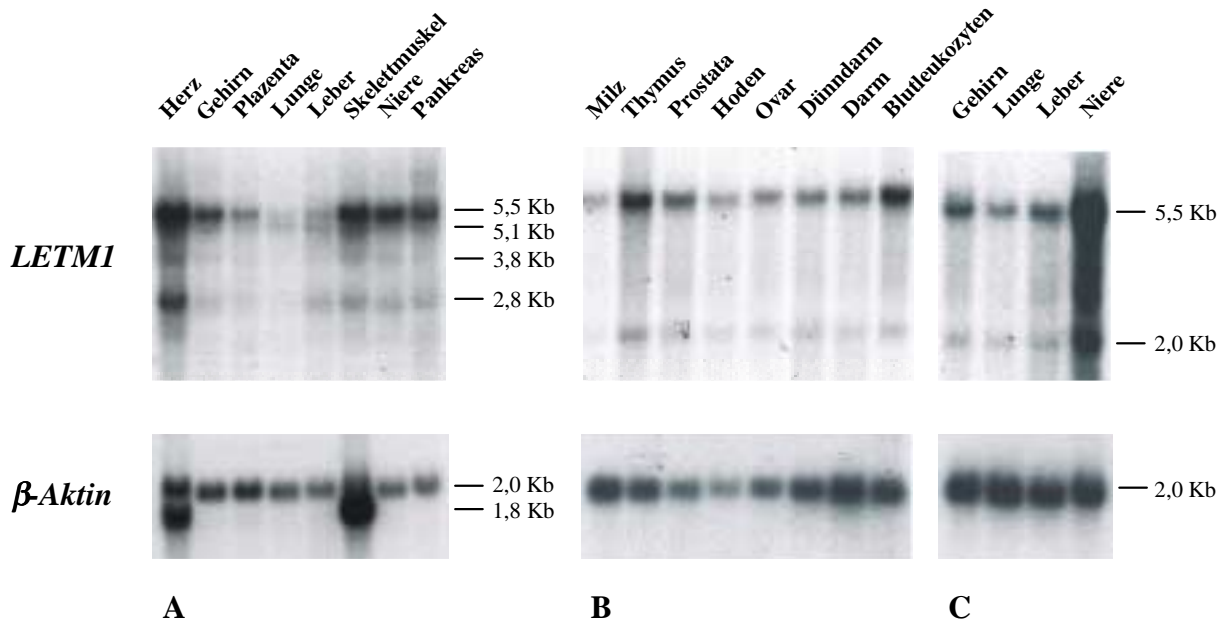


Abb. 3.17: Northern-Blot-Analyse des humanen *LETMI*-Gens. **A:** Zeigt das Autoradiogramm eines humanen adulten „Multiple Tissue Northern Blots“ (2.9.2), der mit einer strangspezifischen Hybridisierungssonde aus dem 5'-Bereich der *LETMI*-cDNA hybridisiert wurde. Die Sonde wurde über lineare PCR mit dem Primer 16-SPL1 an dem PCR-Produkt 16-1F/16-SPL1 generiert. Die dargestellten Autoradiogramme eines adulten (**B**) und fetalen (**C**) Northern-Blots wurden mit einer über „Random Hexamer-Priming“ markierten 0,9 Kb-Sonde hybridisiert (16-6'/AMA11-T7-1), die ausgehend von dem putativen Stoppcodon einen großen Teil der 3'-untranslatierten Region repräsentiert.

Die Hybridisierung der gezeigten Northern-Blots erfolgte nach Angaben des Herstellers. Für eine relative Abschätzung der Hybridisierungssignale wurden die Northern-Blots im Anschluß mit der vom Hersteller mitgelieferten β -Aktin-Sonde hybridisiert.

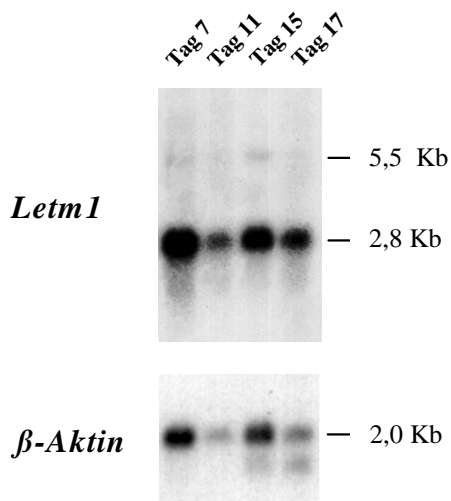


Abb. 3.18: Northern-Blot-Analyse des murinen *Letm1*-Gens. Dargestellt ist das Autoradiogramm eines fetalen murinen Northern-Blots mit poly(A)⁺-RNA der fetalen Stadien Tag 7 bis 17 (post coitus). Als Hybridisierungssonde wurde eine durch lineare PCR generierte, strangspezifische Hybridisierungssonde (m16-21) an dem als Templat verwendeten, linearisierten cDNA-Klon AA107879 verwendet.

Die Hybridisierung eines murinen fetalen Northern-Blots mit poly(A)⁺-RNA aus verschiedenen fetalen Stadien (Tag 7 bis 17 post coitus) ergab unter Verwendung einer aus den Exons 4-5 abgeleiteten Sonde in allen der untersuchten Stadien ein schwaches Hybridisierungssignal von 5,5 Kb und ein deutliches stärkeres Signal von 2,8 Kb (**Abb. 3.18**).

3.3.5.2 RNA-*in situ*-Hybridisierung an Gewebeschnitten

Für eine weitergehende Untersuchung der Expression dieses Gens wurde eine RNA-*in situ*-Hybridisierung (vgl. 2.10.2) an sagittalen Schnitten embryonaler Mäuse (Tag 13,5 und 14,5 post coitus) vorgenommen. Hybridisierungssignale konnten in allen Geweben nachgewiesen werden und reflektieren die ubiquitäre, starke Expression dieses Gens (vgl. **Abb. 3.19**). Ein Vergleich mehrerer Gewebeschnitte (nicht dargestellt) deutete auf eine stärkere Expression in Gehirn und Retina hin. Die scheinbar stärkeren Hybridisierungssignale in der Leber treten hingegen bei zahlreichen Experimenten auf und sind, wie einige der entsprechenden „sense“-Kontrollen zeigen, auf die Struktur dieses Gewebes zurückzuführen.

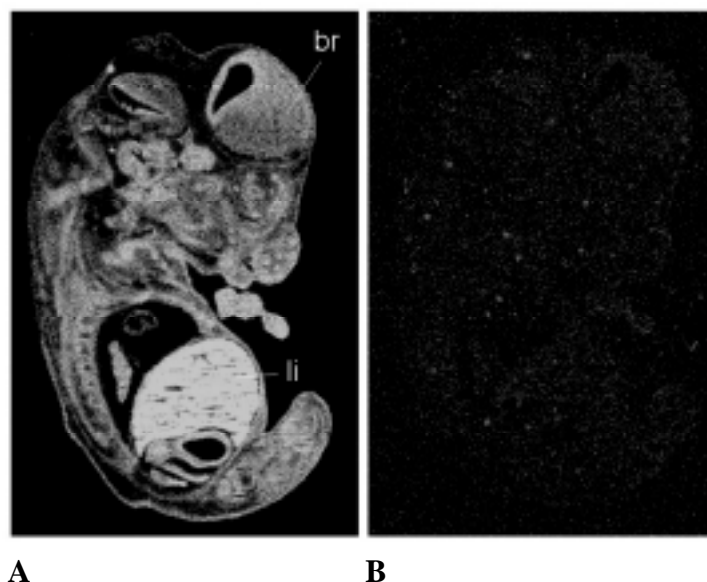


Abb. 3.19: RNA-*in situ*-Hybridisierung an Gewebeschnitten embryonaler Mäuse (Tag 13,5 post coitus). Als Hybridisierungssonden wurden ³⁵S-markierte Oligonukleotide aus der cDNA-Sequenz des murinen *Letm1*-Gens verwendet (m1602 „sense“ und „antisense“: Exon 8/9). Ein ähnliches Bild ergab sich auch für die aus dem Exon 5 abgeleiteten Hybridisierungssonden m1601 „sense“ und „antisense“. Dargestellt sind Autoradiogramme nach 5 tägiger Expositionszeit. Zur Orientierung sind einige Gewebe durch Pfeile markiert (br: Gehirn, li: Leber). **A:** „Antisense“-Hybridisierung. **B:** „Sense“-Hybridisierung.

3.3.6 Identifizierung und Sequenzierung des prozessierten *LETMI*-Pseudogens

Beim Screening der humanen genomischen PAC-Bibliothek mit einer aus dem *LETMI*-Gen abgeleiteten Sonde (6A.B12) konnte ein Klon isoliert werden, der in der Chromosomenregion 8q21-22 kartiert (vgl. **3.1.1**). Durch nachfolgende Direktsequenzierung des Klons mit den Exonprimern 16-5' und 3' konnte eine über das Exon 13 hinausgehende, durchgehende Sequenz mit über 80 %iger Identität zu der cDNA-Sequenz des *LETMI*-Gens ermittelt werden. Die fehlenden Intronbereiche weisen auf ein prozessiertes Pseudogen hin. Durch eine kombinierte Strategie aus Sequenzierung von PCR-Produkten mit cDNA-Primern aus dem *LETMI*-Gen sowie durch Direktsequenzierungen an dem PAC-Klon 167F14 konnte das vollständige Pseudogen ermittelt werden (vgl. **Abb. 3.20**). Die insgesamt 3942 Bp lange Sequenz zeigt eine 85 %ige Sequenzidentität über einen Bereich von Bp 797-Bp 3167 zu Bp 395-Bp 3028 der *LETMI*-cDNA (ca. Exon 2 bis 14). Homologievergleiche des angrenzenden 5'-Bereichs ergaben signifikante Homologien (von Bp 1 bis Bp 535: 91 %ige Identität auf Nukleotidebene bzw. von Bp 1-329: 99 %ige Identität auf Proteinebene) zu dem 3'-Ende aktiver LINE-Elemente (beispielsweise Accession-Nr. U93574) und zu dem C-terminalen Bereich des abgeleiteten, putativen p150-Proteins (ORF2) (Sassaman et al., 1997). In dem angrenzenden 3'-Bereich (Bp 3167-3942) konnten keine signifikanten Homologien in den Datenbanken gefunden werden.

<p> <i>VL</i>LETM1 CTAGAAATACCATTTGACCCAGCCATCCCGTTACTGGGTATATACCCAAAGGATTATAAATCATGCTGCTATAAAGACACATGCACACGT <i>LETM1</i> ATATTTATTGGCGCATTATCACCATAGCAAAGACTTGAACCAACCCAGATGTCCAACAATGATAGACTGGATTAAGAAAATGTGGCAC <i>VL</i>LETM1 GTATACACCATGGAATACTATGCAGCCATAAAAAATGATGAGTTCAITGCTCTTTGTAGGGACAGGATGAACTGGAAACCATTCTCA <i>LETM1</i> GCAAACCTGTCTCAAGAACAAAAACCACACCCGCTATTTCTACTCATAGTGGGAATTAACAATGAGATCACATGGACACAGGAAGG <i>LETM1</i> GGNACATCACACTCCGGGCTGTGTGGGTGGGGGAGGGGAGGATAGCATGGGAGATATACCTAATGCTAGATGACGAGTTAG <i>LETM1</i> TGGGTGCAGCCACCTGTCATGGCACATGTATACATATGTAACCTAACCTGCACAATGTGCACATGTACCTAAAACCTAAAGTATAATAAA <i>LETM1</i> AAAAAAAAACCTCCATACCAACAATAAAAAAAGAAAAATACAGTGTCCCATGATGAAGGAAGAACACTTACAAGGGATGGGACCTT <i>LETM1</i> GAGAGACCTGGGTATCCAGGAGCGAAGCAAAAGATCTGATGTCTGAGGAAGGTGAGAAGGGCCCTGCTCATGAAGATGTAAGATATT <i>LETM1</i> AAGTATGAGTGAATGCTGTATCAATCATGAAGCCTGTGATGAGAAGTCTGGGTAGTCCACCCAGAAGTCCAACTGCTGCTGTCTCACC <i>LETM1</i> CTGCTCATCTCAGC **** * </p>	<p>90 180 270 360 450 540 630 720 810 408</p>
<p> <i>VL</i>LETM1 TGTGCCAACATCTGGGGTTGAGGAGTGGAGCAAATGTTCCATTTGGCTGTTGCACTCCGGTCCACCCCGTGTACCCCTCCCTTCAGAGGC <i>LETM1</i> TGTGCCAGCCCTGGGGTTGAGGAACTCCCTGAATGTTCCATTTGGCTGCTGCACTCCCATCCACCCCTGTATACACATCTCCAGAGGC ***** </p>	<p>900 498</p>
<p> <i>VL</i>LETM1 GATCGCATGCTGTTAGGCCTTGAAGCAGAAGTCCCGCAGCAGGTTGGCGAGAGTGCATGGACCTCTGCCACAGTGGATTTTATGGTT <i>LETM1</i> GATCACTCGCTGTTGGGCTCTGAGGCCGAGTGCCTTCGATGATGTCTGAGAGGCCATGGACCTCTACCTCTGTGGGTTTGTGGCT **** * </p>	<p>990 588</p>
<p> <i>VL</i>LETM1 GCGGGACCTCGGTGCCTTCTGTGCACAGCTGGCACTCTTACGCGCTGTCTCTGATGACTTGGTAGTAGAAGTCCCTCAAGTCCCTC <i>LETM1</i> GTGGACCTCAGTGCCTTCTGTGCGTGGCACTCTTCCGCCCTGTTCCGATGACTCGGTAGTAGAAGTCCCTCAAGTCCCTC * </p>	<p>1080 678</p>
<p> <i>VL</i>LETM1 AAGGACAAGAACAAAGCTTGAGGAAGCGGCCAAGTGTACACCCCAACCCACAGAGGTGGTGGTGAATAAGTCTCGGCAGAGGG <i>LETM1</i> AAGGACAAGAACAAAGCTTGAGGAAGCGGCCAAGTGTACACCCCAACCCACAGAGGTGGTGGTGAATAAGTCTCGGCAGAGGG ***** </p>	<p>1170 766</p>
<p> <i>VL</i>LETM1 AGCTGGATGAGCTGAAGCGCTACTGCTTGGCTTCTGCCTGCTGTGGATCCATGCCAAGATCGTGGTCTCATGCTGTGGCACATCCTTG <i>LETM1</i> TGCTGGACGACTGAAGCCTACTACCATGGCTTCCGCTGCTATGATGATGACACCAAGATCGGGCACGATGCTGTGGCCATCTCA ***** </p>	<p>1260 856</p>
<p> <i>VL</i>LETM1 ACAGCCACACCTGACCAGGCAA-AGCGAGCAGTTTCTCCGGATCCGCGCCGACCTCCTCCGCTTGGTGGCTTCTTGTGTTCTGG <i>LETM1</i> ACGCCACAGCTGACCCCGGGAGCGCAGCAGTTTCTCCGGATCTGCGCTGACCTCTTCCGCTGGTGGCTTCTTGTGTTCTGG * </p>	<p>1349 946</p>
<p> <i>VL</i>LETM1 TTGTCCAGTTTATGGAATTCCTGCGCCCGTGAACATGAAAGCTTCTCCAGCATGCCAGCATCCAGTTTGAAGCCAGTCCATCAAG <i>LETM1</i> TGGTCCCTCAGTGGATTTCTGCTGCTGTTGCTGTGAAGCTTCCCAACATGTTCCCATCCACATTGAGACTCAGTCAAGG * </p>	<p>1439 1036</p>
<p> <i>VL</i>LETM1 AGGAGAGCCTGAAG--GAGCTTGGGTCAACCTGGAAGCTGGCAGAGTTCCTCCAGGACACCATCGAGGAGGACCTGAAGAACATA <i>LETM1</i> AGGAGAGCCTGAAGAGGAGCTTGGGTGAGTGGATTCCTCCAGGACACCATCGAGGAGGACCTGAAGAACATA ***** </p>	<p>1526 1125</p>
<p> <i>VL</i>LETM1 GCAGCCAAAGGCGAGTCCACCA--GACTTCTCCGT--TTTCCAGAAAGATCCGGGGATGGGGAGAGACCCAGCANATGAGAAATCAT <i>LETM1</i> GCAGCCAAAGGCGAGTCCACCA--GACTTCTCCGT--TTTCCAGAAAGATCCGGGGATGGGGAGAGACCCAGCANATGAGAAATCAT ***** </p>	<p>1612 1214</p>
<p> <i>VL</i>LETM1 GCGTTTTTCCAAATTTATTTGAGAAATGAGTGAACCTGGACAACCTGTCCAGGCGCAGCTGGTGGCGCGTGAAGTGTGGAGTGCA <i>LETM1</i> GCGTTTTTCCAAATTTATTTGAGAAATGAGTGAACCTGGACAACCTGTCCAGGCGCAGCTGGTGGCGCGTGAAGTGTGGAGTGCA ***** </p>	<p>1702 1304</p>
<p> <i>VL</i>LETM1 GTCCATCGGCATCAACAATTC-TGCGGTTCCAGCTCACCATGTGACTGCGCTCCATAAAGGCCAATGATAAGTGTGGAGGA-GG <i>LETM1</i> GTCCATCGGCATCAACAATTC-TGCGGTTCCAGCTCACCATGTGACTGCGCTCCATAAAGGCCAATGATAAGTGTGGAGGA-GG ***** </p>	<p>1790 1394</p>
<p> <i>VL</i>LETM1 GGTGGACAACCTGAAATGTCAAGGGGCTGCAGCGCGGTGTGGAAAGTGGAGTGTGGGCT-----CCGCT----- <i>LETM1</i> GGTGGACAACCTGAAATGTCAAGGGGCTGCAGCGCGGTGTGGAAAGTGGAGTGTGGGCT-----CCGCT----- ***** </p>	<p>1859 1484</p>
<p> <i>VL</i>LETM1 -CAGCTGAAGCAGTGTCTGGACCTGCACCTGCATCAGGAGATCCCATATCGCTACTCATCTGTCCAGAGCATGTACCTCCAGAAAC <i>LETM1</i> TCAGCTGAAGCAGTGTCTGGACCTGCACCTGCATCAGGAGATCCCACATCGCTACTCATCTGTCCAGAGCATGTACCTCCAGAAAC ***** </p>	<p>1948 1574</p>
<p> <i>VL</i>LETM1 --TCTTGGCAGCTGACCAGCTCAAGTCCACTGACAGCTCTCCAGAGATTGTGGCAAAGGAAGCAGGTTGAAAGTGGCTGAGGTGGA <i>LETM1</i> CCTCTTCCAGCCAGCTCAAGTCCACTGACAGCTCTCCAGAGATTGTGGCAAAGGAAGCAGGTTGAAAGTGGCTGAGGTGGA ***** </p>	<p>2036 1664</p>
<p> <i>VL</i>LETM1 GGGTGGCAAGTGCACAA--GGTAAAGCTGGAGGTACGCTGCAGGAGGAGGAGCCATCCAGCAGGAGCACCA--CAAGGAGCTGCA <i>LETM1</i> GGGTGGCAAGTGCACAA--GGTAAAGCTGGAGGTACGCTGCAGGAGGAGGAGCCATCCAGCAGGAGCACCA--CAAGGAGCTGCA ***** </p>	<p>2120 1754</p>
<p> <i>VL</i>LETM1 GAAGCCTCGGAGTTGGAAAGGATGTCCGCGCGAAGTGTGGAAGCCACCCCTCGAAGGCGAGGGCCTAACCTCAGCCAGAAGTGGC <i>LETM1</i> GAAGCCTCGGAGTTGGAAAGGATGTCCGCGCGAAGTGTGGAAGCCACCCCTCGAAGGCGAGGGCCTAACCTCAGCCAGAAGTGGC ***** </p>	<p>2210 1844</p>
<p> <i>VL</i>LETM1 TGAAGCAGTCTGTGGTCAAGACCTGAAGGACACTGCCCC--TGCTAGAGGCTTGAAGGAGAAGGAGATAAATGAAGGAGAAATCGA <i>LETM1</i> TGACACAGTCTGTGGTCAAGACCTGAAGGACACTGCCCC--TGCTAGAGGCTTGAAGGAGAAGGAGATAAATGAAGGAGAAATCGA ***** </p>	<p>2300 1934</p>
<p> <i>VL</i>LETM1 CATCTCAGCGATGCTGCTTAAGCTGAAGGAGCAGAAG--TCCGCTACCAAGAGAAGGAGGAGCTGGAGCTGTGAAGGAGGAGCT <i>LETM1</i> CATCTCAGCGATGCTGCTTAAGCTGAAGGAGCAGAAG--TCCGCTACCAAGAGAAGGAGGAGCTGGAGCTGTGAAGGAGGAGCT ***** </p>	<p>2387 2024</p>
<p> <i>VL</i>LETM1 GCAGGACTACAGCGAACAATGCAAGAGATCAAGAAGAACTCTCAAAGACTGGTAGGTGAAGAAAATATGTGAAGACATCAAAGCCA <i>LETM1</i> GCAGGACTACAGCGAACAATGCAAGAGATCAAGAAGAACTCTCAAAGACTGGTAGGTGAAGAAAATATGTGAAGACATCAAAGCCA ***** </p>	<p>2477 2110</p>
<p> <i>VL</i>LETM1 CCAAGAGACTGACAGA--GGGTGACGAGATGATCGGGCAGATCGAGGCTTGTCTCACACCTGAGATGAGACAG--GCGCTGCAAGC <i>LETM1</i> GCAAGAGATTGACAAAAGGGTGCAGCAATGATCGGGCAGATCGATGGCTTGTCTCACACCTGAGATGAGACAG--GCGCTGCAAGC ***** </p>	<p>2563 2200</p>
<p> <i>VL</i>LETM1 TGGGCTGCGCTAGGGCGCGCCGTAGGGGAGAACATCATAGTGTCACTGAGCTCATCAGCACCATGAAGCAAGTCAAGCATTCCAG <i>LETM1</i> TGGGCTGCGCTAGGGCGCGCCGTAGGGGAGAACATCATAGTGTCACTGAGCTCATCAGCACCATGAAGCAAGTCAAGCATTCCAG ***** </p>	<p>2653 2290</p>
<p> <i>VL</i>LETM1 AAAGCAAGCTCACCAGCTGGCCTCAGCACTGGATGAAGACAGGATGGCAAGGTCAACATCAATGACCTCATCAATGTGATGAGTGG <i>LETM1</i> AAAGCAAGCTCACCAGCTGGCCTCAGCACTGGATGAAGACAGGATGGCAAGGTCAACATCAATGACCTCATCAATGTGATGAGTGG ***** </p>	<p>2743 2380</p>

6A.B12

3.4 Detaillierte Analyse der WHS-kritischen Region zwischen D4S43 und dem *LETMI*-Gen

Im Verlauf der vorliegenden Arbeit konnte der WHS-kritische Bereich auf der Basis eines Patienten mit milderem klinischen Befund (Somer et al., 1995) auf einen 165 Kb großen Bereich zwischen dem *FGFR3*-Gen und dem Marker D4S43 eingegrenzt werden (Wright et al., 1997). Daher können die unter 3.1 identifizierten Transkripte und die bekannten Gene der Contigs I und II als mögliche Kandidatengene für die wesentlichen Merkmale dieses Syndroms ausgeschlossen werden. Aus diesem Grund wurde der WHS-kritische Bereich inklusive der flankierenden Regionen (D4S43 bis *FGFR3*) eingehender untersucht. Die genomische Sequenz dieses Bereichs lag durch das Sanger Center vor (siehe 3.1 und Abb. 3.3). Innerhalb dieser ca. 400 Kb großen Region zwischen D4S43 und dem *FGFR3*-Gen waren zu Beginn der Untersuchungen neben dem bereits in Kapitel 3.3 beschriebenen *LETMI*-Gen noch keine weiteren Gene bekannt. Es sollte mittels computergestützter Analysen eine umfassende Transkriptkarte dieser Region erstellt und weiterführende sowie vergleichende Untersuchungen an Mensch und Maus durchgeführt werden.

3.4.1 Identifizierung transkribierter Bereiche durch computergestützte Analysen

Die bereits in den Datenbanken vorliegende genomische Sequenz von D4S43 bis zum *LETMI*-Gen wurde zunächst eingehend mit verschiedenen Softwareprogrammen analysiert (vgl. 2.16). Mit Hilfe des GRAIL II-Programms (Xu et al., 1994) konnten innerhalb der Cosmide HSL108f12 bis 75b9 insgesamt 85 putative Exonsequenzen identifiziert werden. Sie wurden fortlaufend als z1-z85 bezeichnet. 62 dieser Exonsequenzen mit exzellenten oder guten Score-Werten wurden für BLASTX- und BLASTN-Homologievergleiche mit DNA- (nr = „non redundant“) sowie Proteindatenbanken („swissprot“) verwendet. Während 8 dieser Exons auf repetitive Elemente (Alu- und LINE-Elemente) entfielen, zeigten die restlichen 54 Exons keine signifikanten Homologien zu bekannten Genen oder Proteinen und wurden nachfolgend für BLASTN-Analysen der EST-Datenbanken des NCBI (Boguski et al., 1993) verwendet. Aufgrund der häufig fehlerhaften EST-Sequenzen wurden nur solche Suchergebnisse berücksichtigt, die eine Sequenzidentität von über 95 % (human) und ≥ 80 % (murin) auf-

wiesen. Die identifizierten cDNA-Klone wurden wiederum mit den genomischen Cosmid-Sequenzen verglichen und bei entsprechenden Sequenzübereinstimmungen (95-100 % beim Menschen und ≥ 80 % bei der Maus) für weitere Analysen herangezogen (vgl. **Tab. 3.6**).

Sämtliche dieser humanen (AA131767, R16217, N32670, N23294, H51641, A064672, AA159454) und murinen cDNA-Klone (AA118417, AA003395, AA062233, A014730) entfielen auf einen Bereich zwischen den Cosmiden 58b6 und 184d6. Innerhalb des weiter proximal gelegenen Bereichs zwischen den Markern D4S43 und D4S166 (108f12 bis 58b6) konnten hingegen durch diese ersten Analysen keine Transkripte identifiziert werden. Um die Erstellung einer möglichst vollständigen Transkriptkarte zu ermöglichen, wurde die gesamte Region zwischen D4S43 und dem *LETMI*-Gen im Verlaufe dieser Arbeit zusätzlich durch ein weiteres Genvorhersageprogramm („GENSCAN“, Burge und Karlin, 1997) nachanalysiert, das die Erstellung vollständigerer Genmodelle erlaubt (Burge und Karlin, 1997; Burge und Karlin, 1998) (siehe **Abb. 3.21**). Diese Untersuchungen führten zu der Identifizierung weiterer EST-cDNA-Klone des Menschen (AA436996, AI674093, AA675927, AI201252, H30421, A401245, T07860, AA324442, H37943, W28349) und der Maus (AI20069, AW045355, AI475002, W91513, AA230913, W62449). Daneben konnten durch einen gezielten Vergleich genomischer Cosmidsequenzen mit den EST-Sequenzdaten weitere humane und murine EST-Klone identifiziert werden (AI055861, AI841594, R13046, R44930, T03256, M62262, R50359, T85656, A155706, H05215, H05214, A139269, N73027, N53662, AA545768), die nicht durch GRAIL- oder GENSCAN-Analysen identifizierbar waren. Inzwischen konnten insgesamt 82 murine und 260 humane cDNA-Klone mit Sequenzübereinstimmungen zu dieser genomischen Region gefunden werden. Die meisten dieser cDNA-Klone lassen sich 10 humanen bzw. 4 murinen UniGene-Clustern zuordnen (**Tab. 3.6**).

3.4.2 Charakterisierung der identifizierten Transkripte

Zur Verifizierung und Charakterisierung der identifizierten humanen und murinen Transkripte wurden zunächst die Randsequenzen bzw. die vollständigen Integritätssequenzen einzelner humaner und muriner cDNA-Klone bestimmt (vgl. **Tab. 3.6** und **Abb. 3.21**). Hierdurch konnten die Klone verifiziert und Informationen über die genomische Ausdehnung der einzelnen Transkripte gewonnen werden, die wiederum eine weitere Zuordnung überlappender cDNA-Klone zu einer möglichen Transkriptionseinheit ermöglichten. Die Er-

gebnisse sind in **Tab. 3.6** und **Abb. 3.21** dargestellt. Letztlich konnten eine Reihe cDNAs bei Mensch und Maus in verschiedene „Contigs“ angeordnet werden, die aufgrund ihrer Sequenz-übereinstimmungen und ihrer genomischen Lokalisation bzw. ihrer abgeleiteten Transkriptionsorientierung zu mindestens 5 eindeutig unabhängigen Transkriptionseinheiten gehören. Diese wurden zunächst willkürlich als *Gen 51*, *43*, *45* und *57* bzw. *POLAP* (nach einer in die Datenbanken eingegebenen partiellen cDNA-Sequenz) bezeichnet.

Durch die Anzahl und Herkunft der aus unterschiedlichen cDNA-Bibliotheken stammenden cDNA-Klone in den EST-Datenbanken konnten erste Erkenntnisse über die Expression der einzelnen Transkripte gewonnen werden. Die hohe Redundanz der cDNA-Klone einzelner Contigs bzw. ihre Herkunft aus unterschiedlichen fetalen und adulten cDNA-Bibliotheken deutete dabei auf eine ubiquitäre Expression der meisten dieser Transkripte hin, so daß sich aufgrund dieser ersten Expressionsdaten mit Ausnahme des *Gen 45* kein offensichtliches Kandidatengenerale ableiten ließ (vgl. **Tab. 3.6** und **3.4.2.3**).

Homologievergleiche auf DNA und Proteinebene (BLASTN und BLASTX) ergaben für die identifizierten Transkripte (*Gen 51*, *57* und *POLAP*) signifikante Übereinstimmungen zu bekannten Genen oder Proteinen (P- bzw. E-Werte < 0,001) (siehe **3.4.2.1**, **3.4.2.4** und **3.4.2.5**). Die übrigen Transkripte (*Gen 43* und *Gen 45*) sowie Transkripte mit bisher unklarer Zugehörigkeit wiesen keine signifikanten Homologien zu bekannten Genen oder Proteinen auf (P-Werte >0,001) (**3.4.2.2**, **3.4.2.3**, **3.4.2.6**). Einige Transkripte (W62449, T85656, AA155707 und R50359) wiesen Homologien zu repetitiven Elementen auf oder schienen auf Kontaminationen innerhalb der cDNA-Bibliotheken zu beruhen und wurden daher von der weiteren Bearbeitung ausgeschlossen. Diese Ergebnisse werden in Kapitel **3.4.2.6** zusammengefaßt.

Cosmid	Art der Identifizierung und Sequenzidentität in %			Accession	IMAGp998*	Größe (in Bp) bzw. Rand-Sequenz	UniGene-Cluster/ EST-Anzahl	cDNA-Bibliothek	Homologievergleich (BLASTN) (in Bp)	Einteilung in Contigs
	GR	GS	S %							
141a8-58b6	z18	+	-	81-85	(AA118417) siehe Kap. 3.4.2.4	5'	-	<i>Ma, Le, My; Th; Ho, E</i>	58b6 5' 25234- 25664 3'	Gen 57 (erfs)
	z17	+	-	94	AA475002	5'	-	<i>Brd</i>	141a8 5' 6859 - 6924 32404 - 32462 35467 - 35589 3'	Pol4p
	z54	+	-						58b6 5' 19950 - 19985	
	z17	+	-	86-91	AI120069	5'	-		141a8 5' 6859 - 6924 58b6 19939 - 19985 174g8 2065 - 2189 3'	
174g8a	-	-	+	97	T03256	3'	-	<i>fHi</i>	3' 4198 - 4540 5'	
	-	+	-	100	AA436996	5'	-	<i>Ho</i>	5' 4799 - 5128 6066 - 6176 6651 - 6699 3'	POLAP
	z59	+	-	100	AI674093	3'	-	<i>Sb.: (fLu, Ho, Bz), Ho</i>	5' 6851 - 6768 9619 - 9183 3'	
	-	-	+	98	AI055861	3'	Hs.133483	<i>Hi</i>	3' 19012 - 19543 5'	
	-	-	+	95	AI841594	3'	Mm.40583	<i>E</i>	3' 19011 - 19062 5'	
	-	-	+	100	M62262	n	-	<i>Hc, Ho, He</i>	3' 19017 - 19744 5'	
	-	+	-	89	AW045355	n	-	<i>Hi</i>	3' 24283 - 24009 5'	
174g8b	-	-	+	100	R50359/ R50022	3' 5'	*001238	<i>Br</i>	5' 58 - 756 3'	
	-	+	-	95	W91513	5'	*A03985	<i>Ho, Hi, E, Pl, T, Th, Hc</i>	3' 7286 - 7389 8425 - 8513 5'	Gen 45
27h9	-	+	+	99	AA675927 AI341602	5' 5'	-	<i>Ho, Hi, DaT, AnOl, KzT</i>	3' 7286 - 7388 8425 - 8513 8765 - 8921 9180 - 9527 5'	
	-	-	+	98	AA131767	3'	*J071202	<i>U, Me, fHe</i>	5' 9437 - 9533 14584 - 15045 3'	
96a2	-	-	+	97	R13046	5'	*A01143	<i>fhi</i>	27h9 5' 27131 - 27840 33492 - 33564 96a2 636 - 1150 3'	

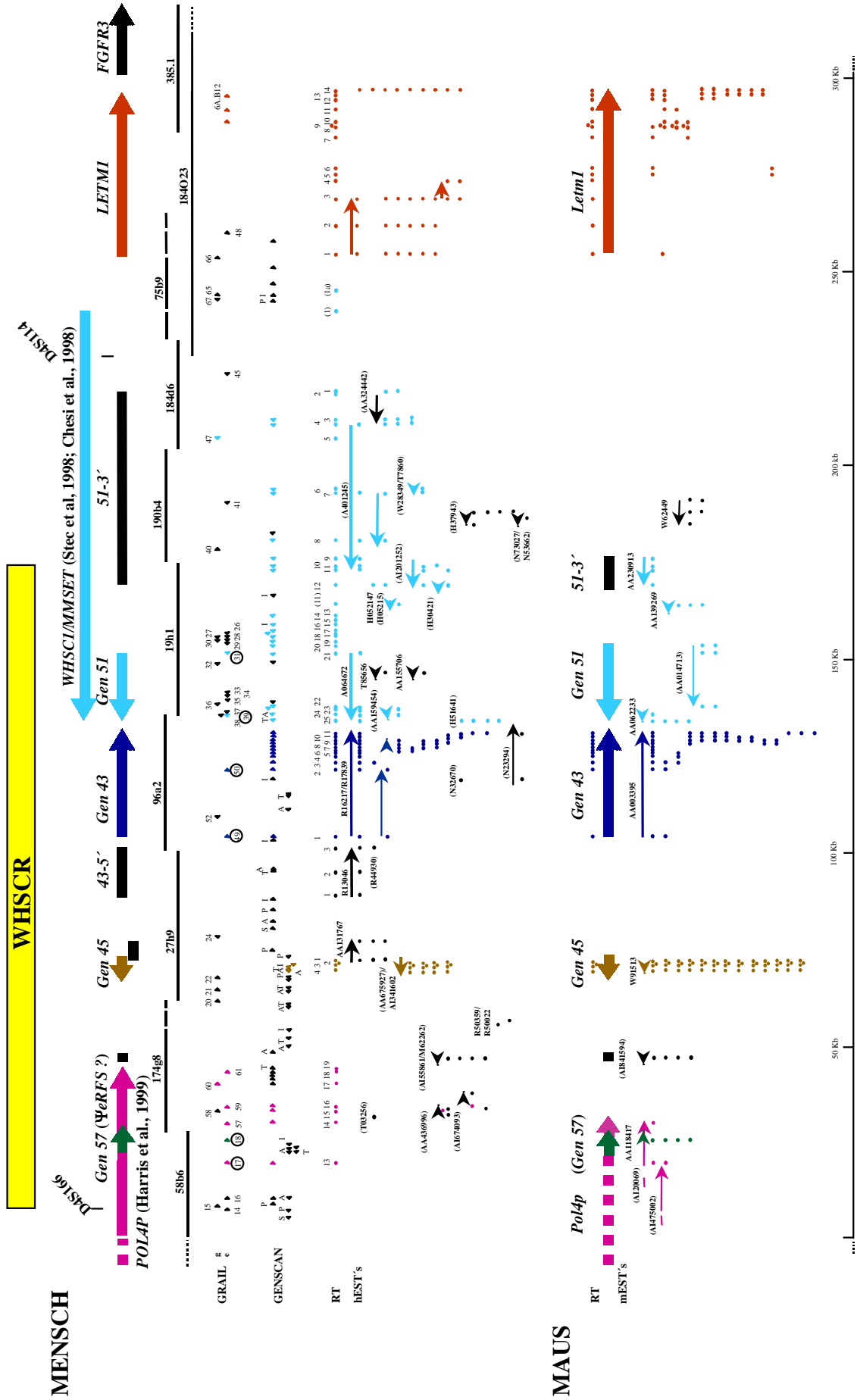


Abb. 3.21: Physikalische Karte der Chromosomenregion 4p16.3 (D4S166 bis *FGFR3*). Legende siehe Seite 104.

Legende zu Tabelle 3.6: Neben der Zuordnung zu den Cosmiden ist die Art der Identifizierung (GR: GRAIL 2; G: GENSCAN; S: Sonstiges: direkter genomischer Sequenzvergleich oder weiterer Homologievergleich mit bereits identifizierten und teilweise sequenzierten Klonen) und der Prozentsatz der Sequenzidentität der identifizierten Klone angegeben. Des Weiteren sind die Accession-Nummern der ESTs und die IMAGE-Klonbezeichnungen der vom RZPD angeforderten und über Rand- oder vollständige Sequenzierung weiter charakterisierten cDNA-Klone aufgeführt. cDNA-Klone der Maus sind in kursiv dargestellt. Identifizierte cDNA-Klone, die aufgrund der Sequenzanalysen oder aufgrund überlappender EST-Sequenzen wahrscheinlich einer Transkriptionseinheit zugeordnet werden konnten (Einteilung in Contigs), sind zur leichteren Orientierung entsprechend **Abb. 3.21** farbig hinterlegt. Farbige Markierungen der UniGene-Cluster kennzeichnen cDNA-Klone, die dem *erfs*-Gen (grün) (Wallrapp et al., 1998), dem *WHSC2*-Gen (Wright et al., 1999b) (dunkelblau), dem *WHCS1*- bzw. *MMSET*-Gen (hellblau) (Stec et al., 1998; Chesi et al., 1998) bzw. dem *POLAP*-Gen (rosa) (Harris et al., 1999; unveröffentlicht) zugeordnet werden konnten.

Ak: Adenokarzinom; Ao: Aorta; Au: Auge; Bl: Blut; B: B-Zellen; Br: Brust, Brd: Brustdrüse; Ei: Eizelle; He: Herz, Hi: Hirn, Da: Darm, DaT: Darm Tumor; E: Embryo; Ha: Haut; He: Herz; Hc: Hippocampus; Kn: Knochen; KzT: Keimzellen-Tumoren; Lu: Lunge; Le: Leber; Ly: Lymphknoten; Ma: Makrophagen; Me: Melanozyten; Ms: Multiple Sklerose-Gewebe; My: Myotuben; Ne: Neuroepithel; AnOl: anaplastische Oligodendrozyten; Pa: Pankreas; Pr: Prostata; Re: Retina, T: T-Zellen; Th: Thymus; To: Tonsillen; U: Uterus; f: fetal; Sb: Subtraktionsbibliothek.

Legende zu Abb. 3.21: Die genomischen Sequenzen des distalen Teils des Cosmid-Contigs (Baxendale et al., 1993) sind schematisch als Linien dargestellt. Unterhalb dieser Linien sind die durch GRAIL- oder GENSCAN-Analysen identifizierten Exons (g = gute oder e = exzellente Score-Werte) als Pfeile gemäß ihrer putativen Transkriptionsorientierung eingezeichnet (I: initiales Exon, T: terminales Exon, A: Polyadenylierungssignal, P: Promotorbereich, S: „single exon gene“). Sequenzierte cDNA-Klone und einige wesentliche EST-Sequenzen sind durch farbige oder schwarze Pfeile markiert und durch die jeweiligen Accession-Nummern entsprechend **Tab. 3.6** gekennzeichnet, wobei nicht selber sequenzierte cDNA-Klone in Klammern angegeben sind. Insgesamt konnten vermutlich 5 eindeutig unabhängige Gene identifiziert werden. Nur einige der identifizierten cDNA-Klone eines überlappenden cDNA-Clusters sind zur Verdeutlichung der Redundanz als farbige Punkte oder bei unklarer Zugehörigkeit, durch schwarze Punkte vermerkt. Exonsequenzen, die sich den identifizierten cDNA-Klonen oder dem *WHSC1*- bzw. *MMSET*-Gen (Stec et al., 1998, Chesi et al., 1998), dem *erfs*-Gen (Wallrapp et al., 1998) und der partiellen cDNA-Sequenz eines DNA-Polymerase ähnlichen Gens (*POLAP*; Harris et al., 1999; unveröffentlicht; Acc.-Nr.: AF044578) zuordnen ließen, sind entsprechend **Tab. 3.6** durch farbige Markierung hervorgehoben. Weiterhin sind oberhalb der EST-cDNA-Klone die in den nachfolgenden Kapiteln bzw. durch Stec et al. (1998), Chesi et al. (1998) (in Klammern vermerkt) und Harris et al. (1999) verifizierte Exonsequenzen durch fortlaufende Nummern vermerkt (siehe Kapitel **3.4.2**). Darüber hinaus sind das humane und murine *LETMI*-Gen sowie das mittels Exonamplifikation isolierte Exon 6A.B12 (roter Pfeil) eingezeichnet (siehe Kapitel **3.3**).

3.4.2.1 Das Gen 51 (*WHSCI*)

cDNA-Sequenzen und Homologievergleiche

Die meisten der durch GRAIL- und GENSCAN-Analysen identifizierten humanen Transkripte innerhalb der Cosmide 19h1 bis 184d6 konnten in insgesamt zwei größere überlappende cDNA-Contigs (*51* und *51-3'*) derselben abgeleiteten Transkriptionsorientierung angeordnet werden. Beide Contigs sind mit inzwischen über 142 Eintragungen innerhalb 3 unabhängiger UniGene-Cluster vertreten (Hs.110457, Hs.226751 und Hs.227240) und stammen aus unterschiedlichen fetalen und adulten cDNA-Bibliotheken (vgl. **Tab. 3.6** und **Abb. 3.21**). Ein Teil dieser cDNA-Klone zeigte in den ersten Homologievergleichen signifikante Übereinstimmungen zu bereits bekannten Proteindomänen, wie beispielsweise Homologien (P-Werte von $3,8 \times e^{-25}$ bis $4,6 \times e^{-6}$) zu der SET-Domäne (Tschiersch et al., 1994) verschiedener Proteine aus *D. melanogaster*, Mensch, Maus und Hefe (AA064672) und der HMG-Box verschiedener Proteine der HMG („high mobility group”)-Superfamilie (AA401245). Beide Contigs sowie einige weitere innerhalb der Cosmide 19h1 bis 184d6 identifizierten cDNA-Klone (Hs.148560, vgl. **Tab. 3.6**) zeigten in den Homologievergleichen eine 100 %ige Identität zu dem in etwa zeitgleich von Stec et al. (1998) und Chesi et al. (1998) beschriebenen *WHSCI*- („Wolf-Hirschhorn syndrome candidate 1”) bzw. *MMSET*- („Multiple Myeloma SET domain”) Gen. Das in der vorliegenden Arbeit als *Gen 51* bezeichnete Gen entspricht somit dem *WHSCI*- bzw. *MMSET*-Gen und wurde daher nicht weiter charakterisiert (vgl. **Tab. 3.6** und **Abb. 3.21**). Das *WHSCI*-Gen erstreckt sich genomisch über einen Bereich von ca. 120 Kb (96a2 bis 75b9) und scheint aufgrund seiner funktionellen Domänen (SET-; PHD-; PWWP-Domäne und HMG-Box) eine entwicklungspezifische Rolle zu spielen.

Zu dem *Gen 51* (*WHSCI*-Gen) konnten insgesamt 13, teilweise überlappende cDNA-Klone der Maus (AA062233 (Mm.19892), A014730, AA230913 und AA139269) identifiziert werden, die aus unterschiedlichen fetalen und adulten cDNA-Bibliotheken stammen (**Tab. 3.6**). Zwei EST-Klone der Maus wurden sequenziert (AA139269 und AA062233). Beide cDNA-Klone weisen eine über 74 %ige (AA139269) bzw. 80 %ige (AA062233) Identität auf Nukleinsäureebene und eine 87,8 %ige Identität auf Proteinebene (AA062233) zu dem *WHSCI*- bzw. *MMSET*-Gen auf und repräsentieren vermutlich tatsächlich das orthologe *Whsc1*-Gen der Maus. Dies wird durch die Analyse der publizierten EST-Sequenzen

bestätigt. Exemplarisch ist die 0,7 Kb lange cDNA-Sequenz des cDNA-Klons AA062233 zusammen mit der abgeleiteten Aminosäuresequenz im Vergleich zu der cDNA- und Protein-Sequenz des humanen *WHSC1*-Gens (Acc.: AF083386) dargestellt (**Abb. 3.22**).

<i>WHSC1</i>	AAAAAGACCAAGAAGAAAACGAGGGCGGCAGAGCAAAGGGGAAGGGAAGGGCAGTCAGAGGACGAGTGCTTCCGCTGCGGTGATGCC	4239
<i>Whsc1</i>	AAAAAGCCAAGAAGAAAACGAGGGCGGCAGAGCAAAGGGTGAAGGAAAACGGGAGTCAGAGGATGAGTGCTTCCGCTGCGGTGATGGT	90
<i>WHSC1</i>	T K A K K K T R R R R A K G E G Q R E S E D E C F R C G D G	1248
<i>Whsc1</i>	K K	30
<i>WHSC1</i>	GGGAGCTGGTGTGTGACCGCAAGTTCTGCACCAAGGCCTACCACCTGCTGCTGGCCTTGGCAAGCGGCCCTTCGGGAAGTGG	4329
<i>Whsc1</i>	GGGAGCTGGTGTGTGACCGCAAGTTCTGCACCAAGGCCTACCACCTATCCTGCTGGTGGCAAGCGGCCCTTCGGGAAGTGG	180
<i>WHSC1</i>	G Q L V L C D R K F C T K A Y H L S C L G L G K R P F G K W	1278
<i>Whsc1</i>	60
<i>WHSC1</i>	GAATGTCCTTGGCATCATTGTGACGTGTGTGGCAAACCTTCGACTTCATTTTGCCACCTCTGCCCAATTCGTTCTGTAAGGAGCACCAG	4419
<i>Whsc1</i>	GAATGTCCTTGGCATCATTGTGATGTATGTGGCAAACCTTCACCTCATTTTGCCACCTCTGCCCAATTCGTTCTGTAAGGAAACACCAA	270
<i>WHSC1</i>	E C P W H H C D V C G K P S T S F C H L C P N S F C K E H Q	1308
<i>Whsc1</i>	90
<i>WHSC1</i>	GACGGGACGACCTTCAGCTGCACCCGGACGGGCGSTCTACTGCTGTGAGCATGACTTAGGGGCGGCATCGTTCAGAAGCACCAGACT	4509
<i>Whsc1</i>	GATGGGACTGCTTCCGTTCCACCAGGATGGGCGTCTACTGCTGTGAGCATGACTTAGGGGCGGCATCGTTCAGAAGCACCAGACT	360
<i>WHSC1</i>	D G T A F S C T P D G R S Y C C E H D L G A A S V R S T K T	1338
<i>Whsc1</i> R S . Q . . Q R . D . S S	120
<i>WHSC1</i>	GAGAAGCCCCCAGAGCCAGGGAAGCCGAAGGGGAAGGGCGGCGGGGCTGGCGGAGAGTCACAGAGGGCAAATAGCCAG	4599
<i>Whsc1</i>	GAGAAGCCCTTCCAGAATTACTGAAGTCAAAGGAAAGGAAAGGCGGTGCTGGCGAAGGTCACAGATGGCAAATAGCCAG---	447
<i>WHSC1</i>	P K P P G E P L K S K R R R K G R R C W R E V T D G K *	1365
<i>Whsc1</i>	E . . F P . L G K . . K R *	147
<i>WHSC1</i>	GCGGCCGCTTGGCCGATCCAGGGCGGTGCAGGGCGCCGCTGCTGCGG-GAGAGGGGAGCATGA-ACTGGCCCGGAGGACCCA	4687
<i>Whsc1</i>	-----TTGGTCAGATCCTGGGTGGTGGGGCAGCCGCTGCAGACAGTGGGAAGGCAAGCACGCCACCCACCCAGGACCCA	529
<i>WHSC1</i>	GCTCGAGCCGCCAGACACAGACGTACAGGCCTCCTGGGAGGGAG-CGCCTCC--CCACCACTGAGCCATCCTCAGCAGCTCCGCTGC	4774
<i>Whsc1</i>	GCTC-----CTAG--CGCTG-TGTATAGCCTGCTTGGTGGGAGGAGCTCCTCCACCACTGAGCCATCCTCAGCAGC-TGCGCTGC	609
<i>WHSC1</i>	GTCTGCACTGATGACCGTCTG---AGCCAGCTCAGCGTCTCTGGACAAACA-GCCTCACTCCTCAGCGTTACCGCCACACTTGAATTT	4859
<i>Whsc1</i>	GTCTGCACTGATGACCGTCTGCTGAGGCAGGCTTGGTGGCCAGGATAAACAGCCCTTACTACAGCGCCACCTCTACACTTGAATCT	699
<i>WHSC1</i>	CT--CCGAATGCAAGTTCCCTCCACTCTATTTTTTAGGTTAAAGTTAA-TTGGCATATGGAATGTTTTAATCTCCTCTGAAATGT	4946
<i>Whsc1</i>	GAGACCCAGCGTCAAGGTGCCCTCCATTTC--TATTTTAGATTAAAGTTATGATTGAAGATTAATAAAAAAAAAAAAAAAAAAAAA	782

Abb. 3.22: Homologievergleich des murinen cDNA-Klons AA062233 (*Whsc1*) mit der cDNA-Sequenz und abgeleiteter Aminosäuresequenz des *WHSC1*-Gens (Stec et al., 1998). Identische Nukleotid- und Aminosäurereste sind durch „*“ bzw. „.“, fehlende durch „-“ gekennzeichnet. Die cDNA-Sequenz der Maus enthält einen durchgehenden offenen Leserahmen (Bp 1 bis Bp 441). Die abgeleitete Proteinsequenz (*Whsc1*) zeigt eine 87,8 %ige Identität zu dem C-terminalen Bereich des 136 kDa *WHSC1*-Proteins (Stec et al., 1998). Putative Kernlokalisierungssignale sind eingerahmt (PSORT), ein putatives PHD-Zinkfingermotiv (PROSITE-Eintragung PS50016) ist fett hinterlegt. Ein mögliches Polyadenylierungssignal der Maus ist unterstrichen.

Expressionsanalyse

Für den Nachweis der Expression des murinen *Whsc1*-Gens wurde eine Hybridisierungssonde aus der Nukleotidsequenz des cDNA-Klons AA062233 (m51-3/m51-4) abgeleitet und für die Hybridisierung eines fetalen murinen Northern-Blots mit poly(A)⁺-RNA aus verschiedenen fetalen Stadien verwendet. Starke Hybridisierungssignale von 5,3 bzw. 8,4 Kb konnten an RNA der Entwicklungsstadien Tag 7 bis 17 (post coitus) der Maus nachgewiesen werden (vgl. **Abb. 3.23**).

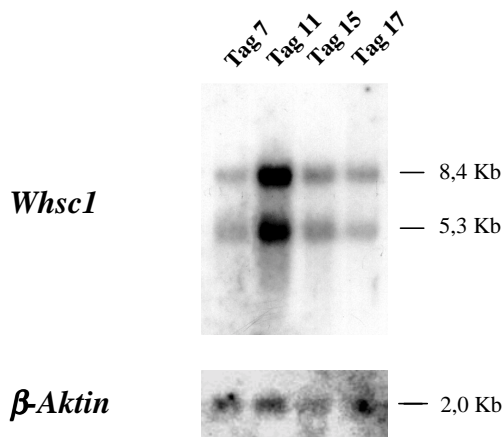


Abb. 3.23: Northern-Blot-Analyse des murinen *Whsc1*-Gens. Im Vergleich zu der Kontrollhybridisierung mit einer β -*Aktin*-Sonde treten die stärksten Signale innerhalb des Stadiums Tag 11 (post coitus) auf. Die Transkriptgrößen entsprechen in etwa den Transkripten von 7,7 und 5,2 Kb des Menschen, die sich scheinbar auf die Nutzung unterschiedlicher Polyadenylierungssignale zurückführen lassen (Chesi et al., 1998).

3.4.2.2 Das *Gen 43* (*WHSC2*)

cDNA-Sequenzen des Menschen und der Maus

Aus dem Cosmid 96a2 konnten mehrere überlappende cDNA-Klone identifiziert werden, die aus verschiedenen fetalen und adulten cDNA-Bibliotheken stammen und mit inzwischen 54 Eintragungen dem UniGene-Cluster Hs.21771 entsprechen (vgl. **Tab. 3.6**). Die vollständige, doppelsträngige Sequenzierung des 2144 Bp großen Integrates eines dieser cDNA-Klone (R16217), der zunächst einem als *Gen 43* bezeichneten Gen zugeordnet wurde, wurde von Pia Hermanns (1999) im Rahmen einer Diplomarbeit vorgenommen. Die cDNA-Sequenz konnte durch RT-PCR mit Hilfe des an der genomischen Sequenz des Cosmids 96a2 abgeleiteten Primers ProTra2 und eines cDNA-Primers (96a2-4) im 5'-Bereich verlängert werden. Die im Anschluß über RT-PCR vollständig überprüfte Sequenz von 2201 Bp enthält einen durchgehenden offenen Leserahmen (+1), der im 3'-Bereich durch ein Stopcodon begrenzt wird (**Abb. 3.24**). Der 5'-Bereich der cDNA (Bp 1-Bp 231) befindet sich innerhalb einer möglichen CpG-Insel mit einem GC-Gehalt von ca. 68 % (Nukleotidsequenz 2342-3995 des Comids 96a2). Er enthält darüber hinaus einen möglichen Translationsstartpunkt (Bp 16-Bp 25, CCCAAGATGG), der weitgehend der Konsensussequenz von eukaryotischen Translationsstartpunkten entspricht (Kozak, 1996) und die konservierten Purinreste (-3, +4) beinhaltet. Die abgeleitete Proteinsequenz besteht aus 528 Aminosäuren. Die Nutzung eines weiter 5'-gelegenen Translationsstartpunkts kann jedoch aufgrund des durchgehenden offenen Leserahmens nicht ausgeschlossen werden.

Zu den humanen cDNA-Klonen konnten auch mehrere cDNA-Klone der Maus identifiziert werden, die mit großer Wahrscheinlichkeit das orthologe murine *Gen 43* repräsentieren. (vgl. **Tab. 3.6**). Die vollständige Nukleotidsequenz eines dieser murinen EST-cDNA-Klone (AA003395) wurde ermittelt und durch die Verwendung eines genomischen Primers (mProTra-2) an genomischer DNA um weitere 62 Bp im 5'-Bereich verlängert. Die 2175 Bp lange, durch RT-PCR an Gesamt-RNA aus verschiedenen murinen Geweben überprüfte Konsensussequenz enthält ebenfalls einen durchgehenden offenen Leserahmen von Bp 1 bis Bp 1629 und kodiert für ein mögliches 530 Aminosäuren großes Protein (vgl. **Abb. 3.24**).

Homologievergleiche der beiden Gene

Das murine und humane Gen wurde in etwa zeitgleich durch Wright et al. (1999b) identifiziert und als *WHSC2*- bzw. *Whsc2h*-Gen (Wolf-Hirschhorn syndrome candidate 2) des Menschen (Accession-Nr.: AF101434) und der Maus (Accession-Nr.: AF101435) bezeichnet. Dabei weist das humane *Gen 43* eine 99 %ige Sequenzidentität zu dem *WHSC2*-Gen nach Wright et al. (1999b) auf. Die cDNA-Sequenz des murinen *Gen 43* zeigt jedoch nur eine 92 %ige Identität (bzw. 98 % Identität auf Proteinebene) zu der um 58 Bp kürzeren cDNA-Sequenz des *Whsc2h*-Gens nach Wright et al. (1999b). Die Unterschiede bei der Maus lassen sich nur durch Sequenzierfehler oder Polymorphismen erklären.

Die cDNA-Sequenzen des humanen und murinen *Gen 43* zeigen im Bereich des offenen Leserahmens eine 84 %ige Sequenzidentität (**Abb. 3.24**). Die von dem jeweiligen ersten putativen Methionin ausgehenden abgeleiteten Aminosäuresequenzen (528 AS und 530 AS mit errechneten Molekulargewichten von 57,3 und 57,5 kDa) sind zu 95 % identisch. Hingegen weisen die durch Wright et al. (1999b) beschriebenen Aminosäuresequenzen (528 und 521 AS) eine 93,3 %ige Sequenzidentität auf. Die hohen Übereinstimmungen des *Gen 43* (bzw. des *WHSC2/Whsc2h*-Gen) des Menschen und der Maus sowohl auf Nukleotidsequenz- als auch auf Proteinebene deuten darauf hin, daß es sich um orthologe Gene bzw. Proteine handelt. BLASTN und BLASTP-Homologievergleiche ergaben keine signifikanten Homologien zu bereits bekannten Genen oder Proteinen. Mit Hilfe verschiedener Vorhersageprogramme wurden die strukturellen Merkmale der Proteine identifiziert (vgl. **Abb. 3.24**). Während durch das Programm „TMPred“ (Hofman und Stoffel, 1993) für beide Proteine eine putative Typ I-Transmembrandomäne identifiziert werden konnte (AS 23 bis AS 45), wurde durch die Programme „SOSUI“ (Hirokawa et al., 1998) und „PSORT“ (Nakai und Kanehisa, 1992) ein mögliches, lösliches Protein vorhergesagt, das durch zwei putative Kernlokalisierungssignale („PSORT“) im Zellkern lokalisiert sein könnte (61 %). Das putative Protein 43 besitzt darüber hinaus mehrere α -helikale Bereiche, von denen einer (AS 109 bis AS 130) eine mögliche „coiled coil“-Domäne darstellt („COILS“ Version 2.2, Lupas et al., 1991). Durch das Programm „PROSITE“ (Bairoch et al., 1997) konnten darüber hinaus mehrere putative Phosphorylierungsstellen für Caseinkinase II (CK2), cAMP/cGMP-abhängige Proteinkinase und Proteinkinase C (PKC) detektiert werden.

RPCIP711P07512 bestimmt (vgl. 3.6). Die vergleichende Darstellung der bisher aufgeklärten Exon-Intron-Struktur beider, aus mindestens 11 Exons bestehender Gene ist in **Tabelle 3.7** ersichtlich. Die Größen der Exons sind mit Ausnahme des Exons 9, das bei der Maus über 6 zusätzliche Basenpaare verfügt und des Exons 11, das den 3'-untranslatierten Bereich beinhaltet, weitgehend identisch. Die Spleißdonor- und akzeptorstellen sind ebenfalls zwischen beiden Genen konserviert.

Exon Nr.	Intron (3')	Exon (5')	Exon (3')	Intron (5')	Exon Größe (Kb)	Intron
hu AS			R T V D E			
hu Exon 1			CGCACGGTGGACGAG gt aaggcagg		≥ 0,231	17,026
mu Exon 1			CGCACCGTGGACGAG gt gaggacgc		≥ 0,249	n.d.
hu AS		M K G A L	E L R E K			
hu Exon 2	tat	tttttgc ag ATGAAGGGCGCCCTA	AACTTAGAGAAAAG gt atgtttgtg		0,172	1,674
mu Exon 2	tt	tcttggc ag ATGAAGGCGGCTCTG	AACTTAGAGAAAAG.....		0,172	n.d.
hu AS		V G E C E	E L L Q K			
hu Exon 3	tt	accggc ag TGGGTGAGTGTGAAG	AGCTGCTGCAGAAG gt gagtgcc		0,162	1,703
mu Exon 3	TGAGTGAGTGCGAGG	AACTGCTGCAGAAAT.....		0,162	n.d.
hu AS		S T E T A	K M D T T			
hu Exon 4	tt	catttcc ag CCACGGAGACCGCCC	AGATGGACACCACC ag taaggcccc		0,090	1,520
mu Exon 4	CTACCGAGACAGCCC	AGATGGACTACTCAA.....		0,090	n.d.
hu AS		T P L K G	E R G V K			
hu Exon 5	cgt	cctcc ag CCCCACTCAAAGGCA	GAACGAGGTGTGAAG gt taggcccgt		0,131	0,088
mu Exon 5	att	tctgt ag CCCCCTCAAAGGCA	GAAAGGGGTGTGAAG gt gagacatc		0,131	0,088
hu AS		L L D I S	R R K T L			
hu Exon 6	gt	tcttggc ag CTGCTGGACATCTCT	GAAGGAAGACTCTCG gt gggggttg		0,070	0,249
mu Exon 6	att	tcttggc ag CTGCTGGATATTTCT	GGAGGAAGACTTTAG gt gtgtggat		0,070	0,308
hu AS		D A E V V	L V S T Q			
hu Exon 7	tt	gtccac ag ATGCGGAGGTGGTGG	CTGGTGTCCACGCAG gt agggtctg		0,089	0,856
mu Exon 7	tt	atttggc ag ACACAGAAGTGGTAG	CTGGTATCCACACAG gt agccactg		0,089	0,661
hu AS		K L G S L	S E T P P			
hu Exon 8	tgt	gttcc ag AAACTTGGGTCCCTG	CCGAGACGCCCCAG gt gagttctg		0,112	0,659
mu Exon 8	tgt	tcttcc ag AAACTTGGGTCTCTG	CTGAGACTCCCCAG gt gagtacc		0,112	n.d.
hu AS		A P S S R	N L S L T			
hu Exon 9	ct	acttgc ag CCCCATCTCCCGGG	AATCTGTCCCTCAC gt aggtgtac		0,266	0,152
mu Exon 9	CTCCACCCTCTCCGGG	AACCTGTCCCTCACT gt aagtacc		0,272	n.d.
hu AS		R E Q M F	M A G S R			
hu Exon 10	ccc	gctgt ag AGAGAGCAGATGTTT	TGGCCGGCTCCCGAG gt atgttcat		0,100	0,920
mu Exon 10	AGAGAACAGATGTTT	TGGCTGGCTCTCGAG gt acggctctg		0,100	0,102
hu AS		E N P C Q				
hu Exon 11	ct	ctctatgc ag AGAACCCGTGCCAGG			0,783	
mu Exon 11	ct	cctatgc ag AGAACCCATGCCCGG			0,728	

Tab. 3.7: Vergleich der Exon/Intron-Struktur des humanen und murinen *Gen 43 (WHSC2)*. Neben der Anzahl und Größe der jeweiligen humanen (hu) und murinen (mu) Exons und Introns sind die Nukleotidsequenzen der Exon-Intron-Grenzen dargestellt. Intronsequenzen sind durch kursive Kleinbuchstaben, Spleißdonor- und Akzeptorstellen durch Fettdruck gekennzeichnet. N.d.: nicht bestimmt.

Expressionsanalyse

Expressionsanalysen des humanen und murinen *Gen 43* (*WHSC2*) mit kommerziell erhältlichen Northern-Blots mit poly(A)⁺-RNA aus mehreren humanen und murinen fetalen und adulten Geweben entsprechen den Ergebnissen von Wright et al. (1999b). Das *Gen 43* (*WHSC2*) des Menschen und der Maus zeigt eine ubiquitäre Expression in allen untersuchten fetalen und adulten Geweben. Die Transkriptgröße beider Gene beträgt ca. 2,4 Kb. In **Abbildung 3.25** ist das Autoradiogramm eines fetalen murinen Northern-Blots dargestellt, der mit einer Hybridisierungssonde (m43-3/m43-4) aus der murinen cDNA (ca. Exon 1-3) hybridisiert wurde. Nach ca. 12 stündiger Exposition konnten starke Hybridisierungssignale von 2,4 Kb und schwache Signale von 4,4 Kb an RNA der fetalen Stadien Tag 7 bis 17 post coitus nachgewiesen werden.

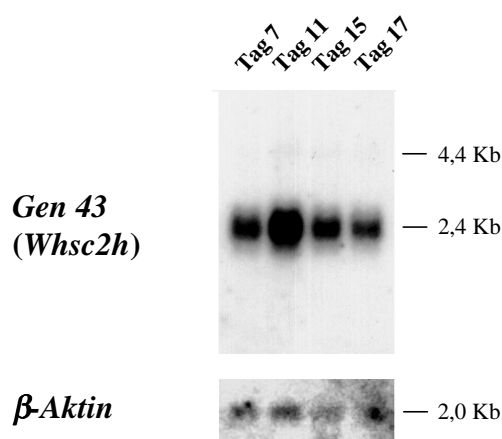


Abb. 3.25: Northern-Blot-Analyse des murinen *Gen 43* (*Whsc2h*). Zeigt die Autoradiogramme eines fetalen murinen Northern-Blots, der mit einer Sonde aus dem murinen *Gen 43* und einer Sonde aus dem *β-Aktin*-Gen hybridisiert wurde.

Eines der ermittelten Hybridisierungssignale von 2,4 Kb des Menschen und der Maus entspricht in etwa den Größen der in dieser Arbeit dargestellten cDNA-Sequenzen von ca. 2,2 Kb und damit der scheinbar nahezu vollständigen Klonierung dieses Transkripts. Das zusätzliche Signal (vgl. **Abb. 3.25**) wurde auch von Wright et al. (1999b) beschrieben und könnte auf mögliche alternative Transkripte desselben Gens (vgl. **Tab. 3.6**: N23294) oder auch auf Kreuzhybridisierungen zurückzuführen sein.

3.4.2.3 Das Gen 45

cDNA-Sequenzen des Menschen und der Maus

Aus der Sequenz des Cosmids 27h9 konnten mehrere überlappende humane EST-cDNA-Klone identifiziert werden, die Bestandteil des UniGene-Clusters Hs.152932 sind. Die meisten dieser cDNA-Klone stammen aus Hoden- (Pawlak et al., 1998) und Hirn-spezifischen sowie aus einigen Tumor-spezifischen cDNA-Bibliotheken, wie zum Beispiel Darmkarzinom- oder Keimzelltumor-Bibliotheken des NCI-CGAPs („Cancer Genome Anatomy Project“). Alle Klone weisen nahezu identische 5'-Enden auf (vgl. **Abb. 3.26**). Die ca. 0,42 Kb lange cDNA-Konsensussequenz der verfügbaren Nukleotidsequenzen dieser EST-Klone wurde durch RT-PCR an fetalem Hoden- und adultem Hirngewebe unter Verwendung der Primer 45-1, 45-2, 45-3 und 45-4 überprüft. Die cDNA-Sequenz konnte weiterhin durch RT-PCR an Gesamt-RNA aus humanen Hirngewebe unter Verwendung der cDNA-Primer 45-4 und 45-1 sowie mittels der aus der genomischen Sequenz des Cosmids 27h9 abgeleiteten Primer h131C sowie 45-6 um insgesamt ca. 314 Bp im 5'-Bereich verlängert werden. Die Häufigkeit und die Stärke der Expression dieses 5'-Bereichs ist unklar und steht im Gegensatz zu den einheitlich endenden cDNA-Klonen des UniGene-Clusters Hs.152932 (vgl. **Abb. 3.26**). Möglicherweise existieren alternative 5'-Enden dieses Gens.

Die cDNA-Konsensussequenz besitzt einen hohen GC-Gehalt von ca. 72 % und liegt innerhalb einer möglichen CpG-Insel (Bp 8500 bis Bp 10880 der Nukleotidsequenz des Cosmids 27h9). Sie enthält einen durchgehenden offenen Leserahmen, der durch Stopcodons begrenzt wird. Innerhalb des offenen Leserahmens befindet sich eine ununterbrochene Folge des Trinukleotids (CTG)₇. Die cDNA-Sequenz enthält zwei mögliche Translationsstartpunkte. Der erste befindet sich innerhalb des Exons 1 (von Bp 274-Bp 283) und entspricht nicht vollständig der Konsensussequenz nach Kozak (1996) (vgl. **Abb. 3.26**). Während die Position +4 durch den konservierten Purinrest (G) eingenommen wird, enthält die Position -3 einen Pyrimidinrest (C) anstelle eines konservierten Purinrestes. Die von diesem möglichen Translationsstart abgeleitete Aminosäuresequenz besteht aus 139 Resten. Der zweite mögliche Translationsstartpunkt befindet sich in Exon 2 (von Bp 406-Bp 414) und enthält an beiden Positionen (+4 und -3) die konservierten Purinreste (G) und erfüllt somit weitgehend die Konsensussequenz. Die von diesem möglichen Translationsstart abgeleitete Proteinsequenz besteht aus 95 Aminosäureresten.

Ein cDNA-Klon (AI341602) einer gepoolten Keimzelltumorbibliothek (Strausberg, „Cancer Genome Anatomy Project“ NCI-CGAP, 1999, unveröffentlicht) enthält im 5'-Bereich der cDNA-Sequenz (Exon 1) eine zusätzliche Nukleotidsequenz von 107 Bp. Sie wird vermutlich durch die Verwendung einer alternativen Spleißdonorstelle innerhalb des Introns 1 hervorgerufen. Dieses Transkript konnte durch RT-PCR an RNA aus adultem humanen Hirngewebe als schwaches zusätzliches Amplifikat verifiziert werden. Durch diese zusätzliche Nukleotidsequenz (107 Bp) wird über eine Leserasterverschiebung im 5'-Bereich der cDNA ein mögliches Stopcodon eingefügt. Die innerhalb des nun kürzeren offenen Leserahmens abgeleitete Proteinsequenz besteht aus 95 Aminosäureresten. Der Leserahmen ändert sich aber nicht (vgl. **Abb. 3.26**).

Zu dem *Gen 45* konnten korrespondierende cDNAs der Maus identifiziert werden (**Tab. 3.6**). Die inzwischen über 32 verfügbaren murinen EST-cDNA-Klone (Mm.4923) mit signifikanten Homologien zu dem Cosmid 27h9 stammen ebenfalls weitgehend aus Hoden- und Hirnspezifischen cDNA-Bibliotheken sowie einigen embryonalen cDNA-Bibliotheken. Die 5'- und 3'-Enden dieser cDNA-Klone sind bis auf wenige Bp identisch (**Abb. 3.26**). Daraus kann geschlossen werden, daß die cDNAs vollständig sind. Die Nukleotidsequenz (0,45 Kb) wurde anhand des cDNA-Klons W91513 (IMAGp998A03985) ermittelt und durch RT-PCR mit Gesamt-RNA aus verschiedenen Geweben embryonaler Mäuse (Tag 14,5) sowie adultem Hoden- und Hirngewebe bestätigt. Die cDNA-Sequenz ist ebenfalls GC-reich (GC-Gehalt von 66,7 %) und enthält eine ununterbrochene Folge des Trinukleotids (CTG)₅ (**Abb. 3.26**). Sie besitzt einen durchgehenden offenen Leserahmen, der sowohl im 3'- als auch im 5'-Bereich durch ein Stopcodon begrenzt wird. Darüber hinaus enthält sie einen möglichen Translationsstart (in Exon 2, Bp 37-Bp 46), der dem innerhalb des Exons 2 des Menschen (Bp 406-Bp 414) entspricht. Die abgeleitete Aminosäuresequenz besteht aus 90 Aminosäureresten.

Unter Verwendung der Primer m45-3 und m45-4 konnte in allen der untersuchten Geweben neben der zu erwartenden Produktgröße von 397 Bp ein weiteres Produkt von 257 Bp erhalten werden. Wie eine Sequenzanalyse zeigte, entspricht dieses Produkt den korrekt gespleißten Exons 1, 3 und 4 des aus mindestens 4 Exons bestehenden murinen *Gens 45*. Damit fehlt das den putativen Translationsstart enthaltende Exon 2. Zudem wird ein mögliches Stopcodon eingefügt, das den offenen Leserahmen im 5'-Bereich begrenzt. Das nächste mögliche Startcodon liegt jedoch innerhalb einer Sequenz, die nicht der Konsensussequenz von Translationsstartpunkten nach Kozak (1996) entspricht (vgl. **Abb. 3.26**). Daher ist unklar, ob diese Spleißvariante für ein funktionstüchtiges Protein (39 AS) kodiert.

Homologievergleiche der beiden Gene

Die Sequenzähnlichkeit der Nukleotidsequenz des murinen und humanen Gens beschränkt sich hauptsächlich auf die mögliche kodierende Sequenz, die man bei beiden Genen findet und beträgt 73 % (vgl. **Abb. 3.26**). Die abgeleitete Proteinsequenz besteht aus 95 (~ 10,2 kDa) bzw. 90 Aminosäureresten (~ 9,8 kDa) und ist zu 83,3 % identisch. Die Proteinsequenzen weisen unter Verwendung der Programme „BLASTN“ und „PROSITE“ keine Homologien zu bekannten Proteinen oder funktionellen Proteindomänen auf. Durch das Programm „PSORT“ (Nakai und Kanehisa, 1992) wurde bei beiden putativen Proteinen (die zunächst als Protein 45 der Maus bzw. 45 (1b) des Menschen bezeichnet wurden) ein mögliches Signalpeptid für eine mutmaßlich extrazelluläre Lokalisation der vermutlich löslichen Proteine identifiziert. Für das putative 139 Aminosäurenprotein des Menschen (45 (1a); ~ 14,7 kDa) konnte hingegen eine mögliche Transmembrandomäne (AS 60-76) mit einer für Typ II-Proteine charakteristischen Orientierung identifiziert werden. Durch das Programm „BLOCKS“ (Henikoff und Henikoff, 1991) konnten Übereinstimmungen zu verschiedenen Proteinen, wie z. B. zu Ly6-uPAR-Proteinen (BLOCKS-Eintragung BL00983) und der Proteinfamilie der Nervenwachstumsfaktoren (BL00248) gefunden werden, wie zum Beispiel dem Sca-2-Protein der Maus (Classon und Coverdale, 1994), dem uPAR (Urokinase Rezeptor) des Menschen und der Maus (Roldan et al., 1990; Kristensen et al., 1991) und dem NGF- β -Protein des Menschen (Ullrich et al., 1983). Die Homologien sind bei allen diesen Proteinen auf den Bereich der möglichen Signalpeptide (BL00983A und BL00248A) beschränkt (vgl. **Abb. 3.26**).

Die Proteine enthalten darüber hinaus zwei mögliche dibasische Schnittstellen (RK und RR), wie man sie bei Proteinen findet, die durch Prohormon-Konvertasen geschnitten werden (Seidah und Chrétien, 1997; Seidah und Chrétien, 1999). Die Bereiche mit einer möglichen Signalpeptidfunktion sowie weitere strukturelle Merkmale, die in den putativen Proteinen ermittelt werden konnten, sind in **Abbildung 3.26** dargestellt.

Genomische Struktur der beiden Gene

Die Exon-Intron-Struktur des humanen *Genes 45* wurde aus einem Vergleich der cDNA mit der genomischen Sequenz des Cosmids 27h9 abgeleitet und mittels PCR mit flankierenden Intron-Primern (45A-F/R bis 45D-F/R) und Sequenzierung überprüft. Das Gen besteht aus 4 Exons und erstreckt sich über einen Bereich von 2,2 Kb. Der 5'-Bereich der ermittelten cDNA-Sequenz des humanen *Genes 45* überlappt interessanterweise mit der cDNA-Sequenz eines durch Sequenzierung überprüften EST-cDNA-Klons AA131767 (IMAGp998J071202). Dieser cDNA-Klon wird inzwischen durch zwei weitere EST-cDNA-Klone aus unterschiedlichen cDNA-Bibliotheken vertreten und ist aufgrund der abgeleiteten Exon-Intron-Struktur auf dem Gegenstrang des *Genes 45* lokalisiert. Es handelt sich hierbei somit um ein mögliches unabhängiges Gen (vgl. **Tab. 3.6**).

Die Exon-Intron-Struktur des murinen *Genes 45* wurde über PCR an genomischer, muriner DNA und nachfolgender Sequenzierung der PCR-Produkte bestimmt. Es erstreckt sich über ca. 1,5 Kb und besteht ebenfalls aus 4 Exons und 3 Introns. Die Lage der Spleißstellen ist ebenfalls zwischen beiden Spezies konserviert (**Tab. 3.8**).

Exon Nr.	Intron (3')	Exon (5')	Exon (3')	Intron (5')	Exon Größe (Kb)	Intron
hu AS			(V P E P R)			
hu Exon 1a			GGTCCCTGAACCGCG g taagggcgg		≥ 0,386	0,258
hu Exon 1b			GTCGGGGAACCGAG g tggggaggg		≥ 0,493	0,151
mu Exon 1			GAGCTCGCGGACAC g tgagtaagg		≥ 0,110	0,117
hu AS		(A A P L P)	A V P A Q			
hu Exon 2	ctggcttc ag GGCGGCCCGCTCCC		CCGTCCCCGCGCAG g taggtgccg		0,157	0,251
mu Exon 2	ctcgtgc ag GGCGCCTGGGCTCC		CTGTCCCCGCTCAG g taggtgcc		0,140	0,155
hu AS		S R P C V	S L D A R			
hu Exon 3	cccccc g agGCCGCCATGCGTGG		AGCCTGGACGCGCG g tgagcgccc		0,089	1,036
mu Exon 3	gtcccc g agGCCGCCGTGCGTGG		AGCCTGGACGCACG g tgagtatac		0,089	0,755
hu AS		T E T L L				
hu Exon 4	tgcttt g agACGGAGACCCCTACTG				0,103	
mu Exon 4	ttcttt g agACGGAGACCCCTCCTG				0,111	

Tab. 3.8: Vergleich der Exon/Intron-Struktur des humanen und murinen *Genes 45*. Neben der Anzahl und Größe der jeweiligen humanen (hu) und murinen (mu) Exons und Introns sind die Nukleotidsequenzen der Exon-Intron-Grenzen dargestellt. Intronsequenzen sind durch kursive Kleinbuchstaben, Spleißdonor- und Akzeptorstellen durch Fettdruck gekennzeichnet. Die abgeleitete Aminosäuresequenz des möglichen, größeren humanen Proteins hu 45 (1a) ist in Klammern angegeben.

Expressionsanalyse

Für den Nachweis der Expression wurden zunächst Northern-Blot- und RNA-Master-Blot-Analysen an der Maus vorgenommen. Die Hybridisierung eines adulten, murinen Northern-Blots mit einer strangspezifischen Hybridisierungssonde aus dem 5'-Bereich der murinen cDNA-Sequenz (lineare PCR des PCR-Produkts der Primer T7 und m45-2 (Exon 1-3) am cDNA-Klon W91513/IMAGp998A03985) ergab starke Hybridisierungssignale an poly(A)⁺-RNA aus adultem Hoden- und Hirngewebe mit einer Transkriptgröße von ca. 0,7 Kb (vgl. **Abb. 3.27A**). Die Transkriptgröße entspricht somit der weitgehend vollständigen klonierten murinen cDNA von ca. 0,45 Kb.

Die nachfolgende Hybridisierung eines „Mouse RNA Master Blots“ der Firma Clontech mit derselben Hybridisierungssonde ergab nach zweitägiger Expositionszeit starke Hybridisierungssignale an RNA fetaler Stadien der Tage 11 und 15 (post coitus) sowie an adultem Hodengewebe. Weitere, deutliche Signale konnten an RNA des fetalen Stadiums Tag 17 (post coitus) und RNA aus adultem Gehirn, Ovar, Prostata, Uterus, Lunge und Auge nachgewiesen werden (**Abb. 3.27B**). Nach einwöchiger Exposition (in **Abb. 3.27B** dargestellt) waren weitere, schwache Signale bei Schilddrüse, Thymus, Speicheldrüse, Milz und dem fetalen Stadium Tag 7 (post coitus) erkennbar.

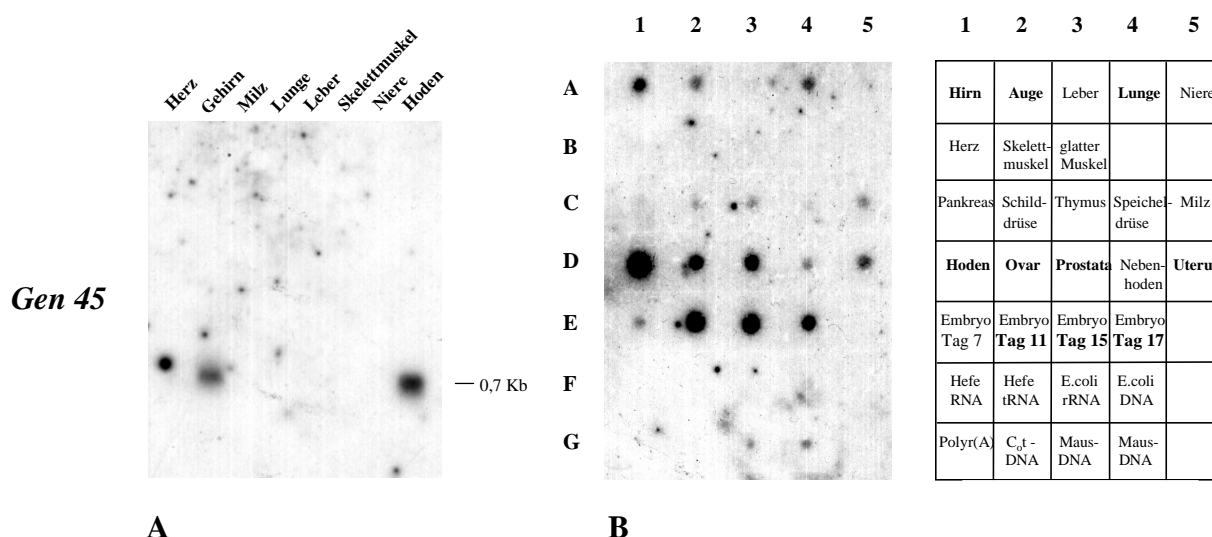


Abb. 3.27: Northern-Blot- (A) und Master-Blot- (B) Analyse des murinen *Gen 45*. (Siehe Text).

Entsprechende Northern-Blot-Analysen mit Hybridisierungssonden aus dem humanen *Gen 45* erbrachten jedoch keine eindeutigen Informationen über die Transkriptgröße dieses Gens. Bei Verwendung eines Northern Blots mit poly(A)⁺-RNA aus adulten humanen Geweben konnten mehrere Hybridisierungssignale von 5,2; 3,8; 2,9; 2,2 und ca 1,0 Kb (unter Verwendung der über lineare PCR an den PCR-Produkten 45-3/45-1 und 45-2/45-4 generierten und somit einzelsträngigen Hybridisierungssonden 45-1 (Exon 1-3) und 45-4 (Exon 2 bis 4)) in allen untersuchten Geweben nachgewiesen werden (vgl. **Abb. 3.28**). Während die Transkriptgröße von ca. 1,0 Kb in etwa der klonierten cDNA entspricht, könnten die zusätzlich nachgewiesenen Signale entweder auf alternative Transkripte desselben Gens oder auf Kreuzhybridisierungen zurückzuführen sein, wobei die Größe der Hybridisierungssignale den einheitlichen EST-Sequenzdaten des Menschen mit einer cDNA-Länge von 0,42 Kb und den nachgewiesenen Transkripten von bis zu 0,74 Kb widerspricht. Die weiteren Expressionsuntersuchungen konzentrierten sich daher auf die Maus.

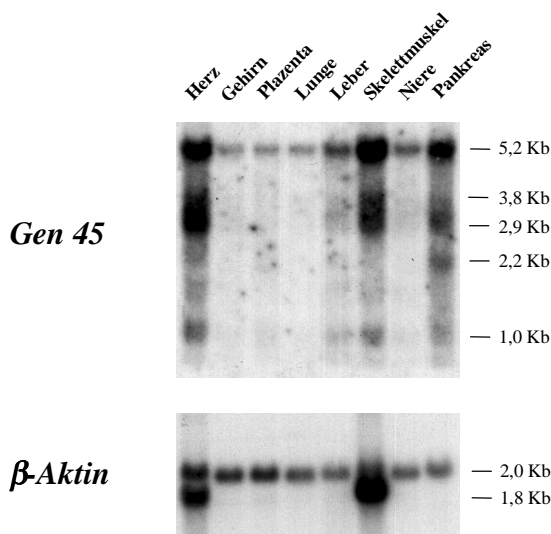


Abb. 3.28: Northern-Blot-Analyse des humanen *Gen 45*. Dargestellt ist das Autoradiogramm eines adulten humanen „Multiple Tissue Northern Blots“, der mit einer Sonde aus dem Bereich der Exons 2-4 (45-4) des humanen *Gen 45* hybridisiert wurde. Zum Vergleich der Signalstärken ist die Autoradiographie desselben Northern-Blots dargestellt, der mit einer Sonde aus dem β -Aktin-Gen hybridisiert wurde.

Für die weitere Untersuchung dieser gewebespezifischen Expression wurden drei verschiedene, 40 Bp lange „antisense“- und „sense“-Oligonukleotide aus unterschiedlichen Bereichen der murinen cDNA abgeleitet (Exon 2: m45O1as und m45O1s; Exon 4: m45O2as, m45O2s, Exon 3/4: m45O3as und m45O3s) und für die Hybridisierung von Gewebeschnitten aus adultem Hoden- und Gehirngewebe verwendet.

Spezifische perinukleäre Hybridisierungssignale konnten nur mit der aus dem Exon 2 abgeleiteten Sonde m45O1as nachgewiesen werden. **Abbildung 3.29B** zeigt einen sagittalen Gehirnschnitt vergleichend zu der etwa in der gleichen Schnittebene durchgeführten Hämatoxylin-Eosin-Färbung (**Abb. 3.29A**) und der entsprechenden „sense“-Hybridisierung zweier weiterer Schnitte (**Abb. 3.29C**). Es hybridisierten Nervenzellen in allen corticalen und subcorticalen Bereichen des Gehirns (**Abb. 3.29B**). Eine spezifische, perinukleäre Färbung findet sich auch innerhalb der Körnerzellen im Endbereich des Hippocampus. Im Bereich des Kleinhirns konnten Signale innerhalb einzelner Körnerzellen in der Körnerschicht und in Zellen der Molekularschicht nachgewiesen werden, wobei es sich hierbei wahrscheinlich um Interneurone handelt. Die Gliazellen und Purkinjezellen des Kleinhirns hingegen zeigen keine Signale.

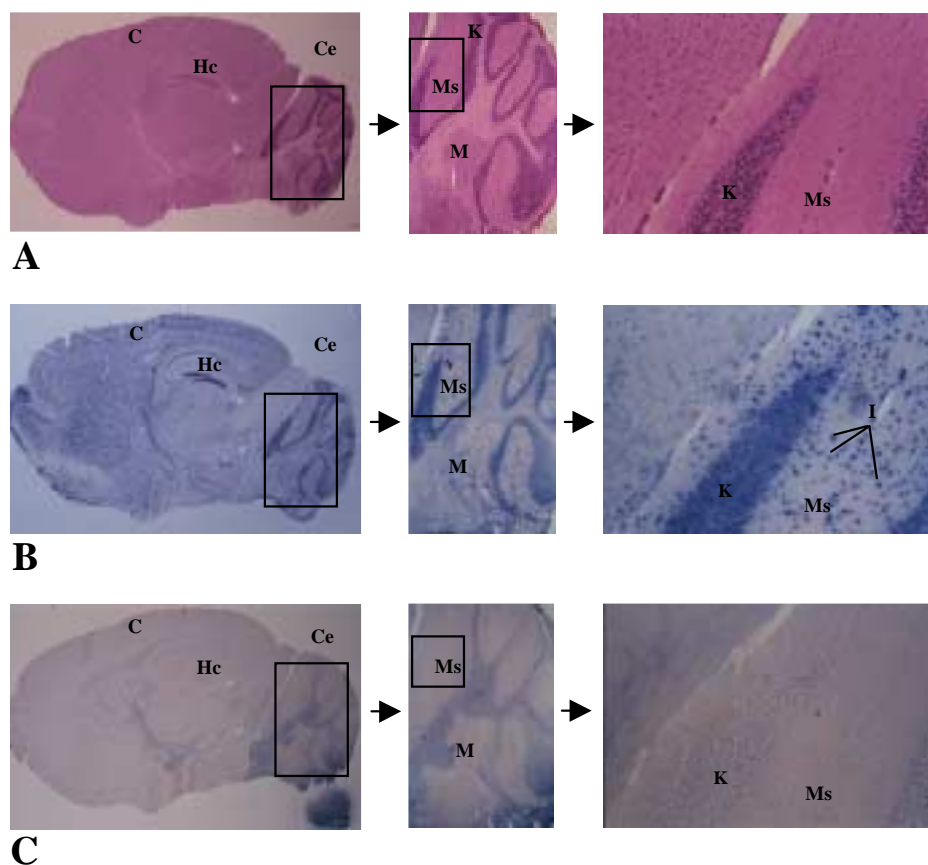


Abb. 3.29: RNA-*in situ*-Hybridisierung an Gehirnschnitten einer adulten Maus. **A:** Zeigt die Hämatoxylin-Eosin-Färbung eines Sagittalschnittes durch ein adultes Mäusegehirn und nachfolgend Vergrößerungen aus dem Bereich des Kleinhirns. C: Cortex, Ce: Cerebellum, Hc: Hippocampus, K: Körnerschicht, Ms: Molekularschicht, M: Mark. **B:** Sagittalschnitt nach Hybridisierung mit der Sonde m45O1as. Die Hybridisierungssignale innerhalb von Nervenzellen betreffen alle Bereiche des Gehirns und sind innerhalb des Kleinhirns (Ce) in der Molekularschicht (Ms) zu finden. Angefärbte Interneurone sind gekennzeichnet (I). Innerhalb des Markbereiches (M) treten keine Signale auf. **C:** Hybridisierung mit der „sense“-Sonde m45O1s. Es sind hier keine spezifischen Hybridisierungssignale erkennbar.

Die Hybridisierung von Gewebeschnitten aus adultem Hodengewebe unter Verwendung derselben Hybridisierungssonde (m45O1as) ergab ausschließlich starke Signale in den lumen-nahen Bereichen einzelner, vermutlich aktiver Hodenkanälchen. Die Hybridisierungssignale betreffen hierbei ausschließlich die Spermatozoen, so daß eine Rolle dieses Gens in der Spermiogenese vermutet werden kann. In **Abbildung 3.30** sind beispielhaft Aufnahmen einzelner transversaler Gewebeschnitte durch adultes Hodengewebe dargestellt.

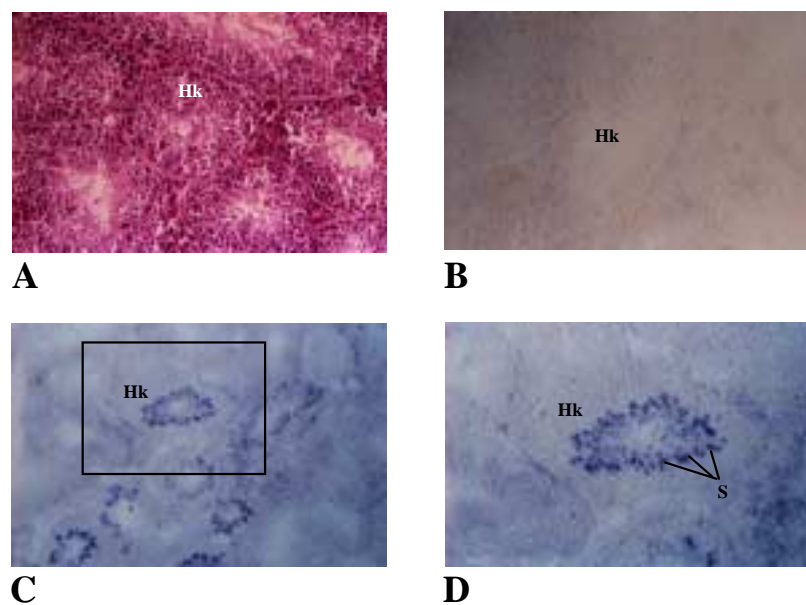


Abb. 3.30: RNA-*in situ*-Hybridisierung an transversalen Gewebeschnitten durch adultes Hodengewebe der Maus. Zur Orientierung ist ein Ausschnitt aus dem Bereich einzelner, transversal angeschnittener Hodenkanälchen (Hk) eines Hämatoxylin-Eosin-gefärbten Schnittes dargestellt (A). Während die Hybridisierung mit der „sense“-Sonde m45O1s keine Hybridisierungssignale zeigt (B), sind nach der Hybridisierung mit der „anti-sense“-Sonde m45O1as spezifische Hybridisierungssignale in den lumen-nahen Bereichen einzelner Hodenkanälchen zu finden (C) und betreffen hierbei vermutlich die Spermatozoen (S), die in **Abb. D** speziell gekennzeichnet sind.

3.4.2.4 Das Gen 57

cDNA Sequenzen der Maus und Homologievergleiche

Mit Hilfe von Sequenzabschnitten aus dem Cosmid 58b6 konnte ein muriner cDNA-Klon (AA118417) identifiziert werden, der eine über 80 %ige Sequenzidentität zu der genomischen Sequenz des Cosmids 58b6 (ca. Bp 25234 bis Bp 25664) besitzt (vgl. **Tab. 3.6**). Die Sequenzhomologien beziehen sich dabei auf eine Region des Cosmids 58b6, die im Intronbereich eines weiteren, unter **3.4.2.5** dargestellten DNA-Polymerase-ähnlichen Gens liegt und zunächst ein weiteres und unabhängiges humanes Gen in dieser Region vermuten lassen.

Die Nukleotidsequenz (0,67 Kb) des murinen cDNA-Klons enthält einen offenen Leserahmen, der im 3'-Bereich durch ein putatives Stopcodon begrenzt wird (vgl. **Abb. 3.31**) und mit einem weiteren EST-cDNA-Klon (AV137499) überlappt. Der putative 3'-untranslatierte Bereich des cDNA-Klons enthält ein mögliches Polyadenylierungssignal und im Abstand von 12 Bp eine Folge von Adeninresten. Das putative korrespondierende humane Gen wurde *Gen 57* genannt. Humane cDNA-Klone zum *Gen 57* (mit 100 %iger Homologie zu dem entsprechenden Bereich des Ausgangscosmids) konnten jedoch nicht identifiziert werden.

BLASTN-Analysen zeigten, das es sich bei der isolierten murinen cDNA wahrscheinlich um das murine *erfs*-Gen („eucaryotic release factor similar gene“; Wallrapp et al., 1998) handelt, das für ein putatives GTP-bindendes Protein von 600 AS kodiert. Die Sequenzidentität (100 % auf Nukleotid- und 100 % auf Proteinebene) beschränkt sich hierbei auf Bp 245 bis Bp 551 des murinen cDNA-Klons AA118417 und schließt damit den möglichen untranslatierten Bereich aus (**Abb. 3.31**). Die cDNA zeigt erwartungsgemäß (Bp 20 bis Bp 551) signifikante Homologien (82 %ige Sequenzidentität auf Nukleinsäureebene und 88 %ige Homologie auf Proteinebene) zu dem 5'-Bereich des orthologen humanen *eRFS*-Gen (Wallrapp et al., 1998). Dieses Gen ist jedoch in der Chromosomenregion 6q24 lokalisiert und kodiert für ein putatives, im Vergleich zu dem murinen *erfs* um 84 Aminosäuren längeres Protein (684 AS). Da die Nukleotidsequenz auf 58b6 keinen offenen Leserahmen zeigt kann davon ausgegangen werden, daß es sich beim humanen *Gen 57* um ein Pseudogen des *eRFS*-Gens handelt, das bei der Maus nicht vorhanden ist. Die isolierte murine cDNA entspricht daher möglicherweise dem murinen *erfs*-Gen (**Abb. 3.31**).

Expressionsanalyse

Northern-Blot-Analyse mit einer aus dem 5'-Bereich des murinen cDNA-Klons AA118417 abgeleiteten Hybridisierungssonde (m57-1/m57-2) ergab ein Hybridisierungssignal von ca. 3,5 Kb an RNA der fetalen Stadien Tag 7 bis 17 (post coitus) der Maus (**Abb. 3.32**). Dieses Signal entspricht in etwa der Transkriptgröße des murinen *erfs*-Gens nach Wallrapp et al. (1998).

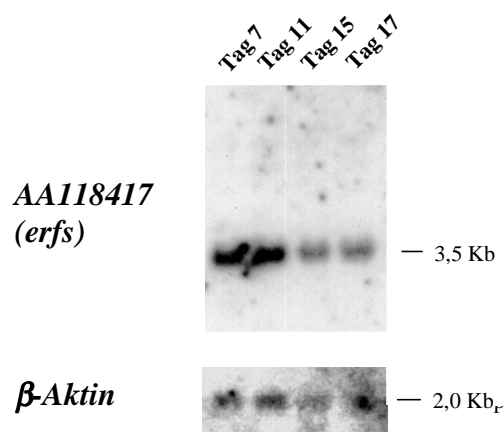


Abb. 3.32: Northern-Blot-Analyse des cDNA-Klons AA118417. Zeigt das Autoradiogramm eines fetalen murinen Northern-Blots mit poly (A)⁺-RNA aus den fetalen Stadien Tag 7 bis 17 (post coitus), der mit einer Sonde aus dem 5'-Bereich der murinen cDNA hybridisiert wurde.

3.4.2.5 Das *POLAP*-Gen

cDNA-Sequenzen des Menschen und der Maus und Homologievergleiche beider Gene

Aus dem Cosmid 58b6 konnte eine partielle cDNA-Sequenz (AF044578) eines DNA-Polymerase-ähnlichen Gens (*POLAP*) identifiziert werden (Harris et al., 1999; unveröffentlicht) (vgl. **Tab. 3.6**). Dieses Gen kann aufgrund von Homologievergleichen der Familie A der DNA-Polymerasen zugeordnet werden (vgl. **Abb. 3.33**), die sich durch Sequenzhomologien zu dem *polA*-Genprodukt aus *E. coli* auszeichnet (Braithwaite und Ito, 1993). Die cDNA-Sequenz dieses bisher aus 19 Exons bestehenden Gens erstreckt sich genomisch über einen Bereich von mehr als 110 Kb zwischen den Cosmiden 108f12 und 174g8 und überspannt somit den proximalen Bruchpunkt der 165 Kb-WHSCR (vgl. **Abb. 3.21**). Diesem Gen konnten zwei humane EST-cDNA-Klone (AA436996 und AI674093) zugeordnet werden, die aus einer Hoden- und einer Subtraktions-cDNA-Bibliothek aus mehreren Geweben stammen (vgl. **Tab. 3.6**). Beide cDNAs weisen eine 100 %ige Sequenzidentität zu den Exons 15 (AA436996) und 16 (AI674093) sowie zu Intronbereichen des *POLAP*-Gens nach Harris et al. (1999) auf (**Tab. 3.6**) und könnten möglicherweise auf alternative Spleißvorgänge zurückzuführen sein.

Zu dem humanen Gen konnten zwei murine EST-cDNA-Klone (AA475002 und AI120069) identifiziert werden, die aus einer Brustdrüsen-spezifischen cDNA-Bibliothek stammen und vermutlich das orthologe Gen der Maus (*Pol4p*) repräsentieren. Sie weisen neben einer 87-94 %igen Sequenzidentität zu verschiedenen Bereichen der Cosmide 58b6, 141a8 und 174g8 Übereinstimmungen zu der *POLAP*-cDNA-Sequenz nach Harris et al. (1999) auf (**Tab. 3.6** und **Abb. 3.21**). Während die cDNA-Sequenz des Klons AA475002 von Bp 61 bis Bp 459 eine durchgehende Sequenzähnlichkeit von 87 % zu den Exonsequenzen 9-13 des *POLAP*-Gens und eine 93 %ige Homologie zu der entsprechenden Proteinsequenz, die das Motiv A der DNA-Polymerasen beinhaltet, aufweist (vgl. **Abb. 3.34**), entspricht die Nukleotidsequenz des cDNA-Klons AI120069 von Bp 119 bis 183 zu 93 % dem Exon 9 und die Sequenz von Bp 185 bis 404 zu ca. 85 % den Exons 13 und 14 der *POLAP*-cDNA und scheint damit ebenfalls auf alternative Spleißvorgänge dieses DNA-Polymerase ähnlichen Gens der Maus hinzuweisen. Die hohen Übereinstimmungen zu der cDNA-Sequenz bzw. der Proteinsequenz des humanen *POLAP*-Gens nach Harris et al. (1999) deuten darauf hin, daß es sich hierbei wahrscheinlich tatsächlich um Transkripte des orthologen Gens der Maus handelt.

```

sp|P00582|DPO1_ECOLI DNA POLYMERASE I (POL I) >gi|67054|pir||DJECI DNA-directed DNA
polymerase (EC 2.7.7.7) I - Escherichia coli
Length = 928

Score = 237 bits (599), Expect = 1e-61
Identities = 148/432 (34%), Positives = 239/432 (55%), Gaps = 29/432 (6%)

      ↓ Exon 3                ↓ Exon 4                ↓ Exon 5
Query: 59 KDYGLWQLFRTLELPLIPIILAVMESHAIQVNKEEMKTSALLGARLKELEQEAHFVAGER 118
      K G +F +E+PL+P+L+ +E +++++ + + S L RL ELE++AH +AGE
Sbjct: 518 KHKGPLNVFENIEMPLVPLSRIERNVGVKIDPKVLHNSSEELTLRLAELEKKAHEIAGEE 577

      ↓ Exon 6                ↓ Exon 7
Query: 119 FLITSNNQLREILFGKLLHLLSQNSLPRGTGLQKYPSTVSEALNALRDLHPLPKIILEY 178
      F ++S QL+ ILF K + L + PST E L L +PLPK+ILEY
Sbjct: 578 FNLSTKQLQTLFEKQGIKPLKKTTPG-----GAPSTSEEVLDELALDYPLPKVILEY 630

      ↓ Exon 8                ↓ Exon 9                ↓ Exon 10
Query: 179 RQVHKIKSTFVDGLLACM--KKSISSTWNQGTGTGRLSAKHPNIQGISKHPIQITTPK 236
      R + K+KST+ D L + K G + ++++Q T TGRLS+ PN+Q I P+ +
Sbjct: 631 RGLAKLKSTYTDKLPMLINPKTGRVHTSYHQAVTATGRLSSTDPNLQNI---PV-----R 682

      ↓ Exon 11                ↓ Exon 12
Query: 237 NFKGKEDKILITISPRAMFVSSKGHTFLAADFSQIELRILTHLSGDPELLKLFQESERDDV 296
      N +G+ R F++ + + ++AD+SQIELRI+ HLS D LL F E + D+
Sbjct: 683 NEEGRRI-----RQAFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTAFAGK--DI 733
      Motiv A

      ↓ Exon 13                ↓ Exon 14
Query: 297 FSTLTSQWKDVPVEQVTHADREQTKKVVYAVVYGAKKERLAACLGVPIQEAAQFLESFLQ 356
      ++ +P+E VT R K + + ++YG LA L +P +EA ++++ + +
Sbjct: 734 HRATAAEVFGFLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMPLYFE 793
      Motiv B

      ↓ Exon 15                ↓
Query: 357 KYKIKDFARAAIAQCHQTGCVVSI MGRRRPLPRIHAHDQQLRAQAERQAVNFVVQGSAA 416
      +Y + ++ AQ + G V ++ GRR LP I + + RA AER A+N +QG+AA
Sbjct: 794 RYPGVLEYMERTRAQAQKQGYVETLDGRRRLYLPDIKSSNGARRAAERAAINAPMQGTAA 853

      Exon 16                ↓ Exon 17                ↓ Exon 18 ↓
Query: 417 DLCKLAMIHVFTAVAASHTLTARLVAQIHDELLFEYEDPQIPECAALVRRRTMESLEQ--V 474
      D+ K AMI V + A R++ QHDEL+FEY + A + + ME+ + V
Sbjct: 854 DIIKRAMIAVDAWLQAEQP-RVRMIMQVHDELVFEYVHKDDVDVAQKQIHQLMENCTRLDV 912
      Motiv C

      Exon 19
Query: 475 PLKVSLSAGRSW 486
      PL V + +G +W
Sbjct: 913 PLLVEVGSGENW 924

```

Abb. 3.33: Proteinvergleich (BLASTX) der partiellen cDNA-Sequenz des humanen *POLAP*-Gens mit der DNA-Polymerase I aus *E. coli* (Joyce et al., 1982). Die Homologie erstreckt sich über den C-terminalen, katalytisch aktiven Bereich des Enzyms aus *E. coli* und beinhaltet die für die Familie der DNA-Polymerasen charakteristischen Motive A, B und C (eingerahmt) (Delarue et al., 1990). Das Motiv B entspricht der Signatur der Familie A der DNA-Polymerasen („PROSITE“-Eintragung PDOC00412: R-x(2)-[GSAV]-K-x(3)-[LIVMFY]-[AGQ]-x(2)-[GS]-x(3)-[LIVMA]). Die ungefähre Lage der Exon-Intron-Grenzen der korrespondierenden cDNA-Sequenz des humanen *POLAP*-Gens (Harris et al., 1999; unveröffentlicht) sind durch kleine Pfeile angedeutet.

```
gb|AAD02338.1| (AF044578) putative DNA polymerase; POL4P [Homo sapiens]
Length = 579
Score = 224 bits (565), Expect = 4e-58
Identities = 106/124 (85%), Positives = 117/124 (93%)
Frame = +1

Query: 61  GSISSTWNQGTGVTGRLSAKHPNIQGISKHPIKISKPWNFKGKEEETVTISPRTLFSSE 240
          GSISSTWNQGTGVTGRLSAKHPNIQGISKHPI+I+ P NFKGKE++ +TISPR +FVSS+
Sbjct: 189 GSISSTWNQGTGVTGRLSAKHPNIQGISKHPIQITTPKNFKGKEDKILTISPRAMFVSSK 248

Query: 241  GHTFLAADFSQIELRILAHLSGDPELLKLFQESERDDVFSTLTSQWKDPIPIERVTHMDRE 420
          GHTFLAADFSQIELRIL HLSGDPELLKLFQESERDDVFSTLTSQWKD+P+E+VTH DRE
Sbjct: 249  GHTFLAADFSQIELRILTHLSGDPELLKLFQESERDDVFSTLTSQWKDVPVEQVTHADRE 308

Query: 421  QPRK 432
          Q +K
Sbjct: 309  QTKK 312
```

Abb. 3.34: Homologievergleich (BLASTX) des murinen cDNA-Klons AA475002 mit der Proteinsequenz des humanen *POL4P*-Gens nach Harris et al. (1999).

3.4.2.6 Weitere mögliche Transkripte des Menschen und der Maus

Eine Reihe weiterer EST-cDNA-Klone des Menschen und der Maus mit Sequenzhomologien zu verschiedenen Regionen der Cosmide 58b6 bis 190b4 konnten bisher keinem der bereits identifizierten Gene zugeordnet werden (vgl. **Tab. 3.6** und **Abb. 3.21**). Die Zugehörigkeit einzelner dieser putativen Transkripte zu einem der in den vorangegangenen Kapiteln beschriebenen Gene bzw. zu möglichen weiteren, unabhängigen Genen kann jedoch nicht ausgeschlossen werden. Einige dieser, teilweise vollständig sequenzierten cDNA-Klone (**Tab. 3.6**) weisen aufgrund ihrer Sequenzhomologien zu den genomischen Sequenzen der entsprechenden Cosmide eine der GT-AG-Regel folgende Exon-Intron-Struktur auf (AA131767, R13046, H37943) und scheinen damit gespleißten Transkripten zu entsprechen. Andere cDNA-Klone lassen sich wahrscheinlich auf Fehler bei der Herstellung der entsprechenden cDNA-Bibliotheken zurückführen (T85656, AA155707, W62449) und stellen genomische Kontaminationen oder unvollständig gespleißte Produkte dar. Weitere cDNA-Klone weisen Homologien zu verschiedenen repetitiven Elementen, wie z. B. Alu-Elementen (R13046, AA155706, N73027, N53662, H37943 und W62449), LINE-Elementen (T85656) oder endogenen Retroviren (AA131767) auf. Dementsprechend erstreckt sich die Homologie bei einigen dieser Klone nur über relativ kurze Bereiche. Obwohl nicht ausgeschlossen werden kann, daß sich einzelne dieser Transkripte einem der bereits dargestellten Gene zuordnen lassen können oder neue Gene repräsentieren, wurde auf eine weitere Untersuchung im Rahmen dieser Arbeit verzichtet.

3.5 Untersuchung eines WHS-Patienten ohne cytogenetisch nachweisbare Deletion

Obwohl das WHS eine klinisch heterogene Erkrankung darstellt, konnten bisher nur einzelne Fälle ohne eine nachweisbare Deletion beschrieben werden (Stec et al., 1998; Marczinek et al., 1995). Allerdings ist bei diesen Patienten unklar, ob es sich um echte WHS-Fälle handelt. Trotzdem ist besonders hierbei die Untersuchung einzelner Gene von Bedeutung.

Untersucht wurde im Rahmen der Arbeit einer dieser Patienten (O. S.), der eine Vielzahl der für WHS-Patienten charakteristischen Merkmale aufweist, wie z. B. die typische kraniofaziale Ausprägung, mentale Retardierung, Wachstumsverzögerungen, Herzdefekte sowie eine Reihe weiterer dysmorpher Merkmale (Marczinek et al., 1995). Genomische DNA aus Blut dieses Patienten (O. S.) und dessen Eltern sowie immortalisierte, EBV-transformierte lymphoblastoide Zelllinien von Mutter und Kind wurden freundlicherweise von P. Nürnberg (Charité, Berlin) zur Verfügung gestellt. Durch die Untersuchung mehrerer polymorpher Marker (D4S125, D4S43, D4S96, D4S111, D4S227, D4S169) konnte bisher kein Verlust genetischen Materials in der distalen Chromosomenregion 4p16.3 festgestellt werden (Marczinek et al., 1995). Der Bereich zwischen D4S43 und D4S111, einschließlich der 165 Kb WHSCR wurde bisher allerdings noch nicht im Detail untersucht (Marczinek et al., 1995; P. Nürnberg, Charité, Berlin; persönliche Mitteilung). Da in diesem Bereich außer dem Mikrosatellitenmarker D4S114 bisher keine weiteren, bekannten polymorphen Marker zur Verfügung standen, wurden die genomischen Sequenzen der Cosmide des Baxendale-Contigs auf mögliche Dinukleotidwiederholungen untersucht. Fünf Dinukleotidsequenzwiederholungen mit einer Anzahl von 12 bis 27 (CA- bzw. GT-) Sequenzwiederholungen wurden identifiziert und entsprechende flankierende Primerpaare abgeleitet (27h9-A_R und B_R, 19h1-A_R und B_R, 75b9B-A_R und B_R, 75b9A-A_R und B_R und 58b6-A_R und B_R), die die Amplifizierung der jeweiligen Dinukleotidrepeats erlauben (58b6: Bp 16832-Bp 17002, (CA)₁₈; 27h9: 444-656, (TG)₂₀; 19h1: (GT)₂₇ 75b9A: 63019-6599, (TC)₂₆, (CA)₁₃; 75b9B: Bp 10003-10341, (TA)₁₈). Diese Primerpaare wurden zunächst unter Verwendung genomischer DNA aus Blut 5 verschiedener Kontrollpersonen auf den Grad an Polymorphie überprüft und im Anschluß für die Untersuchung der DNA des Patienten und dessen Eltern verwendet.

Darüber hinaus wurden in der vorliegenden Arbeit für die Identifizierung möglicher Mikrodeletionen die bereits bekannten polymorphen D-Marker D4S115 (Allitto et al., 1992) und D4S114 (Tagle et al., 1993) verwendet.

Mit Hilfe der Primer D4S115.PCR.2.1 und 2.2 konnten sowohl bei der Mutter (M) als auch beim Vater (V) jeweils zwei PCR-Produkte (M: ca. 610 und 915 Bp, V: ca. 730 und 915 Bp) dargestellt werden. Beim Patienten O. S. (K) konnten ebenfalls zwei Produkte (jeweils eines vom Vater und eines von der Mutter) von ca. 610 und 915 Bp identifiziert werden (vgl. **Abb. 3.35A**). Das durch PCR mit den Primern D4S114-A102.2F und A102.2R erhaltene, fluoreszenzmarkierte PCR-Produkt der Eltern und des Patienten wurde einer „GeneScan“-Analyse unterzogen. Während der Vater nur einen „Allel-Peak“ aufweist und damit homozygot für den in **Abb. 3.35B** gezeigten Marker ist, weist das Kind ein weiteres Allel der heterozygoten Mutter auf und ist damit für den Marker D4S114 heterozygot. Für die Untersuchung der Mikrosatellitenmarker 58b6-A_R/B_R, 27h9-A_R/B_R, 19h1-A_R/B_R, 75b9B-A_R/B_R, und 75b9A-A_R/B_R wurden die an genomischer DNA erhaltenen PCR-Produkte der Eltern und des Patienten radioaktiv markiert und auf denaturierenden Polyacrylamidgelen aufgetrennt. Die Marker 19h1-A_R/B_R, 75b9A-A_R/B_R, und 75b9B-A_R/B_R ergaben bei der Untersuchung der Eltern und des Patienten keine auswertbaren Banden und waren daher nicht informativ (nicht dargestellt). Für die Marker 58b6-A_R/B_R und 27h9-A_R/B_R konnten hingegen bei dem Patienten O.S. jeweils zwei Hauptbanden nachgewiesen werden, von denen jeweils eine eindeutig auch bei der Mutter bzw. beim Vater auftritt (**Abb. 3.35C und D**).

Die Ergebnisse aller untersuchten informativen Marker schließen eine Mikrodeletion in dem untersuchten Bereich weitgehend aus. Die in der vorliegenden Arbeit identifizierten Gene *LETM1*, *Gen 43* (*WHSC2*) und *Gen 45* (vgl. Kapitel 3.3 und 3.4) sowie das *FGFR3*-Gen wurden daraufhin auf mögliche Mutationen untersucht. Aus den Intronbereichen dieser Gene wurden Primerpaare abgeleitet (*LETM1*: LETM1-1F und 1R bis LETM1-14F und 14R, *Gen 43*: 43-1F und 1R bis 43-11F und 11R sowie *Gen 45*: 45A-F und A-R bis 45D-F und D-R), die die Amplifizierung der identifizierten Exons und der angrenzenden Intronbereiche erlauben. Zudem wurden Primerpaare aus den Intronbereichen des *FGFR3*-Gens verwendet (Wüchner et al., 1997). Zuerst wurde neben normalen Kontrollpersonen genomische DNA des Patienten O. S. und dessen Eltern mittels PCR und nachfolgender Sequenzierung der genomischen PCR-Produkte untersucht. Es konnten keine Mutationen identifiziert werden, die die Genfunktion eines dieser Gene beeinträchtigen könnte. Außerdem wurden die ubiquitär exprimierten *LETM1*- und *WHSC2*-Gene durch RT-PCR an Gesamt-RNA aus der

Zelllinie des Patienten O. S. und nachfolgender Sequenzierung der PCR-Produkte überprüft. Die Auswertung der Sequenzdaten ergab auch hier keine Mutation.

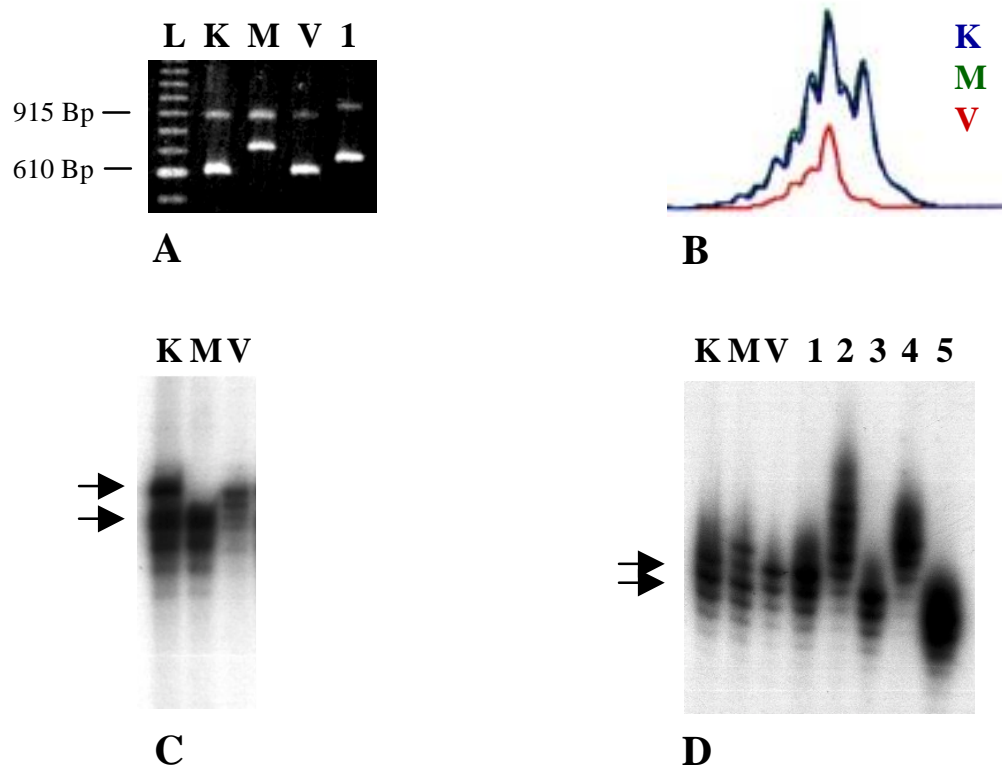


Abb. 3.35A-D: Mini- und Mikrosatellitenanalysen. Die Untersuchungen wurden an genomischer DNA aus Blut des Patienten O. S. (K: Kind) dessen Mutter (M) und Vater (V) sowie verschiedenen Kontrollpersonen (1-5) vorgenommen. Sämtliche Mikrosatelliten-Analysen wurden zum Ausschluß methodischer Fehler doppelt durchgeführt. **A:** Zeigt die gelelektrophoretische Auftrennung der mit Hilfe der Primer D4S115.PCR.2.1 und 2.2 erhaltenen PCR-Produkte. Die zwei informativen 610 und 915 Bp-Banden des Patienten (K) schließen eine mögliche Deletion aus. L: 100 Bp-„Leiter“. **B:** Dargestellt ist das Elektropherogramm der GeneScan-Analyse des Markers D4S114. Während der Vater homozygot für den gezeigten Marker ist (rot dargestellt), weist das heterozygote Kind (blau) ein weiteres Allel der Mutter (grün) auf. **C:** Zeigt die Autoradiographie der auf einem denaturierenden Polyacrylamidgel aufgetrennten radioaktiv markierten PCR-Produkte des Mikrosatellitenmarkers 58b6-A_R/B_R. Die Pfeile kennzeichnen die im Vergleich zu den Nebenbanden stärkeren Hauptbanden, die eindeutig bei dem untersuchten (heterozygoten) Patienten auftreten. **D:** Zeigt die Autoradiographie der Mikrosatelliten-Analyse des Markers 27h9-A_R/B_R des untersuchten Patienten im Vergleich zu verschiedenen Kontrollpersonen. Die beiden Pfeile kennzeichnen die beiden, bei dem Patienten auftretenden Hauptbanden (Allele).

3.6. Charakterisierung der WHSCR-orthologen Region der Maus

Der Genreichtum in der WHS-kritischen Region sowie die zumeist ubiquitäre Expression der identifizierten Gene (vgl. 3.4) erschweren wesentlich die Identifizierung und Charakterisierung aller Gene der WHSCR sowie die weitere Beurteilung dieser Gene hinsichtlich ihrer Bedeutung als Kandidatengene für die Entstehung dieses komplexen Syndroms. Vergleichende Untersuchungen an Mensch und Maus sollten die Identifizierung und Charakterisierung aller beteiligten Gene erleichtern und darüber hinaus die Entwicklung eines Mausmodells durch Einfügen entsprechender Deletionen in das Mausgenom erlauben. Eine genaue Kenntnis der entsprechenden Region im Mausgenom ist hierbei von größter Wichtigkeit. Innerhalb des humanen Chromosoms 4 und des murinen Chromosoms 5 liegen größere durchgehende Syntäniebereiche vor (Lyu und Kozak 1996; Kozak und Stephenson 1997; Kozak und Stephenson 1999). Dies betrifft vermutlich auch die WHS-kritische Region (WHCSR) zwischen dem Marker D4S43 und dem *FGFR3*-Gen, die bei ca. 20 cM auf dem murinen Chromosom 5 (Mmu 5.20) kartieren (Kozak und Stephenson, 1999; Avraham et al., 1994). In der vorliegenden Arbeit wurden zu den identifizierten Genen der Maus murine P1- und PAC-Klone isoliert und relativ zueinander angeordnet. Mittels verschiedenster PCR-Ansätze wurde die Konservierung der Syntänie der bereits identifizierten und scheinbar orthologen murinen Gene innerhalb der WHSCR untersucht.

Für die Isolierung von P1- und PAC-Klonen wurden kommerziell erhältliche Koloniefilter der murinen genomischen P1- und PAC-Bibliotheken 703 und 711 (RZPD, Berlin) verwendet. Als Hybridisierungs sonden dienten radioaktiv markierte RT-PCR- oder genomische PCR-Produkte aus den Exons 5 bis 6 (m16A1/m16A2) und Exon 4 (m16-21/m16-29) des *Letm1*-Gens, aus dem Exon 1 (m43-3/m43-12A) und den Exons 10-11 (m43-11/m43-E2) des *Whsc2h*-Gens (murines *Gen 43*) und Exon 2 bis 3 des murinen *Gens 45* (m45-1/m45-2). Positive Klone wurden nachfolgend durch Koloniefilterhybridisierung und/oder PCR überprüft. Letztlich konnten somit 3 P1- (ICRFP703O20328Q5; ICRFP703D0215Q5 und ICRFP703K232Q5) und 10 PAC-Klone (RPCIP711L14226Q3 und RPCIP711P07512Q3, RPCIPH11421Q3, RPCIP711H20359Q3, RPCIP711C03551Q3, RPCIPH08132Q2, RPCIP711D02264Q2, RPCIPC24474Q2, RPCIP711G02449Q2 und RPCIP711G12129Q2) isoliert und verifiziert werden.

Sämtliche 13 Klone wurden zunächst nach Isolierung der PAC-DNA und vollständigem Bam HI- und Bam HI/Bgl II-Restriktionsverdau durch gelelektrophoretische Auftrennung auf Überlappungen überprüft. Alle Klone konnten in einem durchgehenden Contig angeordnet werden (vgl. **Abb. 3.36A**). Für die Bestimmung der genomischen Ausdehnung dieses Contigs wurden die Integratgrößen einzelner überlappender PAC- und P1-Klone durch Pulsfeldgelelektrophorese (PFGE) der Not I-verdauten DNA abgeschätzt. Durch den Vergleich der ermittelten Integratgrößen (RPCIP711H11421Q3: ca. 120-130 Kb, RPCIP711L1426Q3: ca. 110-120 Kb, RPCIP711P07512Q3: ca. 100-110 Kb, RPCIP711G12129Q2: 120-130 Kb und ICRFP703O20328Q5: ca. 80 bis 100 Kb) mit dem Restriktionsmuster konnte die genomische Ausdehnung dieses Contigs auf ca. 300 Kb geschätzt werden (vgl. **Abb. 3.37**).

Zur Verankerung des Contigs im Mausgenom diente das murine *Fgfr3*-Gen, das bereits in eine Region bei 20 cM auf dem murinen Chromosom 5 kartiert wurde (Avraham et al., 1994). PCR-Analysen an den murinen genomischen Klonen mit Primern aus dem murinen *Letm1*-Gen (vgl. **3.3** und **Abb. 3.37**) und dem humanen *FGFR3*-Gen (8F/9R (Exon 9) und 13F/14R (Exon 14); Wüchner et al., 1997) und die anschließende Sequenzierung der genomischen PCR-Produkte zeigten, daß sowohl das murine *Letm1*- als auch das *Fgfr3*-Gen auf drei überlappenden genomischen Klonen (ICRFP703O20328Q5; ICRFP703D0215Q5 und ICRFP703K232Q5) dieses durchgehenden Contigs lokalisiert sind (**Abb. 3.37**), der somit der Region Mmu 5.20 zugeordnet werden kann.

Um die Reihenfolge der einzelnen murinen Gene bzw. Transkripte innerhalb dieses Contigs bestimmen zu können, wurden sämtliche der isolierten 13 Klone über PCR-Analysen mit Primerpaaren aus verschiedenen Bereichen der einzelnen murinen Gene bzw. Transkripte überprüft. Um falsch positive Signale ausschließen zu können, wurden neben der Sequenzierung der erhaltenen PCR-Produkte die entsprechenden genomischen Bereiche der einzelnen Gene zumeist durch Direktsequenzierung an den entsprechenden PAC-Klonen überprüft. Zusätzlich wurden Southern-Blot-Hybridisierungen vorgenommen (vgl. **Abb. 3.36B**). Letztlich konnte die Reihenfolge sowie die Transkriptionsorientierung der untersuchten Gene mit Ausnahme der *Gene 45* und *Fgfr3*, deren Orientierung aufgrund der genomischen Lokalisation nicht endgültig bestimmt werden konnte, innerhalb des Contigs ermittelt werden. Sie entspricht weitgehend den Verhältnissen beim Menschen (*Pol4p-Gen 45-Gen 43 (Whsc2h)-Whsc1 (Gen 51)-Letm1-Fgfr3*) (siehe **Abb. 3.37**).

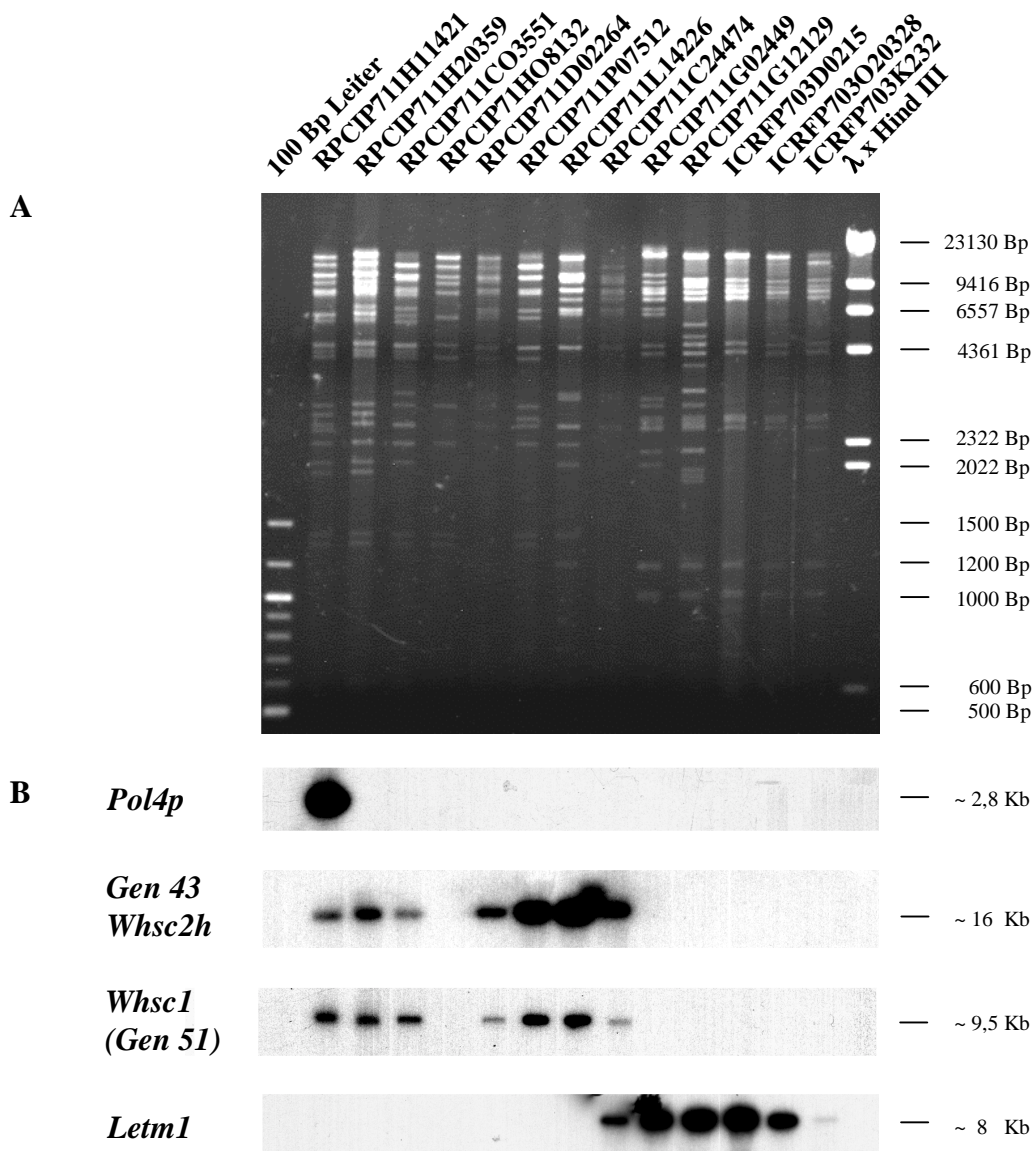


Abb. 3.36: Überlappung muriner P1- und PAC-Klone. **A:** Gelelektrophoretische Auftrennung der Bam HI-verdauten DNA isolierter P1- und PAC-Klone. **B:** Exemplarisch dargestellt sind Autoradiogramme durchgeführter Southern-Blot-Analysen zur Bestimmung der Reihenfolge einzelner Gene der Maus. Als Hybridisierungssonden wurden radioaktiv markierte PCR-Produkte aus den entsprechenden murinen cDNA-Klonen verwendet (vgl. **Abb. 3.37**).

Die Untersuchungen zeigen, daß die Wolf-Hirschhorn-Syndrom-kritische Region und deren flankierenden Bereiche innerhalb eines Syntäniebereichs bei Maus (*Mmu* 5.20) und Mensch liegt, der auch in seiner genomischen Ausdehnung in etwa der des orthologen humanen Bereiches von *FGFR3* bis *POLAP* (**Abb. 3.37**) entspricht.

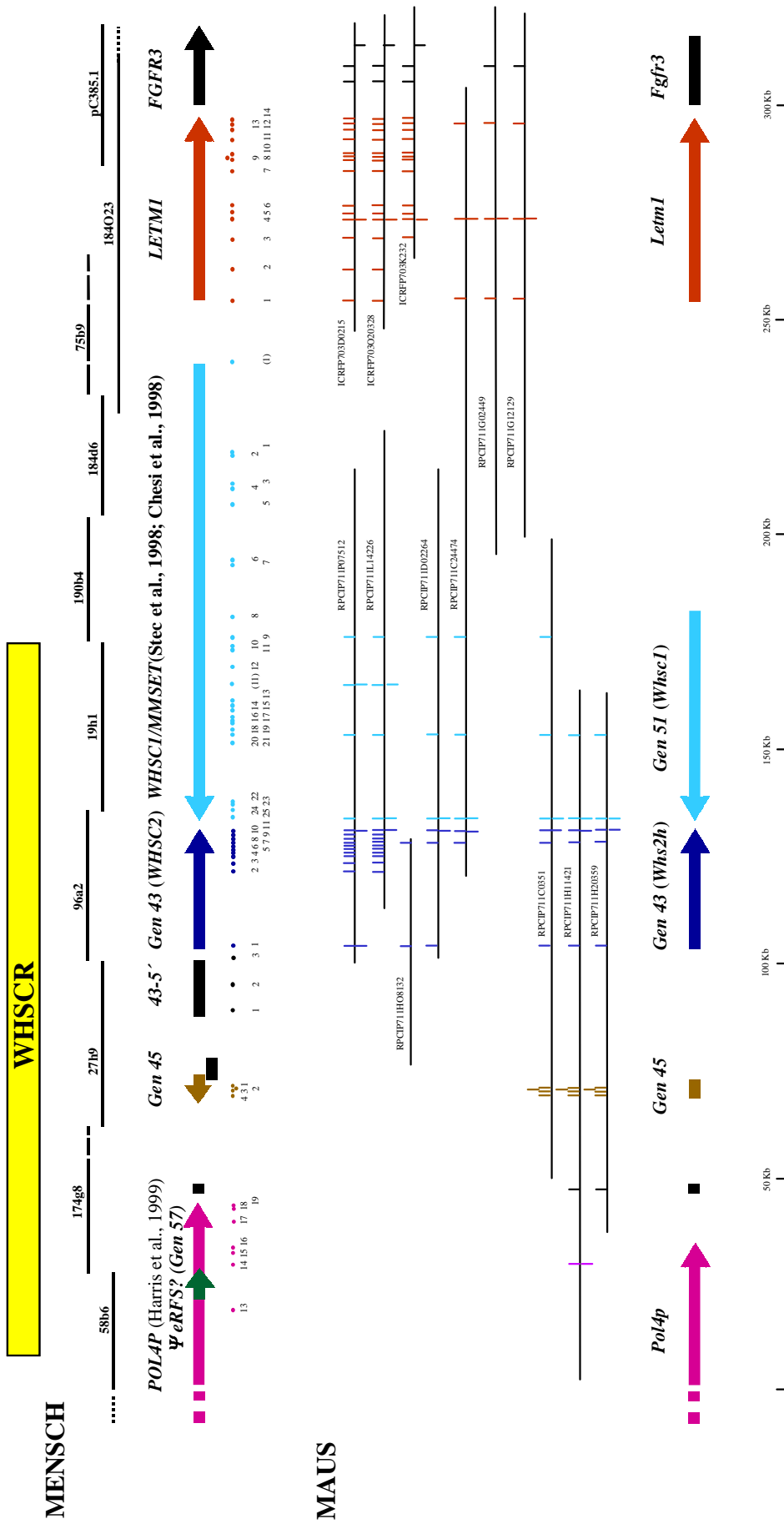


Abb. 3.37: Die WHSCR-Region in der Maus (Mmu 5.20). Legende siehe Seite 137.

Legende zu Abb. 3.37: Alle 13 isolierten P1- bzw. PAC-Klone sind schematisch als Linien dargestellt (vgl. **Abb. 3.21**). Die über PCR und/oder Direktsequenzierung sowie durch Southern-Blot-Analysen überprüften Bereiche einzelner muriner Gene bzw. Transkripte sind als senkrechter Balken oberhalb bzw. unterhalb der 13 Klone bezüglich den Verhältnissen des Menschen verzeichnet und entsprechen daher nicht den tatsächlichen physikalischen Abständen der Maus. Ein putatives trunkiertes Pseudogen des *eRFS*-Gens, wie es beim Menschen in der WHSCR nachweisbar ist, konnte bei der Maus nicht gefunden werden.

Wesentliche Primerpaare bzw. PCR-Produkte, die für Southern-Blot-Analysen (unterstrichen) verwendet wurden, sind nachfolgend aufgeführt: *FGFR3*: 8F/9R (Exon 9), 13F/14R (Exon 14); 16F/17R (Exon 17). *Letm1*: m16-25/m16-19 (Exon 1), m16-21/m16-29 (Exon 4), m16-3'/16-5' (Exon 13), m16-18/m16-38 (Exon 14). *Whsc1* (Gen 51): mWhsc1-9A/B (entspricht Exon 9 des *WHSC1*-Gens), mWhsc1-20A/B (Exon 20), mWhsc1-25A/B (Exon 25), m51-3/m51-4 (Exon 24-25), AA139269-5'/3' (entspricht Exon 11 des *MMSET*-Gens). *Gen 43* (*Whsc2h*): m43-3/m43-12A (Exon 1), m43-8C/m43-10 (Exon 7-8), m43-7A/m43-13 (Exon 11). *Gen 45*: m45-1/m45-2 (Exon 2-3). AA118417 (*erfs*): m57-1/57-2. AI841594: m21-1F/1R. *Pol4p*: mpol-1/mpol-2 (entspricht Exon 14 des *POL4P*-Gens) sowie mpol-3/mpol-4 (Exon 12).

4. Diskussion

Das Wolf-Hirschhorn-Syndrom (Wolf et al., 1965; Hirschhorn et al., 1965; OMIM-Nr. 194190) ist ein komplexes Fehlbildungs-Retardierungssyndrom, das durch den Verlust von genetischem Material der distalen Region 4p16.3 hervorgerufen wird und dessen Ätiologie und Pathogenese bisher weitgehend unverstanden sind. Nach dem aktuellen Kenntnisstand wird davon ausgegangen, daß es sich bei diesem Syndrom um ein echtes „contiguous gene“-Syndrom handelt, d. h. um ein Syndrom bei dem mehrere, benachbarte Gene oder auch gemeinsame regulatorische Kontrollelemente deletiert sind und das je nach dem Grad der Deletionsgröße zu variablen Phänotypen führen kann (Altherr et al., 1997; Wright et al., 1997; Estabrooks et al., 1995a; Petit et al., 1996; Wieczorek et al., 2000). Der Versuch einer Genotyp/Phänotyp-Korrelation wurde mehrfach unternommen, aber die Daten sind oftmals kontrovers und lassen sich nur schwer interpretieren (Estabrooks et al., 1995a; Wieczorek et al., 2000; Zollino et al., 2000). Die Identifizierung und Charakterisierung der durch Haploinsuffizienz betroffenen Gene sowie ihre Beurteilung als mögliche Kandidatengene stellt daher einen wesentlichen Aspekt zum Verständnis der Ätiologie und Pathogenese dieser komplexen Erkrankung dar.

Zu Beginn der vorliegenden Arbeit konnte die kleinste kritische Region, die zu der Ausbildung dieses Syndroms führt, auf einen zwischen den Markern D4S43 und D4S142 ca. 2 Mb umfassenden Bereich eingegrenzt werden (Gandelmann et al., 1992; Johnson et al., 1994; Estabrooks et al., 1994), in der zunächst nur wenige Gene bekannt waren (Thompson et al., 1991; Weber et al., 1991b; McCombie et al., 1992; Collins et al., 1992; Tommerup et al., 1993). Während der proximale Bereich dieser Region zwischen D4S43 und D4S114 Bestandteil eines bereits durch Baxendale et al. (1993) erstellten Cosmid-Contigs war, lagen in der distalen Region zunächst nur wenige kleinere Cosmid-Contigs vor (Bucan et al., 1990; Whaley et al.; 1991; Weber et al., 1991a). Durch Untersuchungen an Patienten mit erkennbarem klassischem Phänotyp konnte der kritische Bereich für die wesentlichen Merkmale des WHS jedoch im Verlaufe dieser Arbeit vom distalen Ende her auf einen 165 Kb großen Bereich zwischen D4S43 und dem *FGFR3*-Gen (D4166 bis D4S3327) eingegrenzt werden (Reid et al., 1996; Altherr et al., 1997; Wright et al., 1997), der inzwischen vom Sanger Center nahezu vollständig sequenziert wurde. Offensichtliche Kandidatengene, die mit der Entstehung dieses Syndroms in Verbindung gebracht werden können, waren in dieser Region

jedoch zunächst noch nicht bekannt.

Die Aufgabe in der vorliegenden Arbeit bestand in der Identifizierung und vorläufigen Charakterisierung von Genen, die zur Entstehung des WHS-Syndroms beitragen könnten. Vor dem Hintergrund der zu Beginn dieser Arbeit noch ca. 2 Mb großen WHSCR, in der noch keine genomischen Sequenzdaten zur Verfügung standen, wurde zunächst die in den folgendem Kapitel dargestellte Methode der Exonamplifikation für die Genidentifizierung verwendet. Diese Methode diene nicht nur der Identifizierung transkribierter Bereiche, sondern konnte zudem für die weitere Kartierung der gesamten Region verwendet werden. Durch die weitere Eingrenzung der WHSCR auf einen 165 Kb großen Bereich (zwischen D4S43 und *FGFR3*) wurde jedoch auf die Erstellung eines durchgehenden genomischen Contigs der distalen Region 4p16.3 sowie auf eine weitere Charakterisierung der bereits identifizierten Transkripte/bzw. Gene der Region distal des *FGFR3*-Gens verzichtet. Die weiteren Analysen konzentrierten sich auf die detaillierte Untersuchung dieser 165 Kb großen WHSCR, für die aufgrund der inzwischen vorliegenden genomischen Sequenzdaten computergestützte Analysen verwendet werden konnten. Von anderen Arbeiten ist bekannt, daß chromosomale Rearrangements Gene in mehreren 100 Kb Entfernung beeinflussen können (Foster et al., 1994; Wagner et al., 1994; Belloni et al., 1996; Sutherland et al., 1996; Kleinjahn und van Heyningen, 1998). Die Beteiligung angrenzender Gene an der Entstehung dieses Syndroms kann daher nicht ausgeschlossen werden. Es wurde deshalb der gesamte, ca. 400 Kb große Bereich zwischen D4S43 und *FGFR3* bei der detaillierten Analyse dieser Region berücksichtigt. Um letztlich ein Mausmodell für diese Erkrankung vorbereiten zu können, wurden darüber hinaus vergleichende Untersuchungen an der Maus durchgeführt.

4.1 Identifizierung transkribierter Bereiche

4.1.1 Exonamplifikation

Durch die Entwicklung der „high throughput“-Sequenzierung ist es in den letzten Jahren möglich geworden, Gene durch direkte Sequenzanalyse großer Bereiche - und nun letztlich des gesamten menschlichen Genoms - zu identifizieren (Claverie, 1997; Borsani et al., 1998). Zu Beginn der vorliegenden Arbeit bestand diese Möglichkeit nur sehr begrenzt, so daß

andere Methoden verwendet werden mußten. Die durch Buckler et al. (1991) entwickelte und durch Church et al. (1994) verbesserte Methode der Exonamplifikation ermöglicht eine rasche und direkte Identifizierung transkribierter Bereiche aus genomischer DNA ohne Kenntnis der DNA-Sequenz. Diese Methode wurde bereits mehrfach erfolgreich für die Suche nach Genen bzw. die Erstellung von Transkriptkarten eingesetzt, wie beispielsweise für die Suche nach dem Gen für die Chorea Huntington auf 4p16.3, der Klonierung des humanen *PDE3A*-Gens oder der Erstellung eines Genkatalogs für das humane Chromosom 22 (Ambrose et al., 1992; Taylor et al., 1992; The Huntington's Disease Collaborative Research Group, 1993; Trofatter et al., 1995, Löbbert et al., 1996). Auch noch in neuerer Zeit wird sie neben der genomischen Sequenzierung vergleichend verwendet (Pribill et al., 1997; Hisama et al., 1998; Church und Buckler, 1999; Smith et al., 2000). So konnten mit dieser Methode selbst innerhalb bereits sequenzierter und durch computergestützte Analysen untersuchter Bereiche neue Exonsequenzen identifiziert werden, die mit den herkömmlichen Genvorhersageprogrammen nicht erkannt wurden (Pribill et al., 1997). Durch die Entwicklung besserer Softwareprogramme und das Vorliegen der humanen Genomsequenz wird die Methode jedoch zunehmend überflüssig.

Für die Identifizierung transkribierter Bereiche aus der distalen Region 4p16.3 wurden zunächst auf der Grundlage von 18 durch S. Baxendale und M. MacDonald (MIT, Cambridge) und J. Wasmuth (Irvine, Californien) zur Verfügung stehenden Cosmidklone drei größere genomische Cosmid/PAC-Contigs (I-III) erstellt. Während zwei dieser Contigs (I und II) Bereiche um die Marker D4S90/D4S142 und D4S111/D4S133 abdecken, stellt der Contig III den Anschluß an den von Baxendale et al. (1993) erstellten Cosmid-Contig dar und erlaubt somit eine Vervollständigung der physikalischen Karte zwischen D4S114 und dem *HDA*-Gen. Aus diesen insgesamt etwa 700-800 Kb umfassenden Contigs (I-III) konnten letztlich 67 unterschiedliche putative Exons isoliert werden (Contig: I/27; Contig II/35; Contig III/5). Hierbei wurden durch zahlreiche Kontrollen falsch positive „Exons“ weitestgehend ausgeschlossen (siehe **3.1.2.**).

Obwohl durch die weitere Eingrenzung der WHS-kritischen Region (Wright et al., 1997) Bereiche distal des *FGFR3*-Gens vermutlich nicht zu den wesentlichen Merkmalen des WHS beitragen, könnten einzelne dieser neuen bzw. erstmals lokalisierten Gene der Contigs I bis II bei WHS-Fällen mit großen terminalen oder auch interstitiellen Deletionen eine Rolle spielen. Die wesentlichen Ergebnisse, die mit Hilfe der Exonamplifikation erzielt werden konnten sowie die Beurteilung dieser Gene werden nachfolgend kurz zusammengefaßt und diskutiert.

Die 67 isolierten Exonsequenzen wurden aufgrund ihrer Homologievergleiche mit bereits bekannten Genen oder Proteindomänen in drei verschiedene Kategorien (**Gruppe A** bis **C**) eingeteilt (vgl. **3.1.3**). Dabei konnten insgesamt 21 (31,3 %) der Exonsequenzen 6 bereits bekannten Genen zugeordnet werden (**Gruppe A**: Contig I: *ZNF141*; Contig II: *PDEB*, *MYL5*, *DAGK4* und *GAK*; Contig III: *FGFR3*). Zwei dieser Gene konnten im Rahmen der vorliegenden Arbeit erstmals (*DAGK4*) bzw. genauer (*GAK*) innerhalb der distalen Region 4p16.3 kartiert werden.

DAGK4: Das *DAGK4*-(Diacylglycerinkinase 4)-Gen konnte in eine Region zwischen D4S111 (*IDUA*) und D4S227/D4S228 (*MYL5/PDEB*) kartiert werden (Endele et al., 1996). Dieses Gen gehört zu einer inzwischen aus 9 Mitgliedern bestehenden Genfamilie, deren einzelne Mitglieder aufgrund ihrer strukturellen Diversität und ihrer komplexen gewebespezifischen Expression vermutlich eine Rolle in verschiedenen Signalkaskaden spielen (Topham und Prescott, 1999). Die Expression des *DAGK4*-Gens konnte wie bei einigen weiteren der bekannten Gene dieser Familie jedoch auch in der Retina nachgewiesen werden (Houssa et al., 1997). Seitdem gezeigt werden konnte, daß Defekte eines retinaspezifischen *DAGK*-Gens (*rdgA*) retinale Degeneration bei *D. melanogaster* verursachen, wurden *DAGK*-Gene als Kandidatengene für Erkrankungen des menschlichen Auges angesehen (Pilz et al., 1995; Stohr et al., 1999). Das WHS ist jedoch nicht mit einer Retinopathie assoziiert und es konnten zudem bisher auch keine entsprechenden Erkrankungen des Auges in dieser Region kartiert werden. Das *DAGK4*-Gen scheint daher wahrscheinlich keine Rolle für das WHS zu spielen.

GAK: Das *GAK* (Cyclin G assoziierte Kinase)-Gen kartiert in eine Region distal des *DAGK4*-Gens zwischen D4S111 und D4S227/D4S228 und weist ein ubiquitäres Expressionsmuster auf (Kimura et al., 1997). Dieses Gen scheint eine Vielzahl an Funktionen aufzuweisen und wurde bereits im Zusammenhang mit der zellulären Apoptose und bei Signaltransduktionswegen des Zytoskeletts diskutiert (Kanaoka et al., 1997; Kimura et al., 1997). Obwohl diesem Gen inzwischen auch eine Rolle bei der rezeptorvermittelten Endozytose in nicht-neuronalen Geweben zugewiesen werden konnte, wurde es bisher mit keiner Erkrankung in Verbindung gebracht (Greener et al., 2000). Für eine Beteiligung an phänotypischen Merkmalen bei WHS-Patienten gibt es keine Anhaltspunkte.

9 (13,4 %) weitere Exonsequenzen aus dem Bereich des Contig I wurden in eine Kategorie (**Gruppe B**) eingeteilt, die signifikante Homologien zu verschiedenen, bekannten Genen bzw. funktionellen Proteindomänen (*WEE1hu* und *ZNF141/ZNF1*) (P-Werte < 0,001) aufweisen, aber aufgrund von Sequenzunterschieden sowie ihrer genauen Kartierung bezüglich der vor-

liegenden Contigs nicht mit ihnen identisch sind.

Exons mit Homologien zu *WEE1hu*: Die genauere Untersuchung dieser Exons ergab, daß vermutlich ein prozessiertes sowie ein möglicherweise unprozessiertes Pseudogen des *WEE1hu*-Gens im telomernahen Bereich lokalisiert sind (Pak, 1999). Das *WEE1hu*-Gen kodiert für eine nukleäre Tyrosinkinase, die an der negativen Regulation des Zellzyklus beteiligt ist und wird als mögliches Tumorsuppressorgen diskutiert (Igarashi et al., 1991; Pendergast, 1996).

Exons mit Homologien zu Zinkfingerengen: 5 Exons korrespondieren vermutlich zu zwei neuen Zinkfingerengen oder auch Pseudogenen der Zinkfingerengenfamilie des Krüppeltyps. Sie werden von zwei weiteren, bereits bekannten und funktionellen Mitgliedern derselben Genfamilie flankiert (*ZNF141* und *ZNF1*; Tommerup et al., 1993; Flint et al., 1997). Sie scheinen damit Bestandteile eines aus bisher 4 Mitgliedern bestehenden Zinkfingerengencusters dieser telomernahen Region 4p16.3 zu sein. Aufgrund der hohen Homologien untereinander könnten diese Gene entsprechend einer Reihe weiterer, bekannter und in Clustern organisierten Zinkfingerengene (Shannon et al., 1996; Eichler et al., 1998) durch Genduplikation auseinander hervorgegangen sein.

Da Zinkfingerengene als Transkriptionsfaktoren einen großen Einfluß auf zelluläre Differenzierungs- und Entwicklungsprozesse ausüben können, wurden sie bereits mehrfach mit entwicklungsspezifischen Erkrankungen und auch dem WHS (*ZNF141*) in Verbindung gebracht (Tommerup et al., 1993; Tommerup und Vissing, 1995). Obwohl Deletionen, die einen Bereich um den Marker D4S142 und damit auch die zwei möglichen neuen Zinkfingerengene betreffen, zu zahlreichen mildereren Fehlbildungen führen und auch Wachstums- und mentale Retardierung verursachen können (Estabrooks et al., 1992; Estabrooks et al., 1995a; Anderson et al., 1997), konnte bisher keine Verbindung zwischen den in dieser Region kartierten Zinkfingerengen (*ZNF141*, *ZNF1*) und dem klinischen Bild hergestellt werden.

Die **Gruppe C** enthält 28 Exons (41,8 %), die keine bzw. nur schwache Homologien zu bekannten Genen oder Proteinen (P-Werte >0,001) aufweisen. In einigen Fällen zeigen sie aber signifikante Homologien zu EST-cDNA-Klonen aus den Datenbanken des NCBI und können daher neue Gene repräsentieren. Eines dieser Exons aus dem Bereich des Contig III führte zu der Identifizierung und Klonierung des in den nachfolgendem Kapitel (4.2) dargestellten humanen *LETMI*-Gens. Die restlichen 9 „Exons“ (13,4 %) dieser Gruppe, die alle aus dem Bereich des Contigs I stammen, weisen jedoch Sequenzübereinstimmungen zu mehreren chromosomalen Regionen auf und können daher ohne eine weitere Charakteri-

sierung nur schwer als „echte“ Bestandteile neuer Gene klassifiziert werden.

4.1.2 Detaillierte Untersuchung der WHSCR durch computergestützte Analysen

Die Sequenzierung großer genomischer Bereiche erlaubt die Identifizierung transkribierter Abschnitte (Exons) bzw. ganzer Gene mit Hilfe spezieller Computerprogramme. Diese Strategie wurde auch in der vorliegenden Arbeit für die Analyse der WHSCR benutzt. Es wurden die Programme „GRAIL II“- und „GENSCAN“ verwendet (Xu et al., 1994; Burge und Karlin, 1997; Burge und Karlin, 1998), mit denen bis zu 66 % bzw. 80 % bekannter Exons korrekt erkannt werden können. Die Programme beruhen auf der Erkennung von konservierten Bereichen wie Spleißdonor- und Akzeptorstellen sowie von offenen Leserahmen und dem GC-Gehalt der DNA (Xu et al., 1994; Burge und Karlin, 1997; Burge und Karlin, 1998). Das Programm „GENSCAN“, das als eines der zuverlässigsten Programme für diese Anwendung gilt, kann zudem durch die Einbeziehung von 5'- und 3'-terminaler Exonsequenzen sowie der Einbeziehung offener Leserahmen benachbarter Exons auch für die Erstellung vollständigerer Genmodelle verwendet werden. Die grundsätzliche Effizienz dieses Programmes verdeutlicht die Analyse der inzwischen vollständig aufgeklärten genomischen Sequenz des humanen Chromosoms 22 (Dunham et al., 1999). Hierbei wurden 94 % der annotierten Gene zumindest teilweise erkannt, wobei jedoch nur bei 20 % dieser Gene alle Exons korrekt vorhergesagt werden konnten (Dunham et al., 1999). Die vergleichende Anwendung verschiedener Vorhersageprogramme an den genomischen Sequenzen bereits bekannter Gene zeigte, daß neben einer Vielzahl „falsch“ positiver oder auch nicht erkannter Exons oftmals falsche Spleißstellen vorhergesagt werden und letztlich somit keine Vorhersage vollständiger Transkriptionseinheiten möglich ist (Lopez et al., 1994; Burset und Guigio, 1996; Burge und Karlin, 1998; Dunham et al., 1999). Unzuverlässig ist auch die Vorhersage von regulatorischen Bereichen wie beispielsweise Promotorbereichen oder Polyadenylierungssignalen, die oftmals nur aus wenigen Basen bestehen und damit schwierig sicher vorherzusagen sind (Fickett und Hatzigeorgiou, 1997; Burge und Karlin, 1998). Letztlich stellen somit selbst die besten dieser Programme keine verlässliche Grundlage für die vollständige Identifizierung von Genen dar (Borsani et al., 1998; Dunham et al., 1999). Die meisten Vorhersageprogramme werden daher zumeist vergleichend und in Kombination mit weiteren Programmen wie „BLASTN“ und „BLASTX“ (zur Identifizierung funktioneller

Proteindomänen und korrespondierenden EST-cDNA-Klone) verwendet. Die erhaltenen Ergebnisse müssen in jedem Fall aber durch experimentelle Analysen bestätigt werden. So wurden beispielsweise auch bei der Analyse der inzwischen vollständig aufgeklärten Sequenz des humanen Chromosoms 21 auch die Vorhersageprogramme „GRAIL“ und „GENSCAN“ in Kombination mit Homologievergleichen und EST-Datenbanksuchen verwendet (Hattori et al., 2000). Letztlich wurden mehr als 95% der bekannten Gene durch die Anwendung computergestützter Analysen erkannt und 98 neue Gene identifiziert (Hattori et al., 2000).

Mit dem Ziel eine möglichst umfassende Transkriptkarte dieser Region zu erstellen, wurden in der vorliegenden Arbeit ebenfalls zwei Genvorhersageprogramme („GRAIL“ und „GENSCAN“) verwendet sowie nachfolgende Datenbanksuchen zur Identifizierung von EST-cDNA-Klonen und bereits bekannten Genen bzw. Proteindomänen durchgeführt. Um eine weitere Einteilung dieser Vielzahl an identifizierten Klonen (82 murine und 260 humane EST-cDNA-Klone) in unterschiedliche Transkriptionseinheiten zu ermöglichen und falsch positive Klone ausschließen zu können, wurden zunächst einzelne cDNA-Klone durch Rand- oder vollständige Sequenzierung überprüft (vgl. **Tab. 3.6**). Sie konnten letztlich in mehrere Contigs eingeteilt werden, die zu vermutlich mindestens 5 eindeutig unabhängigen Genen (*Gen 51*, *43*, *45*, *57* und *POLAP*) korrespondieren. Durch die Redundanz der EST-Homologien sowie über Homologievergleiche der einzelnen Transkripte zu bereits bekannten Genen oder Proteinen konnten zudem erste Erkenntnisse über Expressionsmuster und putative funktionelle Proteindomänen dieser Gene gewonnen werden. Darüber hinaus wurden die orthologen murinen Gene isoliert, die zur vergleichenden Kartierung der WHSCR bei der Maus sowie für detaillierte Expressionsstudien verwendet werden sollten. Mit einer ähnlichen Strategie wurde der Bereich bereits grob von Wright et al. (1997) untersucht. Wright et al. (1997) beschrieben 9 unabhängige Transkripte in der Region zwischen D4S166 und dem *FGFR3*-Gen. In der vorliegenden Arbeit konnten aber allein drei dieser Transkripte einem einzelnen in der vorliegenden Arbeit identifizierten Gen (*Gen 51*) zugeordnet werden. Wie schwierig die Einteilung der einzelnen transkribierten Bereiche in unabhängige bzw. gemeinsame Transkriptionseinheiten ist, zeigt, daß allein das *Gen 51* durch 4 verschiedene unabhängige UniGene-Cluster (**Tab. 3.6**) in den Datenbanken vertreten ist. Die Unterschiede in der Anzahl und Zugehörigkeit der einzelnen Transkripte zwischen den beiden Untersuchungen lassen sich auf die Verwendung weiterentwickelter Computerprogramme und auf die Verwendung unterschiedlicher Berechnungsparameter zurückführen. Obwohl im Rahmen dieser Arbeit letztlich nicht alle dieser Transkripte charakterisiert werden konnten, wurde zum

ersten Mal eine ausführliche und vergleichende Transkriptkarte der Region zwischen D4S43 und dem *FGFR3*-Gen erstellt.

4.2 Das *LETMI*-Gen

In der vorliegenden Arbeit konnte durch Exonamplifikation ein neues Gen identifiziert und kloniert werden, das aufgrund der abgeleiteten Proteinstruktur *LETMI* („Leucin-Zipper und EF-Hand enthaltendes Transmembranprotein 1“) genannt wurde. Das Gen ist ca. 40 Kb groß und besteht aus 14 Exons. Es zeigte sich, daß das Exon 1 des Gens innerhalb einer Sequenzierlücke des Cosmids HSL75b9 (zw. 75b9B und C) lokalisiert ist. Das 5´-Ende des *LETMI*-Gens und die regulatorischen Sequenzen sind somit weniger als 80 Kb von der durch Wright et al. (1997) definierten 165 Kb großen WHSCR entfernt und damit bei fast allen WHS-Fällen deletiert (**Abb. 4.1**).

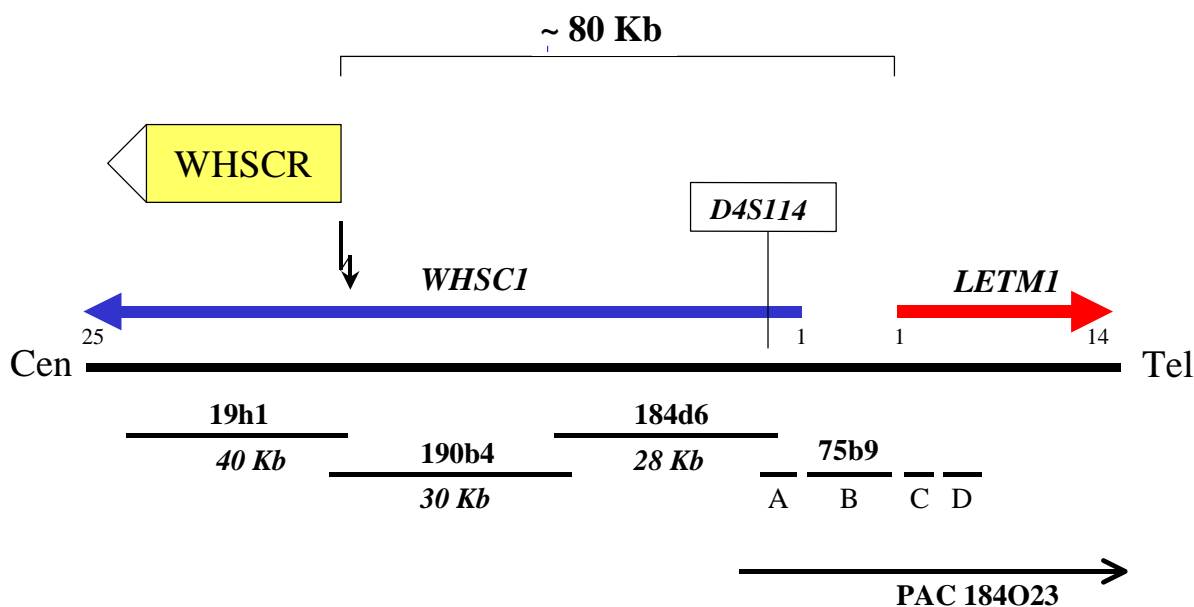


Abb. 4.1: Genomische Lokalisation des humanen *LETMI*-Gens. Das *LETMI*-Gen ist vollständig in dem PAC-Klon 184O23 enthalten, der den Anschluß an den Cosmid-Contig von Baxendale et al. (1993) darstellt. Das Exon 1 des *WHSCI/MMSET*-Gens konnte durch Chesi et al. (1998) ebenfalls in eine Sequenzierlücke des Cosmids 75b9 (zwischen HSL75b9A und B) kartiert werden.

Bisher wurde nur ein WHS-Patient beschrieben, bei dem das *LETMI*-Gen nicht durch Deletion betroffen ist und der dementsprechend auch für die distale Eingrenzung der 165 Kb WHSCR verwendet wurde (Wright et al., 1997). Bei diesem Patient konnte eine über 4,5 Mb große interstitielle Deletion nachgewiesen werden, wobei der distale Deletionsbruchpunkt durch die Cosmide 19h1/190b4 definiert werden konnte. Über diesen milder betroffenen Patienten liegen jedoch nur wenig klinische Daten vor. Neben dem klassischen „griechischen Helm“-Gesicht und Wachstumsretardierung wurde eine nur moderate psychomotorische Retardierung sowie moderate Entwicklungsverzögerungen beschrieben (Somer et al., 1995). Informationen über weitere Merkmale wie Krampfanfälle oder Hypotonie liegen jedoch nicht vor.

Chromosomale Rearrangements können auch noch die Aktivität von Genen beeinflussen, die einige 100 Kb entfernt lokalisiert sind. Dieser Effekt kommt möglicherweise durch Veränderungen der Chromatinstruktur zustande (Kleinjahn und van Heyningen, 1998). Ein Verlust cis-aktiver Elemente, die das Expressionsmuster einiger, benachbarter Gene ändern, wurde beispielsweise für das ebenfalls durch Haploinsuffizienz verursachte DiGeorge/velokardio-faziale-Syndrom (Sutherland, 1996) sowie für das WHS diskutiert (Wright et al., 1997). Ein entsprechender Effekt auf das *LETMI*-Gen wäre auch bei dem von Wright et al. (1997) beschriebenen Patienten möglich. Dementsprechend muß *LETMI* als unmittelbares Kandidatengen aus dem WHS-kritischen Bereich angesehen werden.

Ein Vergleich des murinen und humanen *LETMI*-Gens zeigt, daß beide Gene (mit einer Aminosäureidentität von 84 %) konserviert sind. Dies bezieht sich auch auf die putativen Proteindomänen sowie die übereinstimmenden Transkriptgrößen von 2,8 und 5,5 Kb, die in allen der untersuchten fetalen und adulten Geweben vorhanden sind. Anhand der vorliegenden Sequenzdaten sowie anhand von Northern-Blot-Analysen konnte gezeigt werden, daß sich diese unterschiedlichen Transkriptgrößen von 2,8 und 5,5 Kb vermutlich auf die Nutzung alternativer Polyadenylierungssignale zurückführen lassen.

Die abgeleitete Proteinsequenz von *LETMI* weist neben einer möglichen Typ II-Transmembrandomäne ein Leucin-Zipper-Motiv auf sowie mehrere Bereiche, die zu der Ausbildung von „coiled coils“ und damit möglicherweise zu Protein-Protein-Interaktionen befähigt sein könnten (Kohn et al., 1997). Daneben verfügt LETM1 über zwei EF-Hand Motive (EF-Hand I und II). Die EF-Hand II des LETM1 Proteins des Menschen und der Maus erfüllt die Konsensussequenz des hoch konservierten, kanonischen EF-Hand-Motivs und wird dabei durch zwei α -Helices flankiert (Kawasaki und Kretsinger, 1994; Celio, 1996). Sie ist daher

vermutlich in der Lage, ein Calciumion zu binden. Das zweite EF-Hand-Motiv (I) zeigt jedoch mehrere Abweichungen zu dieser Konsensussequenz. Es enthält zudem keine flankierenden α -Helices und scheint daher seine Fähigkeit zur Calciumbindung verloren zu haben. Diese Ergebnisse konnten durch den Tufty-Kretsinger-Test bestätigt werden, der Hinweise auf die Calciumbindungsfähigkeit EF-Hand bindender Proteine liefert (Tufty und Kretsinger, 1975; Babitch, 1990). Interessanterweise zeigt die Aminosäuresequenz des LETM1-Proteins (AS 297-317) Homologien (57 %) zu der Linker Region zwischen EF-Hand II und III bei der Caltractin (Centrin)-Familie der Pflanzen. Dies läßt die Vermutung zu, daß LETM1 aus einem Protein hervorgegangen ist, das insgesamt über mindestens 4 EF-Hände verfügte, die dann im Verlaufe der Evolution bis auf eines sukzessive verloren gegangen sind. Das LETM1-Protein weist eine signifikante Homologie zu verschiedenen, putativen und bisher nur wenig charakterisierten Transmembranproteinen verschiedener Spezies auf. Die Homologien sind dabei jedoch hauptsächlich auf einen Bereich um die Transmembrandomäne beschränkt und beinhalten eine mögliche Phosphorylierungsstelle für Proteinkinase C (PKC) und Caseinkinase 2 (CK2), die bei allen diesen Proteinen konserviert ist (siehe **Abb. 3.15**) und in einem definierten Abstand zur Transmembrandomäne vorkommt. Obwohl diese Phosphorylierungsstelle auf eine gemeinsame Funktion dieser Proteine hindeuten könnte, sind die weiteren Proteindomänen zwischen diesen Proteinen nicht konserviert und entsprechen nicht denen von LETM1. Eine Ausnahme stellt ein putatives Protein aus *C. elegans* dar (Z81094/CAB03156.1), das ebenfalls ein mögliches funktionelles EF-Hand-Motiv aufweist und damit eine ähnliche Funktion wahrnehmen könnte. Dieses putative Protein wurde allerdings bisher nicht weiter charakterisiert.

Während die meisten funktionellen EF-Hand-Motive calciumbindender Proteine (wie beispielsweise bei Calmodulin oder dem Troponin C) paarweise angelegt sind, sind nur wenige Proteine mit nur einem funktionellen EF-Hand-Motiv bekannt. Hierzu gehören Osteonectin, Glycerol-3-Phosphat-Dehydrogenase, Myosin regulatory light chain (PROSITE: PDOC00018) und Ninein (Bouckson-Castaing et al., 1996). Dabei spielen Interaktionen zwischen EF-Hand-Motiven eine wesentliche Rolle für die Funktionsfähigkeit EF-Hand enthaltender Proteine (Nelson und Chazin, 1998). So besitzt die durch Calcium induzierte Glycerol-3-Phosphat-Dehydrogenase, ähnlich dem LETM1-Protein zwei EF-Hand-Motive von denen jedoch nur eines funktionell ist. Die Ausbildung hydrophober Interaktionen zwischen diesen beiden EF-Händen scheint jedoch eine maßgebliche Rolle für die Funktion des Proteins zu spielen (Brown et al., 1994), ein Mechanismus, der auch für die Funktion des

LETM1-Proteins denkbar ist. Ähnlich dem Ninein, enthält das LETM1-Protein eine Leucin-Zipper-Domäne und mehrere „coiled-coil“-Regionen, die eine wesentliche Rolle bei der Ausbildung von Protein-Protein-Interaktionen spielen (Bouckson-Castaing et al., 1996; Kohn et al., 1997). Vorläufige Ergebnisse zeigten, daß das Ninein Oligomere ausbilden kann und damit Bestandteil makromolekularer Komplexe bildet, die somit über mehrere EF-Hände verfügen und damit eine wichtige Rolle bei der Calciumbindung und Signaltransduktion spielen. Das LETM1-Protein könnte daher wie das Ninein ebenfalls calciuminduzierte Protein-Protein-Interaktionen ausbilden. Zusätzlich zu den beschriebenen Domänen enthält das LETM1-Protein eine Transmembrandomäne. LETM1 stellt somit das erste Mitglied einer neuen Subfamilie von EF-Hand-Proteinen dar, die membrangebunden sind und sich durch die Fähigkeit zur Ausbildung von Protein-Protein-Interaktionen auszeichnen. Bisher konnten nur wenige Transmembranproteine beschrieben werden, die zur Familie der EF-Hand Ca^{2+} -bindenden Proteine gehören. Einige dieser Proteine stellen Untereinheiten von Kalium- oder Calciumkanälen dar (Babitch, 1990; Czempinski et al., 1997; Salinas et al., 1999), andere sind an der Regulation der Ca^{2+} -Konzentration in der Zelle beteiligt (Beeler et al., 1994) oder spielen eine Rolle bei der Zelladhäsion (Gale et al., 1996). Die Familie der Ca^{2+} -bindenden EF-Hand-Proteine umfaßt inzwischen über 45 Subfamilien, wobei die meisten Mitglieder dieser Subfamilien entweder eine direkte calciumabhängige enzymatische Aktivität, eine Calcium-Transport- bzw. Pufferfunktion, eine Rolle bei der Weiterleitung calciumabhängiger Signale oder bei der Calcium-Homeostase spielen (Kawasaki und Kretsinger, 1994; Kawasaki et al., 1998). Aufgrund seiner putativen Proteineigenschaften und seines offensichtlichen Mangels an enzymatischen Domänen scheint es wahrscheinlich, daß das LETM1-Protein in einer calciumabhängigen Weise mit anderen Proteinen interagiert und an der Weiterleitung Ca^{2+} -abhängiger Signale und/oder der Ca^{2+} -Homeostase beteiligt ist. Dies würde auch die ubiquitäre Expression dieses Gens in allen untersuchten fetalen und adulten Geweben erklären.

Calciumionen kontrollieren als sogenannte „first“ oder „second messenger“ eine Reihe verschiedener zellulärer Prozesse, so beispielsweise auch in Nerven- oder Muskelzellen (Brini und Carafoli, 2000; Berchtold et al., 2000). In Nervenzellen kontrollieren Calciumionen eine Vielzahl an Prozessen, wie beispielsweise den schnellen axonalen Transport von Substanzen, Synthese und Freisetzung von Neurotransmittern sowie die Membranerregbarkeit (Heizmann und Braun, 1992). Calciumionen werden auch eine zentrale Rolle bei der Entwicklung und Plastizität des Nervensystems zugeschrieben (Ghosh und Greenberg, 1995). Die deregulierte

Expression calciumbindender Proteine wurde darüber hinaus mit verschiedenen neurologischen und auch neoplastischen Erkrankungen in Verbindung gebracht (Heizman und Braun, 1992; Schäfer und Heizman, 1996). Eine gestörte Ca^{2+} -Homeostase in Nervenzellen scheint darüber hinaus mit neurodegenerativen Erkrankungen wie Alzheimer, neuronalem Altern und auch Krampfanfällen assoziiert zu sein (Heizman und Braun, 1992; Burgess et al., 1997; Letts et al., 1998; Wheal et al., 1998). Auch in Muskelzellen wurde eine gestörte Ca^{2+} -Homeostase mit Krankheitsbildern in Verbindung gebracht, so z. B. bei der malignen Hyperthermie sowie der Hypotonie (McCarthy et al., 2000). Obwohl die Haploinsuffizienz und damit die reduzierte Proteinmenge eines calciumbindenden Gens nicht unbedingt zu einer Störung des basalen Stoffwechsels führen muß, kann sie besonders in spezialisierten Zellen wie Nerven- oder Muskelzellen eine Rolle spielen und daher in einem klinischen Phänotyp resultieren. Das LETM1-Protein stellt somit ein Kandidatengen dar, dessen Haploinsuffizienz zu den neuromuskulären Merkmalen (mentale Retardierung, Krampfanfälle, Hypotonie) der WHS-Patienten beitragen könnte.

Im Rahmen eines „yeast two hybrid“-Screenings nach SMAD-interagierenden Proteinen konnte von Prof. Dr. G. Gross (BFG, Braunschweig) das LETM1-Protein isoliert werden. Weiterführende Untersuchungen zeigten, daß vermutlich der N-terminale Bereich von LETM1 mit dem C-terminalen Bereich von SMAD interagiert. SMAD-Proteine spielen eine wesentliche Rolle bei der Weiterleitung von Signalen von der Zelloberfläche zum Zellkern, wie sie durch Mitglieder der TGF β -Superfamilie, den TGF β s, BMPs und Aktivinen induziert werden (Übersichtsartikel: Massague, 1998; Wrana, 2000). Diese vorläufigen Ergebnisse stellen eine interessante weitere Möglichkeit für die Funktion des LETM1-Proteins dar, insbesondere da diese Signalwege die zelluläre Proliferation und Differenzierung betreffen und einen großen Einfluß auf die Musterbildung, Entwicklung und Homeostase ausüben (Massague, 1998).

Trotz des Vorliegens einer möglichen mitochondrialen Lokalisation (wie es durch die Programme PSORT und PSORT II vorgeschlagen wurde) deuten erste vorläufige Ergebnisse darauf hin, daß das LETM1-Protein keine mitochondriale Lokalisation aufweist. Interessanterweise konnte das LETM1-Protein auch kürzlich durch M. Dreger (Institut für Biochemie, Freie Universität Berlin) aus einer Kernmembranfraktion isoliert werden. Die Interaktion mit spezifischen SMAD-Proteinen und damit eine mögliche Rolle für das LETM1-Protein bei einem dieser Signalwege sowie seine mögliche zelluläre Lokalisation muß in weiterführenden Experimenten nachgewiesen werden.

Das *LETMI*-Gen ist in einem Bereich lokalisiert, der mit der Entstehung von Multiplen Myelomen in Verbindung gebracht wird. Diese malignen B-Zelltumore sind mit Translokationen am *IgH*-Lokus auf Chromosom 14q32.2 assoziiert (Chesi et al., 1997; Richelda et al., 1997). Etwa 25 % der Translokationen sind vom Typ t(4;14) (p16.3;q32), wobei die Bruchpunkte in einem großen Teil der Tumore bis zu 120 Kb proximal des *FGFR3*-Gens und damit in der Nähe des *LETMI*-Gens lokalisiert sind (Richelda et al., 1997, Chesi et al., 1997). Durch Vergleich der in der Literatur beschriebenen klonierten und sequenzierten Bruchpunkte mit den ermittelten Sequenzdaten der Exon- und Intronbereiche des *LETMI*-Gens konnte einer dieser Bruchpunkte (LB1017, Richelda et al., 1997) in das Intron 1 des *LETMI*-Gens lokalisiert werden. Die genaue Lage dieses Bruchpunktes ist in **Abb. 4.2** dargestellt.

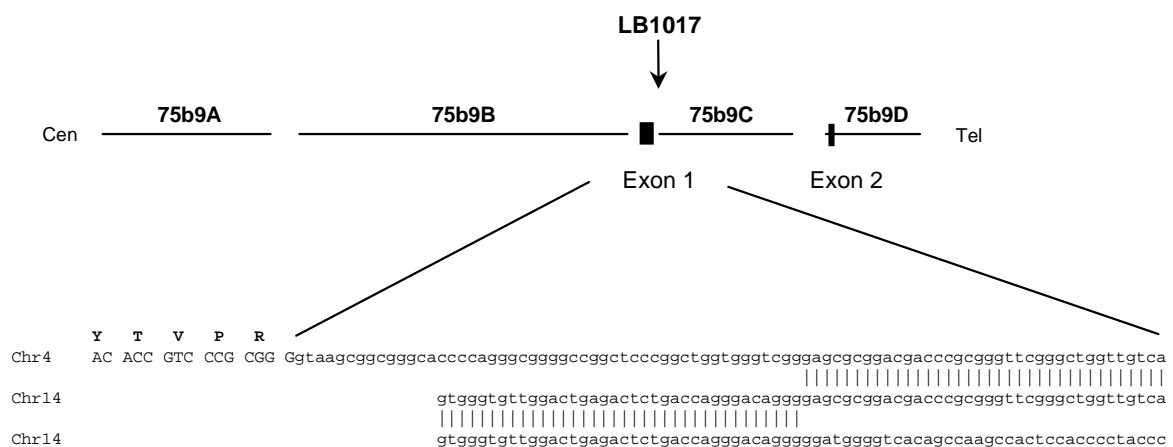


Abb. 4.2: Lage des Translokationsbruchpunktes der Multiplen Myelom-Zelllinie LB1017 innerhalb des Introns 1 des *LETMI*-Gens. Durch die Translokation gelangen Exon 2 bis 14 in die Nähe von *IgH*-Enhancern, wobei Exon 1 mit dem putativen Translationsstart sowie die regulatorischen Elemente des *LETMI*-Gens fehlen. Der Bruchpunkt betrifft die „Switch“-Region des *C α* -Gens (Richelda et al., 1997).

Hierbei könnte das *LETMI*-Gen unter die Kontrolle *IgH*-regulatorischer Elemente, wie Enhancer oder LCRs („Locus control regions“) gelangen. Da aber die 5'-regulatorischen Bereiche sowie das Exon 1 mit dem putativen Translationsstart des *LETMI*-Gens fehlen, ist bei diesen Tumorzellen wahrscheinlich keine Hochregulierung des *LETMI*-Gens zu erwarten. Darüber hinaus können durch die genomische Orientierung des *LETMI*-Gens keine Fusionsproteine entstehen, wie sie z. B. zwischen *IgH* und *WHSC1* nachgewiesen wurden (Stec et al., 1998; Chesi et al., 1998). Das *LETMI*-Gen ist daher wahrscheinlich nicht an der Pathogenese der Multiplen Myelome beteiligt.

4.3 Das *Gen 51 (WHSC1)*

Die in der vorliegenden Arbeit als *Gen 51* bezeichnete Transkriptionseinheit entspricht dem in etwa zeitgleich von Stec et al. (1998) bzw. Chesi et al. (1998) beschriebenen *WHSC1 (MMSET)*-Gen und wurde daher nicht im Detail charakterisiert. *WHSC1* besteht aus 25 Exons, erstreckt sich über einen Bereich von ca. 120 Kb (Cosmid 75b9 bis 96a2) und beinhaltet damit den distalen Deletionsbruchpunkt (Cosmid 19h1/190b4) der WHSCR. Wie die Untersuchungen von Stec et al. (1998) und Chesi et al. (1998) beim Menschen sowie meine eigenen Untersuchungen bei der Maus zeigten, wird *WHSC1* ubiquitär als 5,3 und 8,4 Kb- (Maus) bzw. als 5,2 und 7,7 Kb- (entspricht in etwa den von Stec et al. (1998) entsprechenden ermittelten 9 und 6 Kb-Transkripten) in fetalen und adulten Geweben exprimiert.

Das humane *WHSC1*-Gen kodiert für ein aus 1365 (136 kDa) Aminosäuren bestehendes Protein, das neben einer SET-Domäne vier PHD-Motive, zwei PWWP-Domänen sowie zwei nukleäre Lokalisierungssignale aufweist (Stec et al., 1998; Chesi et al., 1998). Das humane *WHSC1/MMSET*-Gen zeigt ein sehr komplexes alternatives Spleißmuster, so daß vermutlich mehrere funktionelle Proteine (62 kDa, 64 kDa und 136 kDa) entstehen können (Stec et al., 1998; Chesi et al., 1998; Stec et al., 2000). Die Proteindomänen konnten auch bei einer Vielzahl entwicklungspezifischer Gene beschrieben werden. So kommt die SET-Domäne bei einer Reihe von eukaryotischen Proteinen vor, die einen chromatinvermittelten Einfluß auf die Expression von Genen ausüben. Hierzu gehören beispielsweise das Trithorax-Gen (*trx*) von *D. melanogaster*, das homologe *All1*- oder *HRX*-Gen des Menschen und der Maus, das *ash1*-Gen von *D. melanogaster*, das *SET1*-Gen der Hefe sowie ein weiteres identifiziertes und als *Nsdl* bezeichnetes Gen der Maus (Mazo et al., 1990; Gu et al., 1992; Nislow et al., 1997; Tripoulas et al., 1996; Huang et al., 1998). Beispielsweise sind sowohl *trx* als auch *ash1* an der Beibehaltung der Genexpressionsmuster homeotischer Gene beteiligt. Für das *ash1*-Gen von *D. melanogaster* („absent small or homeotic discs“), das die höchsten Homologien zu der SET-Domäne des *WHSC1*-Gen zeigt, konnte gezeigt werden, daß Mutationen dieses Gens zu homeotischen Transformationen (wie die Umwandlung von Halteren in Flügel, Antennen in Beine oder den dritten Beinen zu den zweiten Beinen oder auch der Genitalien in Beine) führen (Tripoulas et al., 1996). Dabei wird auch dieses Gen wie das *WHSC1*-Gen ubiquitär in der frühen Entwicklung exprimiert. Der Einfluss SET-Domänen-enthaltender Proteine auf die Expression homeotischer Gene konnte auch für mehrere Mitglieder der TrxG und Polycomb-

Gruppe der Maus gezeigt werden, die entsprechend den Verhältnissen bei *D. melanogaster* antagonistisch auf die Expression dieser Entwicklungsgene wirken (Übersichtsartikel: Hanson et al., 1999).

Aufgrund seiner möglichen Transkriptionsfaktor- bzw. Koregulatorfunktion wurde *WHSC1* als gutes Kandidatengen für viele der entwicklungspezifischen und möglicherweise auch neurologischen Merkmale des WHS-Syndroms diskutiert (Stec et al., 1998; Chesi et al., 1998). Allerdings wird das Gen ubiquitär exprimiert. Zudem konnten durch Stec et al. (1998) bei der Untersuchung von insgesamt sieben Patienten, die spezifische Merkmale des WHS aber keine detektierbaren Deletionen aufweisen, keine Mutationen nachgewiesen werden, so daß dieses Gen wenn, dann mit anderen Genen gemeinsam durch Haploinsuffizienz zu dem klinischen WHS-Phänotyp beiträgt.

Das *WHSC1*-Gen ist häufig durch Translokationen bei Patienten mit Multiplen Myelomen betroffen, wobei hierbei Fusionsproteine mit dem Translokationspartner (*IgH*) auf Chromosom 14q entstehen (Chesi et al., 1998; Stec et al., 1998). Zudem führen entsprechende Translokationen t(4;14) auch zu einer erhöhten Expression dieses Gens, für das (aufgrund seiner funktionellen Eigenschaften als Transkriptionsfaktor) ein onkogenes Potential vermutet wird. *WHSC1* wird somit eine Rolle bei der Pathogenese dieser B-Zelltumoren zugeschrieben (Chesi et al., 1998; Stec et al., 1998).

4.4 Das Gen 43 (*WHSC2*)

Das humane *Gen 43* ist zentral in der 165 Kb großen WHSCR lokalisiert (Cosmid 96a2). Das *Gen 43* des Menschen und der Maus weisen eine nahezu identische Genstruktur auf und kodieren für ein putatives, zu 95 % identisches Protein von 528 Aminosäuren (Mensch) bzw. 530 Aminosäuren (Maus). Obwohl bei beiden Genen die Nutzung eines weiteren 5'-gelegenen Translationsstarts nicht vollständig ausgeschlossen werden konnte, entsprechen die ermittelten cDNA-Sequenzen von jeweils ca. 2,2 Kb in etwa der über Northern-Blot-Analysen ermittelten Transkriptgröße von ca. 2,4 Kb. Zudem liegt zumindest das erste Exon des humanen Gens innerhalb einer möglichen CpG-Insel, die im 5'-Bereich bei ca. 60 % aller Haushaltsgene vorkommt (Antequera und Bird, 1993; Cross and Bird, 1995). Das humane und murine Gen wurden zeitgleich von Wright et al. (1999b) isoliert und als *WHSC2* (Wolf-

Hirschhorn sndrome candidate 2-Gen bezeichnet. Ein Vergleich der Daten zeigt Unterschiede bei der murinen, jedoch nicht bei der humanen cDNA. Die in der vorliegenden Arbeit identifizierte murine cDNA-Sequenz ist 58 Bp länger und nur zu 92 % identisch zu der von Wright et al. (1999b) publizierten. Die Unterschiede führen auch zu Änderungen in der abgeleiteten Aminosäuresequenz (98 % Identität) und lassen sich neben möglichen Sequenzierfehlern zum Teil durch Polymorphismen erklären. In Anbetracht der Publikation durch Wright et al. (1999b) wurde auf eine detaillierte Aufklärung der zumeist auch funktionell nicht relevanten Sequenzunterschiede sowie auf eine weiterführende Analyse des Gens verzichtet.

Das *WHSC2*-Gen (*Gen 43*) des Menschen und der Maus zeigt keine Homologien zu bereits bekannten Genen oder Proteindomänen. Die abgeleiteten, putativen Proteine verfügen jedoch über zwei mögliche Kernlokalisierungssignale sowie eine mögliche „coiled-coil“-Domäne und mehrere putative Phosphorylierungsstellen. Während durch Wright et al. (1999b) aufgrund einer möglichen Typ I-Transmembrandomäne im N-terminalen Bereich des Proteins eine kernmembranständige Lokalisation vermutet wurde, ergaben sich in der vorliegenden Arbeit durch die Anwendung verschiedener Vorhersageprogramme jedoch Hinweise darauf, daß es sich möglicherweise um lösliche Proteine handeln könnte. Dieser Befund wird durch neuere Untersuchungen unterstützt. So konnte durch Yamaguchi et al. (1999a) gezeigt werden, daß das putative *WHSC2*-Protein wahrscheinlich Bestandteil eines als NELF („negative elongation factor“) bezeichneten Proteinkomplexes ist, der an der negativen Regulation der RNA Polymerase II Transkriptionselongation beteiligt ist. Bisher konnte erst eine Untereinheit dieses aus vermutlich 5 verschiedenen Untereinheiten (NELFA bis E mit Molekulargewichten von 46 bis 66 kDa) bestehenden NELF-Komplexes kloniert werden (Yamaguchi et al., 1999a). Diese Untereinheit weist wie auch das *WHSC2*-Gen keine Homologien zu bereits bekannten Proteinen auf und wird ebenfalls nahezu in allen untersuchten Geweben exprimiert (Levi-Strauss et al., 1988; Yamaguchi et al., 1999a). Für dieses Protein, das neben einer möglichen Leucin-Zipper-Dimerisierungsdomäne auch eine putative RNA-Bindedomäne aufweist, wurde eine Funktion als RNA-bindender Bestandteil des NELF-Komplexes vorgeschlagen. Obwohl bisher letztlich nicht zweifelsfrei *in vivo* nachgewiesen werden konnte, daß das putative *WHSC2*-Protein wirklich einen Bestandteil dieses Proteinkomplexes darstellt, stimmen sowohl die strukturellen Merkmale dieses Proteins (zwei mögliche Kernlokalisierungssignale, „coiled coil“-Domäne), sein errechnetes Molekulargewicht von 57,3 kDa sowie seine ubiquitäre Expression und die Daten von Yamaguchi et al. (1999a) gut mit einer Funktion innerhalb dieses Komplexes überein. Anhand von *in vitro*-

Untersuchungen konnte jedoch gezeigt werden, daß der NELF-Komplex vermutlich mit einem weiteren negativ wirkenden Elongationsfaktor DSIF („DRB sensitivity inducing factor“) ausschließlich einen Arrest der hypophosphorylierten RNA-Polymerase II zu bewirken scheint und vermutlich eine reversible Blockade entstehender RNA-Polymerase II-Transkripte bewirkt (Yamaguchi et al., 1999a). Somit könnte eine allgemeine Rolle dieser Faktoren bei der Transkriptionselongation angenommen werden, die im Zusammenspiel mit anderen Faktoren zusammen über die prozessive oder abortive Elongation einer Vielzahl von Transkripten entscheiden (Yamaguchi et al., 1999b). Inzwischen wurden eine Reihe von Elongationsfaktoren beschrieben, die auf verschiedene Weise an der Regulation der Transkriptionselongation proteinkodierter Gene beteiligt sind (Übersichtsartikel: Shilatifard et al., 1998; Reines et al., 1999; Conaway et al., 2000). Es ist jedoch nur wenig über ihre Funktion *in vivo* bekannt. Einige dieser Faktoren, denen jedoch eine Rolle als positive Faktoren zugeschrieben wurde, konnten bereits mit verschiedenen Erkrankungen in Verbindung gebracht werden (Übersichtsartikel: Shilatifard et al., 1997; Conaway und Conaway, 1999). Sie wurden jedoch zumeist, wie beispielsweise auch das Elongin, ELL oder auch der Faktor P-TEFB mit dem deregulierten Wachstum von Zellen (P-TEFB), der Onkogenese (Elongin) oder auch der Entstehung akuter Leukämien (ELL) in Verbindung gebracht. Ein weiterer Faktor (CSB), dem eine Rolle als positiver Elongationsfaktor zugeschrieben werden konnte (Selby und Sancar, 1997), wurde zudem mit Entwicklungsdefekten wie mentaler Retardierung oder auch Skelettanomalien in Verbindung gebracht. Obwohl für die negativ wirkenden Faktoren DSIF oder NELF eine gestörte Expression einzelner Untereinheiten bereits mit der Deregulation der Expression von Genen diskutiert wurde (Yamaguchi et al., 1999a; Yamaguchi et al., 1999b), bleibt bisher unklar, ob sie zu spezifischen klinischen Merkmalen führen. Es kann daher letztlich nur schwer beurteilt werden, ob eine Haploinsuffizienz des *WHSC2*-Gens als möglicher Bestandteil eines regulatorischen Proteinkomplexes zu einer Störung der Transkription und damit vermutlich zu einer Überexpression spezifischer RNA-Polymerase II transkribierter Gene führen und so an spezifischen Merkmalen des WHS beteiligt sein könnte.

4.5 Das Gen 45

Das *Gen 45* kartiert auf dem Cosmid 27h9 und ist in der Mitte des WHS-kritischen Bereichs lokalisiert. Inzwischen liegen etwa 17 humane und 50 murine EST-Einträge in den Daten-

banken vor (UniGene-Cluster Hs.152932 und Mm.4923), die alle in ihrem 3'- und 5'-Ende nur um wenige Basen differieren. Die aus diesen Einträgen abgeleitete Nukleotidsequenz wurde durch RT-PCR-Analysen bestätigt. Die Daten deuten darauf hin, daß die vollständige humane und murine cDNA ermittelt wurde. Die cDNAs haben eine Länge von bis zu ca. 740 Bp (bzw. durch alternatives Spleißen zusätzliche 107 Bp) beim Menschen und 449 Bp bei der Maus. Bei den ermittelten cDNA-Sequenzen liegen mögliche Translationsstarts vor, die weitgehend der Konsensussequenz von Translationsstartpunkten entsprechen (Kozak, 1996). Dabei wird der offene Leserahmen in allen Fällen durch ein Stopcodon begrenzt, was darauf hindeutet, daß der gesamte kodierende Bereich ermittelt wurde. *Gen 45* kodiert für ein putatives Protein von 95 und 139 Aminosäuren beim Menschen bzw. 90 Aminosäuren bei der Maus, mit einer Identität von 83,8 % zwischen beiden Spezies. Computerauswertungen zeigten, daß die putativen Proteine vermutlich alle über ein Signalpeptid verfügen, das an Position 32/33 (95 AS-Protein), 76/77 (139 AS) bzw. 28/29 (90 AS) abgespalten wird.

Die Genstruktur bestehend aus jeweils 4 Exons ist zwischen beiden Spezies konserviert. Bei beiden Genen konnten aber unterschiedliche alternativ gespleißte Transkripte nachgewiesen werden, wobei bisher keine Informationen über die Häufigkeit bzw. über die Expressionsstärke vorliegen. So scheint bei dem humanen *Gen 45* aufgrund der EST- und RT-PCR-Daten vorwiegend ein aus den Exons 1 bis 4 bestehendes Transkript vorzuliegen (*hu 45 (1a)*), das zwei mögliche Translationsstartpunkte in Exon 1 und Exon 2 enthält. Ob beide Translationsstartpunkte genutzt werden, kann letztendlich nicht entschieden werden. Der stromaufwärts gelegene zusätzliche Translationsstart liegt entsprechend der Klassifizierung von Kozak (1996) nur in einem mäßigen Kontext (Purin an Position -3 fehlt) während der zweite Startpunkt in einem starken Kontext liegt. Denkbar ist, das beide Startstellen genutzt werden: Startpunkt eins mit schwacher Translation in Transkriptvariante 1 und Startpunkt 2 mit starker Transkription in Transkriptvariante zwei. Demnach könnte es sich hierbei um einen zusätzlichen Regulationsmechanismus handeln. Die zusätzlichen Aminosäuren verlängern vermutlich das Signalpeptid. Es konnte ein weiteres Transkript nachgewiesen werden, bei dem eine alternative Spleißdonorstelle im Intron 1 verwendet wird. Durch das Einfügen von zusätzlichen 107 Bp entsteht ein Stopcodon, das eine Verkürzung des offenen Leserahmens zur Folge hat und das damit den ersten Translationsstartpunkt eliminiert.

Ähnliche Verhältnisse liegen beispielsweise bei dem murinen *NGF β* -Gen vor. Auch bei diesem Gen konnten zwei Translationsstartpunkte beschrieben werden, wobei der erste ähnlich dem *Gen 45* in demselben offenen Leserahmen vor einem weiteren möglichen

Translationsstartpunkt liegt, der zudem unmittelbar vor dem Signalpeptid des abgeleiteten Proteins liegt (Edwards et al., 1986; Selby et al., 1987). Obwohl zuerst eine unterschiedliche zelluläre Lokalisation der beiden möglichen Proteine vermutet wurde, konnte gezeigt werden, daß *in vitro*-Systemen beide Translationsstartpunkte verwendet werden und zudem beide Proteine sezerniert werden (Edwards et al., 1986; Edwards et al., 1988). Zudem konnte auch hier ein alternativ gespleißtes Transkript beschrieben werden, das nur zu der Verwendung des zweiten Translationsstartpunkts und damit dem kürzeren Protein führt (Selby et al., 1987; Edwards et al., 1986). Obwohl beide Transkripte in einer Vielzahl von Geweben exprimiert werden, scheinen Unterschiede in der Expressionsstärke in verschiedenen Geweben vorzuliegen und damit eine Rolle bei der Regulation der Expression dieser Gene zu spielen.

Eine Translation vom zweiten Startcodon des humanen *Gen* 45 aus entspricht dem putativen 90 Aminosäuren großen Protein der Maus, das stark in adultem Hoden- und Hirngewebe und in embryonalen Stadien exprimiert wird. Ein entsprechendes längeres Transkript (und damit ein entsprechend größeres Protein), das dem des alternativ gespleißten humanen Transkripts entspricht, wurde aber bei der Maus bisher nicht nachgewiesen, so daß die funktionelle Bedeutung dieser Spleißvariante unklar bleibt. Demgegenüber scheint bei der Maus ein alternativ gespleißtes Transkript vorzuliegen, bei dem Exon 2 fehlt. Auch hier ist unklar ob und welche Bedeutung diese Variante haben könnte.

Entsprechend der ermittelten cDNA ergaben Northern-Blot-Analysen des murinen *Gen* 45 ein ca. 0,7 Kb großes Hybridisierungssignal an RNA aus adultem Hoden- und Hirngewebe der Maus. Weitere Northern-Blot-Analysen zeigten aber, daß jedoch auch bei der Maus vermutlich weitere Transkripte (ca. 4,4 Kb, 2,0 und 1,7 Kb) vorhanden sind. Bisher liegen jedoch keine eindeutigen Daten über die gewebespezifische Expression dieser Transkripte vor, die möglicherweise einer variablen Expression in verschiedenen Geweben unterliegen könnten (Ergebnisse nicht dargestellt). Die Northern-Blot-Analysen des humanen *Gen* an humanem Gewebe ergaben hingegen mehrere Hybridisierungssignale (5,2; 3,8; 2,9; 2,2 und ca. 1,0 Kb) in allen bisher untersuchten Geweben. Sie widersprechen der bisher ermittelten cDNA-Größe von ca. 0,74 Kb sowie der durch die vorliegenden EST-Daten scheinbar vorwiegenden Hoden- und Gehirn-spezifischen Expression dieses *Gen*s. Obwohl bei diesem *Gen* möglicherweise variable 5'-Enden (vgl. 3.4.2.3) und alternativ gespleißte Transkripte vorliegen, gibt es bisher keine konkreten Hinweise, die die vorliegenden und scheinbar ubiquitär exprimierten Transkripte von bis zu 5,2 Kb erklären könnten. Möglicherweise liegen Kreuzhybridisierungen mit anderen Genen vor oder das humane *Gen* unterliegt einem wesentlich

komplexeren Spleißen als bisher angenommen und ist Teil eines möglichen größeren, ubiquitär exprimierten Transkripts. Weitergehende direkte Untersuchungen an menschlichem Gewebe sind nötig, um die Northern-Blot-Analysen zu interpretieren. Die weiteren Expressionsanalysen wurden aufgrund dieser widersprüchlichen Ergebnisse und aufgrund des schwierig zu beschaffenden humanen Gewebes an der Maus vorgenommen. Dabei konnte die starke Expression dieses Gens in adultem Hoden- und Hirngewebe der Maus durch eine nachfolgend durchgeführte RNA-Master-Blot-Hybridisierung weitgehend bestätigt werden. Es zeigten sich zudem starke Hybridisierungssignale an poly(A)⁺-RNA aus weiteren murinen Geweben wie beispielsweise Uterus, Ovar und Prostata sowie eine starke Expression in fetalen Entwicklungsstadien (Tag 11 bis 17 post coitus). RNA-*in situ*-Hybridisierungen an Schnitten aus adultem Hoden- und Hirngewebe der Maus bestätigten diese Ergebnisse und zeigten zudem sehr spezifische Expressionsmuster, so u. a. in Nervenzellen verschiedener Gehirnbereiche und in Spermatozoen.

Die abgeleitete Aminosäuresequenz des Proteins bei Mensch und Maus weist keine signifikante Homologie zu bereits bekannten Proteinen auf. Sie enthalten jedoch kurze Bereiche mit Ähnlichkeit zu Mitgliedern der Proteinfamilie der Nervenwachstumsfaktoren wie NGF- β (Ullrich et al., 1983) oder auch Mitgliedern der Ly6/uPAR-Familie (z. B. Sca-2-Protein und uPAR-Rezeptoren des Menschen und der Maus) (Classon und Coverdale et al., 1994; Kristensen et al., 1991). Die Homologie beschränkt sich jedoch ausschließlich auf den Bereich des Signalpeptids dieser Proteine und umfaßt vor allen Dingen hydrophobe Leucinreste, die bei einer Vielzahl extrazellulärer Proteine mit unterschiedlichen Funktionen auftreten (Ladunga, 2000). Auf der Grundlage dieser Befunde und der Vorhersage eines Signalpeptids kann das Protein 45 vermutlich in die Gruppe der sezernierten (extrazellulären) Proteine eingeordnet werden.

Protein 45 enthält 4 konservierte Cysteinreste. Entsprechende Cysteine nehmen eine wichtige Rolle bei der Faltung, Stabilität und Tertiärstruktur vieler sekretorischer Proteine wahr. Ihre Anordnung ist daher bei vielen Proteinfamilien, wie beispielsweise auch bei der Familie der Nervenwachstumsfaktoren (PROSITE PDOC00221) oder der Chemokine (PROSITE PDOC00434), konserviert (Ibáñez, 1998; Asensio und Campbell, 1999). Obwohl das Protein 45 ähnlich einiger Chemokine vier Cysteinreste enthält, entspricht ihre Anordnung jedoch keinem bisher bekannten Protein dieser Klasse.

Ein weiteres Charakteristikum der Aminosäuresequenz von *Gen 45* ist das Vorhandensein von zwei möglichen, dibasischen Schnittstellen (RK und RR) für Konvertasen, wie man sie

beispielsweise auch bei einer Reihe von Wachstumsfaktoren (z. B. NGF-beta), Prohormonen, verschiedenen Oberflächenrezeptoren oder auch bei Vorläufermolekülen von Neuropeptiden findet (Seidah und Chrétien, 1997; Seidah und Chrétien, 1999; Steiner, 1998). Inzwischen sind einige Prohormonkonvertasen (PC) bekannt, die Vorläuferproteine unter anderem an dibasischen Stellen (Konsensussequenz R/K-X_n-R/K↓, n=0, 2, 4, 6) schneiden und dabei ein gewebe- und entwicklungsspezifisches Expressionsmuster zeigen. Zwei dieser Konvertasen (PC1 und PC2), die eine vorwiegende Expression in endokrinen Drüsen und in Nervenzellen des Gehirns aufweisen, wurden mit der Prozessierung von Neuropeptid- und Neurohormon-Vorläufermolekülen an dibasischen Schnittstellen in Verbindung gebracht (Seidah und Chrétien, 1997; Seidah und Chrétien, 1999). Hierzu gehört beispielsweise auch das in Nervenzellen des Cortex und des Hippocampus exprimierte Preprocortistatin des Menschen, der Ratte und der Maus, das ebenfalls mehrere dieser putativen Schnittstellen enthält (de Lecea et al., 1997a; de Lecea et al., 1997b).

Auch bei einem Vergleich der Expressionsmuster zeigen sich Analogien zu Neuropeptiden. RNA-*in-situ*-Hybridisierungen an Gewebeschnitten der Maus zeigten, daß das murine *Gen 45* in Nervenzellen des Gehirns (evtl. Interneurone) exprimiert wird. Hybridisierungssignale konnten in allen kortikalen und subkortikalen Bereichen erhalten werden, scheinen aber im Gegensatz zu vielen Neuropeptiden nicht auf spezifischen Areale des Gehirns beschränkt zu sein. Die vorliegenden Daten deuten auf eine erhöhte Expression im Endbereich des Hippocampus hin (**Abb. 3.29B**). Interessanterweise wurden Nervenzellen im Bereich des Hippocampus und des cerebralen Cortex mit der Expression zahlreicher Neuropeptide in Verbindung gebracht (Seidah und Chrétien, 1999). Darüber hinaus konnte die Expression vieler Neuropeptidvorläufer ähnlich wie bei *Gen 45* auch in völlig anderen Geweben, wie beispielsweise Hodengewebe, nachgewiesen werden (DiMeglio et al., 1998). Bisher ist jedoch unklar, ob dem *Gen 45* allein aufgrund seiner Expression in Nervenzellen des Gehirns und putativer Konvertase-Schnittstellen eine Rolle als Neuropeptidvorläufer zugeordnet werden kann. Weiterführende zellbiologische Untersuchungen, wie sie beispielsweise auch bei dem *NGFβ*-Gen der Maus durchgeführt wurden (Seidah et al., 1996) könnten zeigen, ob das Protein 45 tatsächlich durch solche Konvertasen geschnitten wird und kleinere und möglicherweise biologisch aktive Peptide entstehen.

Obwohl die Funktion des *Gens 45* nicht ohne weiterführende Untersuchungen beurteilt werden kann, wäre es zumindest denkbar, daß es aufgrund seiner Eigenschaften (ähnlich wie Wachstumsfaktoren oder auch Neuropeptiden) eine Signalfunktion ausüben und unter

anderem für die Differenzierung, Proliferation und die Lebensfähigkeit von Nervenzellen sowie Zell-Zell-Interaktionen und weitere physiologische Prozesse in nicht-neuronalem Gewebe (Levi-Montalcini et al., 1996; Tessarollo, 1998; Hökfelt et al., 2000; Levine et al., 2000) verantwortlich sein kann.

Interessant ist, daß die Expression von *Gen 45* auch solche Gewebe betrifft (Gehirn, Hoden, Uterus, Prostata und Nebenhoden), die zum Teil in das Wolf-Hirschhorn-Syndrom involviert sind. So weisen WHS-Patienten klinische Merkmale wie Uterusfehlbildungen, Kryptorchismus, Hypospadie, Kleinhirnhypoplasie oder -agenesis auf. Zudem konnte auch eine gestörte Migration von Nervenzellen in Bereichen des cerebralen Cortex sowie Heterotopien von Purkinjezellen des Cerebellums beschrieben werden (Johnson et al., 1976). Aufgrund der vorliegenden Expressionsdaten sowie seiner möglichen funktionellen Eigenschaften stellt das *Gen 45* ein Kandidatengen dar, das durch Haploinsuffizienz an der Entstehung des WHS beteiligt sein könnte.

4.6 Die Gene *POLAP* und 57

Das *POLAP*-Gen umfaßt den proximalen Deletionsbruchpunkt der WHSCR (Wright et al., 1997). Es erstreckt sich über einen Bereich von mehr als 110 Kb zwischen dem Marker D4S43 und dem Cosmid 174g8. Die partielle cDNA-Sequenz dieses Gens, die von Harris et al. (1999) in die Datenbanken eingegeben wurde, kodiert für ein aus bisher mehr als 579 Aminosäuren bestehendes putatives Protein (Leserahmen noch nicht vollständig ermittelt). Das Proteinfragment weist Homologien zu der Familie A der DNA-Polymerasen auf (Delarue et al., 1990; Braithwaite und Ito, 1993) und scheint daher ein weiteres Mitglied dieser Familie darzustellen (Harris et al., 1999; unveröffentlicht). Es beinhaltet die hoch konservierten Motive A, B (Signatur der Familie A) und C, die als Bestandteil der katalytisch aktiven Polymerasedomäne eine Rolle bei der dNTP- sowie der Mg^{2+} -Bindung spielen (Steitz, 1999). Die Homologien der bisher bekannten Aminosäuresequenz des putativen *POLAP*-Proteins sind jedoch ausschließlich auf die Polymerasedomäne beschränkt, so daß bisher keine Hinweise auf weitere mögliche Proteindomänen bzw. Proteinmotive vorliegen. Das *POLAP*-Gen ist bei mehreren Patienten betroffen, die Deletionen proximal der WHSCR aufweisen und damit keine der WHS-spezifischen klinischen Merkmale zeigen (Gandelmann et al.,

1992; Wright et al., 1997; White et al., 1995). Dieses Gen scheint daher zunächst allein aufgrund seiner genomischen Lokalisation kein gutes Kandidatengen für spezifische Merkmale des WHS darzustellen und wurde daher nicht weiter charakterisiert.

In einem ca. 10 Kb großen Intron (Intron 13 der bisher 19 bekannten Exons des Gens) liegt eine mögliche Transkriptionseinheit, die ich zunächst als *Gen 57* bezeichnet habe und das in die Nähe des Markers D4S166 und damit nahe der proximalen Grenze der WHSCR kartiert. Das Gen zeigt zudem keine Übereinstimmungen zu dem *POLAP*-Gen und ließ daher ein weiteres unabhängiges Gen derselben Transkriptionsorientierung vermuten. Die Existenz von Genen in Intronbereichen anderer Gene konnte bereits mehrfach beschrieben werden, so z. B. bei dem Faktor VIII-Gen, dem Neurofibromatose 1-Gen, dem *HIRA*-Gen sowie dem Dystrophin-Gen (Levinson et al., 1992; Levinson et al., 1990; Viskochil et al., 1991; Pizzuti et al., 1999, Ferlini und Muntoni, 1998). Bisher sind jedoch nur wenige Gene bekannt, die dabei dieselbe Transkriptionsorientierung aufweisen (Levinson et al., 1992; Ferlini und Muntoni, 1998). Die weitere Charakterisierung des *Gens 57* zeigte jedoch, daß es sich bei der entsprechenden chromosomalen Region auf dem Cosmid 58b6 möglicherweise um ein trunkiertes Pseudogen des humanen *eRFS*-Gen auf Chromosom 6q24 handeln könnte (Wallrapp et al., 1998).

4.7 Mutationsscreening von Kandidatengen bei einem WHS-Patienten ohne Deletion

Bei klassischen „contiguous gene“-Syndromen ist es nicht möglich, Kandidatengene aufgrund von Mutationen zu identifizieren. Eine Ausnahme bildet die Untersuchung von Patienten, bei denen einzelne oder mehrere der spezifischen klinischen Merkmale eines Syndroms isoliert vorkommen und dementsprechend keine oder nur sehr kleine Deletionen vorliegen. So konnte beispielsweise auch bei dem Williams-Beuren- und dem Miller-Dieker-Syndrom gezeigt werden, daß Mutationen einzelner Gene (Elastin bzw. *LIS1*) für spezifische, isolierte Merkmale dieser Syndrome verantwortlich sind (Ewart et al., 1993; Tassabehji et al., 1997; Chong et al., 1996). Ein Zusammenhang der Chromosomenregion 4p16.3 mit isoliert vorkommenden klinischen Merkmalen des WHS konnte bisher nicht beschrieben werden (z. B. Epilepsie). Auf der anderen Seite gibt es aber Patienten, die spezifische Merkmale des WHS

aufweisen und bei denen keine Deletionen nachgewiesen werden konnten (Marczinek et al., 1995; Stec et al., 1998). Obwohl bei diesen Patienten unklar ist, ob es sich um echte WHS-Fälle handelt (Stec et al., 1998; Marczinek et al., 1995), stellt ihre weitere Untersuchung einen wichtigen Aspekt dar um die WHSCR weiter einzuschränken. Aus diesem Grund wurde einer dieser Fälle näher untersucht. Es handelt sich um einen Patienten mit charakteristischen Merkmalen des WHS, wie Wachstumsretardierung, Entwicklungsverzögerungen sowie dem typischen „griechischem Helm“-Gesicht (Marczinek et al., 1995). Zur Identifizierung von Mikrodeletionen wurden 5 neu erstellte Mikrosatellitenmarker etabliert und zusammen mit den bereits bekannten polymorphen Markern D4S115 und D4S114 für die Untersuchung der Region zwischen D4S43 und dem *LETMI*-Gen (und damit der WHSCR) eingesetzt. Bei den durchgeführten Untersuchungen ergaben sich aber keine Hinweise auf eine mögliche Deletion bei dem Patienten, die die WHSCR beeinträchtigen könnte. Die Untersuchungen wurden durch Mutationsanalysen der kodierenden Regionen des *LETMI*-, *WHSC2*-Gens sowie des *Gens 45* ergänzt. Erwartungsgemäß scheint keines dieser Gene alleine für den Phänotyp verantwortlich zu sein. Obwohl das *WHSCI*-Gen nicht auf Mutationen überprüft wurde, konnten jedoch bereits durch Stec et al. (1998) Mutationen bei der Untersuchung entsprechender Patienten (Phänotyp ohne Mikrodeletionen) nicht nachgewiesen werden. Aus den Untersuchungen ergibt sich, daß es sich bei dem Patienten entweder nicht um einen Wolf-Hirschhorn-Syndrom-Fall handeln könnte oder in cis wirkende Veränderungen (Inversion, Punktmutation etc.) die Aktivität von Genen in der WHSCR über eine größere Entfernung beeinflussen (inhibieren). Weitere Untersuchungen sind zur Beurteilung dieser Möglichkeiten notwendig.

4.8 Zusammenfassende Beurteilung der identifizierten Gene

Die vorliegenden Daten zeigen, daß es sehr schwierig ist, eine direkte Beteiligung eines der isolierten Gene an der Ätiologie und Pathogenese des WHS unmittelbar nachzuweisen. Dies hat zwei Gründe. Zum einen wird das Syndrom durch Haploinsuffizienz einzelner oder mehrerer Gene hervorgerufen, so daß Mutationen zur Identifizierung des Krankheitsgens nicht gefunden werden können. Zum anderen sind aufgrund der mangelnden Beschreibung von Patienten mit kleinen Deletionen bzw. partiellen Phänotypen bisher alle Versuche gescheitert,

im Sinne einer Genotyp/Phänotyp-Korrelation die Hemizygotie einzelner Gene einer bestimmten phänotypischen Komponente zuzuordnen.

Betrachtet man die einzelnen Gene, so kommt eine Beteiligung fast aller am Phänotyp in Frage. Die ubiquitär exprimierten Gene *WHSC1* und *WHSC2* sind vermutlich Bestandteile regulatorischer Proteinkomplexe, die einen Einfluß auf die Expression einer Vielzahl von Genen ausüben könnten. Gendosiseffekte könnten somit breit gefächerte Auswirkungen haben. Für das *WHSC1*-Gen wird eine Beteiligung an den Entwicklungsdefekten und möglicherweise auch den neurologischen Merkmalen dieses Syndroms diskutiert (Stec et al., 1998; Chesi et al., 1998). Während dem ubiquitär exprimierten *LETMI*-Gen eine Rolle bei der Entstehung der mentalen Retardierung, Hypotonie und Krampfanfällen und möglicherweise auch eine Rolle in der frühen Entwicklung zugeschrieben werden könnte, könnte das *Gen 45* aufgrund seines Expressionsmusters an Merkmalen, die die Genital- und Gehirnentwicklung (Kleinhirnhypoplasie und neurologische Defekte, Kryptorchismus, Uterusfehlbildungen) betreffen, beteiligt sein. Lediglich das *POLAP*-Gen scheint aufgrund seiner möglichen Funktion als DNA-Polymerase sowie seiner genomischen Lokalisation zunächst nicht als Kandidatengene für spezifische Merkmale des WHS in Frage zu kommen.

Vermutlich gibt es aber weitere Gene oder auch regulatorische Elemente in der WHSCR die bisher noch nicht identifiziert wurden und an Merkmalen des Syndroms beteiligt sind, wobei letztendlich erst der Verlust mehrerer Gene gemeinsam zum Vollbild der Erkrankung führt.

Vergleichbar schwierig ist die Situation bei einer Reihe anderer „contiguous gene“-Syndrome, wie beispielsweise dem Williams-Beuren- und dem DiGeorge/velo-kardio-fazialen-Syndrom. Auch bei diesen Syndromen konnten jeweils zahlreiche Gene beschrieben werden, die aufgrund ihrer Expression oder ihrer möglichen Funktion als Kandidatengene für spezifische Merkmale dieser Syndrome diskutiert werden, für die aber ebenfalls zumeist keine direkte Beteiligung am Krankheitsbild nachgewiesen werden konnte (Übersichtsartikel: Lindsay und Baldini, 1998; Osborne, 1999). Die meisten dieser Arbeitsgruppen haben daher ihre weiteren Untersuchungen auf die Maus ausgedehnt. Ziel ist es letztlich, durch Sequenzierungen des orthologen Bereichs bei der Maus und anschließendem Vergleich mit der humanen Sequenz die Identifizierung aller beteiligten Gene und regulatorischen Bereiche zu erleichtern sowie detaillierte Expressionsstudien durchzuführen (Galili et al., 1997; Hardison et al., 1997; Sutherland et al., 1998; DeSilva et al., 1999; Winterpacht et al., 1998; Lund et al., 2000). So ergab beispielsweise die vergleichende Sequenzanalyse einer 634 Kb großen Region des DiGeorge/VCFS-syntänen Bereichs auf Chromosom Mmu 16, daß die meisten Gene eine

hohe Ähnlichkeit aufweisen und somit aufgrund ihrer Homologie durch Interspeziesvergleich sicher identifiziert werden können (Lund et al., 2000). Zudem sind auch nichtkodierende Bereiche (ca. 6 % der genomischen Sequenz) konserviert, was auf das Vorhandensein zahlreicher regulatorischer Regionen hindeutet.

Die Analyse des orthologen Bereichs der Maus liefert darüber hinaus wichtige Hinweise auf die Syntanie beider Regionen. Dies ist eine Voraussetzung für die Etablierung von Mausmodellen für die Erkrankungen durch Einfügen von Deletionen in das Mausgenom. Erst die Herstellung und detaillierte Charakterisierung dieser Mäuse wird die Aufklärung von „contiguous gene“-Syndromen (und letztendlich auch dem WHS) ermöglichen. So konnten beispielsweise auch bei dem DiGeorge/VCF-Syndrom entsprechende Deletionen im Mausgenom erzeugt werden (Kimber et al., 1999; Lindsay et al., 1999). Während hemizygot k.o.-Mäuse mit einer ca. 150 Kb großen Deletion für den proximalen Bereich der minimalen kritischen Region (MDGCR) keinen spezifischen Phänotyp aufweisen (Kimber et al., 1999), zeigen Mäuse mit einer ca. 1,2 Mb umfassenden Deletion, die den distalen Bereich der MDGCR beinhaltet, spezifische Merkmale (Herzdefekte) des DGS und können für weitere Untersuchungen herangezogen werden (Lindsay et al., 1999).

Um die Identifizierung und Charakterisierung sämtlicher der am WHS beteiligten Gene erleichtern zu können und die Etablierung eines Mausmodells vorbereiten zu können, wurde in der vorliegenden Arbeit auch die WHSCR auf die Konservierung der Syntanie bei der Maus untersucht.

4.9 Charakterisierung der WHSCR-orthologen Region der Maus

Untersuchungen einiger Arbeitsgruppen zeigten, daß viele Gene des humanen Chromosoms 4 (annähernd 60) auf dem Mausechromosom 5 kartieren, wobei hier größere durchgehende Syntaniebereiche vorliegen (Übersichtsartikel: siehe Grosson et al., 1994; Lyu und Kozak 1996, Kozak und Stephenson 1997 und Kozak und Stephenson 1999). Hierzu gehört eine ganze Gruppe von Genen, die um das Huntington-Gen (*HD*) kartieren, das *FGFR3*-Gen und das erst kürzlich beschriebene *TACC3*-Gen, das in einer Region distal des *FGFR3*-Gens kartiert (Still et al., 1999). Es wurde vermutet, daß der Syntaniebereich die WHS-kritische Region (WHCSR) zwischen den Markern D4S43 und dem *FGFR3*-Gen bei ca. 20 cM

einschließt, wobei detaillierte Untersuchungen innerhalb dieses Bereichs und der hier lokalisierten murinen Gene bisher jedoch nicht vorlagen (vgl. **Abb. 4.3**). In der vorliegenden Arbeit konnte durch die Erstellung und Charakterisierung eines genomischen P1/PAC-Contigs bei der Maus gezeigt werden, daß die murinen Gene *Fgfr3*, *Letm1*, *Whsc1*, *Gen 43* (*Whsc2h*), *Gen 45* und *Pol4p* sowie einige weitere der überprüften EST-cDNA-Klone der Maus in einem durchgehenden Syntänieblock zwischen Mensch (*POLAP* bis *FGFR3*) und Maus (Mmu 5.20) enthalten sind (vgl. **3.6**). Sie stellen somit die orthologen murinen Gene der WHSCR-Region (zwischen *POLAP* und dem *FGFR3*-Gen) dar. Lediglich das mögliche trunkierte *eRFS*-Pseudogen des Menschen scheint nicht in beiden Spezies vorzuliegen, was für ein Pseudogen allerdings nicht ungewöhnlich ist (Lund et al., 2000). Es konnte zudem nachgewiesen werden, daß nicht nur die Reihenfolge dieser Gene sondern auch die Transkriptionsorientierung (mit Ausnahme von *Gen 45*) zwischen beiden Spezies konserviert ist (siehe **Abb. 3.37**).

Bisher sind keine Mausmodelle innerhalb dieser Region bekannt, die für die Entstehung des Wolf-Hirschhorn-Syndroms herangezogen werden können. Durch Goodwin et al. (1998) konnten im Rahmen eines groß angelegten Mutagenese-Screenings auf rezessive Mutationen heterozygote Mäuse mit großen Deletionen auf dem Chromosom 5 hergestellt werden, die das Huntington-Disease-Gen aber auch den putativen WHS-kritischen Bereich beinhalten und einige Merkmale des Wolf-Hirschhorn-Syndroms aufweisen (Schimenti, 1998; Schimenti et al., 2000). Durch diese Deletionen (zumeist mehrere Mb) werden jedoch eine Vielzahl an Genen betroffen, die diese Stämme nicht zuletzt aufgrund der sehr komplexen Phänotypen für die Identifizierung der am WHS beteiligten Gene als ungeeignet erscheinen lassen. In jedem Fall zeigen diese Stämme aber, daß ein WHS-ähnlicher Phänotyp bei der Maus induziert werden kann. Geeignete Mausmodelle für die Erkrankung, wie sie durch die Einführung entsprechender, großer Deletionen in das Mausgenom erzeugt werden könnten (Ramirez-Solis et al., 1995), konnten jedoch bisher aufgrund fehlender, detaillierter Untersuchung dieser Region nicht hergestellt werden. Mit Hilfe der vorliegenden Untersuchungen ist es nun möglich, die Maus als möglichen Modellorganismus für das WHS zu verwenden und durch das Einfügen gezielter Deletionen und die detaillierte Analyse der entstehenden Phänotypen einen wesentlichen Beitrag zum Verständnis dieses komplexen Syndroms zu leisten.

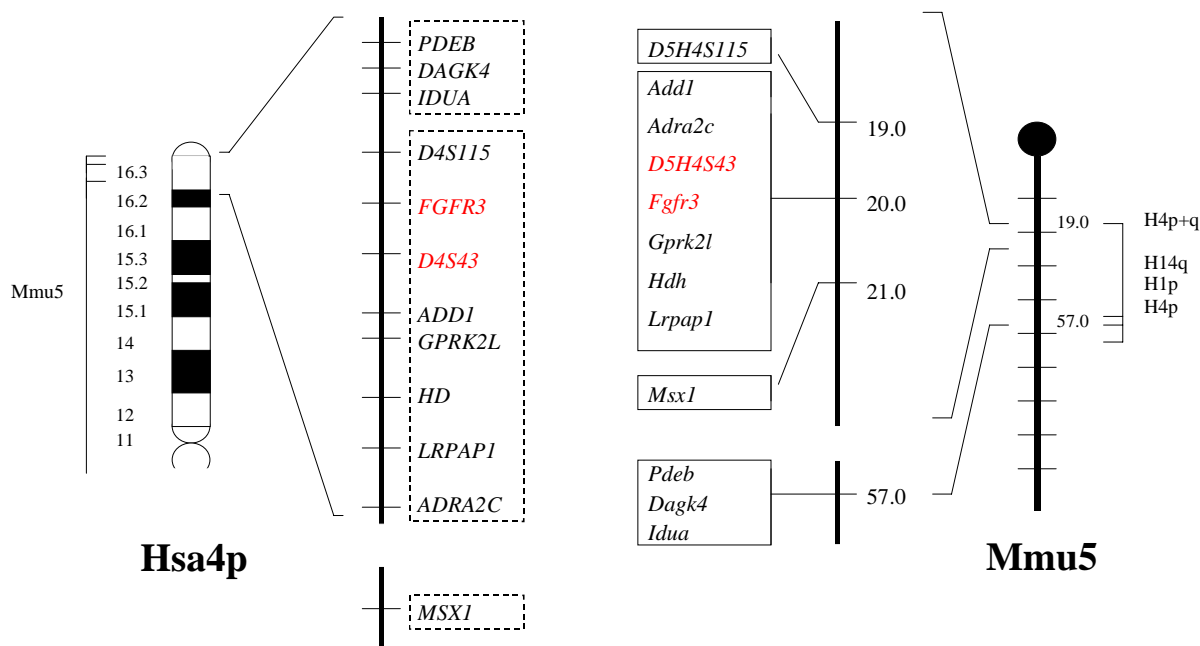


Abb. 4.3: Vergleichende Kartierung humaner und muriner orthologer Gene in konservierten Bereichen auf Chromosom 4p16.3 und Mmu 5 nach Kozak und Stephenson (1999). Ein größerer Block von ca. 60 Genen von Chr. 4 (ca. 4p16.3 bis 4q21) ist innerhalb des murinen Chromosoms 5 (19-57 cM) lokalisiert, wobei Umordnungen wie Unterbrechungen der Genanordnung sowie eine Inversion zu finden sind. Dies betrifft auch die humanen Gene *PDEB*, *IDUA* und *DAGK4* der distalen Region 4p16.3, die innerhalb eines kleineren Bereiches auf dem Mmu 5 (57 cM) kartieren. Eine Reihe von Genen um das Huntington-Disease-Gen sind innerhalb einer Region von 19-21 cM lokalisiert, wobei keine detaillierte Untersuchungen über die genaue Abfolge der hier dargestellten Gene und Marker vorlagen. Ebenso lagen keine Untersuchungen über den WHSCR-orthologen Bereich der Maus zwischen D4S43 und dem *FGFR3*-Gen bzw. zwischen D5H4S43 und *Fgfr3* und der hier lokalisierten Gene vor.

5. Zusammenfassung

Das Wolf-Hirschhorn-Syndrom (WHS) ist ein komplexes und variables Fehlbildungs- Retardierungssyndrom, das durch Deletion in der distalen Chromosomenregion 4p16.3 hervorgerufen wird und dessen Ätiologie und Pathogenese bisher weitgehend unverstanden sind. Die Zielsetzung in der vorliegenden Arbeit bestand in der Identifizierung und vorläufigen Charakterisierung neuer Gene, die an der Entstehung des Syndroms beteiligt sein könnten.

Die Wolf-Hirschhorn-Syndrom-kritische Region (WHSCR) konnte zu Beginn der vorliegenden Arbeit auf einen ca. 2 Mb großen Bereich zwischen den Markern D4S43 und D4S142 eingegrenzt werden. Für die Identifizierung neuer Gene wurden zunächst drei größere genomische Cosmid-/PAC-Contigs (I-III) im Bereich der Marker D4S114 bis D4S142 erstellt und mittels Exonamplifikation auf transkribierte Bereiche (Exons) untersucht. Es konnten insgesamt 67 putative „Exons“ isoliert werden, von denen einige bereits bekannten Genen (*ZNF141*, *PDEB*, *MYL5*, *GAK*, *DAGK4* und *FGFR3*) entsprechen. Zwei dieser Gene konnten im Rahmen dieser Arbeit erstmals (*DAGK4*) bzw. genauer (*GAK*) in die distale Region 4p16.3 kartiert werden. Die restlichen Exons können aufgrund von Homologievergleichen und/oder EST-cDNA-Homologien vermutlich neuen Genen oder auch Pseudogenen (z. B. *ΨWEE1hu*) zugeordnet werden. Durch die im Verlaufe der vorliegenden Arbeit publizierte weitere Eingrenzung der WHSCR auf einen 165 Kb-großen Bereich proximal des *FGFR3*-Gens konzentrierten sich weitere Untersuchungen auf die detaillierte Analyse der WHSCR zwischen dem Marker D4S43 und *FGFR3*. Mit Hilfe von Exonamplifikation bzw. computergestützter Auswertung vorliegender Sequenzdaten aus diesem Bereich („GRAIL“, „GENSCAN“ und Homologievergleiche in den EST-Datenbanken des NCBI) konnten mehrere neue Gene identifiziert werden. In distaler-proximaler Reihenfolge handelt es sich dabei um die Gene *LETM1*, *51*, *43*, *45*, *57* und *POLAP*. *LETM1* kodiert für ein putatives Transmembran-Protein mit einem Leucin-Zipper- und zwei EF-Hand-Motiven und könnte aufgrund seiner möglichen Beteiligung an der Ca^{2+} -Homeostase und/oder der Signaltransduktion zu Merkmalen des WHS (Krampfanfällen, mentale Retardierung und muskuläre Hypotonie) beitragen. Das *Gen 51* entspricht einem in etwa zeitgleich durch Stec et al. (1998) und Chesi et al. (1998) als *WHSC1* bzw. *MMSET* bezeichnetem Gen und wurde daher nicht weiter charakterisiert. Es wird genauso wie das *Gen 43*, das zeitgleich von Wright et al. (1999b) als *WHSC2* beschrieben werden konnte und eine mögliche Rolle bei der

Transkriptionselongation spielt, ubiquitär exprimiert. Das in der vorliegenden Arbeit identifizierte *Gen 45* zeigt demgegenüber ein ausgesprochen spezifisches Expressionsmuster (in Nervenzellen des Gehirns sowie in Spermatozyten). Dies stellt zusammen mit der strukturellen Ähnlichkeit des putativen Genprodukts zu Signalmolekülen einen interessanten Zusammenhang zu Merkmalen des WHS (beispielsweise Kryptorchismus, Uterusfehlbildungen oder auch neurologische Defekte) her. Demgegenüber handelt es sich bei dem *Gen 57* möglicherweise um ein trunkiertes Pseudogen des *eRFS*-Gens auf Chromosom 6q24 (Wallrapp et al., 1998). Das *POLAP*-Gen schließlich stellt allein aufgrund seiner genomischen Lokalisation sowie seiner möglichen Funktion (als DNA-Polymerase-ähnliches Gen) kein gutes Kandidatengens für spezifische Merkmale des Syndroms dar und wurde daher nicht im Detail charakterisiert.

Um die Beteiligung der Gene an der Ätiologie und Pathogenese des Syndroms zu verstehen, ist die Entwicklung eines Mausmodells (über das Einfügen gezielter Deletionen in das Mausgenom) geplant. Um dies zu ermöglichen, wurde in der vorliegenden Arbeit die Charakterisierung der orthologen Region bei der Maus vorgenommen. Zunächst wurden die orthologen Gene der Maus (*Letm1*, *Whsc1*, *Gen 43 (Whsc2h)*, *Gen 45* und *Pol4p*) identifiziert. Durch die Erstellung sowie die genaue Kartierung eines murinen genomischen P1/PAC-Klon-Contigs konnte gezeigt werden, daß die murinen Gene *Fgfr3*, *Letm1*, *Whsc1*, *Gen 43 (Whsc2h)*, *Gen 45* und *Pol4p* sowie einige weitere der überprüften EST-cDNA-Klone der Maus in einem durchgehenden Syntänieblock zwischen Mensch (*POLAP* bis *FGFR3*) und Maus (Mmu 5.20) enthalten sind, der in seiner genomischen Ausdehnung in etwa den Verhältnissen beim Menschen (zwischen *POLAP* und *FGFR3*) entspricht.

6. Literaturverzeichnis

- AASLAND, R.; GIBSON, T. J.; STEWART, A. F. (1995). The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.* 20: 56-9.
- ALLITTO, B. A.; MCCLATCHEY, A. I.; BARNES, G.; ALTHERR, M.; WASMUTH, J.; FRISCHAUF, A. M.; MACDONALD, M. E.; GUSELLA, J. (1992). Assay by polymerase chain reaction (PCR) of multi-allele polymorphisms in the Huntington's disease region of chromosome 4. *Mol. Cell Probes.* 6: 513-20.
- ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W. ; LIPMAN, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- ALTSCHUL, S. F.; BOGUSKI, M. S.; GISH, W.; WOOTTON, J. C. (1994). Issues in searching molecular sequence databases. *Nat. Genet.* 6: 119-129.
- ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- ALTHERR, M. R.; WRIGHT, T. J.; DENISON, K.; PEREZ-CASTRO, A. V.; JOHNSON, V. P. (1997). Delimiting the Wolf-Hirschhorn Syndrome critical region to 750 kilobase pairs. *Am. J. Med. Genet.* 71: 47-53.
- AMBROSE, C.; JAMES, M.; BARNES, G.; LIN, C.; BATES, G.; ALTHERR, M.; DUYAO, M.; GROOT, N.; CHURCH, D.; WASMUTH, J. J.; LEHRACH, H.; HOUSMAN, D.; BUCKLER, A.; GUSELLA, J. F.; MACDONALD, M. E. (1992). A novel G protein-coupled receptor kinase gene cloned from 4p16.3. *Hum. Molec. Genet.* 1: 697-703.
- ANDERSON, C. E.; WALLERSTEIN, R.; ZAMEROWSKI, S. T.; WITZLEBEN, C.; HOYER, J. R.; GIBAS, L.; JACKSON, L. G. (1997). Ring chromosome 4 mosaicism coincidence of Oligomeganephronia and signs of Seckel syndrome. *Am. J. Med. Genet.* 72: 281-285.
- ANTEQUERA, F. AND BIRD, A. (1993). Number of CpG islands and genes in human and in mouse. *Proc. Natl. Acad. Sci. USA* 90: 11995-11999.
- ANVRET, M.; NORDENSKJOLD, M.; STOLPE, L.; JOHANSSON, L.; BRONDUM-NIELSEN, K. (1991). Molecular analysis of 4p deletion associated with Wolf-Hirschhorn syndrome moving the 'critical segment' towards the telomere. *Hum. Genet.* 86: 481-483.
- ASENSIO, V. C.; CAMPBELL, I. L. (1999). Chemokines in the CNS: plurifunctional mediators in diverse states. *Trends Neurosci.* 22: 504-12.
- AUGUSSEAU, S.; JOUK, S.; JALBERT, P.; PRIEUR, M. (1986). DiGeorge syndrome and 22q11 rearrangements. *Hum. Genet.* 74: 206.
- AVET-LOISEAU, H.; LI, J.-Y.; FACON, T.; BRIGAUDEAU, C.; MORINEAU, N.; MALOISEL, F.; RAPP, M.-J.; TALMANT, P.; TRIMOREAU, F.; JACCARD, A.; HAROUSSEAU, J.-L.; BATAILLE, R. (1998). High incidence of translocations t(4;14)(q13;q32) and t(4;14)(p16;q32) in patients with plasma cell malignancies. *Cancer Res.* 58: 5640-5645.
- AVRAHAM, K. B.; GIVOL, D.; AVIVI, A.; YAYON, A.; COPELAND, N. G.; JENKINS, N. A. (1994). Mapping of murine fibroblast growth factor receptors refines regions of homology between mouse and human chromosomes. *Genomics* 21: 656-658.
- BABITCH, J. (1990). Channel hands. *Nature* 346: 321-2.
- BAIROCH, A.; BUCHER, P.; HOFMANN, K. (1997). The PROSITE database, its status in 1997. *Nucleic Acids Res.* 25: 217-221.
- BATES, G. P.; MACDONALD, M. E.; BAXENDALE, S.; SEDLACEK, Z.; YOUNGMAN, S.; ROMANO, D.; WHALEY, W. L.; ALLITTO, B. A.; POUSTKA, A.; GUSELLA, J. F.; LEHRACH, H. (1990). A yeast artificial chromosome telomere clone spanning a possible location of the Huntington Disease gene. *Am. J. Hum. Genet.* 46: 762-775.

- BATES, G. P.; VALDES, J.; HUMMERICH, H.; BAXENDALE, S.; LE PASLIER, D. L.; MONACO, A. P.; TAGLE, D.; MACDONALD, M. E.; ALTHERR, M.; ROSS, M. (1992).** Characterization of a yeast artificial chromosome contig spanning the Huntington's disease gene candidate region. *Nat. Genet.* 1: 180-187.
- BATTAGLIA, A.; CAREY, J. C. (1998).** Wolf-Hirschhorn syndrome and Pitt-Rogers-Danks syndrome. *Am. J. Med. Genet.* 75: 541.
- BATTAGLIA, A.; CAREY, J. C.; CEDERHOLM, P.; VISKOCHIL, D. H.; BROTHMAN, A. R.; GALASSO, C. (1999).** Natural history of Wolf-Hirschhorn syndrome: experience with 15 cases. *Pediatrics* 103: 830-836.
- BAYES, M.; GIORDANO, M.; BALCELLS, S.; GRINBERG, D.; VILAGELIU, L.; MARTINEZ, I.; AYUSO, C.; BENITEZ, J.; RAMOS-ARROYO, M.A.; CHIVELET, P.; SOLANS, T.; VALVERDE, D.; AMSELEM, S.; GOOSSENS, M.; BAIGET, M.; GONZALEZ-DUARTE, R.; BESMOND, C. (1995).** Homozygous tandem duplication within the gene encoding the beta-subunit of rod phosphodiesterase as a cause for autosomal recessive retinitis pigmentosa. *Hum. Mutat.* 5: 228-234.
- BAXENDALE, S.; MACDONALD, M. E.; MOTT, R.; FRANCIS, F.; LIN, C.; KIRBY, S. F.; JAMES, M.; ZEHETNER, G.; HUMMERICH, H.; VALDES, J.; COLLINS, F. S.; DEAVEN, L. J.; GUSELLA, J. F.; LEHRACH, H.; BATES, G. P. (1993).** A cosmid contig and high resolution restriction map of the 2 megabase region containing the Huntington's disease gene. *Nat. Genet.* 4: 181-186.
- BEELER, T.; GABLE, K.; ZHAO, C.; DUNN, T. (1994).** A novel protein, CSG2p, is required for Ca²⁺ regulation in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 269: 7279-7284.
- BELL, S. M.; ZUO, J.; MYERS, R. M.; KNOWLES, M. A. (1996).** Fluorescence in situ hybridization deletion mapping at 4p16.3 in bladder cancer cell lines refines the localisation of the critical interval to 30 kb. *Genes Chromosomes Cancer* 17:108-117.
- BELL, S. M.; SHAW, M.; JOU, Y.-S.; MYERS, R. M.; KNOWLES, M. A. (1997).** Identification and characterization of the human homologue of SH3BP2, an SH3 binding domain protein within a common region of deletion at 4p16.3 involved in bladder cancer. *Genomics* 44: 163-170.
- BELLONI, E.; MUENKE, M.; ROESSLER, E.; TRAVERSO, G.; SIEGEL-BARTELT, J.; FRUMKIN, A.; MITCHELL, H. F.; DONIS-KELLER, H.; HELMS, C.; HING, A. V.; HENG, H. H. Q.; KOOP, B.; MARTINDALE, D.; ROMMENS, J. M.; TSUI L.-C.; SCHERER, S. W. (1996).** Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly. *Nat. Genet.* 14: 353-356.
- BELLUS, G. A.; MCINTOSH, I.; SMITH, E. A.; AYLSWORTH, A. S.; KAITILA, I.; HORTON, W. A.; GREENHAW, G. A.; HECHT, J. T.; FRANCOMANO, C. A. (1995).** A recurrent mutation in the tyrosine kinase domain of fibroblast growth factor receptor 3 causes hypochondroplasia. *Nat. Genet.* 10: 357-359.
- BENTON, W. D. UND DAVIS, R. W. (1977).** Screening λ gt recombinant clones by hybridization to single plaques in situ. *Science* 196: 180-182.
- BERCHTOLD, M. W.; BRINKMEIER, H.; MUNTENER, M. (2000).** Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease. *Physiol. Rev.* 80: 1215-1265.
- BERGSAGEL, P. L.; CHESI, M.; NARDINI, E.; BRENTS, L. A.; KIRBY, S. L.; KUEHL, W. M. (1996).** Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc. Natl. Acad. Sci. USA* 93: 13931-13936.
- BIRD, A. P. (1987).** CpG islands as gene markers in vertebrate nucleus. *Trends Genet.* 3: 342-347.
- BIRNBOIM, H. C. UND DOLY, J. (1979).** A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7: 1513-1523.
- BIRNSTIEL, M. L.; BUSSLINGER, M.; STRUB, K. (1985).** Transcription termination and 3' processing: the end is in site! *Cell* 41: 349-359.
- BOGUSKI, M. S.; LOWE, T. M.; TOLSTOSHEV, C. M. (1993).** dbEST-database for "expressed sequence tags". *Nat. Genet.* 4: 332-333.
- BORSANI, G., BALLABIO, A.; BANFI, S. (1998).** A practical guide to orient yourself in the labyrinth of genome databases. *Hum. Mol. Genet.* 7: 1641-1648.
- BOUCKSON-CASTAING, M., MOUDJOU, M.; FERGUSON, D. J. P.; MUCKLOW, S.; BELKAID, Y.; MILON, G.; CROCKER, P. R. (1996).** Molecular characterization of Ninein, a new coiled-coil protein of the centrosome. *J. Cell Science* 109: 179-190.

- BRAITHWAITE, D. K. UND ITO, J.** (1993). Compilation, alignment and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res.* 21: 787-802.
- BRINI, M. UND CARAFOLI, E.** (2000). Calcium signalling: a historical account, recent developments and future perspectives. *Cell. Mol. Life Sci.* 57: 354-370.
- BROWN, L. J.; MACDONALD, M. J.; LEHN, D. A.; MORAN, S. M.** (1994). Sequence of rat mitochondrial glycerol-3-phosphate dehydrogenase cDNA. Evidence for EF-hand calcium-binding domains. *J. Biol. Chem.* 269: 14363-14366.
- BUCAN, M.; ZIMMER, M.; WHALEY, W. L.; POUSTKA, A.; YOUNGMAN, S.; ALLITTO, B. A.; ORMONDROYD, E.; SMITH, B.; POHL, T. M.; MACDONALD, M.; BATES, G. P.; RICHARDS, J.; VOLINA, S.; GILLMAN, T. C.; SEDLACEK, Z.; COLLINS, F. S.; WASMUTH, J. J.; SHAW, D. J.; GUSELLA, J. F.; FRISCHAUF, A.-M.; LEHRACH, H.** (1990). Physical maps of 4p16.3, the area expected to contain the Huntington Disease mutation. *Genomics* 6: 1-15.
- BUCKLER, A. J.; CHANG, D. D.; GRAW, S.L.; BROOK, J. D.; HABER, D. A.; SHARP, P. A.; HOUSMAN, D. E.** (1991). Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci. USA* 88: 4005-4009.
- BUDARF, M. L.; COLLINS, J.; GONG, W.; ROE, B.; WANG, Z.; SELLINGER, B.; MICHAUD, D.; DRISCOLL, D.; EMANUEL, B. S.** (1995). Cloning a balanced translocation associated with DiGeorge syndrome and identification of disrupted candidate genes. *Nat. Genet.* 10: 269-288.
- BUDARF, M. L.; EMANUEL, B. S.** (1997). Progress in the autosomal aneusomy syndromes (SASs): single or multi-locus disorders? *Hum. Mol. Genet.* 6: 1657-1665.
- BUITING, K.; KAYA-WESTERLOH, S.; HORSTHEMKE, B.** (1996). A pseudogene for the human ribosomal protein L5 (RPL5P1) maps within an intron of the SNRPN transcription unit on chromosome 15. *Cytogenet. Cell Genet.* 75: 224-226.
- BUNGE, S.; KLEIJER, W. J.; STEGLICH, C.; BECK, M.; ZUTHER, C.; MORRIS, C. P.; SCHWINGER, E.; HOPWOOD, J. J.; SCOTT, H. S.; GAL, A.** (1994). Mucopolysaccharidosis type I: identification of 8 novel mutations and determination of the frequency of the two common alpha-L-iduronidase mutations (W402X and Q70X) among European patients. *Hum. Molec. Genet.* 3: 861-866.
- BURFOOT, M. S.; CAMPBELL, R. D.** (1994). Improved method of gene detection using exon amplification. *Nucleic Acids Res.* 22: 5510-5511.
- BURGE, C. AND KARLIN, S.** (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268: 78-94.
- BURGE, C. B. AND KARLIN, S.** (1998) Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.* 8: 346-354.
- BURGESS, D. L.; JONES, J. M.; MEISLER, M. H.; NOEBELS, J. L.** (1997). Mutation of the Ca²⁺ channel-subunit gene Cchb4 is associated with ataxia and seizures in the lethargic (lh) mouse. *Cell* 88: 385-392.
- BURSET, M. UND GUIGO, R.** (1996). Evaluation of gene structure prediction programs. *Genomics* 34: 353-67.
- CARON, H.; VAN SLUIS, P.; BUSCHMAN, R.; PEREIRA, D. O.; TANQUE, R.; MAES, P.; BEKS, L.; DE KRAKER, J.; VOUTE, P. A.; VERGNAUD, G.; WESTERVELD, A.; SLATER, R.; VERSTEEG, R.** (1996). Allelic loss of the short arm of chromosome 4 in neuroblastoma suggests a novel tumour suppressor gene locus. *Hum. Genet.* 97: 834-837.
- CELIO, M. R.** (1996). Introduction to EF-hand calcium-binding proteins. Guidebook to the Calcium-Binding Proteins, M. R. Celio, T. L. Pauls, B. Schwaller, Eds., Oxford Univ. Press New York: 15-20.
- CHEN, K. S.; MANIAN, P.; KOEUTH, T.; POTOCKI, L.; ZHAO, Q.; CHINAULT, A. C.; LEE, C. C. ET AL.** (1997). Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat. Genet.* 17: 154-163.
- CHESI, M.; NARDINI, E.; BRENTS, L. A.; SCHROCK, E.; RIED, T.; KUEHL, W. M.; BERGSAGEL, P. L.** (1997). Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. *Nat. Genet.* 16: 260-264.
- CHESI, M.; NARDINI, E.; LIM, R. S. C.; SMITH, K. D.; KUEHL, W. M.; BERGSAGEL, P. L.** (1998). The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood* 92: 3025-3034.

- CHITAYAT, D.; RUVALCABA, R. H.; BABUL, R.; TESHIMA, I. E.; POSNICK, J. C.; VEKEMANS, M. J.; SCARPELLI, H.; THULINE, H. (1995). Syndrome of proximal interstitial deletion 4p15: report of three cases and review of the literature. *Am. J. Med. Genet.* 55: 147-154.
- CHONG, S. S.; LO NIGRO, C.; ROSCHKE, A. V.; TANIGAMI, A.; PACK, S. D.; SMITH, A. C. M.; CARROZZO, R.; DOBYNS, W. B.; LEDBETTER, D. H. (1996). Point mutations and an intragenic deletion in three ILS patients confirm LIS1 as the lissencephaly causative gene in isolated lissencephaly sequence and Miller-Dieker syndrome. (Abstract) *Am. J. Hum. Genet.* 59 (suppl.): A23.
- CHOU, P. Y. UND FASMAN, G. D. (1974). Prediction of protein conformation. *Biochemistry* 13: 222-245.
- CHURCH, G. M. AND GILBERT, W. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81: 1991-1995.
- CHURCH, D. M.; STOTLER, C. J.; RUTTER, J. L., MURRELL, J. R.; TROFATTER, J. A.; BUCKLER, A. J. (1994). Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nat. Genet.* 6: 98-104.
- CHURCH, D. M.; BUCKLER, A. J. (1999). Gene identification by exon amplification. *Methods Enzymol.* 303: 83-99.
- CLASSON, B. J. UND COVERDALE, L. (1994). Mouse stem cell antigen Sca-2 is a member of the Ly-6 family of cell surface proteins. *Proc. Natl. Acad. Sci. USA.* 91: 5296-5300.
- CLAVERIE, J. M. (1997). Computational methods for the identification of genes in vertebrate genomic sequences. *Hum. Mol. Genet.* 6: 1735-1744.
- CLEMENS, M.; MARTSOLF, J. T.; ROGERS, J.G.; MOWERY-RUSHTON, P.; SURTI, U.; MCPHERSON, E. (1996). Pitt-Rogers-Danks syndrome: the result of a 4p microdeletion. *Am. J. Med. Genet.* 66: 95-100.
- COLLINS, C.; SCHAPPERT, K.; HAYDEN, M. R. (1992). The genomic organization of a novel regulatory myosin light chain gene (MYL5) that maps to chromosome 4p16.3 and shows different patterns of expression between primates. *Hum. Molec. Genet.* 1: 727-733.
- CONAWAY, J. W.; UND CONAWAY, R. C. (1999). Transcription elongation and human disease. *Annu. Rev. Biochem.* 68: 301-319.
- CONAWAY, J. W.; SHILATIFARD, A.; DVIR, A.; CONAWAY, R. C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* 25: 375-380.
- CROSS, S. H. UND BIRD, A. P. (1995): CpG islands and genes. *Curr. Opin. Genet. Dev.* 5: 309-314.
- CZEMPINSKI, K.; ZIMMERMANN, S.; EHRHARDT, T.; MULLER-ROBER, B. (1997). New structure and function in plant K⁺ channels: KCO1, an outward rectifier with a steep Ca²⁺ dependency. *EMBO J.* 16: 2565-2575.
- DALLAPICCOLA, B.; MASTROIACOVO, P. P.; MONTALI, E.; SOMMER, A. (1977). Trisomy 4p: five new observations and overview. *Clin. Genet.* 12: 344-356.
- DALLAPICCOLA, B.; MANDICH, P.; BELLONE, E.; SELICORNI, A.; MOKIN, V.; AJMAR, F.; NOVELLI, G. (1993). Parental Origin of chromosome 4p deletion in Wolf-Hirschhorn Syndrome. *Am. J. Med. Genet.* 47: 921-924.
- DELARUE, M.; POCH, O.; TORDO, N.; MORAS, D.; ARGOS, P. (1990). An attempt to unify the structure of polymerases. *Protein Eng.* 3: 461-467.
- DE LECEA, L.; DEL RIO, J. A.; CRIADO, J. R.; ALCANTARA, S.; MORALES, M.; DANIELSON, P. E.; HENRIKSEN, S. J.; SORIANO, E.; SUTCLIFFE, J. G. (1997a). Cortistatin is expressed in a distinct subset of cortical interneurons. *J. Neurosci.* 17: 5868-5880.
- DE LECEA, L.; RUIZ-LOZANO, P.; DANIELSON, P. E.; PEELLE-KIRLEY, J.; FOYE, P. E.; FRANKEL, W. N.; SUTCLIFFE, J. G. (1997b). Cloning, mRNA expression, and chromosomal mapping of mouse and human preprocortistatin. *Genomics* 42: 499-506.
- DESILVA, U.; MASSA, H.; TRASK, B. J.; GREEN, E. D. (1999). Comparative mapping of the region of human chromosome 7 deleted in williams syndrome. *Genome Res.* 9: 428-436.
- DIMEGLIO, L. A.; STEINMETZ, R.; PESCOVITZ, O. H. (1998). Variations on a theme: testis-derived neuropeptide hormones. *Eur. J. Endocrinol.* 139: 361-362.
- DONNAI, D. (1996) Editorial comment: Pitt-Rogers-Danks syndrome and Wolf-Hirschhorn syndrome. *Am. J. Med. Genet.* 66: 101-103.

- DUNHAM, I.; SHIMIZU, N. M.; ROE, B. A.; CHISSOE, S.; HUNT, A. R.; COLLINS, J. E.; BRUSKIEWICH, R.; BEARE, D. M.; CLAMP, M.; SMINK, L. J.; AINSCOUGH, R.; ALMEIDA, J. P.; BABBAGE, A.; BAGGULEY, C.; BAILEY, J.; BARLOW, K.; BATES, K. N.; BEASLEY, O.; BIRD, C. P.; BLAKEY, S.; BRIDGEMAN, A. M.; BUCK, D.; BURGESS, J.; BURRILL, W. D.; O'BRIEN, K. P.; ET AL. (1999). The DNA sequence of human chromosome 22. *Nature* 402: 489-495.
- EDELMANN, L.; PANDITA, R. K.; SPITERI, E.; FUNKE, B.; GOLDBERG, R.; PALANISAMY, N.; CHAGANTI, R. S.; MAGENIS, E.; SHPRINTZEN, R. J.; MORROW, B. E. (1999). A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum. Mol. Genet.* 8: 1157-1167.
- EDWARDS, R. H.; SELBY, M. J.; RUTTER, W. J. (1986). Differential RNA splicing predicts two distinct nerve growth factor precursors. *Nature* 319: 784-787.
- EDWARDS, R. H.; SELBY, M. J.; GARCIA, P. D.; RUTTER, W. J. (1988). Processing of the native nerve growth factor precursor to form biologically active nerve growth factor. *J. Biol. Chem.* 263: 6810-6815.
- EHRICH, E.; CRAIG, A.; POUSTKA, A.; FRISCHAUF, A.-M.; LEHRACH, H. (1987). A family of cosmid vectors with the multi-copy R6K origin. *Gene* 57: 229-237.
- EICHLER, E. E.; HOFFMAN, S. M.; ADAMSON, A. A.; GORDON, L. A.; MCCREADY, P.; LAMERDIN, J. E.; MOHRENWEISER, H. W. (1998). Complex β -satellite repeat structures and the expansion of the zinc gene finger cluster in 19p12. *Genome Res.* 8: 791-808.
- ELDER, P. A.; BELL, S. M.; KNOWLES, M. A. (1994). Deletion of two regions on chromosome 4 in bladder carcinoma: definition of a critical 750kB region at 4p16.3. *Oncogene* 9: 3433-3436.
- ELLIOTT, J.; COULTER-MACKIE, M. B.; JUNG, J. H.; RODENHISER, D. I.; SINGH, S. M. (1991). A method for transforming lymphocytes from very small blood volumes suitable for pediatric samples. *Hum. Genet.* 86: 615-616.
- ENDELE, S.; ZABEL, B.; WINTERPACHT, A. (1996). Assignment of the human diacylglycerol kinase 4 (DAGK4) gene to chromosome 4p16.3. *Genomics* 33: 145-146.
- ENDELE, S.; FUHRY, M.; PAK, S. J.; ZABEL, B. U.; WINTERPACHT, A. (1999). LETM1, a novel gene encoding a putative EF-hand Ca(2+)-binding protein, flanks the Wolf-Hirschhorn syndrome (WHS) critical region and is deleted in most WHS patients. *Genomics* 60: 218-225.
- ESTABROOKS, L. L.; LAMB, A. N.; KIRKMAN, H. N.; CALLANAN, N. P.; RAO, K. W. (1992). A molecular deletion of distal chromosome 4p in two families with a satellited chromosome 4 lacking the Wolf-Hirschhorn syndrome phenotype. *Am. J. Hum. Genet.* 51: 971-978.
- ESTABROOKS, L. L.; LAMB, A. N.; AYLSWORTH, A. S.; CALLANAN, N. P.; RAO, K. W. (1994). Molecular characterisation of chromosome 4p deletions resulting in Wolf-Hirschhorn syndrome. *J. Med. Genet.* 31: 103-107.
- ESTABROOKS, L. L.; RAO, K. W.; DRISCOLL, D. A.; CRANDALL, B. F.; DEAN, C.S.; IKONEN, E.; KORF, B.; AYLSWORTH, A. S. (1995a). Preliminary Map of Chromosome 4p16 based on 4p deletions. *Am. J. Med. Genet.* 57: 581-586.
- ESTABROOKS, L. L.; BREG, W. R.; HAYDEN, M. R.; LEDBETTER, D. H.; MYERS, R. M.; WYANDT, H. E.; YANG-FENG, T. L.; HIRSCHHORN, K. (1995b). Summary of the 1993 ASHG Ancillary Meeting „Recent Research on Chromosome 4p syndromes and genes“. *Am. J. Med. Genet.* 55: 453-458.
- EWART, A. K.; MORRIS, C. A.; ATKINSON, D.; JIN, W.; STERNES, K.; SPALLONE, P.; STOCK, A. D.; LEPPERT, M.; KEATING, M. T. (1993). Hemizygoty at the elastin locus in a developmental disorder, Williams syndrome. *Nature Genet.* 5: 11-16.
- FEINBERG, A. P. UND VOGELSTEIN, B (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specificity activity. *Anal. Biochem.* 132: 6-13.
- FERLINI, A.; MUNTONI, F. (1998). The 5' region of intron 11 of the dystrophin gene contains target sequences for mobile elements and three overlapping ORFs. *Biochem. Biophys. Res. Commun.* 242: 401-406.
- FICKETT, J. W.; HATZIGEORGIOU, A. G. (1997). Eukaryotic promoter recognition. *Genome Res.* 7: 861-878.
- FLINT, J.; BATES, G. P.; CLARK, K.; DORMAN, A.; WILLINGHAM, D.; ROE, B., A.; MICKLEM, G., HIGGS, D. R.; LOUIS, E. J. (1997). Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains. *Hum. Mol. Gen.* 6: 1305-1314.

- FOSTER, J. W.; DOMINGUEZ-STEGLICH, M. A.; GUIOLI, S.; KWOK, C.; WELLER, P. A.; STEVANOVIC, M.; WEISSENBACH, J.; MANSOUR, S.; YOUNG, I. D.; GOODFELLOW, P. N. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in a SRY-related gene. *Nature* 372: 525-530.
- FRACCHIOLLA, N. S.; LUMINARI, S.; BALDINI, L.; LOMBARDI, L.; MAIOLO, A. T.; NERI, A. (1998). FGFR3 gene mutations associated with human skeletal disorders occur rarely in Multiple Myeloma. *Blood* 92: 2987-2989.
- FRANCKE, U.; ARIAS, D. E.; NYHAN, W. L. (1977). Proximal 4p-deletion: phenotype differs from classical 4p-syndrome. *J. Pediatrics* 90: 250-252.
- FRANCKE, U. (1999). Williams-Beuren syndrome: genes and mechanisms. *Hum. Mol. Genet.* 8: 1947-1954.
- FRANCOMANO, C. A.; ORTIZ DE LUNA, R. I.; HEFFERON, T. W.; BELLUS, G. A.; TURNER, C. E.; TAYLOR, E.; MEYERS, D. A.; BLANTON, S. H.; MURRAY, J. C.; MCINTOSH, I.; ET AL. (1994). Localization of the achondroplasia gene to the distal 2.5 Mb of human chromosome 4p. *Hum. Mol. Genet.* 3: 787-792.
- GAHLMANN, R.; WADE, R.; GUNNING, P.; KEDES, L. (1988). Differential expression of slow and fast skeletal muscle troponin C. Slow skeletal muscle troponin C is expressed in human fibroblasts. *J. Mol. Biol.* 201: 379-391.
- GAL, A.; ORTH, U.; BAEHR, W.; SCHWINGER, E.; ROSENBERG, T. (1994). Heterozygous missense mutation in the rod cGMP phosphodiesterase beta-subunit gene in autosomal dominant stationary night blindness. *Nature Genet.* 7: 64-68.
- GALE, C.; FINKEL, D.; TAO, N.; MEINKE, M.; MCCLELLAN, M.; OLSON, J.; KENDRICK, K.; HOSTETTER, M. (1996). Cloning and expression of a gene encoding an integrin-like protein in *Candida albicans*. *Proc. Natl. Acad. Sci. USA* 93: 357-361.
- GALILI, N.; BALDWIN, H. S.; LUND, J.; REEVES, R.; GONG, W.; WANG, Z.; ROE, B. A.; EMANUEL, B. S.; NAYAK, S.; MICKANIN, C.; BUDARF, M. L.; BUCK, C. A. (1997). A region of mouse chromosome 16 is syntenic to the DiGeorge, velocardiofacial syndrome minimal critical region. *Genome Res.* 7: 17-26.
- GANDELMANN, K.; GIBSON, L.; STEPHEN-MEYN, M.; YANG-FENG, T. L. (1992). Molecular definition of the smallest region of overlap in the Wolf-Hirschhorn syndrome. *Am. J. Hum. Genet.* 51: 571-578.
- GAVEL, Y. UND VAN HEINJE, G. (1990). Cleavage-site motifs in mitochondrial targeting peptides. *Protein Eng.* 4: 33-37.
- GHOSH, A. UND GREENBERG, M. E. (1995). Calcium signaling in neurons: molecular mechanisms and cellular consequences. *Science* 268: 239-247.
- GLUZMAN, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* 23: 175-182.
- GONG, W.; EMANUEL, B. S.; COLLINS, J.; KIM, D. H.; WANG, Z.; CHEN, F.; ZHANG, G.; ROE, B.; BUDARF, M. L. (1996). A transcription map of the DiGeorge and velo-cardio-facial syndrome minimal critical region on 22q11. *Hum. Mol. Genet.* 5: 789-800.
- GONZALEZ, C. H.; SOMMER, A.; MEISNER, L. F.; ELEJALDE, B. R.; OPITZ, J. M. (1977). The trisomy 4p syndrome: case report and review. *Am. J. Med. Genet.* 1: 137-156.
- GOODWIN, N. C.; KNEECE, J. C.; SCHIMENTI, J. C. (1998). An embryonic stem cell deletion bank for producing deletion complexes in mice. 12th International Mouse Genome Conference, Garmisch-Partenkirchen 29.09.-03.10.98, G6.
- GREENER, T.; ZHAO, X.; NOJIMA, H.; EISENBERG, E.; GREENE, L. E. (2000). Role of cyclin G-associated kinase in uncoating clathrin-coated vesicles from non-neuronal cells. *J. Biol. Chem.* 275: 1365-1370.
- GRIMM, T. (1998). Isolierung und Charakterisierung von Transkripten der Chromosomenregion 4p16.3. Diplomarbeit, Fachbereich Biologie, Johannes-Gutenberg Universität Mainz.
- GROSSON, C. L. S.; MACDONALD, M. E.; DUYAO, M. P.; AMBROSE, C. M.; ROFFLER-TARLOV, S.; GUSELLA, J. F. (1994). Synteny conservation of the Huntington's disease gene and surrounding loci on mouse chromosome 5. *Mamm. Genome* 5: 424-428.
- GU, Y.; NAKAMURA, T.; ALDER, H.; PRASAD, R.; CANAANI, O.; CIMINO, G.; CROCE, C. M.; CANAANI, E. (1992). The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila trithorax*, to the AF-4 gene. *Cell* 71: 701-708.

- GUSELLA, J. F.; ALTHERR, M. R.; MCCLATCHEY, A. I.; DOUCETTE-STAMM, L. A.; TAGLE, D.; PLUMMER, S.; GROOT, N.; BARNES, G.; HUMMERICH, H.; COLLINS, F. S.; HOUSMAN, D. E.; LEHRACH, H.; MACDONALD, M. E., BATES, G.; WASMUTH, J. J. (1992). Sequence-tagged sites (STSs) spanning 4p16.3 and the Huntington Disease candidate region. *Genomics* 13: 75-80.
- GUSELLA, J. F.; WEXLER, N. S.; CONNEALLY, P. M.; NAYLOR, S. L.; ANDERSON, M. A.; TANZI, R. E.; WATKINS, P. C.; OTTINA, K.; WALLACE, M. R.; SAKAGUCHI, A. Y.; YOUNG, A. B.; SHOULSON, I.; BONILLA, E.; MARTIN, J. B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 306: 234-238.
- HANAHAN, D. (1985). Techniques for transformation in *Escherichia coli*. In: DNA cloning I; Glover, D. M. (ed.). IRL Press, Washington D. C.: 109-135.
- HANLEY-LOPEZ, J.; ESTABROOKS, L. L.; STIEHM, R. (1998). Antibody deficiency in Wolf-Hirschhorn syndrome. *J. Pediatr.* 133: 141-143.
- HANSON, R. D.; HESS, J. L.; YU, B. D.; ERNST, P.; VAN LOHUIZEN, M.; BERNS, A.; VAN DER LUGT, N. M.; SHASHIKANT, C. S.; RUDDLE, F. H.; SETO, M.; KORSMEYER, S. J. (1999). Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc. Natl. Acad. Sci USA*. 96: 14372-14377.
- HARDISON, R. C.; OELTJEN, J.; MILLER, W. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res.* 7:959-966.
- HARTMANN, E.; RAPOPORT, T. A.; LODISH, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci. USA* 86: 5786-5790.
- HATTORI, M., FUJIYAMA, A.; TAYLOR, T. D.; WATANABE, H.; YADA, T.; PARK, H. S.; TOYODA, A.; ISHII, K.; TOTOKI, Y.; CHOI, D. K.; SOEDA, E.; OHKI, M.; TAKAGI, T.; SAKAKI, Y.; TAUDIEN, S.; BLECHSCHMIDT, K.; POLLEY, A.; MENZEL, U.; DELABAR, J.; KUMPF, K.; LEHMANN, R.; PATTERSON, D.; REICHWALD, K.; RUMP, A.; SCHILLHABEL, M.; SCHUDY, A. (2000). The DNA sequence of human chromosome 21. The chromosome 21 mapping and sequencing consortium. *Nature* 405: 311-319.
- HEIZMANN, C. W., AND BRAUN K. (1992). Changes in Ca²⁺-binding proteins in human neurodegenerative disorders. *Trends Neurosci.* 15: 259-264.
- HENIKOFF, S. UND HENIKOFF, J. G. (1991) Automated assembly of protein blocks for database searching, *Nucleic Acids Res.* 19: 6565-6572.
- HERMANN, P. (1999). Charakterisierung von cDNAs aus dem Wolf-Hirschhorn kritischen Bereich (WHSCR) auf dem Chromosom 4p16.3. Diplomarbeit, Fachbereich Biologie, Johannes-Gutenberg Universität Mainz.
- HILLIER, L.; LENNON, G.; BACKER, M.; BONALDO, M. F.; CHIAPPELLI, S.; CHISSOE, N.; DIETRICH, N.; DUBUQUE, T.; FAVELLO, A.; GISH, W. et al. (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Res.* 6: 807-828.
- HIROTSUNE, S.; PACK S. D.; CHONG, S. S.; ROBBINS, C. M.; PAVAN, W. J.; LEDBETTER, D. H.; WYNshaw-BORIS, A. (1997). Genomic organization of the murine Miller-Dieker/Lissencephaly region: Conservation of linkage with the human region. *Genome Res.* 7: 625-634.
- HIROKAWA, T.; BOON-CHENG, S.; MITAKU, S. (1998). SOSUI: Classification and secondary structure prediction system for membrane proteins. *Bioinformatics formerly* 14: 378-379.
- HIRSCHHORN, K.; COOPER, H. L.; FIRSCHEIN, I. L. (1965). Deletion of short arms of chromosome 4-5 in a child with defects of midline fusion. *Hum. Genet.* 1: 479-482.
- HISAMA, F. M.; OSHIMA, J.; YU, C.-E.; FU, Y.-H.; MULLIGAN, J.; WEISSMAN, S. M.; SCHELLENBERG, G. D. (1998): Comparison of methods for identifying transcription units and transcription map of the Werner syndrome gene region. *Genomics* 52: 352-357.
- HÖKFELT, T.; BROBERGER, C.; DAVID XU, Z.-Q.; SERGEYEV, V.; UBINK, R.; DIEZ, M. (2000). Neuropeptides-an overview. *Neuropharmacology* 39: 1337-1356.
- HOFMANN, K. UND STOFFEL, W. (1993). TMBASE- a database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler* 374: 166.
- HOLLWAY, G. E.; PHILLIPS, H. A.; ADES, L. C.; HAAN, E. A.; MULLEY, J. C. (1995). Localization of craniosynostosis Adelaide type to 4p16. *Hum. Molec. Genet.* 4: 681-683.

- HOLLWAY, G. E.; SUTHERS, G. K.; BATTESE, K. M.; TURNER, A. M.; DAVID, D. J.; MULLEY, J. C. (1998): Deafness due to pro250-to-arg mutation of FGFR3. *Lancet* 351: 877-878.
- HONG, Y.; TAKANO, M.; LIU, C. M.; GASCH, A.; CHYE, M. L.; CHUA, N. H. (1996). Expression of three members of the calcium-dependent protein kinase gene family in *Arabidopsis thaliana*. *Plant. Mol. Biol.* 30: 1259-1275.
- HOUSSA, B.; SCHAAP, D.; VAN DER WAL, J.; GOTO, K.; KONDO, H.; YAMAKAWA, A.; SHIBATA, M., TAKENAWA, T.; VAN BLITTERSWIJK, W. J. (1997). Cloning of a Novel Human Diacylglycerol Kinase (DGK) Containing Three Cysteine-rich Domains, a Proline-rich Region, and a Pleckstrin Homology Domain with an Overlapping Ras-associating Domain. *J. Biol. Chem.* 272: 10422-10428.
- HU, N.; ROTH, M. J.; POLYMERPOLOUS, M.; TANG, Z.-Z.; EMMERT-BUCK, M. R.; WANG, Q.-H.; GOLDSTEIN, A. M.; FENG, S.-S.; DAWSEY, S. M.; DING, T.; ZHUANG, Z.-P.; HAN, X.-Y.; RIED, T.; GIFFEN, C.; TAYLOR, P. R. (2000). Identification of novel regions of allelic loss from a genomewide scan of esophageal squamous-cell carcinoma in a high-risk Chinese population. *Genes, Chromosomes and Cancer* 27: 217-228.
- HUANG, N.; VOM BAUR, E.; GARNIER, J. M.; LEROUGE, T.; VONESCH, J. L.; LUTZ, Y.; CHAMBON, P.; LOSSON, R. (1998). Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. *EMBO J.* 17: 3398-3412.
- IBANEZ, C. F. (1998). Emerging themes in structural biology of neurotrophic factors. *Trends Neurosci.* 21: 438-444.
- IGARASHI, M.; NAGATA, A.; JINNO, S.; SUTO, K.; OKAYAMA, H. (1991). Wee1(+)-like gene in human cells. *Nature* 353: 80-83.
- IOANNOU, P. A.; AMEMIYA, C. T.; GARNES, J.; KROISEL, P. M.; SHIZUYA, H.; CHEN, C.; BATZER, M. A.; DE JONG, P. J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.* 6: 84-89.
- JOHNSON, V. P.; MULDER, R. D.; HOSEN, R. (1976). The Wolf-Hirschhorn (4p-) syndrome. *Clin. Genet.* 10: 104-112.
- JOHNSON, V. P.; ALTHERR, M. R.; BLAKE, J. M.; KEPPEL, L. D. (1994). FISH detection of Wolf-Hirschhorn syndrome: exclusion of D4F26 as critical site. *Am. J. Med. Genet.* 52: 70-74.
- JOYCE, C. M.; KELLEY, W. S.; GRINDLEY, N. D. (1982). Nucleotide sequence of the *Escherichia coli* polA gene and primary structure of DNA polymerase I. *J. Biol. Chem.* 257: 1958-1964.
- KANAOKA, Y.; KIMURA, S. H.; OKAZAKI, I.; IKEDA, M.; NOJIMA, H. (1997). GAK: a cyclin associated kinase contains a tensin/auxilin-like domain. *FEBS Letters* 402: 73-80.
- KANT, S. G.; HAERINGEN, A. V.; BAKKER, E.; STEC, I.; DONNAI, D.; MOLLEVANGER, P.; BEVERSTOCK, G. C.; LINDEMANN-KUSSE, M. C.; VAN OMMEN, G.-J. (1997). Pitt-Rogers-Danks syndrome and Wolf-Hirschhorn syndrome are caused by a deletion in the same region on chromosome 4p16.3. *J. Med. Genet.* 34: 569-572.
- KAWASAKI, H. AND KRETSINGER, R. H. (1994). Calcium-binding proteins. 1. EF-hands. *Protein Profile* 1: 343-517.
- KAWASAKI, H.; NAKAYAMA, S.; KRETSINGER, R. H. (1998). Classification and evolution of EF-hand proteins. *BioMetals* 11: 277-295.
- KEEGAN, K.; JOHNSON, D. E.; WILLIAMS, L. T.; HAYMAN, M. J. (1991). Isolation of an additional member of the fibroblast growth factor receptor family, FGFR-3. *Proc. Nat. Acad. Sci. USA* 88: 1095-1099.
- KIMBER, W. L.; HIROTSUNE, S.; SUTHERLAND, H. F.; PACK, S.; GARRETT, L.; CHEN, A.; SCAMBLER, P. J.; WYNshaw-BORIS, A. (1997). Using mouse models for understanding the genetic basis of human microdeletion syndromes. *Am. J. Hum. Genet.* 61: Suppl., A155.
- KIMBER, W. L.; HSIEH, P.; HIROTSUNE, S.; YUVA-PAYLOR, L.; SUTHERLAND, H. F.; CHEN, A.; RUIZ-LOZANO, P.; HOOGSTRATEN-MILLER, S. L.; CHIEN, K. R.; PAYLOR, R.; SCAMBLER, P. J.; WYNshaw-BORIS, A. (1999). Deletion of 150 kb in the minimal DiGeorge/velocardiofacial syndrome critical region in mouse. *Hum. Mol. Genet.* 8: 2229-2237.
- KIMURA, S. H.; TSURUGA, H.; YABUTA, N.; ENDO, Y.; NOJIMA, H. (1997). Structure, expression, and chromosomal location of human GAK. *Genomics* 44: 197-187.

- KLEINJAHN, D.-J. UND VAN HEYNINGEN, V.** (1998). Position effect in human genetic disease. *Hum. Mol. Genet.* 7: 1611-1618.
- KLEVER, M.; GROND-GINSBACH, C.; SCHERTHAN, H.; SCHROEDER-KURTH, T. M.** (1991). Chromosomal in situ suppression hybridization after Giemsa banding. *Hum. Genet.* 86: 484-486.
- KOHN, W. D.; MANT, C. T.; HODGES, R. S.** (1997). α -helical protein assembly motifs. *J. Biol. Chem.* 272: 2583-2586.
- KOZAK, M.** (1996): Interpreting sequences: some insights from studies on translation. *Mamm. Genome* 7: 563-574.
- KOZAK, C. A.; STEPHENSON, D. A.** (1993): Encyclopedia of the mouse genome III. October 1993. Mouse chromosome 5. *Mamm. Genome* (Suppl.) 4: 72-87.
- KOZAK, C. UND STEPHENSON, D.** (1997). Mouse chromosome 5. *Mamm. Genome* 7: 80-99.
- KOZAK, C. A. UND STEPHENSON, D. A.** (1999). Mouse chromosome 5. *Mamm. Genome.* 10: 944.
- KOZMA, C.; CHONG, S. S.; MECK, J. M.** (1999). Interstitial deletion of 4p15.32p16.3 in a boy with minor anomalies, hearing loss, borderline intelligence, and oligodontia. *Am. J. Med. Genet.* 86: 316-320.
- KRISTENSEN, P.; ERIKSEN, J.; BLASI, F.; DANOE, K.** (1991). Two alternatively spliced mouse urokinase receptor mRNAs with different histological localization in the gastrointestinal tract. *J. Cell Biol.* 115: 1763-1771.
- LADUNGA, I.** (2000). Large-scale predictions of secretory proteins from mammalian genomic and EST sequences. *Curr. Opin. Biotechnol.* 11: 13-18.
- LEDBETTER, D. H.; BALLABIO, A.** (1995). Molecular cytogenetics of contiguous gene syndrome: Mechanisms and consequences of gene dosage imbalance. In: *The metabolic basis of inherited disease* (ed. C. Scriver, A.L. Beaudet, W.S. Sly und D.Valle), 811-839. McGraw-Hill, New York, NY.
- LEE, P. L.; CONNELL, C. R.; WOO, S. L.; CHENG, R. D.; MCARDLE, B. F.; FULLER, C. W.; HALLORAN, N. D.; WILSON, R. K.** (1992). DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effects of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments. *Nucleic Acids Res.* 20: 2471-2483.
- LEMIEUX, N.; DUTRILLAUX, B.; VIEGAS-PEQUINOT, E.** (1992). A simple method for simultaneous R- or G-Banding and fluorescence in situ hybridization of small copy genes. *Cytogenet. Cell Genet.* 59: 311-312.
- LESPERANCE, M. M.; GRUNDFAST, K. M.; ROSENBAUM, K. N.** (1998). Otologic manifestations of Wolf-Hirschhorn syndrome. *Arch. Otolaryngol. Head Neck Surg.* 124: 193-196.
- LETTS, V. A.; FELIX R.; BIDDLECOME, G. H.; ARIKATH, J.; MAHAFFEY, C. L.; VALENZUELA, A.; BARTLETT, F. S.; MORI, Y.; CAMPBELL, K. P.; FRANKEL, W. N.** (1998). The mouse stargazer gene encodes a neuronal Ca^{2+} -channel subunit. *Nat. Genet.* 19: 340-347.
- LEVI-MONTALCINI, R.; SKAPER, S. D.; DAL TOSO, R.; PETRELLI, L.; LEON, A.** (1996). Nerve growth factor: from neurotrophin to neurokinin. *Trends Neurosci.* 19: 514-520.
- LEVINE, E.; CUPP, A. S.; SKINNER, M. K.** (2000). Role of neurotrophins in rat embryonic testis morphogenesis (cord formation). *Biol. Reprod.* 62: 132-142.
- LEVI-STRAUSS, M.; CARROLL, M. C.; STEINMETZ, M.; MEO, T.** (1988). A previously undetected MHC gene with an unusual periodic structure. *Science* 240: 201-204.
- LEVINSON, B.; KENWRICK, S.; LAKICH, D.; HAMMONDS, G. JR.; GITSCHIER, J.** (1990). A transcribed gene in an intron of the human factor VIII gene. *Genomics* 7: 1-11.
- LEVINSON, B.; KENWRICK, S.; GAMEL, P.; FISHER, K.; GITSCHIER, J.** (1992). Evidence for a third transcript from the human factor VIII gene. *Genomics* 14: 585-589.
- LICHTER, P.; TANG, C. C.; CALL, K.; HERMANSON, G.; EVANS, G. A.; HOUSMAN, D.; WARD, D. C.** (1990). High resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. *Science* 247: 64-69.
- LINDSAY, E. A.; BOTTA, A.; JURECIC, V.; CARATTINI-RIVERA, S.; CHEAH, Y.-C.; ROSENBLATT, H. M.; BRADLEY, A.; BALDINI, A.** (1999). Congenital heart disease in mice deficient for the DiGeorge syndrome region. *Nature* 401: 379-383.

- LINDSAY, E. A., BALDINI. (1998): Congenital heart defects and 22q11 deletions: which genes count? *Mol. Med. Today* 4: 350-357.
- LÖBBERT, R. W.; WINTERPACHT, A., SEIPEL, B., ZABEL, B. U. (1996). Molecular cloning and chromosomal assignment of the human homologue of the rat cGMP-inhibited phosphodiesterase 1 (PDE3A)- a gene involved in fat metabolism located at 11p 15.1. *Genomics*. 37: 211-218.
- LÖBBERT, R. W. (1996). Exon-Amplifikation und PCR-Screening hierarchisch angeordneter cDNA-Banken- eine neue Strategie zur Analyse von Genen in der chromosomalen Region 11p15.1-3. Dissertation, Fachbereich Biologie der Johannes-Gutenberg-Universität Mainz.
- LOPEZ, R.; LARSEN, F., PRYDZ, H. (1994). Evaluation of the exon predictions of the GRAIL software. *Genomics* 24: 133-136.
- LÜDECKE, H. J.; WAGNER, M. J.; NARDMANN, J.; LA PILLO, B.; PARRISH, J. E.; WILLEMS, P. J.; HAAN, E. A.; FRYDMAN, M.; HAMERS, G. J.; WELLS, D. E.; ET AL. (1995). Molecular dissection of a contiguous gene syndrome: localization of the genes involved in the Langer-Giedion syndrome. *Hum. Mol. Genet.* 4: 31-36.
- LUND, J.; CHEN, F.; HUA, A.; ROE, B.; BUDARF, M.; EMANUEL, B. S.; REEVES, R. H. (2000). Comparative sequence analysis of 634 kb of the mouse chromosome 16 region of conserved synteny with the human velocardiocardial syndrome region on chromosome 22q11.2. *Genomics* 63: 374-383.
- LUPAS, A.; VAN DYKE, M.; STOCK, J. (1991). Predicting coiled coils from protein sequences. *Science* 252: 1162-1164.
- LURIE, I. W.; LAZJUK, Y. I.; USSOVA, E. B.; PRESMAN, E. B.; GUREVICH, D. B. (1980). The Wolf-Hirschhorn syndrome. *Clinical Genet.* 17: 375-383.
- LYU, M. S.; KOZAK, C. A. (1996). Linkage and synteny homologies between chromosome 4 and 6 mouse chromosomes. 4th Single Chromosome 4 Workshop, Bochum 18th-19th March (Abstract).
- MACGREGOR, H. C.; MIZUNO, S. (1976). In situ hybridization of "nick-translated" 3H-ribosomal DNA to chromosomes from salamanders. *Chromosoma* 54: 15-25.
- MACPHERSON, I. ; STOKER, M. (1962). Polyoma transformation of hamster cell clones -an investigation of genetic factors affecting cell competence. *Virology* 16: 147.
- MANGION, J.; RAHMAN, N.; EDKINS, S.; BARFOOT, R.; NGUYEN, T.; SIGURDSSON, A.; TOWNEND, J. V.; FITZPATRICK, D. R.; FLANAGAN, A. M.; STRATTON, M. R. (1999). The gene for cherubism maps to chromosome 4p16.3. *Am. J. Hum. Genet.* 65: 151-157.
- MARCZINEK, K.; NÜRNBERG, P.; TRÄGER, T.; RIEB, O.; LEHMANN, K.; THIEL, G.; TINSCHERT, S. (1995). Deleted or not deleted? - The critical region at 4p16.3 in a patient with Wolf-Hirschhorn syndrome. *Med. Genet.* 2: 140.
- MARTSOLF, J. T.; CHASE, T. R.; JALAL, S. M.; WASDAHL, W. A. (1987). Familial transmission of Wolf syndrome resulting from specific deletion 4p16 from t(4;8)(p16;p21) mat. *Clin. Genet.* 31: 366-369.
- MASSAGUE, J. (1998). TGF-beta signal transduction. *Annu. Rev. Biochem.* 67: 753-791.
- MAZO, A. M.; HUANG, D. H.; MOZER, B. A.; DAWID, I. B. (1990). The trithorax gene, a trans-acting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains. *Proc. Natl. Acad. Sci. USA* 87: 2112-2116.
- MCCARTHY, T. V.; QUANE, K. A.; LYNCH, P. J. (2000). Ryanodine receptor mutations in malignant hyperthermia and central core disease. *Hum. Mutat.* 15: 410-417.
- MCCOMBIE, W. R.; MARTIN-GALLARDO, A.; GOCAYNE, J. D.; FITZGERALD, M. G.; DUBNICK, M.; KELLY, J. M.; CASTILLA, L.; LIU, L.- I.; WALLACE, S.; TRAPP, S.; TAGLE, D.A.; WHALEY, W. L.; CHENG, S.; GUSELLA, J.; FRISCHAUF, A.- M.; POUSTKA, A.; LEHRACH, H.; COLLINS, F. S.; KERLAVAGE, A. R.; FIELDS, C.; VENTER, J. C. (1992). Expressed genes, Alu repeats and polymorphisms in cosmids sequenced from chromosome 4p16.3. *Nat. Genet.* 1: 348-353.
- MCLAUGHLIN, M. E.; EHRHART, T. L.; BERSON, E. L.; DRYJA, T. P. (1995). Mutation spectrum of the gene encoding the beta subunit of rod phosphodiesterase among patients with autosomal recessive retinitis pigmentosa. *Proc. Nat. Acad. Sci USA* 92: 3249-3253.

- MEYERS, G. A.; ORLOW, S. J.; MUNRO, I. R.; PRZYLEPA, K. A.; JABS, E. W. (1995). Fibroblast growth factor receptor 3 (FGFR3) transmembrane mutation in Crouzon syndrome with acanthosis nigricans. *Nat. Genet.* 11: 462-464.
- MOMENI, P.; GLOCKNER, G.; SCHMIDT, O.; VON HOLTUM, D.; ALBRECHT, B.; GILLESSEN-KAESBACH, G.; HENNEKAM, R.; MEINECKE, P.; ZABEL, B.; ROSENTHAL, A.; HORSTHEMKE, B.; LUDECKE, H. J. (2000). Mutations in a new gene, encoding a zinc-finger protein, cause tricho-rhino-phalangeal syndrome type I. *Nat. Genet.* 24: 71-74.
- MUNROE, D. J.; LÖBBERT, R.; BRIC, E.; WHITTON, T.; PRAWITT, D.; VU, D.; BUCKLER, A.; WINERPACHT, A.; ZABEL, B.; HOUSMAN, D. E. (1995). Systematic screening of an arrayed cDNA library by PCR. *Proc. Natl. Acad. Sci. USA* 92: 2209-2213.
- NAKAI, K.; KANEHISA, M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14: 897-911.
- NEITZEL, H. (1986). A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum. Genet.* 73: 320-326.
- NELSON, M. R. UND CHAZIN, W. J. (1998). Structures of EF-hand Ca(2+)-binding proteins: diversity in the organization, packing and response to Ca²⁺ binding. *Biomaterials* 11: 297-318.
- NING, Y.; WEBER, J. L.; KILLARY, A. M.; LEDBETTER, D. H.; SMITH, J. R.; PEREIRA-SMITH, O. M. (1991). Genetic analysis of indefinite division in human cells: evidence for a cell senescence-related gene(s) on human chromosome 4. *Proc. Natl. Acad. Sci. USA* 88: 5635-5639.
- NISHIDA, K.; TAMURA, A.; NAKAZAWA, N.; UEDA, Y.; ABE, T.; MATSUDA, F.; KASHIMA, K.; TANIWAKI, M. (1997). The Ig heavy chain gene is frequently involved in chromosomal translocations in multiple myeloma and plasma cell leukemia as detected by in situ hybridization. *Blood* 90: 526-534.
- NISLOW, C.; RAY, E.; PILLUS, L. (1997): SET1, a yeast member of the Thrithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* 8: 2421-2436.
- OSBORNE, L. R.; MARTINDALE, D.; SCHERER, S. W.; SHI, X.-M.; HUIZENGA, J.; HENG, H. H. Q.; COSTA, T.; POBER, B.; LEW, L.; BRINKMAN, J.; ROMMENS, J.; KOOP, B.; TSUI, L.-C. (1996). Identification of genes from a 500-kb region at 7q11.23 that is commonly deleted in Williams syndrome patients. *Genomics* 36: 328-336.
- OSBORNE, L. R. (1999). Williams-Beuren syndrome: unraveling the mysteries of a microdeletion disorder. *Molec. Genet. Metab.* 67: 1-10.
- PAK, S.-J. (1999). Charakterisierung neuer cDNAs aus der distalen Region 4p16.3 des menschlichen Chromosoms 4 mit Bezug zum Wolf-Hirschhorn-Syndrom und der Craniosynostose Typ Adelaide. Inaugural-dissertation, Fachbereich Medizin, Johannes Gutenberg-Universität Mainz.
- PARMACEK, M. S.; BENGUR, A. R.; VORA, A. J.; LEIDEN, J. M. (1990): The structure and regulation of expression of the murine fast skeletal troponin C gene. Identification of a developmentally regulated, muscle-specific transcriptional enhancer. *J. Biol. Chem.* 265: 15970-15976.
- PARTINGTON, M. W.; FAGAN, K.; SOUBJAKI, V.; TURNER, G. (1997). Translocations involving 4p16.3 in three families: deletion causing the Pitt-Rogers-Danks syndrome and duplication resulting resulting in a new overgrowth syndrome. *J. Med. Genet.* 34: 719-728.
- PARTINGTON, M.; TURNER G. (1999). Wolf-Hirschhorn and Pitt-Rogers-Danks syndromes. *Am. J. Med. Genet.* 82: 88-90.
- PAWLAK, A., CHIANNIKULCHI, N.; ANSORGE, W.; BULLE, F.; WEISSENBACH, J.; GYPAPAY, G.; GUELLAEN, G. (1998). Identification and mapping of 26 human testis mRNAs containing CAG/CTG repeats. *Mamm. Genome* 9: 745-748.
- PEARSON, W. R. ; LIPMAN, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85: 2444-2448.
- PENDERGAST, A. M. (1996). Nuclear protein tyrosine kinases: from Abl to WEE1. *Curr. Opin. Cell Biol.* 8: 174-181.
- PEOPLES, R.; FRANKE, Y.; WANG, Y. K.; PEREZ-JURADO, L.; PAPERNA, T.; CISCO, M.; FRANCKE, U. (2000). A physical map, including a BAC/PAC clone contig, of the Williams-Beuren syndrome-deletion region at 7q11.23. *Am. J. Hum. Genet.* 66: 47-68.

- PEREIRA-SMITH, O. M.; SMITH, J. R.** (1988). Genetic analysis of indefinite division in human cells: identification of four complementation groups. *Proc. Natl. Acad. Sci. USA*. 85: 6042-6046.
- PETIT, P.; SCHMIT, J.; VAN DEN BERGHE, H.; FRYNS, J. P.** (1996). On two patients with and without the classical Wolf-Hirschhorn syndrome (WHS) sharing the same chromosome 4p16.3 specific probe deletion: evidence of a contiguous gene deletion syndrome. *Clin. Genet.* 50: 19-22.
- PIERCE, J. C.; SAUER, B.; STERNBERG, N.** (1992): A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: improved cloning efficacy. *Proc. Natl. Acad. Sci. USA* 89: 2056-2060.
- PILZ, A.; SCHAAP, D.; HUNT, D.; FITZGIBBON, J.** (1995). Chromosomal localization of three mouse Diacylglycerol Kinase (DAGK) genes: genes sharing sequence homology to the Drosophila retinal degeneration A (rdgA) gene. *Genomics* 26: 599-601.
- PITT, D. B.; ROGERS, J. G.; DANKS, D. M.** (1984). Mental retardation, unusual face, and intrauterine growth retardation: A new recessive syndrome? *Am. J. Med. Genet.* 19: 307-313.
- PIZZUTI, A.; NOVELLI, G.; RATTI, A.; AMATI, F.; BORDONI, R.; MANDICH, P.; BELLONE, E.; CONTI, E.; BENGALA, M.; MARI, A.; SILANI, V.; DALLAPICCOLA, B.** (1999). Isolation and characterization of a novel transcript embedded within HIRA, a gene deleted in DiGeorge syndrome. *Mol. Genet. Metab.* 67: 227-235.
- PLOWRIGHT, E. E.; LI, Z.; BERGSAGEL, P. L.; CHESI, M.; BARBER, D. L.; BRANCH, D. R.; HAWLEY, R. G.; STEWART, A. K.** (2000). Ectopic expression of fibroblast growth factor receptor 3 promotes myeloma cell proliferation and prevents apoptosis. *Blood* 95: 992-998.
- POLASCIK, T. J.; CAIRNS, P.; CHANG, W. Y.; SCHOENBERG, M. P.; SIDRANSKY, D.** (1995): Distinct regions of allelic loss on chromosome 4 in human primary bladder carcinoma. *Cancer Res.* 55: 5396-5399.
- PRAWITT, D.** (1999). Identifizierung genetischer Faktoren der Tumörprädisposition in der Chromosomenregion 11p15. Dissertation, Fachbereich Biologie der Johannes-Gutenberg-Universität Mainz.
- PRIBILL, I., BARNES, G. T., CHEN, J., CHURCH, D., BUCKLER, A., BAXENDALE, S., BATES, G. P., LEHRACH, H., GUSELLA, M. J., DUYAO, M. P., AMBROSE, C.M. GUSELLA, J. F., MACDONALD, M- E.** (1997): Exon Trapping and Sequence-Based methods of gene finding in Transcrit Mapping of Human 4p16.3. *Som. Cell. Mol. Genet.* 23: 413-427.
- PROUDFOOT, N. J.; BROWNLEE, G. G.** (1976). 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* 263: 211-214.
- PUECH, A.; SAINT-JORE, B.; FUNKE, B.; GILBERT, D. J.; SIROTKIN, H.; COPELAND, N. G.; JENKINS, N. A.; KUCHERLAPATI, R.; MORROW, B.; SKOULTCHI, A. I.** (1997). Comparative mapping of the human 22q11 chromosomal region and the orthologous region in mice reveals complex changes in gene organization. *Proc. Natl. Acad. Sci. USA* 94: 14608-146013.
- RABBITS, T. H.** (1994). Chromosomal translocations in human cancer. *Nature* 372: 143-149.
- RAMIREZ-SOLIS, R.; LIU, P.; BRADLEY, A.** (1995). Chromosome engineering in mice. *Nature* 378: 720-724.
- REID, E.; MORRISON, N.; BARRON, L.; BOYD, E.; COOKE, A.; FIELDING, D.; TOLMIE, J. L.** (1996). Familial Wolf-Hirschhorn syndrome resulting from a cryptic translocation: a clinical and molecular study. *J. Med. Genet.* 33: 197-202.
- REINES, D.; CONAWAY, R. C.; CONAWAY, J. W.** (1999). Mechanism and regulation of transcriptional elongation by RNA polymerase II. *Curr. Opin. Cell. Biol.* 11: 342-346.
- RICHELDA, R.; RONCHETTI, D.; BALDINI, L.; CRO, L.; VIGGIANO, L.; MARZELLA, R.; ROCCHI, M.; OTSUKI, T.; LOMBARDI, L.; MAIOLO, A. T.; NERI, A.** (1997). A novel chromosomal translocation t(4;14)(p16.3;q32) in multiple myeloma involves the fibroblast growth-factor receptor 3 gene. *Blood* 90: 4062-4070.
- RIESS, O.; KOZAK, C.; VAN OMMEN, G.-J.** (1996). Report of the fourth international workshop on human chromosome 4 mapping 1996. *Cytogenet. Cell Genet.* 74: 57-69.
- ROLDAN A. L., CUBELLIS M.V., MASUCCI M.T., BEHRENDT N., LUND L.R., DANOE K., APPELLA E., BLASI F.** (1990). Cloning and expression of the receptor for human urokinase plasminogen activator, a central molecule in cell surface, plasmin dependent proteolysis. *EMBO J.* 9: 467-474.

- ROUSSEAU, F.; BONAVENTURE, J.; LEGEAI-MALLET, L.; PELET, A.; ROZET, J.-M.; MAROTEAUX, P.; LE MERRER, M.; MUNNICH, A. (1994).** Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature* 371: 252-254.
- SAIKI, R. K.; GELFAND, D. H.; STOFFEL, S.; SCHARF, S. J.; HIGUCHI, R.; HORN, G. T.; MULLIS, K. B.; EHRLICH, H. A. (1988).** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487-491.
- SALINAS, M.; REYES, R.; LESAGE, F.; FOSSET, M.; HEURTEAUX, C.; ROMÉY, G.; LAZDUNSKI, M. (1999).** Cloning of a new mouse two-P domain channel subunit and a human homologue with a unique pore structure. *J. Biol. Chem.* 274: 11751-11760.
- SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. (1989).** Molecular cloning: A laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. (1977).** DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.
- SASSAMAN, D. M.; DOMBROWSKI, B. A.; MORAN, J. V.; KIMBERLAND, M. L.; NAAS, T. P.; DEBERARDINIS, R. J.; GABRIEL, A.; SWERGOLD, G. D.; KAZAZIAN, H. H. (1997).** Many L1 elements are capable of retrotransposition. *Nat. Genet.* 16: 37-43.
- SCHÄFER, B. W. UND HEIZMAN, C. W. (1996).** The S100 family of EF-hand calcium-binding proteins: functions and pathology. *Trends Biochem. Sci.* 21: 134-140.
- SCHIMENTI, J. (1998).** Phenotype-based saturation mutagenesis in mice. Presentation at the 48th Annual Meeting of the American Society of Human Genetics, Denver October 27-31.
- SCHIMENTI, J. C.; LIBBY, B. J.; BERGSTROM, R. A.; WILSON, L. A.; NAF, D.; TARANTINO, L. M.; ALAVIZADEH, A.; LENGELING, A.; BUCAN, M. (2000):** Interdigitated deletion complexes on mouse chromosome 5 induced by irradiation of embryonic stem cells. *Genome Res.* 10: 1043-1050.
- SCHMICKEL, R. D. (1986).** Contiguous gene syndromes: A component of recognizable syndromes. *J. Pediatrics* 109: 231-241.
- SCOTT, H. S.; ASHTON, L. J.; EYRE, H. J.; BAKER, E.; BROOKS, D. A.; CALLEN, D. F.; SUTHERLAND, G. R.; MORRIS, C. P.; HOPWOOD, J. J. (1990).** Chromosomal localization of the human alpha-L-iduronidase gene (IDUA) to 4p16.3. *Am. J. Hum. Genet.* 47: 802-807.
- SCOTT, H. S.; LITJENS, T.; NELSON, P. V.; THOMPSON, P. R.; BROOKS, D. A.; HOPWOOD, J. J.; MORRIS, C. P. (1993).** Identification of mutations in the alpha-L-iduronidase gene (IDUA) that cause Hurler and Scheie syndromes. *Am. J. Hum. Genet.* 53: 973-986.
- SEBASTIO, G.; PERONE, L.; GUZZETTA, V.; SEBASTIO, L.; VICARI, L.; DELLA CASA, R.; GURRIERI, F.; ZAPPATA, S.; POMPONI M. G.; MAZZEI, A.; NERI, G.; ANDRIA, G.; BRAHE, C. (1996).** Molecular and cytogenetic characterization of a recurrent unbalanced translocation (4;21)(p16.3;q22.1): relevance to the Wolf-Hirschhorn and Down syndrome critical regions. *Am. J. Med. Genet.* 63: 366-72.
- SEIDAH, N. G.; BENJANNET, S.; PAREEK, S.; SAVARIA, D.; HAMELIN, J.; GOULET, B.; LALIBERTE, J.; LAZURE, C.; CHRETIEN, M.; MURPHY, R. A. (1996).** Cellular processing of the nerve growth factor precursor by the mammalian pro-protein convertases. *Biochem. J.* 14: 951-960.
- SEIDAH, N. G. UND CHRETIEN, M. (1997).** Eukaryotic protein processing: endoproteolysis of precursor proteins. *Curr. Opin. Biotechnol.* 8: 602-607.
- SEIDAH, N. G. UND CHRETIEN, M. (1999).** Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res.* 848: 45-62.
- SELBY, M. J.; EDWARDS, R.; SHARP, F.; RUTTER, W. J. (1987).** Mouse nerve growth factor gene: structure and expression. *Mol. Cell Biol.* 7: 3057-3064.
- SELBY, C. P. UND SANCAR, A. (1997).** Cockayne syndrome group B protein enhances elongation by RNA polymerase II. *Proc. Natl. Acad. Sci. USA.* 94: 11205-11209.
- SGRO, V.; RIVA, E.; CANEVINI, M. P.; COLAMARIA, V.; ROTTOLI, A.; MINOTTI, L.; CANGER, R.; DALLA BERNARDINA, B. (1995).** 4p(-) syndrome: a chromosomal disorder associated with a particular EEG pattern. *Epilepsia* 36: 1206-1214.

- SHAIKH, T. H.; KURAHASHI, H.; SAITTA, S. C.; MIZRAHY O'HARE, A.; HU, P.; ROE, B. A.; DRISCOLL, D. A.; McDONALD-MCGINN, D. M.; ZACKAI, E. H.; BUDARF, M. L.; EMANUEL, B. S. (2000). Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum. Mol. Genet.* 9: 489-501.
- SHANNON, M.; ASHWORTH, L. K.; MUCENSKI, M. L.; LAMERDIN, J. E.; BRANSCOMB, E.; STUBBS, L. (1996). Comparative analysis of a conserved zinc finger gene cluster on human chromosome 19q and mouse chromosome 7. *Genomics* 33: 112-120.
- SHIANG, R.; THOMPSON, L. M.; ZHU, Y.-Z.; CHURCH, D. M.; FIELDER, T. J.; BOCIAN, M.; WINOKUR, S. T.; WASMUTH, J. J. (1994). Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell* 78: 335-342.
- SHILATIFARD, A.; CONAWAY, J. W.; CONAWAY, R. C. (1997): Mechanism and regulation of transcriptional elongation and termination by RNA polymerase II. *Curr. Opin. Genet. Dev.* 7: 199-204.
- SHILATIFARD, A. (1998). Factors regulating the transcriptional elongation activity of RNA polymerase II. *FASEB J.* 12: 1437-1446.
- SIMMONS, A. D.; OVERHAUSER, J.; LOVETT, M. (1997). Isolation of cDNAs from the Cri-du-Chat Critical region by direct screening of a chromosome 5-specific cDNA library. *Genome Res.* 7: 118-127.
- SMITH, J. S.; TACHIBANA, I.; POHL, U.; LEE, H. K.; THANARAJASIGAM, U.; PORTIER, B. P.; UEKI, K.; RAMASWAMY, S.; BILLINGS, S. J.; MOHRENWEISER, H. W.; LOUIS, D. N.; JENKINS, R. B. (2000). A transcript map of the chromosome 19q-arm glioma tumor suppressor region. *Genomics* 64: 44-50.
- SOARES, M. B.; DE FATIMA BONALDO, M.; JELENE, P. , SU, L.; LAWTON, L.; EFSTRATIDIS, A. (1994). Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci. USA* 91: 9228-9232.
- SOMER, M.; PEIPPO, M.; KEINÄNEN, M. (1995). Controversial findings in two patients with commercial available probe D4S96 for the Wolf-Hirschhorn syndrome. *Am. J. Hum. Genet.* 57 (Suppl.): 127.
- SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.
- STEC, I.; WRIGHT, T. J.; VAN OMMEN, G.-J. B.; DE BOER, P. A. J.; VAN HAERINGEN, A.; MOORMAN, A. F. M.; ALTHERR, M. R.; DEN DUNNEN, J. T. (1998). WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphia gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. *Hum. Molec. Genet.* 7: 1071-1082.
- STEC, I.; NAGL, S. B.; VAN OMMEN, G. J.; DEN DUNNEN, J. T. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett.* 473: 1-5.
- STEINER, D. F. (1998). The proprotein convertases. *Curr. Opin. Chem. Biol.* 2: 31-39.
- STEITZ, T. A. (1999). DNA Polymerases: Structural diversity and Common Mechanisms. *J. Biol. Chem.* 274: 17395-17398.
- STENGEL-RUTKOWSKI, S.; WARKOTSCH, A.; SCHIMANEK, P.; STENE, J. (1984). Familial Wolf's syndrome with a hidden 4p deletion by translocation of an 8p segment. Unbalanced inheritance from a maternal translocation (4;8)(p15.3;p22). Case report, review and risk estimates. *Clin. Genet.* 25: 500-521.
- STILL, I. H.; VINCE, P.; COWELL, J. K. (1999). The third member of the transforming acidic coiled coil-containing gene family, TACC3, maps in 4p16, close to translocation breakpoints in Multiple Myeloma, and is upregulated in various cancer cell lines. *Genomics* 58: 165-170.
- STOHR, H.; KLEIN, J.; GEHRIG, A.; KOEHLER, M. R.; JURKLIES, B.; KELLNER, U.; LEO-KOTTLER, B.; SCHMID, M.; WEBER, B. H. (1999). Mapping and genomic characterization of the gene encoding diacylglycerol kinase gamma (DAGK3): assessment of its role in dominant optic atrophy (OPA1). *Hum. Genet.* 104: 99-105.
- SUTHERLAND, H. F.; WADEY, R.; MCKIE, J. M.; TAYLOR, C.; ATIF, U.; JOHNSTONE, K. A.; HALFORD, S.; KIM, U. J.; GOODSHIP, J.; BALDINI, A.; SCAMBLER, P. J. (1996). Identification of a novel transcript disrupted by a balanced translocation associated with DiGeorge syndrome. *Am. J. Hum. Genet.* 59: 23-31.
- SUTHERLAND, H. F.; KIM, U. - J.; SCAMBLER, P. J. (1998). Cloning and comparative mapping of the DiGeorge syndrome critical region in the mouse. *Genomics* 52: 37-43.

- TAGLE, D. A.; BLANCHARD-MCQUATE, K. L.; VALDES, J.; CASTILLA, L.; COLLINS, F. S. (1993).** Dinucleotide repeat polymorphisms at the D4S126 and D4S114 loci. *Hum. Mol. Genet.* 2: 1077.
- TAMURA, G.; SAKATA, K.; NISHIZUKA, S.; MAESAWA, C.; SUZUKI, Y.; TERASHIMA, M.; EDA, Y.; SATODATE, R. (1996).** Allelotype of adenoma and differentiated adenocarcinoma of the stomach. *J. Pathol.* 180: 371-377.
- TASSABEHJI, M.; METCALFE, K.; DONNAI, D.; HURST, J.; REARDON, W.; BURCH, M.; READ, A. P. (1997).** Elastin: genomic structure and point mutations in patients with supravalvular aortic stenosis. *Hum. Mol. Genet.* 6: 1029-1036.
- TAVIAUX, S. A. AND DEMAILLE, J. G. (1993).** Localization of Human cell cycle regulatory genes CDC25C to 5q31 and Wee1 to 11p15.3-11p15.1 by fluorescence in situ hybridization. *Genomics* 15: 194-196.
- TAVORMINA, P. L.; SHIANG, R.; THOMPSON, L. M.; ZHU, Y.-Z.; WILKIN, D. J.; LACHMAN, R. S.; WILCOX, W. R.; RIMOIN, D. L.; COHN, D. H.; WASMUTH, J. J. (1995).** Thanatophoric dysplasia (types I and II) caused by distinct mutations in fibroblast growth factor receptor 3. *Nat. Genet.* 9: 321-328.
- TAVORMINA, P. L.; BELLUS, G. A.; WEBSTER, M. K.; BAMSHAD, M. J.; FRALEY, A. E.; MCINTOSH, I.; SZABO, J.; JIANG, W.; JABS, E. W.; WILCOX, W. R.; WASMUTH, J. J.; DONOGHUE, D. J.; THOMPSON, L. M.; FRANCOMANO, C. A. (1999).** A novel skeletal dysplasia with developmental delay and acanthosis nigricans is caused by a Lys650Met mutation in the fibroblast growth factor receptor 3 gene. *Am. J. Hum. Genet.* 64: 722-731.
- TAYLOR, S. A. M.; SNELL, R. G.; BUCKLER, A.; AMBROSE, C.; DUYAO, M.; CHURCH, D.; LIN, C. S.; ALTHERR, M.; BATES, G. P.; GROOT, N.; BARNES, G.; SHAW, D. J.; LEHRACH, H.; WASMUTH, J. J.; HARPER, P. S.; HOUSMAN, D. E.; MACDONALD, M. E.; GUSELLA, J. F. (1992).** Cloning of the α -adducin gene from the Huntington's disease candidate region of chromosome 4 by exon amplification. *Nat. Genet.* 2: 223-227.
- TESSAROLLO, L. (1998).** Pleiotropic functions of neurotrophins in development. *Cytokine Growth Factor Rev.* 9: 125-137.
- THE HUNTINGTON'S DISEASE COLLABORATIVE RESEARCH GROUP. (1993).** A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72: 971-983.
- THOMPSON, L. M.; PLUMMER, S.; SCHALLING, M.; ALTHERR, M. R.; GUSELLA, J. F.; HOUSMAN, D. E.; WASMUTH, J. J. (1991).** A gene encoding a fibroblast growth factor receptor isolated from the Huntington disease gene region of human chromosome 4. *Genomics* 11: 1133-1142.
- THOMSON, P. (1998).** Wolf-Hirschhorn Syndrome. Review of the literature and three case studies. *J. Am. Podiatr. Med. Assoc.* 88: 192-197.
- TIZIANI, V.; REICHENBERGER, E.; BUZZO, C. L.; NIAZI, S.; FUKAI, N.; STILLER, M.; PETERS, H.; SALZANO, F. M.; RAPOSO, D. O.; AMARAL, C. M.; OLSEN, B. R. (1999).** The gene for cherubism maps to chromosome 4p16. *Am. J. Hum. Genet.* 65: 158-166.
- TOMMERUP, N.; AAGAARD, L.; LUND, C. L.; BOEL, E.; BAXENDALE, S.; BATES, G. P.; LEHRACH, H.; VISSING, H. (1993).** A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate gene for the Wolf-Hirschhorn (4p-) syndrome. *Hum. Mol. Genet.* 2: 1571-1575.
- TOMMERUP, N.; VISSING, H. (1995):** Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identify putative candidate genes for developmental and malignant disorders. *Genomics* 27: 259-264.
- TOPHAM, M. K.; PRESCOTT, S. M. (1999).** Mammalian diacylglycerol kinases, a family of lipid kinases with signaling functions. *J. Biol. Chem.* 274: 11447-11450.
- TRIPOULAS, N.; LAJEUNESSE, D.; GILDEA, J.; SHEARN, A. (1996).** The Drosophila ash1 gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger. *Genetics* 143: 913-928.
- TROFATTER, J. A.; LONG, K. R.; MURRELL, J. R.; STOTLER, C. J.; GUSELLA, J. F.; BUCKLER, A. J. (1995).** An expression-independent catalog of genes from human chromosome 22. *Genome Res.* 5: 214-224.
- TSCHIRSCH, B.; HOFMANN, A.; KRAUSS, V.; DORN, R.; KORGE, G.; REUTER, G. (1994).** The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* 13: 3822-3831.

- TUFTY, R. M.; KRETSINGER, R. H. (1975). Troponin and parvalbumin calcium binding regions predicted in myosin light chain and T4 lysozyme. *Science* 187: 167-169.
- UBERBACHER, E. C.; MURAL, R. J. (1991). Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Natl. Acad. Sci. USA* 88: 11261-11265.
- ULLRICH, A.; GRAY, A.; BERMAN, C.; DULL, T. J. (1983). Human beta-nerve growth factor gene sequence highly homologous to that of mouse. *Nature* 303: 821-825.
- VALERO, M. C.; CRUCES, J.; PEREZ-JURADO, L. A. (1998). Syntenic organization of the Williams-Beuren syndrome deleted region in chromosome 7q11.23 and the mouse chromosome 5 band C. *Am. J. Hum. Genet.* 63 (Suppl.): A259.
- VELINOV, M.; SLAUGENHAUPT, S. A.; STOILOV, I.; SCOTT, C. I. JR, GUSELLA, J. F.; TSIPOURAS, P. (1994). The gene for achondroplasia maps to the telomeric region of chromosome 4p. *Nat. Genet.* 6: 314-317.
- VISKOCHIL, D.; CAWTHON, R.; O'CONNELL, P.; XU, G. F.; STEVENS, J.; CULVER, M.; CAREY, J.; WHITE, R. (1991). The gene encoding the oligodendrocyte-myelin glycoprotein is embedded within the neurofibromatosis type 1 gene. *Mol. Cell. Biol.* 11: 906-912.
- WADEY, R.; MCKIE, J., PAPAPETROU, C.; SUTHERLAND, H.; LOHMAN, F., OSINGA, J., FROHN, I., HOFSTRA, R., MEIJERS, C., AMATI, F., CONTI, E., PIZZUTI, A., DALLAPICCOLA, B.; NOVELLI, G., SCAMBLER, P. (1999). Mutations of UFD1L are not responsible for the majority of cases of DiGeorge Syndrome/Velocardiofacial Syndrome without deletions within chromosome 22q11. *Am. J. Hum. Genet.* 65: 247-249.
- WAGNER, T.; WIRTH, J. ; MEYER, J.; ZABEL, B.; HELD, M.; ZIMMER, J.; PASANTES, J.; BRICARELLI, F. D.; KEUTEL, J.; HUSTERT, E.; WOLF, U.; TOMMERUP, N.; SCHEMPP, W.; SCHERER G. (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* 79: 1111-1120.
- WAHL, G. M.; LEWIS, K. A.; RUIZ, J. C.; ROTHENBERG, B.; ZHAO, B.; EVANS, G. A. (1987). Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. *Proc. Natl. Acad. Sci. USA* 84: 2160-2164.
- WALDHERR, M.; RAGNINI, A.; JANK, B.; TEPLY, R.; WIESENBERGER, G.; SCHWEYEN, R. J. (1993). A multitude of suppressors of group II intron-splicing defects in yeast. *Curr. Genet.* 24: 301-306.
- WALLRAPP, C.; VERRIER, S.-B.; ZHOURAVLEVA, G.; PHILIPPE, H.; PHILIPPE, M.; GRESS, T. M.; JEAN-JEAN, O. (1998). The product of the mammalian orthologue of the *Saccharomyces cerevisiae* HBS1 gene is phylogenetically related to eukaryotic release factor 3 (eRF3) but does not carry eRF3-like activity. *FEBS Letters* 440: 387-392.
- WEBER, B.; COLLINS, C.; KOWBWL, D.; RIESS, O.; HAYDEN, M. R. (1991a). Identification of multiple CpG islands and associated conserved sequences in a candidate region for the Huntington disease gene. *Genomics* 11: 1113-1124.
- WEBER, B.; RIESS, O.; HUTCHINSON, G.; COLLINS, C.; LIN, B. Y.; KOWBEL, D.; ANDREW, S.; SCHAPPERT, K.; HAYDEN, M. R. (1991b). Genomic organization and complete sequence of the human gene encoding the beta-subunit of the cGMP phosphodiesterase and its localisation to 4p16.3. *Nucleic Acids Res.* 19: 6263-6268.
- WEBER, B.; RIESS, O.; DANESHVAR, H.; GRAHAM, R.; HAYDEN, M. R. (1993). (CA)_n-dinucleotide repeat at the PDEB locus in 4p16.3. *Hum. Mol. Genet.* 2: 827.
- WHALEY, W. L.; BATES, G. P.; NOVELLETTA, A.; SEDLACEK, Z.; CHENG, S.; ROMANO, D.; ORMONDROYD, E.; ALLITTO, B.; LIN, C.; YOUNGMAN, S.; BAXENDALE, S.; BUCAN, M.; ALTHERR, M.; WASMUTH, J.; WEXLER, N. S.; FRONTALI, M.; FRISCHAUF, A.-M.; LEHRACH, H.; MACDONALD, M. E.; GUSELLA, J. F. (1991): Mapping of Cosmid Clones in Huntingtons Disease Region of Chromosome 4. *Som. Cell. Mol. Genet.* 17: 83-91.
- WHEAL, H. V.; BERNARD C.; CHAD, J. E.; CANNON, R. C. (1998). Pro-epileptic changes in synaptic function can be accompanied by pro-epileptic changes in neuronal excitability. *Trends Neurosci.* 21: 167-174.
- WHITE, D. M.; PILLERS, D. A.; REISS, J. A.; BROWN, M. G.; MAGENIS, R. E. (1995). Interstitial deletions of the short arm of chromosome 4 in patients with a similar combination of multiple minor anomalies and mental retardation. *Am. J. Med. Genet.* 57: 588-597.

- WIECZOREK, D.; KRAUSE, M.; MAJEWSKI, F.; ALBRECHT, B.; HORN, D.; RIESS, O.; GILLESSEN-KAESBACH G.** (2000). Effect of the size of the deletion and clinical manifestation in Wolf-Hirschhorn syndrome: analysis of 13 patients with a de novo deletion. *Eur. J. Hum. Genet.* 8: 519-526.
- WILMING, L. G.; SNOEREN, C. A. S.; VAN RIJSWIJK, A.; GROSVELD, F.; MEIJERS, C.** (1997). The murine homologue of HIRA, a DiGeorge syndrome candidate gene, is expressed in embryonic structures affected in human CATCH22 patients. *Hum. Mol. Genet.* 6: 247-258.
- WINTERPACHT, A.; AMID, C.; BAHR, A.; BRUECKMANN, T.; CICHUTEK, A.; HANKELN, T.; SCHMIDT, E. R.; SEIPEL, B.; ZABEL, B.** (1998). Analysis of a 1 Mb region on chromosome 11p15.3 by comparative sequencing between man and mouse. 10. Jahrestagung der Deutschen Gesellschaft für Humangenetik, Jena, 25.03. - 28.03. *Med. Genetik* 10: 113.
- WOLF, U.; REINWEIN, H.; PORSCH, R.; SCHROTTER, R.; BAITSCH, H.** (1965). Defizienz an den kurzen Armen eines Chromosoms Nr. 4. *Hum. Genet.* 1: 397-413.
- WRANA, J. L.** (2000). Regulation of Smad Activity. *Cell* 100: 189-192.
- WRIGHT, T. J.; RICKE, D. O.; DENISON, K.; ABMAYR, S.; COTTER, P. D.; HIRSCHHORN, K.; KEINÄNEN, M.; McDONALD-MCGINN, D.; SOMER, M.; SPINNER, N.; YANG-FENG, T.; ZACKAI, E.; ALTHERR, M. R.** (1997). Transcription map of the newly defined 165 kb Wolf-Hirschhorn syndrome critical region. *Hum. Mol. Genet.* 6: 317-324
- WRIGHT, T. J.; CLEMENS, M.; QUARRELL, O.; ALTHERR, M. R.** (1998). Wolf-Hirschhorn and Pitt-Rogers-Danks syndromes caused by overlapping 4p deletions. *Am. J. Med. Genet.* 75: 345-350.
- WRIGHT, T. J.; ALTHERR, M. R.; CALLEN, D.; HIRSCHHORN, K.** (1999a). Reply to the letter to the editor by Partington and Turner-'Wolf-Hirschhorn and Pitt-Rogers-Danks syndrome'. *Am. J. Med. Genet.* 82: 89-90.
- WRIGHT, T. J.; COSTA, J. L.; NARANJO, C.; FRANCIS-WEST, P.; ALTHERR, M. R.** (1999b). Comparative analysis of a novel gene from the Wolf-Hirschhorn/Pitt-Rogers-Danks syndrome critical region. *Genomics* 59: 203-212.
- WÜCHNER, C.; HILBERT, K.; ZABEL, B.; WINTERPACHT, A.** (1997). Human fibroblast growth factor receptor 3 gene (FGFR3): genomic sequence and primer set information for gene analysis. *Hum. Genet.* 100: 215-219.
- WYANDT, H. E.; MILUNSKY, J.; LERNER, T.; GUSELLA, J. F.; HOU, A.; MACDONALD, M.; ADEKUNLE, S.; MILUNSKY, A.** (1993). Characterization of a duplication in the terminal band of 4p by molecular cytogenetics. *Am. J. Med. Genet.* 46: 72-76.
- YAMAGISHI, H.; GARG, V.; MATSUOKA, R.; THOMAS, T.; SRIVASTAVA, D.** (1999). A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects. *Science* 283: 1158-1161.
- YAMAGUCHI, Y.; TAKAGI, T.; WADA, T.; YANO, K.; FURUYA, A.; SUGIMOTO, S.; HASEGAWA, J.; HANDA, H.** (1999a). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97: 41-51.
- YAMAGUCHI, Y.; WADA, T.; WATANABE, D.; TAKAGI, T.; HASEGAWA, J.; HANDA, H.** (1999b). Structure and function of the human transcription elongation factor DSIF. *J. Biol. Chem.* 274: 8085-8092.
- XU, Y.; MURAL, R.; SHAH, M.; UBERBACHER, E.** (1994). Recognizing exons in genomic sequence using GRAIL II. *Genetic Engineering* 16: 241-253.
- ZOLLINO, M.; DI STEFANO, C.; ZAMPINO, G.; MASTROIACOVO, P.; WRIGHT, T. J.; SORGE, G.; SELICORNI, A.; TENCONI, R.; ZAPPALA, A.; BATTAGLIA, A.; DI ROCCO, M.; PALKA, G.; PALLOTTA, R.; ALTHERR, M. R.; NERI, G.** (2000). Genotype-phenotype correlations and clinical diagnostic criteria in wolf-hirschhorn syndrome. *Am. J. Med. Genet.* 94: 254-261.

7 Anhang