

## 4. Ergebnisse und Auswertung

### 4.1. Polymorphusanalyse und Definition von OR-Haplotypen

Ziel des ersten Abschnittes dieser Arbeit war es, herauszufinden, ob HLA-gekoppelte olfaktorische Rezeptoren polymorph sind. Die gefundenen Polymorphismen sollten die Grundlage für die Definition der OR-Haplotypen darstellen, die bestimmten HLA-Haplotypen zugeordnet werden sollten. Deshalb wurden HLA-homozygote oder -hemizygoten Zelllinien für diese Analyse gewählt.

Einige OR Gene aus diesem Cluster waren bereits seit Mitte der 90er Jahre bekannt (Fan et al., 1995; Gruen et al., 1996), wobei zu erwähnen ist, daß das beschriebene Gen FAT11 (Fan et al., 1995) artifiziell ist, weil es Sequenzen von zwei verschiedenen, nah verwandten Genen enthält (Ziegler et al., 2000b). Die Sequenzen der anderen HLA-gekoppelten OR Gene wurden vom Sanger-Center in Hinxton ermittelt und der NCBI-Gen-Datenbank zur Verfügung gestellt. Insgesamt sind 34 + 5 OR Gene in diesem Cluster identifiziert worden (s. Abb. 4) (Ehlers et al., 2000; Younger et al., 2001), von denen 34 zur M-OR Familie und fünf zum V1R Typ gehören. 15 der M-OR Gene (*hs6M1-1, -3, -6, -10, -12, -15, -16, -17, -18, -20, -21, -27, -28, -32, -35*) zeigten ein offenes Leseraster (ORF, open readingframe) mit der erwarteten Länge, während die restlichen 19 (*hs6M1-2P, -4P, -5P, -7P, -8P, -9P, -13P, -14P, -19P, -22P, -23P, -24P, -25P, -26P, -29P, -30P, -31P, -33P, -34P*) Pseudogene zu sein schienen.

Im Rahmen der Polymorphusanalyse wurden die folgenden 22 M-OR Gene in zehn verschiedenen Zelllinien (s. Tab. 1) sequenziert<sup>1</sup>: *hs6M1-1; -2P; -3; -4P; -5P; -6; -7P; -8P; -10; -12; -13P; -15; -16; -17; -18; -19P; -20; -21; -23P; -24P; -25P; -26P*. Mit Ausnahme von *hs6M1-10*, liegen alle analysierten Gene im Haupt-OR-Cluster zwischen RFP und HLA-F (s. Abb. 4); (Ehlers et al., 2000; Younger et al., 2001). *hs6M1-10* liegt im Neben-Cluster distal von RFP. Zusätzlich wurden drei OR Gene auf den Chromosomen 7, 17 und 19 (*hs7M1-1, hs17M1-20* und *hs19M1-4*) sequenziert, um einen möglichen Unterschied im Grad des Polymorphismus zwischen HLA-gekoppelten und nicht HLA-gekoppelten Genen zu ermitteln. Bei den 10 getesteten Zelllinien (s. Tab. 1) handelte es sich um 8 HLA-homozygote und 2 HLA-hemizygoten (BM19.7; BM28.7) Linien. Von den HLA-hemizygoten Linien fehlt BM19.7 ein komplettes Chromosom 6, während BM28.7 eine Deletion in der Region 6p24.2-p21.31 aufweist (Volz et al., 1992). Somit wurden insgesamt 18 Chromosomen (19 einschließlich des PAC's aus der NCBI-Datenbank) analysiert. Die Auswahl von Zelllinien unterschiedlicher ethnischer Abstammung und sehr unterschiedlicher HLA-Haplotypen (s. Tab. 1) sollte die Chancen maximieren, OR Polymorphismen zu entdecken.

---

<sup>1</sup> *hs6M1-8P, -12, -13P, -17, -19P, -20, -21, -23P, -24P* und *hs7M1-1, hs17M1-20* und *hs19M1-4* wurden von Dr. Simon Forbes bzw. Ruth Younger sequenziert, die ein etwas anderes Sequenzierverfahren anwendeten (Ehlers et al., 2000).

Zunächst wurde genomische DNA aus den zehn verschiedenen Zelllinien (s. Tab. 1) isoliert (s. Abschnitt 3.2). Anhand der bekannten Sequenzen wurden für die jeweiligen OR Gene spezifische Primer festgelegt (s. Tab. 2). Da die einzelnen Gene z.T. eine hohe Nukleotid-Sequenz-Homologie aufwiesen (z.T. > 90% wie z.B. *hs6M1-16*, *-12* und *-13P*), wurde bei der Primer-Herstellung besonders darauf geachtet, daß die Oligonukleotide ausschließlich spezifisch für das entsprechende Gen waren. Durch ein Alignment (Corpet, 1988) aller bekannten nah verwandten OR Gene konnten spezifische Positionen für die jeweiligen Oligonukleotide gefunden werden. Pro Gen wurden zwei Paar Primer festgelegt, die jeweils den vorderen bzw. hinteren Teil der kodierenden Region umfaßten. Die Primer wurden so gelegt, daß jeweils noch 50-100 bp in der 5' bzw. 3' UTR mit amplifiziert wurden und zusätzlich eine Überlappung der PCR-Produkte in der Mitte des Gens gegeben war. Die Amplifikate wurden getrennt sequenziert und die Daten anschließend anhand der Sequenzüberlappung zusammengefügt. Dieser Weg wurde eingeschlagen, weil die gesamte kodierende Region eines Gens aufgrund der limitierten Leseweite (700 bp) des für die Sequenzierung eingesetzten Gerätes nicht in einem Ansatz sequenziert werden konnte. Sämtliche Primer, die für die Polymorphismusanalyse verwendet wurden, besaßen einen M13-Anhang (s. Tab. 2). Somit konnte der PCR-Ansatz direkt, d.h. ohne zu klonieren, sequenziert werden, indem die entsprechenden M13-IRD-markierten Sequenzier-Primer verwendet wurden (s. Tab. 11). Diese Methode hat den Vorteil, daß Fehler der *Taq*-Polymerase vernachlässigbar sind, da sich Schreibfehler der *Taq*-Polymerase in dem komplexen DNA-Pool, der sequenziert wurde, herausmitteln. Zusätzlich ist die Methode weniger zeitaufwendig. Ein Nachteil besteht allerdings darin, daß die Qualität der Sequenzen, die direkt aus PCR-Produkten stammen, nicht mit der Qualität der Sequenzen konkurrieren kann, die aus Plasmid-Präparationen stammen.

Die Rohdaten wurden mit dem Programm Base ImagIR auf OS2-Ebene ausgewertet und anschließend mit dem Programm AlignIR überarbeitet. Der Vergleich der Sequenzen der verschiedenen Zelllinien erfolgte mit MultAlin (Corpet, 1988) (<http://www.toulouse.inra.fr/multalin.html>).

#### 4.1.1. OR Pseudogene

Die genomischen Sequenzen der PACs und BACs zeigten, daß 19 Gene in diesem OR Cluster aufgrund von Rastermutationen, Stop-Codons oder eines fehlenden Start-Codons vermutlich Pseudogene sind (s. Abb. 5).

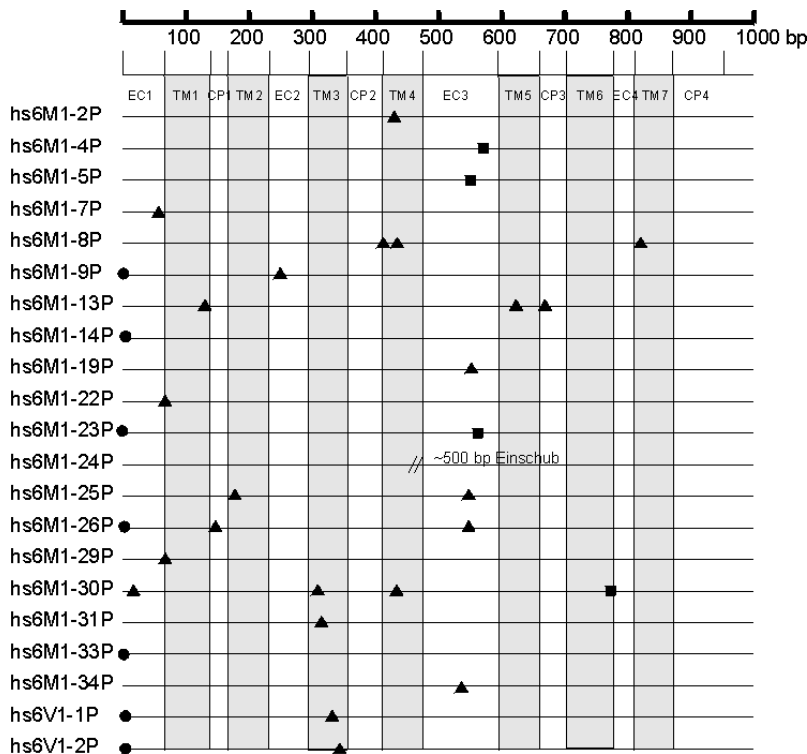


Abb. 5: OR Pseudogene in der Nähe des HLA-Komplexes. Aufgrund von Rastermutationen, Stop-Codons und fehlenden Start-Codons sind 19 Gene vom M-OR Typ und 2 Gene vom V1-OR Typ als Pseudogene deklariert worden (s. NCBI Gen Datenbank). Die erst kürzlich gefundenen Gene *hs6V1-3P*, *-4P* und *-5P* sind noch nicht in dieser Abbildung enthalten. Dreiecke: Rastermutation; Viereck: Stop-Codon; Kreis: kein Start-Codon (Ehlers et al., 2000; Younger et al., 2001)

Es sollte untersucht werden, ob diese Mutationen, die aller Wahrscheinlichkeit nach zu einem Funktionsverlust der jeweiligen Gene führen, in allen zehn Zelllinien zu finden sind.

Die folgenden 9 Pseudogene wurden analysiert: *hs6M1-2P*, *-5P*, *-7P*, *-8P*; *-13P*; *-23P*, *-24P*, *-25P* und *-26P*. Während von *hs6M1-2P*, *-7P*, und *-13P* das gesamte Gen sequenziert wurde, wurden von den restlichen Genen nur die Abschnitte sequenziert, die zu einem Abruch des Leserasters führten. Die Analyse ergab, daß die Rastermutationen und Stop-Codons in allen zehn Zelllinien zu finden sind. Diese Gene besitzen damit in allen analysierten Zelllinien kein offenes Leseraster und sind daher unabhängig von der Herkunft funktionslos.

Bei den komplett sequenzierten Genen wurden vereinzelt Mutationen in Form von Einzelbasenparaustauschen gefunden. Bei *hs6M1-2P*, *hs6M1-7P* und *hs6M1-13P* wurden zum Beispiel ein bzw. zwei Austausche gefunden. Für *hs6M1-2P* und *hs6M1-13P* konnten zwei, beide ohne Auswirkung auf die Aminosäuresequenz, und für *hs6M1-7P* drei Allele definiert werden, die alle zu einer veränderten Aminosäuresequenz führen (s. Tab. 14). Für die restlichen Gene können aufgrund der nicht vollständigen Sequenzanalyse keine Angaben über die Anzahl der Allele gemacht werden.

Domäne	BM28.7	BM19.7	LG2	KR3598	H2LCL	WT51	SA	YAR	OLGA	AMAI	PAC/BAC			
	A1, B35	A2, B13	A2, B27	A2, B44	A3, B7	A23, B65	A24, B7	A26, B38	A31, B62	A68, B53				
	Schwarzafr.	Schwarzafr.	Kaukasier	Kaukasier	Kaukasier	Kaukasier	Japaner	Jude	Am. Indian.	Algerier				
EC2	Leu CTA *02	Leu CTG *01	Leu CTG *01	Leu CTG *01	hs6M1-2P Leu CTG *01		Leu CTG *01	Leu CTG *01	Leu CTG *01	Leu CTG *01	Leu CTG *01			
	Allele-Bezeichnung													
TM3	Cys TGT	Cys TGT	Tyr TAT	Cys TGT	Tyr TAT	hs6M1-7P Tyr TAT		Cys TGT	Tyr TAT	Tyr TAT	Cys TGT	Tyr TAT	Cys TGT	Tyr TAT
TM6	Tyr TAT *02	His CAT *03	Tyr TAT *01	Tyr TAT *02	Tyr TAT *01	Tyr TAT *01	Cys TGT *02	Tyr TAT *01	Tyr TAT *01	Tyr TAT *02	Tyr TAT *01	Tyr TAT *02	Tyr TAT *02	Tyr TAT *01

Tab. 14 : Nukleotid- und Aminosäuresubstitutionen (theoretische) sowie stumme Mutationen in den verschiedenen Domänen zweier HLA-gekoppelter Pseudogene. Angegeben sind auch Alleldefinitionen, sowie die HLA-Klasse I Haplotypen und ethnische Abstammung der analysierten Zelllinien. *hs6M1-13P* wurde von Dr. Simon Forbes sequenziert und ist hier nicht dargestellt.

#### 4.1.2. OR Gene mit einem offenen Leseraster (ORF)

Die folgenden potentiell funktionstüchtigen Gene wurden in allen zehn Zelllinien sequenziert: *hs6M1-1*, *-3*, *-6*, *-10*, *-12*, *-15*, *-16*, *-17*, *-18*, *-20*; *-21*, *hs7M1-2*, *hs17M1-20* und *hs19M1-4*. Alle besitzen in der NCBI Datenbank einen ORF. Zusätzlich wurden zwei Gene sequenziert, die in den bekannten genomischen Sequenzen der PACs und BACs keinen ORF besitzen (*hs6M1-4P* und *-19P*). Diese beiden Gene und *hs6M1-17* weisen eine Besonderheit auf: In einigen Zelllinien stellen sie Pseudogene dar, während sie in anderen Zelllinien potentiell funktionstüchtige Gene repräsentieren.

Insgesamt zeigen die 13 analysierten HLA-gekoppelten M-OR Gene mit ORF 52 verschiedene Mutationen (s. Tab. 15): 16 (31%) sind stumm, d.h. sie haben keine Auswirkung auf die Aminosäuresequenz, während 36 (69%) zu einer veränderten Aminosäuresequenz führen. Mit Ausnahme des *hs6M1-19P* Genes, das in einigen Zelllinien eine das Leseraster verschiebende Deletion von 16 bp aufweist, handelt es sich bei diesen Mutationen um einzelne Basenpaaraustausche, die innerhalb des Codons an allen drei Positionen gefunden wurden. Während das *hs6M1-1* Gen nur eine einzelne stumme Mutation aufweist, zeigen alle anderen analysierten Gene mindestens eine Mutation, die zu einer veränderten Aminosäuresequenz führt. Die Polymorphismen innerhalb der Gene treten an unterschiedlichen Positionen auf (s. Tab. 15, s. Abb. 6). So zeigt *hs6M1-3* Variationen in den Domänen TM3, CP3 und TM6, während bei *hs6M1-6* nur in den Transmembran Domänen (TM2, 3, 4 und 5) Aminosäure (AS) Austausch gefunden wurden. *hs6M1-4P* wiederum weist nur in den Extrazellulär Domänen (EC1 und EC3) Polymorphismen auf. Besonders auffällig sind zwei Mutationen, die bei unterschiedlichen Genen an analogen Positionen lokalisiert sind. Ein Alignment (s. Abb. 7) zeigt, daß die Variationen in *hs6M1-3* (AA 113) und in *hs6M1-6* (AA 111) an korrespondierenden Aminosäurepositionen auftreten. Das gleiche Phänomen findet man bei *hs6M1-20* (AA 104) und *hs6M1-21* (AA106), wobei die Mutation in diesem Falle allerdings stumm ist.

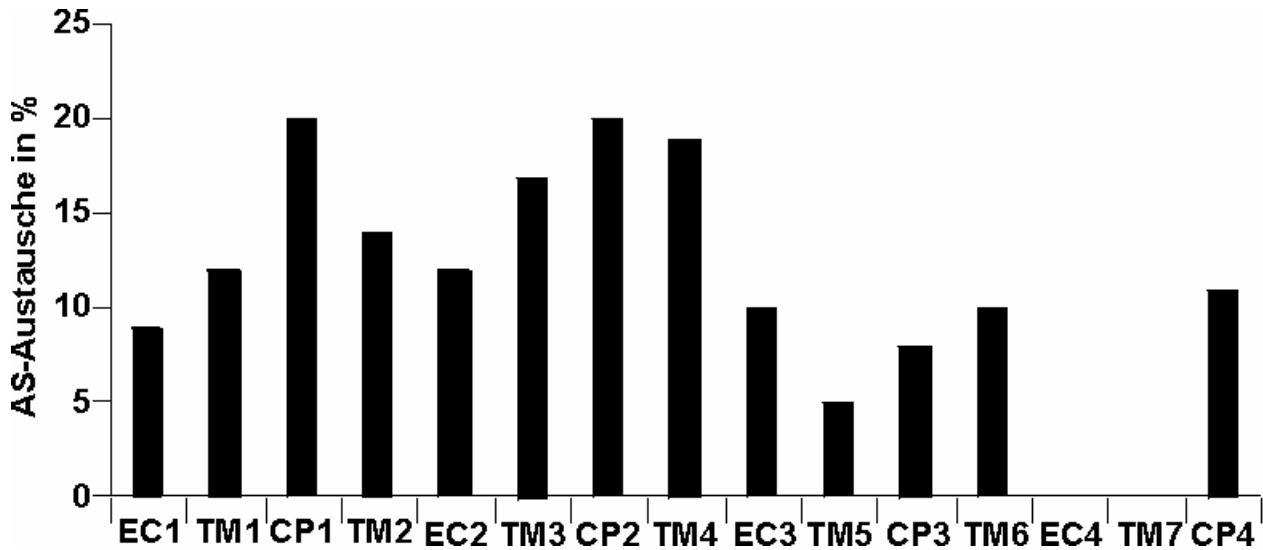


Abb. 6: Prozentuale Häufigkeit der Mutationen in den einzelnen Domänen bei allen analysierten HLA-gekoppelten OR Genen mit ORF. Angegeben sind nur die Mutationen, die zu einem Aminosäure (AS) Austausch führen. Die Domänen der Proteine sind gegen die prozentualen AS Austausche aufgetragen.

	EC1		TM1		
hs6M1-3:BM19.7	MNDDGKVNAS	SEGYFILVGF	SNWPHLEVVI	FVVVLIFLYL	TLIGNLFIII
hs6M1-3: WT51	..DDG.V...	..GY...V..	....H...I	...V.....	..I.....
hs6M1-6: LG2	MMI.K...	..DF...L..	....Q....L	...I.....	..T.....
hs6M1-6: WT51	MMI.K...	..DF...L..	....Q....L	...I.....	..T.....

	CP1	TM2	EC2
hs6M1-3:BM19.7	LSYLDLHLHT	PMYFFLSNLS	FLLDLCYTTSS
hs6M1-3: WT51	...L.....	.....Y...	.....W.....
hs6M1-6: LG2	..V.....	.....H....	.....R.....
hs6M1-6: WT51	..V.....	.....Y....	.....R.....

	EC2	TM3	CP2	TM4
hs6M1-3:BM19.7	MIQLYFVLAL	GTAECVLLV	MSYDRYA AVC	RPLHYTVLMH
hs6M1-3: WT51	.I.....	.T.....	....A...	.....AV.
hs6M1-6: LG2	.V.....	.I.....	....V...	.....AA.
hs6M1-6: WT51	.V.....	.I.....	....V...	.....VA.

	TM4	EC3
hs6M1-3:BM19.7	SWVSGFTNSA	LHSSFTFWVP
hs6M1-3: WT51	...S...N..	.....Q....
hs6M1-6: LG2	...I...I..	.....L....
hs6M1-6: WT51	...I...I..	.....L....

	EC3	TM5	CP3	TM6
hs6M1-3:BM19.7	TLMITSSIFV	LIPLILILTS	YGAIVQAVLR	MQSTTGLQKV
hs6M1-3: WT51	...IT.....	.....S	...VR...R	.....G.....
hs6M1-6: LG2	...VM.....	.....A	...AR.V.S	.....R.....
hs6M1-6: WT51	...VM.....	.....T	...AR.V.S	.....R.....

	TM6	EC4	TM7	CP4
hs6M1-3:BM19.7	VSLFFIPAMC	IYLPSPSGNS	QDQGFIALF	YTVVTPSLNP
hs6M1-3: WT51	.....A..	M.....G..	Q.....	.....V..
hs6M1-6: LG2	.....V..	M.....E..	P.....	.....H..
hs6M1-6: WT51	.....V..	M.....E..	P.....	.....H..

	CP4
hs6M1-3:BM19.7	RGAVKRLMGW E
hs6M1-3: WT51	R..V...M... .
hs6M1-6: LG2	K..A...L.. .WGK
hs6M1-6: WT51	K..A...L.. .WGK

Abb. 7: AS Alignment von *hs6M1-3* und *-6*. Für die beiden Gene sind exemplarisch jeweils nur zwei Zelllinien aufgeführt. Die grauen vertikalen Boxen zeigen Mutationen innerhalb eines Gens; die schwarze vertikale Box zeigt die Mutation, die bei beiden Genen an einer korrespondierenden Position auftritt.

#### 4.1.3. Aminosäure (AS) Austausch in den OR

Die Häufigkeit der einzelnen AS Austausch in den zehn Zelllinien ist sehr unterschiedlich. Zum Beispiel wurde der Ala146Val Austausch in TM4 von *hs6M1-6* in 11 von 19 analysierten Chromosomen gefunden, während der Ala111Thr Austausch in TM3 von *hs6M1-6* nur in der Zelllinie WT51 beobachtet wurde. Es wurden konservative Substitutionen entdeckt, in der beide AS ähnliche Eigenschaften besitzen, wie z.B. die Ala182Val Substitution in EC3 von *hs6M1-4P*. Ebenso häufig wurden aber auch drastische Veränderungen gefunden, wie z.B. die Gln234Arg Substitution in CP3 von *hs6M1-10*, die zu einem Wechsel der Polarität der Aminosäure führt.

Sowohl die HLA-gekoppelten als auch die OR Gene auf den Chromosomen 7, 17 und 19 sind polymorph und zeigen in gleicher Weise konservative wie auch nicht konservative Aminosäure Austausch. Bei den nicht HLA-gekoppelten Genen wurden die Mutationen nur in TM1, 3 und 4 und in EC2 und 3 gefunden. Da aber nur drei Gene analysiert wurden, fehlt dieser Aussage die statistische Relevanz.

#### 4.1.4. OR Gene mit potentiell funktionstüchtigen Allelen und nicht funktionstüchtigen Allelen

Für die beiden Gene *hs6M1-4P* und *hs6M1-17* wurden in einigen Zelllinien intakte Allele gefunden, während in anderen Zelllinien (lediglich eine Zelllinie bei *hs6M1-17*) Allele gefunden wurden, die ein Stop-Codon aufweisen. In dem *hs6M1-4P* Gen ist das Stop-Codon an Position 194 in der EC3 Domäne lokalisiert (s. Tab. 15) und wurde in 11 von 19 analysierten Chromosomen gefunden. Die beiden Zelllinien KR3598 und OLGA sind bezüglich dieser AS Position heterozygot, d.h. sie besitzen sowohl ein intaktes Allel als auch ein Allel ohne offenes Leseraster.

In dem *hs6M1-17* Gen wurde das Stop-Codon an Position 55 in CP1 gefunden, allerdings nur in einem (BM19.7) von 19 Chromosomen. Im Gegensatz zu den oben genannten Einzelbasenaustauschen von *hs6M1-4P* und *-17*, weist *hs6M1-19P* in 12 analysierten Chromosomen 6 eine Deletion von 16 bp auf (in EC3), die zu einer Verschiebung des Leserasters führt. Die übrigen sieben Chromosomen weisen für dieses Gen keine Deletion auf und besitzen daher ein intaktes Allel.

Tab. 15: Nukleotid- und Aminosäuresubstitutionen sowie stumme Mutationen in den verschiedenen Domänen der HLA-gekoppelten und nicht HLA-gekoppelten OR Gene. Angegeben sind auch Allelbezeichnungen sowie die HLA-Klasse I Haplotypen und ethnische Abstammung der Personen, von denen die Linien abstammen.

Domäne, Aminosäure- Position	Zelllinie, HLA-Typ, ethnische Abstammung											PAC/BAC	
	BM28.7	BM19.7	LG2	KR3598	H2LCL	WT51	SA	YAR	OLGA	AMAI			
	A1, B35	A2, B13	A2, B27	A2, B44	A3, B7	A23, B65	A24, B7	A26, B38	A31, B62	A68, B53			
	Schwarzafzr.	Schwarzafzr.	Kaukasier	Kaukasier	Kaukasier	Kaukasier	Japaner	Jude	Am. Indian.	Algerier			
TM3 107	Leu CTA *01	Leu CTG *02	Leu CTA *01	Leu CTA *01	hs6M1-1							Leu CTA *01	Leu CTA *01
TM3 113	Ala GCA	Ala GCA	Thr ACA	Ala GCA	Thr ACA	Thr ACA	Thr Ala ACA GCA	Thr ACA	Thr Ala Thr Ala GCA	Thr Ala Thr Ala GCA	Thr ACA	Thr ACA	
CP3 226	Gln CAA	Gln CAA	Arg CGA	Gln CAA	Arg CGA	Arg CGA	Arg Gln CGA CAA	Arg CGA	Arg Gln Arg Gln CGA CAA	Arg Gln Arg Gln CGA CAA	Arg CGA	Arg CGA	
CP3 228	Val GTA	Val GTA	Val GTA	Val GTA	Val GTA	Ile ATA	Val GTA	Val GTA	Val GTA	Val GTA	Val GTA	Val GTA	
TM6 261	Ile ATA *02	Ile ATA *02	Ile ATA *01	Ile ATA *02	Ile ATA *01	Ile ATA *03	Met ATG *01 *02	Met ATA *01	Met ATA *01 *02	Met ATA *01 *02	Met ATA *04	Met ATA *01	
EC1 14	Ile ATT	Ile ATT	Ile ATT	Ile ATT	Ile ATT	Leu CTT	Ile ATT	Ile ATT	Ile ATT	Ile ATT	Leu CTT	Ile ATT	
EC2 84	Val GTG	Val GTG	Val GTG	Val GTG	Val GTG	Val GTG	Val GTG	Val GTG	Val GTG	Val GTG	Val GTC	Val GTG	
EC2 99	Thr ACG	Thr ACG	Thr ACA	Thr ACA	Thr ACA	Thr ACG	Thr Ala ACA GCA	Thr ACA	Thr Ala Thr Ala ACA	Thr Ala Thr Ala ACA	Thr ACG	Thr ACA	
EC3 182	Val GTT	Val GTT	Ala GCT	Ala GCT	Val GTT	Ala GCT	Ala GCT	Ala GCT	Ala GCT	Ala GCT	Ala GTT	Ala GCT	
EC3 194	Gln CAG	Gln CAG	STOP TAG	STOP TAG	Gln CAG	STOP TAG	Gln CAG	STOP TAG	STOP TAG	STOP TAG	Gln CAG	STOP TAG	
TM5 206	Ile ATT *05	Ile ATT *05	Ile ATT *01	Ile ATT *01 *02	Ile ATT *01	Ile ATA *03	Ile ATA *01	Ile ATA *01	Ile ATA *01 *02	Ile ATA *01 *02	Ile ATA *04	Ile ATT *01	
TM2 74	Tyr TAC	Tyr TAC	His CAC	Tyr TAC	His CAC	His CAC	Tyr TAC	Tyr TAC	His CAC	Tyr TAC	His CAC	Tyr TAC	
TM3 111	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Thr ACA	Thr GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	
TM3 120	Ser TCG	Ser TCG	Ser TCG	Ser TCG	Ser TCG	Ser TCA	Ser TCA	Ser TCG	Ser TCG	Ser TCG	Ser TCG	Ser TCG	
TM4 146	Val GTT	Val GTT	Ala GCT	Val Ala GCT	Ala GCT	Val GTT	Val Ala GTT	Ala GCT	Val Ala GTT	Ala GCT	Val Ala GTT	Ala GCT	
TM5 214	Leu CTC	Leu CTC	Leu CTG	Leu CTC	Leu CTG	Leu CTC	Leu CTC	Leu CTG	Leu CTC	Leu CTC	Leu CTC	Leu CTC	
TM5 218	Thr ACC *02	Thr ACC *02	Ala GCC *01	Thr Ala ACC GCC *02 *01	Ala GCC *01	Thr Ala GCC ACC *03	Thr Ala GCC ACC *02 *01	Thr Ala GCC ACC *01	Thr Ala GCC ACC *01 *02	Thr Ala GCC ACC *01 *02	Thr Ala GCC ACC *02	Thr Ala GCC ACC *01	
CP3 234	Gln CAA *01	Gln CAA *01	Gln CAA *01	Arg CGA *02	Gln CAA *01	Gln CAA *01	Gln CAA *01	Gln CAA *01	Gln CAA *01	Gln CAA *01	Gln CAA *01	Gln CAA *01	
EC1 19	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	Pro CCA	
TM1 30	Phe TTC	Phe TTC	Leu CTC	Phe Leu TTC CTC	Leu CTC	Phe TTC	Phe Leu TTC CTC	Leu CTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	
TM1 37	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	Leu CTA	
TM1 48	Ala GCG	Val GTG	Ala GCG	Val Ala GTG GCG	Ala GCG	Ala GCG	Ala Val GTG GCG	Ala GCG	Ala Val GTG GCG	Ala Val GTG GCG	Ala GCG	Ala GCG	
EC2 78	Gln CAA *01	Gln CAA *02	Gln CAA *03	Gln CAA *02 <sub>a</sub> *03 <sub>a</sub>	Gln CAA *03	Gln CAA *01 <sub>a</sub> *04 <sub>a</sub>	Gln CAA *02 <sub>a</sub> *03 <sub>a</sub>	Gln CAA *03	Gln CAA *01 <sub>a</sub> *04 <sub>a</sub>	Gln CAA *01 *04 <sub>a</sub>	Gln CAA *01	Gln CAA *01	
EC2 81	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Val GTG	Met ATG	
TM7 279	Thr ACC	Thr ACC	Thr ACC	Thr ACC	Thr ACC	Thr ACT	Thr ACC	Thr ACC	Thr ACC	Thr ACC	Thr ACT	Thr ACC	
CP4 296	Asp GAC *01	Asp GAC *01	Asp GAC *01	Asp GAC *01	Asp GAC *01	Asn AAC *02	Asp GAC *01	Asp GAC *01	Asp GAC *01	Asp GAC *01	Asp GAC *03	Asp GAC *01	
TM2 62	Ser TCC	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	Ser TCT	
TM2 63	Asn AAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	Asp GAC	
TM5 208	Pro CCT *02	Pro CCC *03	Pro CCT *01	Pro CCT *01	Pro CCT *01	Pro CCT *01	Pro CCT *01	Pro CCT *01	Pro CCT *01	Pro CCT *01	Pro CCT *03	Pro CCT *01	

Fortsetzung von Tab.15:

	BM28.7	BM19.7	LG2	KR3598	H2LCL	WT51	SA	YAR	OLGA	AMAI	PAC/BAC	
<b>hs6M1-17</b>												
CP1 55	Gln CAG	STOP TAG	Gln CAG	Gln CAG	Gln CAG	Gln CAG	Gln CAG	Gln CAG	Gln CAG	Gln CAG	Gln CAG	
TM2 61	Phe TTC	Phe TTT	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	
EC2 89	Arg CGC	Ser AGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	
CP2 121	Arg CGC	Cys TGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	Arg CGC	
TM4 138	Arg CGG	Trp TGG	Arg CGG	Arg CGG	Arg CGG	Arg CGG	Arg CGG	Arg CGG	Arg CGG	Arg CGG	Arg CGG	
TM4 160	Pro CCT	Ser TCT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Ser TCT	Pro CCT	
EC3 174	Pro CCG	Pro CCG	Pro CCG	Gln CAG	Pro CCG	Pro CCG	Pro CCG	Pro CCG	Gln CAG	Pro CCG	Pro CCG	
TM6 246	Val GTG	Val GTG	Val GTG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Val GTG	
TM6 254	Ala GCA	Ala GCC	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	Ala GCA	
CP4 310	Met ATG	Arg AGG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	Met ATG	
	<b>*01</b>	<b>*05</b>	<b>*01</b>	<b>*03</b>	<b>*02</b>	<b>*02</b>	<b>*02</b>	<b>*02</b>	<b>*03</b>	<b>*04</b>	<b>*01</b>	
<b>hs6M1-18</b>												
TM4 165	Ala GCC	Ala GCC	Ala GCC	Thr ACC	Ala GCC	Ala GCC	Ala GCC	Ala GCC	Ala GCC	Ala GCC	Ala GCC	
	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	
<b>hs6M1-19P</b>												
EC3 186-190	del. 16 bp	-	del. 16 bp	del. 16 bp	-	-	del. 16 bp	del. 16 bp	-	-	del. 16 bp	del. 16 bp
	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	<b>*02</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	<b>*02</b>	<b>*01</b>	<b>*01</b>
<b>hs6M1-20</b>												
TM1 47	Val GTC	Phe TTC	Val GTC	Phe TTC	Phe TTC	Val GTC	Val GTC	Phe TTC	Val GTC	Val GTC	Val GTC	
TM2 56	Leu CTT	Pro CCT	Leu CTT	Pro CCT	Pro CCT	Leu CTT	Leu Pro CTT CCT	Pro CCT	Leu Pro CTT CCT	Leu CTT	Leu CTT	
TM3 104	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Ser TCC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	
TM3 113	Phe TTC	Leu TTG	Phe TTC	Leu TTG	Leu TTG	Phe TTC	Phe Leu TTC TTG	Leu TTG	Phe Leu TTC TTG	Phe TTC	Phe TTC	
CP2 120	Leu CTC	Arg CGC	Leu CTC	Arg CGC	Arg CGC	Leu CTC	Leu Arg CTC CGC	Arg CGC	Leu Arg CTC CGC	Leu CTC	Leu CTC	
CP2 121	Ser TCT	Cys TGT	Ser TCT	Cys Ser TGT TCT	Cys TGT	Ser TGT	Ser Cys TCT TGT	Cys TGT	Ser Cys TCT TGT	Ser TCT	Ser TCT	
TM4 159	Val GTA	Ile ATA	Val GTA	Ile ATA	Ile ATA	Val GTA	Ile Val ATA GTA	Ile ATA	Ile ATA	Val GTA	Val GTA	
TM6 255	Leu CTT	Leu CTC	Leu CTT	Leu CTC	Leu CTC	Leu CTT	Leu CTC	Leu CTC	Leu CTC	Leu CTT	Leu CTT	
	<b>*01</b>	<b>*03</b>	<b>*01</b>	<b>*03<sub>a</sub></b>	<b>*04<sub>a</sub></b>	<b>*03</b>	<b>*02</b>	<b>*05<sub>a</sub></b>	<b>*06<sub>a</sub></b>	<b>*03</b>	<b>*05<sub>a</sub></b>	<b>*07<sub>a</sub></b>
										<b>*01</b>	<b>*01</b>	
<b>hs6M1-21</b>												
EC1 23	Leu TTG	Leu TTG	Leu TTG	Leu TTG	Leu TTG	Leu TTG	Leu TTG	Leu TTG	Leu Trp TTG TGG	Leu TTG	Leu TTG	
TM3 106	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTC	Phe TTT	Phe TTC	
CP3 233	Gly GGA	Gly GGA	Gly GGA	Gly GGA	Gly GGA	Gly GGA	Gly GGA	Gly GGA	Gly GGA	Arg AGA	Gly GGA	
CP3 238	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	Phe TTT	
	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*01<sub>a</sub></b>	<b>*03<sub>a</sub></b>	
										<b>*04</b>	<b>*01</b>	
<b>hs7M1-2</b>												
TM4 137	His CAT	His CAT	His CAT	His Arg CAT CGT	His CAT	His CAT	His CAT	His Arg CAT CGT	His CAT	His Arg CAT CGT	His CAT	
EC3 170	Arg AGA	Arg AGA	Arg Thr AGA ACA	Arg AGA	Arg AGA	Arg AGA	Arg AGA	Arg AGA	Arg AGA	Arg AGA	Arg AGA	
	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*03</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*03</b>	<b>*01</b>	
										<b>*01</b>	<b>*03</b>	
										<b>*01</b>	<b>*03</b>	
<b>hs17M1-20</b>												
TM1 42	Ile Val ATT GTT	Ile Val ATT GTT	Ile ATT	Ile ATT	Ile ATT	Ile ATT	Ile ATT	Ile ATT	Ile Val ATT GTT	Ile ATT	Ile ATT	
EC2 86	Ser Arg AGC AGG	Ser Arg AGC AGG	Ser AGC	Ser AGC	Ser AGC	Ser AGC	Ser AGC	Ser AGC	Ser Arg AGC AGG	Ser AGC	Ser AGC	
TM3 120	Ile Val ATC GTC	Ile Val ATC GTC	Ile ATC	Ile ATC	Ile ATC	Ile ATC	Ile ATC	Ile ATC	Ile Val ATC GTC	Ile ATC	Ile ATC	
EC3 168	Pro Thr CCT ACT	Pro Thr CCT ACT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Pro CCT	Pro Thr CCT ACT	Pro Thr CCT ACT	Pro CCT	
EC3 175	Asn Ser AAT AGT	Asn Ser AAT AGT	Asn AAT	Asn AAT	Asn AAT	Asn AAT	Asn AAT	Asn AAT	Asn Ser AAT AGT	Asn AAT	Asn AAT	
EC3 193	Ala Ser GCC TCC	Ala Ser GCC TCC	Ala Ser GCC TCC	Ala Ser GCC TCC	Ala Ser GCC TCC	Ser TCC	Ala Ser GCC TCC	Ser TCC	Ala Ser GCC TCC	Ala Ser GCC TCC	Ala GCC	
	<b>*01<sub>a</sub></b>	<b>*03<sub>a</sub></b>	<b>*01<sub>a</sub></b>	<b>*03<sub>a</sub></b>	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*02</b>	<b>*02</b>	<b>*01<sub>a</sub></b>	<b>*03<sub>a</sub></b>	
										<b>*01<sub>a</sub></b>	<b>*04<sub>a</sub></b>	
											<b>*01</b>	
<b>hs19M1-4</b>												
TM1 46	Ile ATC	Ile ATC	Ile ATC	Ile Thr ATC ACC	Ile Thr ATC ACC	Ile ATC	Ile ATC	Ile ATC	Ile Thr ATC ACC	Ile ATC	Ile ATC	
	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*02</b>	<b>*01</b>	<b>*01</b>	<b>*01</b>	<b>*02</b>	
											<b>*01</b>	

a: Allel-Bezeichnung nur spekulativ



#### 4.1.5. Anzahl der OR Allele

Bei den analysierten Genen ist die Zahl der Allele teilweise sehr unterschiedlich (s. Tab. 16). *hs6M1-1*, *-10*, *-18* und *-19P* besitzen z.B. nur zwei verschiedene Allele, wobei *hs6M1-1* in beiden Allelen die gleiche Aminosäuresequenz aufweist. Die anderen OR Gene haben entweder drei (*hs6M1-6*, *-15* und *-16*), vier (*hs6M1-3*, *-12*, *-21*), fünf (*hs6M1-4P* und *-17*) oder sogar sieben (*hs6M1-20*) Allele. Die nicht HLA-gekoppelten Gene zeigen ebenfalls eine unterschiedliche Allel-Anzahl. Bei diesen Genen wurden zwei (*hs19M1-4*), drei (*hs7M1-2*) oder vier (*hs17M1-20*) Allele gefunden. Diese Gene sind daher ähnlich polymorph wie die HLA-gekoppelten.

Die Alleldefinition ist bei einigen Genen nur spekulativ, da z.T. mehr als eine heterozygote Nukleotidposition für eine bestimmte Zelllinie gefunden wurde (z.B. *hs6M1-12*, *-20* und *hs17M1-20*). Somit konnte nicht eindeutig festgelegt werden, welche der verschiedenen Austausch zu einem Allel gehören. Es wurde aber immer die geringst mögliche Anzahl an Allelen angenommen, d.h. in einer heterozygoten Zelllinie wurden die Aminosäuren – wenn möglich - immer so zu einem Allel kombiniert, wie es auch in einer in Bezug auf dieses Gen homozygoten Zelllinie zu finden war. Durch die Wahl von HLA-homozygoten und -hemizygoten Zelllinien hätte dieses Problem eigentlich umgangen werden sollen. Die Ergebnisse zeigen aber, daß sich die Homozygotie bezüglich der HLA-Gene (typisiert bis einschließlich HLA-A) nicht bis auf den OR Cluster erstreckt, der ca. 500 kb telomer von HLA-A beginnt.

Einige der analysierten OR Gene, die heterozygote Positionen aufwiesen, wurden in voller Länge kloniert und sequenziert (*hs6M1-3*, *-4P* und *-6*). Da ein Klon nur eines der beiden möglichen Allele enthalten kann, war es nun möglich, die Allele trotz mehrerer Mutationen eindeutig zu definieren.

#### 4.1.6. OR Haplotypen

Der hier analysierte OR Cluster liegt in einem Gebiet, das ein starkes Kopplungsungleichgewicht zum HLA-Komplex aufweist (Malfroy et al., 1997). Es ist daher anzunehmen, daß die HLA-gekoppelten OR Gene "en bloc" vererbt werden. Somit kann man erweiterte Haplotypen für diese OR Gene annehmen (s. Tab. 16), die eine Distanz von ca. 2000 kb umfassen – von *hs6M1-10* bis *hs6M1-12* (s. Abb. 4). Von den analysierten zehn Zelllinien und den Kombinationen der 13 HLA-gekoppelten OR Allele (s. Tab. 15 u. Tab. 16) können mindestens 13 verschiedene Haplotypen abgeleitet werden. Zwei Zelllinien (H2LCL und YAR) besitzen für alle analysierten 13 OR Gene identische Allele, obwohl sie unterschiedliche HLA-Klasse I Haplotypen aufweisen. Alle anderen ausgedehnten OR Haplotypen unterscheiden sich. Bei den heterozygoten Zelllinien kann man unterschiedliche Haplotypen mütterlicher- und väterlicherseits festlegen. Allerdings sind bestimmte

Kombinationen von Allelen benachbarter Gene bei einigen Zelllinien konserviert. Zum Beispiel ist ein Haplotyp der Zelllinie SA – mit Ausnahme von zwei OR Genen (*hs6M1-20* und *hs6M1-21*) – identisch mit den Haplotypen von H2LCL und YAR. Ebenso haben die beiden Zelllinien BM19.7 und BM28.7 identische Allele für die folgenden OR Gene: *hs6M1-10*, *-15*, *-4P*, *-3*, *-6*, *-21* und *-18*. Jeweils ein Haplotyp von OLGA und ein Haplotyp von KR3598 besitzen gemeinsame Allele für die *hs6M1-15*, *-1*, *-4P*, *-3*, *-6*, *-19P*, *-17* und *-16* Gene. Die OR Haplotypen von WT51 und AMAI unterscheiden sich von den restlichen Zelllinien am stärksten. Die drei verschiedenen HLA-A2 Haplotypen (in BM19.7, LG2 und KR3598) sind mit vier verschiedenen HLA-gekoppelten OR Haplotypen assoziiert.

Zelllinie	OR Gen - hs6M1-													HLA		OR
	10	15	1	4P	3	6	21	20	19P	18	17	16	12	A	B	Haplotyp
H2LCL	*01	*01	*01	*01	*01	*01	*01	*03	*02	*01	*02	*01	*03	A3	B7	<b>1</b>
YAR	*01	*01	*01	*01	*01	*01	*01	*03	*02	*01	*02	*01	*03	A26	B38	<b>1</b>
LG2	*01	*01	*01	*01	*01	*01	*01	*01	*01	*01	*01	*01	*03	A2	B27	<b>2</b>
OLGA-1	*01	*01	*01	*01	*01	*01	*01	*05	*01	*01	*03	*01	*01	A31	B62	<b>3</b>
OLGA-2	*01	*01	*01	*02	*02	*02	*03	*07	*01	*01	*03	*01	*04	A31	B62	<b>4</b>
SA-1	*01	*01	*01	*01	*01	*01	*02	*05	*02	*01	*02	*01	*03	A24	B7	<b>5</b>
SA-2	*01	*01	*01	*01	*02	*02	*02	*06	*01	*01	*02	*01	*02	A24	B7	<b>6</b>
BM28.7	*01	*01	*01	*05	*02	*02	*01	*01	*01	*01	*01	*02	*01	A1	B35	<b>7</b>
BM19.7	*01	*01	*02	*05	*02	*02	*01	*03	*02	*01	*05	*03	*02	A2	B13	<b>8</b>
WT51-1	*01	*02	*01	*03	*03	*03	*01	*02	*01	*01	*02	*01	*01	A23	B65	<b>9</b>
WT51-2	*01	*02	*01	*03	*03	*03	*01	*02	*01	*01	*02	*01	*04	A23	B65	<b>10</b>
KR3598	*02	*01	*01	*01	*02	*01	*01	*03	*02	*02	*03	*01	*02	A2	B44	<b>11</b>
KR3598	*02	*01	*01	*02	*02	*02	*01	*04	*01	*02	*03	*01	*03	A2	B44	<b>12</b>
AMAI	*01	*03	*01	*04	*04	*02	*04	*01	*01	*01	*04	*03	*01	A68	B53	<b>13</b>
Anzahl der Allele	<b>2</b>	<b>3</b>	<b>2</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>4</b>	<b>7</b>	<b>2</b>	<b>2</b>	<b>5</b>	<b>3</b>	<b>4</b>			

Tab. 16: OR Allele und Haplotypen in den zehn analysierten Zelllinien. Für die heterozygoten Zelllinien Olga, SA, WT51 und KR3598 sind jeweils beide Haplotypen angegeben.

Grau: Identische Allel-Kombination benachbarter Gene, die mindestens sieben Gene umfassen.

## 4.2. Expressionsanalyse

Ein weiteres Ziel dieses Projektes war es, zu untersuchen, in welchen Geweben bzw. Zellen HLA-gekoppelte OR Gene exprimiert werden. Die Gewebezuordnung sollte durch Hybridisierung radioaktiv markierter Sonden mit einem Dotblot erreicht werden, auf den mRNA verschiedener Gewebe aufgebracht war. Des Weiteren sollte durch die Methode der *In-situ* Hybridisierung mit Gewebeschnitten die Expression in den einzelnen Zelltypen eines Organs untersucht werden.

### 4.2.1. Spezifität von OR Sonden

Für die radioaktive Markierung des mRNA-dot-Blot's mußten zunächst spezifische OR Sonden hergestellt werden. Dies erfolgte mittels PCR aus genomischer DNA. Die genomische DNA stammte von den zehn verschiedenen Zelllinien aus Tab. 1 (meist LG2) und die verwendeten Primer stammten aus Tab. 2 (Polymorphismusanalyse). In Abhängigkeit von der Spezifität der Sonde wurde wahlweise der 5' oder 3' Bereich des jeweiligen Gens amplifiziert (s.u.). Von zehn HLA-gekoppelten OR Genen mit ORF (*hs6M1-1*, *-3*, *-4P*, *-6*, *-10*, *-15*, *-16*, *-17*, *-18*, *-21*) wurden Genfragmente (500-700 bp) amplifiziert, auf ein PA-Gel aufgetragen und anschließend aus diesem isoliert und gereinigt (s. Abschnitt 3.7.1.1). Wie in der Einleitung bereits erwähnt, sind die OR Gene z.T. sehr homolog zueinander. Deshalb mußten die Sonden zunächst auf ihre Spezifität hin getestet werden. Dafür wurden vier identische Southern-Blots mit restriktionsverdauter humaner genomischer DNA verschiedener Individuen erstellt und dann jeweils mit verschiedenen <sup>32</sup>P-markierten OR Sonden hybridisiert (s. Abschnitt 3.13). Die Anzahl der gefundenen Banden gab dann Aufschluß über die Anzahl ähnlicher Kopien im Genom. War nur eine Bande in der entsprechenden Spur zu sehen, wurde diese Sonde als spezifisch für das jeweilige Gen angesehen. Wurden mehrere Banden gefunden, wurde eine neue Sonde mit dem anderen Primerpaar aus Tab. 2 hergestellt und ebenfalls getestet. Die Hybridisierungslösung der Sonden wurde bei -20°C gelagert, und, falls spezifisch, für die Hybridisierung des mRNA-dot-Blot's verwendet. Für einige Gene (*hs6M1-3*, *-4P*, *-6* und *-16*) war es nicht möglich, spezifische Sonden zu erstellen, da diese OR Gene z.T. eine Sequenzhomologie von über 90% mit anderen HLA-gekoppelten OR Genen aufwiesen. Die Gene *hs6M1-3*, *-4P* und *-6* weisen untereinander und zu *hs6M1-5P* eine starke Homologie auf (s. Abb. 8), während *hs6M1-16* sehr ähnlich zu *-12* und *-13P* ist (s. Abb. 9). Diese Proben wurden dennoch für den mRNA-dot-Blot verwendet, wobei aber später vermerkt wurde, mit welchen Genen sie vermutlich eine Kreuzreaktion aufgewiesen hatten.

	1		80
<b>hs6M1-3</b>	ATGAATGATG ATGGAAAAGT CAATGCTAGC TCTGAGGGGT ACTTTATTTT AGTTGGATT TCTAATTGGC CTCATCTGGA		
<b>hs6M1-5P</b>	.....T.....		
<b>hs6M1-4P</b>	...T... ..A...A- -....A..T .T...A.AC. T.....C. .C..... .C.....		
<b>hs6M1-6</b>	.....TA...A- -....A..T ..G..A.AC. T.....C. .C..... ..G.....		
	81		160
<b>hs6M1-3</b>	AGTAGTTATC TTTGTGGTTG TCTTGATCTT CTACTIONGATG ACACTIONGATG GAAACCTGTT CATCATCATC CTGTCATACC		
<b>hs6M1-5P</b>	.....C... ..A .T..... ..G..... ..A.G....		
<b>hs6M1-4P</b>	.....C... ..A ..A.....A ..A.....		
<b>hs6M1-6</b>	.....C... ..A ..C..... ..C..... ..G.....		
	161		240
<b>hs6M1-3</b>	TGGACTCCCA TCTGCACACA CCAATGTACT TCTTCCTTTC AAACCTCTCA TTTCTGGATC TCTGCTACAC CACCAGCTCT		
<b>hs6M1-5P</b>	.....C..T..T ..CT...T. ....T.....		
<b>hs6M1-4P</b>	.....C...T ..C..... ..T.....		
<b>hs6M1-6</b>	.....C..... ..C..... ..C..... ..C.....		
	241		320
<b>hs6M1-3</b>	ATCCCTCAGT TGCTGGTCAA TCTCTGGGGC CCGAAAAGA CCATCTCTTA TGCTGGTGC ATGATTCAAC TTTACTTTGT		
<b>hs6M1-5P</b>	.....G .....T GT..... ..T..... ..G..... ..T.....		
<b>hs6M1-4P</b>	.....G.. .....T .....CAG.....		
<b>hs6M1-6</b>	.....G.. ..C..... ..G..... ..G.....		
	321		400
<b>hs6M1-3</b>	TCTCGCACTG GGAACCACAG AGTGTGCCT ACTGGTGGTG ATGTCCTATG ACCGTTATGC AGCTGTGTGT AGACCTTTGC		
<b>hs6M1-5P</b>	...A.....		
<b>hs6M1-4P</b>	.....G..... ..T.....		
<b>hs6M1-6</b>	...T..... ..T.G... ..G..... ..T.....T ..T.....		
	401		480
<b>hs6M1-3</b>	ATTACACTGT CCTCATGCAC CCTCGTTTCT GCCACCTGCT GGCTGTGGCT TCTTGGGTAA GTGGTTTTAC CAACTCAGCA		
<b>hs6M1-5P</b>	.....T..... ..T..... ..T..... ..A...C....		
<b>hs6M1-4P</b>	.....G..T..T. ....C..... ..A.C.....		
<b>hs6M1-6</b>	.....T..T. ....C..... ..T..... ..C..... ..T..... ..T.....		
	481		560
<b>hs6M1-3</b>	CTTCATTCCT CCTTCACCTT CTGGGTACCT CTGTGTGGAC ACCGCCAAGT AGATCACTTT TTCTGTGAAG TTCCAGCACT		
<b>hs6M1-5P</b>	.....A..... ..G.....		
<b>hs6M1-4P</b>	.....T..T.. ..A...C ..A...A... ..T...T... G.....C ..C.....		
<b>hs6M1-6</b>	.....T..T.. .....C ..T..... ..T...T... G.....C ..C.....		
	561		640
<b>hs6M1-3</b>	TCTGCGATTA TCGTGTGTG ATACCCATGT CAATGAGCTG ACCCTCATGA TCACAAGCTC CATATTGTT CTCATACCTC		
<b>hs6M1-5P</b>	.T.AT..... ..A.T...CA .....G..A A...A.A... ..T..... ..GC...T..		
<b>hs6M1-4P</b>	.....A..... ..T.G.C A..... ..G ..TG..... ..T.....		
<b>hs6M1-6</b>	.....T... ..A..... ..C.....C A..... ..G ..TG..... ..T.....		
	641		720
<b>hs6M1-3</b>	TCATCCTCAT TCTCACTTCT TATGGTGCCA TCGTCCGAGC TGTACTGAGG ATGCAGTCAA CCACTGGGCT TCAGAAAAGT		
<b>hs6M1-5P</b>	...C..... ..T.....C ..T.C..AG.. ..A.....		
<b>hs6M1-4P</b>	.....C..... ..C..... ..T.C...G... ..C ..A.....		
<b>hs6M1-6</b>	...T..G... ..G.C ..T.C...G... ..C ..A.....		
	721		800
<b>hs6M1-3</b>	TTTGAACAT GTGGAGCTCA TCTTATGGCT GTATCTCTCT TTTTCATTCC GGCCATGTGC ATATATCTCC AGCCACCATC		
<b>hs6M1-5P</b>	.....A.....T. ....G.....		
<b>hs6M1-4P</b>	C..A.G... ..C... ..T..... ..A.T..... ..G.....		
<b>hs6M1-6</b>	...A.G... ..C... ..T..... ..A.T..... ..G.....		
	801		880
<b>hs6M1-3</b>	AGGAAATTCT CAAGATCAAG GCAAGTTCAT TGCCCTCTTT TATACTGTG TCACACCTAG TCTTAACCT CTAACTACA		
<b>hs6M1-5P</b>	...G..... ..T..... ..T.....		
<b>hs6M1-4P</b>	..A..... ..C.....		
<b>hs6M1-6</b>	..A..... ..CT...G. ....G..... ..T.....		
	881		945
<b>hs6M1-3</b>	CCCTCAGAAA CAAAGTTGTA AGAGGGGCG TGAAGAGACT AATGGGGTGG GAA		
<b>hs6M1-5P</b>	.....A..... ..T..... ..G..... ..G.....		
<b>hs6M1-4P</b>	..TT..... ..G.A... ..TGGGGGA TG		
<b>hs6M1-6</b>	..T..... ..GCA... ..A..... ..C..... ..T..... ..GTGGGGGA AG		

Abb. 8: Alignment von *hs6M1-3*, *-4P*, *-5P* und *-6*. Diese vier Gene weisen eine Homologie von ~90% auf.

```

1
hs6M1-12 ATGGTTAACC AAAGCTCCAC ACCGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGGG CTGGAAAGGA CTCCTCTCGT
hs6M1-13P .....G. ...A....T .....CA .....T..
hs6M1-16 .....C. CAT..... .....CA .....T..

81
hs6M1-12 GGTGTCTTC ACTTCCTACC TCCTAACCCCT AGTGGGCAAC ACACTCATCA TCCTGCTGTC TGCCTGGAC CCCAAGCTCC
hs6M1-13P A.....C G.....T.....G.....
hs6M1-16 .....T.G.... G..... ..TA..T.. ..G.....

161
hs6M1-12 ACTCTCCAAT GTACTTTTTC CTCTCCAACC TCTCCTTCTT GGACCTCTGT TTCACCACGA GTTGTGTTCC CCAAATGCTG
hs6M1-13P .....TA. ....C.. .GGG.....
hs6M1-16 .....TG... .....C ..T...A. ....C.. ..G.....

241
hs6M1-12 GTCAACCTCT GGGGCCCAAA GAAGACCATC AGCTTCCTGG ACTGCTCTGT CCAGATCTTC ATCTTCTCTGT CCCTGGGGAC
hs6M1-13P .....AG..... ..T...A... G..... ..T.....
hs6M1-16 .....GA..... ..C.....

321
hs6M1-12 AACTGAGTGC ATCCTCTTGA CAGTGATGGC TTTTGATCGC TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA
hs6M1-13P C.....C... .G..... C....C... ..A....A ..T.A... ..G.G.C... ..G
hs6M1-16 C.....C... ..C....C..A .....

401
hs6M1-12 TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGCATTGGG CTAGTGGAGT CAGTGGTCCA GACACCATCC
hs6M1-13P .....T.T. ....G..A....
hs6M1-16 .....T..GA.T ..G..TC.A. .GA.A....

481
hs6M1-12 ACCCTGCACC TGCCCTTCTG CCCCGATCGG CAGGTGGATG ATTTTGTCTG TGAGGTCCCA GCTCTAATTC GACTCTCCTG
hs6M1-13P .....G.. ..T..... ..C...A. ....
hs6M1-16 .....C...T .....T...C.C.A. ...A.A... .C...T.A.. .....T...G....

561
hs6M1-12 TGAAGACACC TCCTACAATG AGATCCAGGT GGCTGTTGCC AGTGTCTTCA TCTTGGTGTG GCCTCTCAGC CTCATCCTTG
hs6M1-13P .....A. ....C... ..
hs6M1-16 ..G...T... ..A.....T. ....GT. ....A..T ..G.....

641
hs6M1-12 TCTCTTACGG AGCCATTACC TGGGCAGTGC TGAGGATTAA CTCTGCAAAA GGGCGGAGGA AAGCTTTTGG GACCTGCTCC
hs6M1-13P .....T.. .....G.. .....A...C... ..A.....
hs6M1-16 C.....T.. .....C.G.. CA..... ..C.C. .CAT...A. .G..C....

721
hs6M1-12 TCCATCTCA CTGTGGTCAC CCTCTTCTAC AGCTCAGTCA TTGCTGTCTA CCTCCAGCCC AAAAATCCCT ATGCCCAAGA
hs6M1-13P .....
hs6M1-16 .....G. ....G

801
hs6M1-12 GAGGGGCAAG TTCTTTGGTC TCTTCTATGC AGTGGGCACT CCTTCACTTA ACCCTCTCAT ATACACCCTG AGGAACAAGG
hs6M1-13P .....
hs6M1-16 .....G. ....

881
hs6M1-12 AGGTAACCAG GGCATTACAG AGATTGCTGG GGAAGGAAAT GGGGCTCACA CAAAGCTGA
hs6M1-13P .....C.....T.....T...
hs6M1-16 ..A...AGC. A...C..... ..G..A..A. ....G A.ACTC..GG G.....GA GAGCTGCTTA A

```

Abb. 9: Alignment von *hs6M1-12*, *-13P* und *-16*. Diese drei Gene weisen eine Homologie von ~90% auf.

Um Informationen darüber zu erhalten, ob diese HLA-gekoppelten OR Gene auch in anderen Spezies auftreten, wurde zusätzlich restriktionsverdauter genomische DNA anderer Tierarten auf die Southern-Blots aufgebracht: Schimpanse, Pavian, Grüne Meerkatze, Ratte, Rind und Schwein (s. Abb. 10) sowie E.coli DNA als Negativ-Kontrolle. Weiterhin wurden zwei zusätzliche Pseudogene analysiert (*hs6M1-2P* und *-7P*), die nicht für den mRNA-Blot

verwendet wurden. Abb. 10 zeigt, daß nur die Gene *hs6M1-1*, *-2P*, *-10*, *-15*, *-17*, *-18* und *-21* in den humanen Zelllinien als Einzelkopie vorkommen. Interessanterweise wurden beim Schwein für die Gene *hs6M1-1*, *-2P* und *-15* mehrere Banden gefunden. Alle anderen analysierten Gene zeigen zwei (z.B. *hs6M1-16*) bis sieben (z.B. *hs6M1-3*) Kopien im menschlichen Genom und bis zu 10 Kopien (*hs6M1-6*) in den Genomen der anderen Spezies. Bis auf *hs6M1-17* sind alle Gene auch beim Schimpansen und Pavian präsent, während im Genom der Grünen Meerkatze nur ein Teil der OR Gene zu finden ist (*hs6M1-1*, *-3*, *-6*, *-7P*, *-10*, *-15*, *-16* und *-21*). Im Genom des Schweins kommt ebenfalls der größte Teil der analysierten Gene vor; nur *hs6M1-7P* und *-21* sind dort nicht zu finden. Beim Rind sind kaum Banden erkannt worden (nur *hs6M1-3*, *-6*, *-15* und *-16*), aber noch seltener sind sie bei der Ratte zu finden (nur *hs6M1-16*). Bei Schwein, Rind und Ratte sind alle Banden nur sehr schwach. Die Negativ-Kontrolle – verdaute E.coli DNA – zeigte bei allen Sonden keine positiven Signale. Die Sequenz von *hs6M1-3*, *-4P*, *-6* und *-5P* weisen eine Sequenzhomologie von ca. 90% auf (s. Abb. 8), was wahrscheinlich zu der beobachteten Kreuzreaktion führte. Eine ähnliche Situation findet man für *hs6M1-16*, das starke Homologie zu *hs6M1-12* (90%) und *-13P* (87%) aufweist (s. Abb. 9).

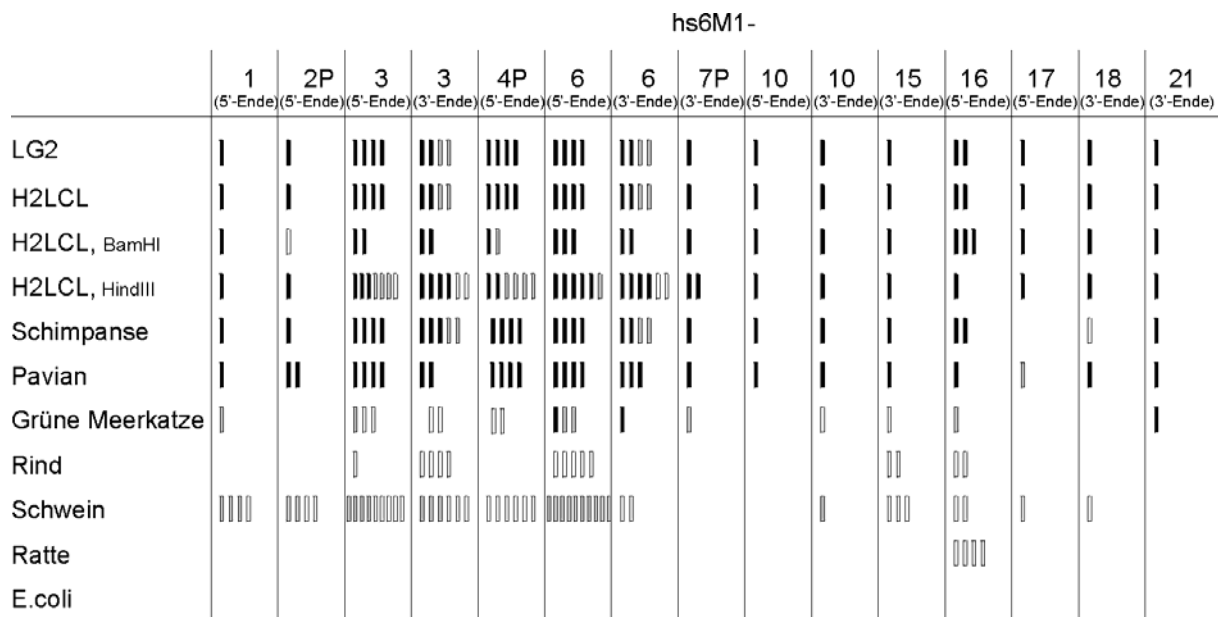
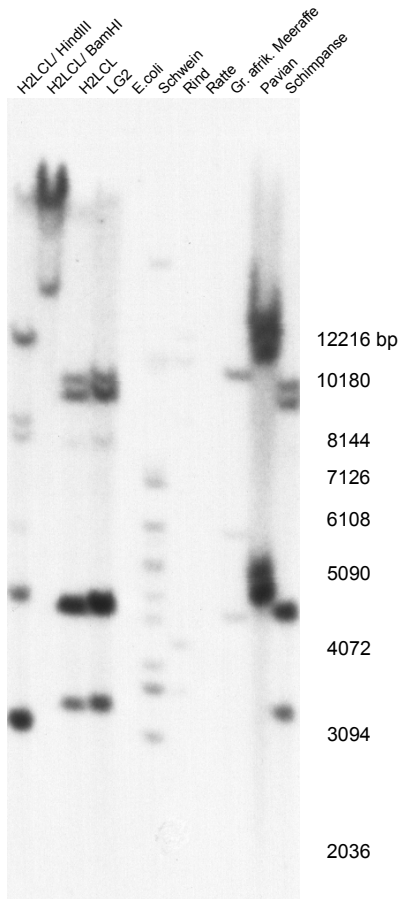


Abb. 10: Darstellung der Hybridisierungsergebnisse eines Southern-Blots mit unterschiedlichen <sup>32</sup>P-markierten HLA-gekoppelten OR Gen Sonden. Die Sonden umfaßten entweder die erste Hälfte des jeweiligen Gens mit dem 5' flankierenden Bereich oder die zweite Hälfte des Gens mit dem 3' flankierenden Bereich. Die genomische DNA der verschiedenen Spezies (humane Zelllinien: LG2 und H2LCL) wurde mit EcoRI bzw. z.T. mit BamHI und Hind III verdaut. Gezeigt ist die Anzahl und Intensität der gefundenen Banden bei der jeweiligen Spezies. Schwarz: starkes Signal; grau: mittleres Signal; weiß: schwaches Signal

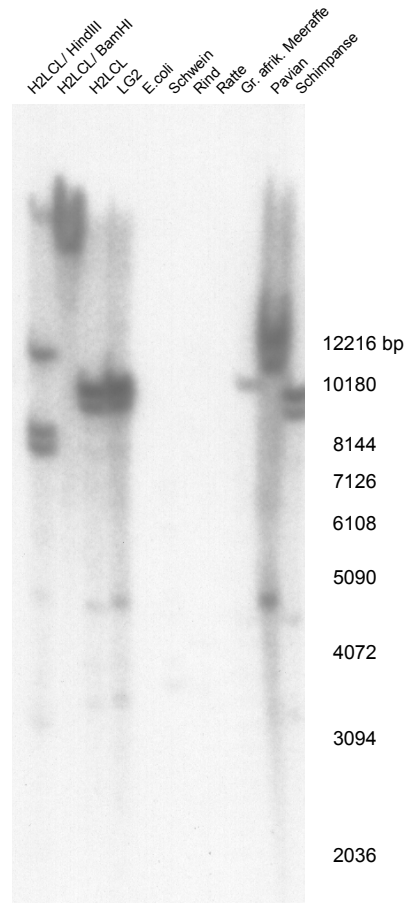
Vier ausgewählte Autoradiographien der Hybridisierungsergebnisse sind in Abb. 11 dargestellt. *hs6M1-17* zeigt jeweils nur eine einzige Bande bei Mensch, Schwein und Pavian, wobei die Bande beim Schwein nur sehr schwach ist. Die beiden menschlichen mit EcoRI verdauten Zelllinien zeigen erwartungsgemäß jeweils eine Bande gleicher Größe (~3000 bp), beim Pavian findet man ein positives Signal bei ~2000 bp und beim Schwein bei ~13000 bp. *hs6M1-16* reagiert beim Schimpansen und den menschlichen Zelllinien mit jeweils zwei EcoRI Fragmenten, die bei beiden Spezies, die gleiche Größe aufweisen. Dies deutet daraufhin, daß die *hs6M1-16*-Sonde zwei OR Gene beim Menschen und Schimpansen erkennt, die in gleicher Weise genomisch organisiert sind. Beim Pavian wird nur eine positive Bande erkannt, die etwa auf gleicher Höhe mit der oberen Bande von Mensch und Schimpanse liegt. Ein ähnliches Muster findet man auch bei *hs6M1-3* und -6. Beide Autoradiographien zeigen jeweils zwei EcoRI Fragmente von ~10.000 kb bei Mensch und Schimpanse. Beim Pavian findet man ebenfalls Doppelbanden mit ähnlichem Abstand, die aber ~2500 bp größer sind. Da diese Spur jedoch etwas überladen ist, könnte der Größenunterschied auch einfach auf einem veränderten Laufverhalten der DNA beruhen. Weitere Banden, die bei Mensch und Schimpanse ein ähnliches Muster aufweisen, findet man noch bei ~4500 bp und ~3200 bp. Die *hs6M1-3* Sonde reagiert schwach mit etwa zehn EcoRI Fragmenten des Schweins (kaum zu sehen), während derartige Banden mit der 3' Sonde von *hs6M1-6* nicht auftauchen. Verwendet man jedoch die 5' Sonde von *hs6M1-6* werden auch bei diesem OR etwa 10 EcoRI Fragmente erkannt (s. Abb. 10). Insgesamt gesehen zeigen diese beiden Blots von *hs6M1-3* und -6 ein sehr ähnliches Bandenmuster. Wie bereits erwähnt, weisen diese beiden Gene (*hs6M1-3* und -6) und die Gene *hs6M1-4P* und -5P eine Sequenzhomologie von über 90% auf. Wahrscheinlich repräsentieren auf beiden Blots die vier kräftigen Banden bei den EcoRI-verdauten humanen und Schimpansen DNA diese vier sehr homologen Sequenzen.

Abb. 11 Autoradiographie von vier exemplarisch ausgewählten Southern-Blots. Die <sup>32</sup>P-markierten Sonden umfaßten entweder die erste Hälfte des jeweiligen HLA-gekoppelten OR Gens mit dem 5' flankierenden Bereich oder die zweite Hälfte des Gens mit dem 3' flankierenden Bereich. Die genomische DNA von verschiedenen humanen Zelllinien, sowie verschiedenen Säugern und E.coli wurde mit EcoRI bzw. z.T. mit BamHI und HindIII verdaut. Die *hs6M1-3*, -6 und -16 Sonden detektieren mehrere OR Kopien in verschiedenen Genomen, während die *hs6M1-17* Sonde nur jeweils eine einzige Kopie detektiert.

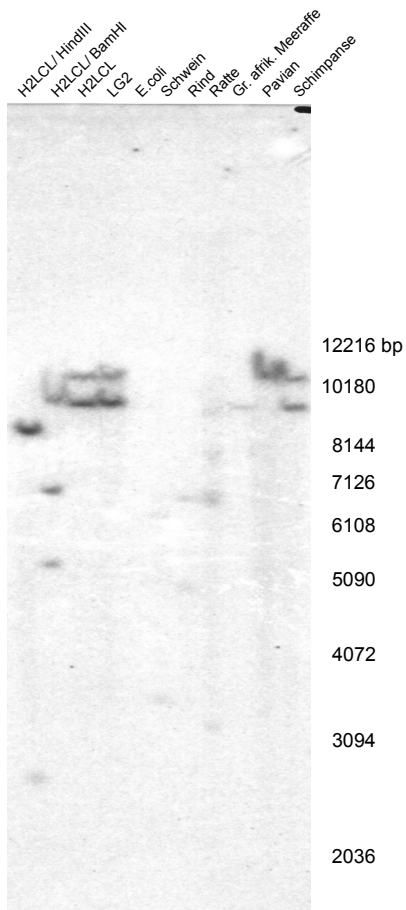




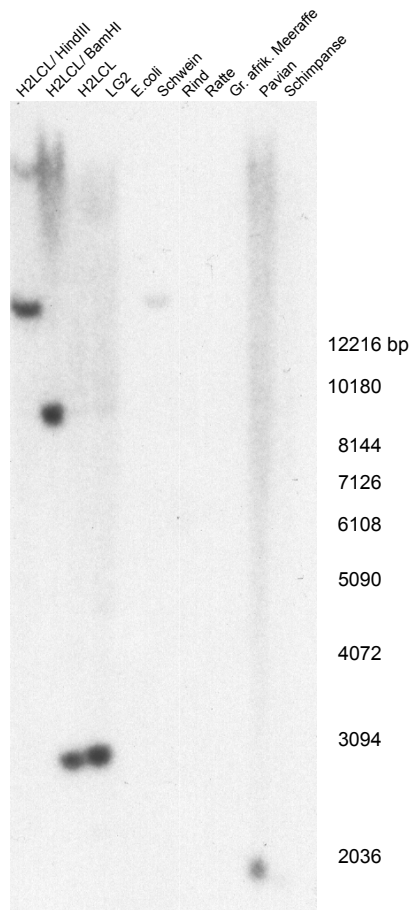
hs6M1-3 (5'Ende)



hs6M1-6 (3'Ende)



hs6M1-16 (5'Ende)



hs6M1-17 (5'Ende)

#### 4.2.2. HLA-gekoppelte OR Transkripte in verschiedenen Organen

Um die Gewebeverteilung der OR zu untersuchen, wurde ein Dot-Blot der Fa. Clontech verwendet, der PolyA<sup>+</sup> mRNA aus 50 verschiedenen menschlichen Geweben (s. Tab. 17), sowie sechs verschiedene Kontroll RNAs und DNAs enthielt. Die RNA war gepoolt von mehreren Individuen, die zwischen 16 und 80 Jahren alt waren. Da nur ein mRNA-Blot zur Verfügung stand, konnten nicht alle HLA-gekoppelten Gene getestet werden. Als Sonden wurden in der Regel die gereinigten PCR-Produkte verwendet, die bei der Hybridisierung des Southern-Blots (s. Abschnitt 4.2.1) eine einzige Bande im humanen Genom zeigten. In Ausnahmefällen wurden allerdings auch OR-Sonden verwendet, die mehr als ein Gen im humanen Genom detektierten (s. Abschnitt 4.2.1 und Abb. 11), was aber bei den Ergebnissen ausdrücklich vermerkt wurde. Folgende Gene wurden analysiert: *hs6M1-16*, *-10*, *-6*, *-1*, *-17* und *-18*.

Die Positiv-Kontrolle (humane DNA) zeigt bei allen analysierten Genen das korrekte Ergebnis (s. Tab. 17), während die Negativ-Kontrollen (*E.coli* DNA und RNA sowie Hefe RNA) bei *hs6M1-10* leider ein Signal ergaben. Da eine BLAST-Suche mit der *hs6M1-10* Sequenz keine positiven Ergebnisse bei *E.coli* ergab, läßt sich dieses Resultat nur als Artefakt erklären. Daher sind die gezeigten Ergebnisse von *hs6M1-10* mit Vorsicht zu betrachten.

Tab. 17 zeigt, daß zwei der getesteten Gene (*hs6M1-16* (-12) und *hs6M1-18*) nur in drei Geweben transkribiert werden, während die anderen Gene (*hs6M1-10*, *-6* (-3, -4P), *-1* und *-17*) (Die in den Klammern angegebenen Gene zeigen vermutlich eine Kreuzreaktion mit der jeweiligen Sonde) in fast allen auf dem Blot repräsentierten Organen transkribiert werden. Dies verdeutlicht auch Abb. 12, wo exemplarisch einige der Autoradiographien des mRNA-Blot's dargestellt sind. Anhand der Graustufe der jeweiligen Punkte kann man nur eine ungefähre Aussage über die Quantität der mRNA machen, da die Schwärzung des Filmes nicht linear mit der Menge an mRNA korreliert ist. *hs6M1-16* (-12) zeigt ein besonders starkes positives Signal im Hoden (s. Abb. 12, Tab. 17), ein weniger intensives Signal in der Niere und ein relativ schwaches Signal in der Lunge. Ob jetzt *hs6M1-16* oder *hs6M1-12* oder vielleicht beide Gene in diesen Geweben transkribiert werden, läßt sich anhand dieser Untersuchung nicht sagen.

*hs6M1-18* wird besonders stark in der Niere (s. Tab. 17), weniger intensiv in der Lunge und schwach in der Leber transkribiert. Die anderen Gene, die in sehr vielen Geweben positive Signale hervorrufen, zeigen ebenfalls in manchen Organen eine massive Expression, während sie in vielen anderen eher schwach transkribiert werden. *hs6M1-10*, *hs6M1-6* (-3, -4P) und *-1* z.B. zeigen in Niere, Leber und Lunge kräftige Signale. *hs6M1-6* (-3, -4P) weist jedoch in relativ vielen Geweben eine massive Transkription auf, während man bei *hs6M1-17* nur im Herzen eine verstärkte Transkription findet. Alle anderen Signale sind wesentlich

schwächer.

Einige Organe exprimieren mehr HLA-gekoppelte OR Gene als andere. In der Niere und Lunge werden alle getesteten OR Gene transkribiert. Im Herzen sind bis auf *hs6M1-16(-12)* und *-18* von allen getesteten HLA-gekoppelten OR Genen Transkripte gefunden worden. Im Hoden werden ebenfalls bis auf *hs6M1-18* alle getesteten Gene exprimiert, wobei aber nur das *hs6M1-16* Gen ein starkes Signal erzeugte. Im Gegensatz dazu wurden in den Hirnarealen in der Regel nur sehr wenige Transkripte gefunden. Im Nucleus caudatus, Cerebellum und im cerebralen Cortex wurden gar keine Transkripte nachgewiesen. Auch in der Aorta, der Gebärmutter, der Brustdrüse und den peripheren Leukozyten waren kaum Transkripte nachzuweisen.

Zusammenfassend läßt sich sagen, daß einige HLA-gekoppelte OR Gene nur in wenigen spezifischen Organen exprimiert werden, während andere in zahlreichen Geweben des humanen Organismus zu finden sind. Umgekehrt gilt, daß einige Organe nur ausgewählte OR Gene stärker transkribieren, während andere ein eher breites Spektrum an OR Genen gleich stark oder gleich schwach exprimieren.

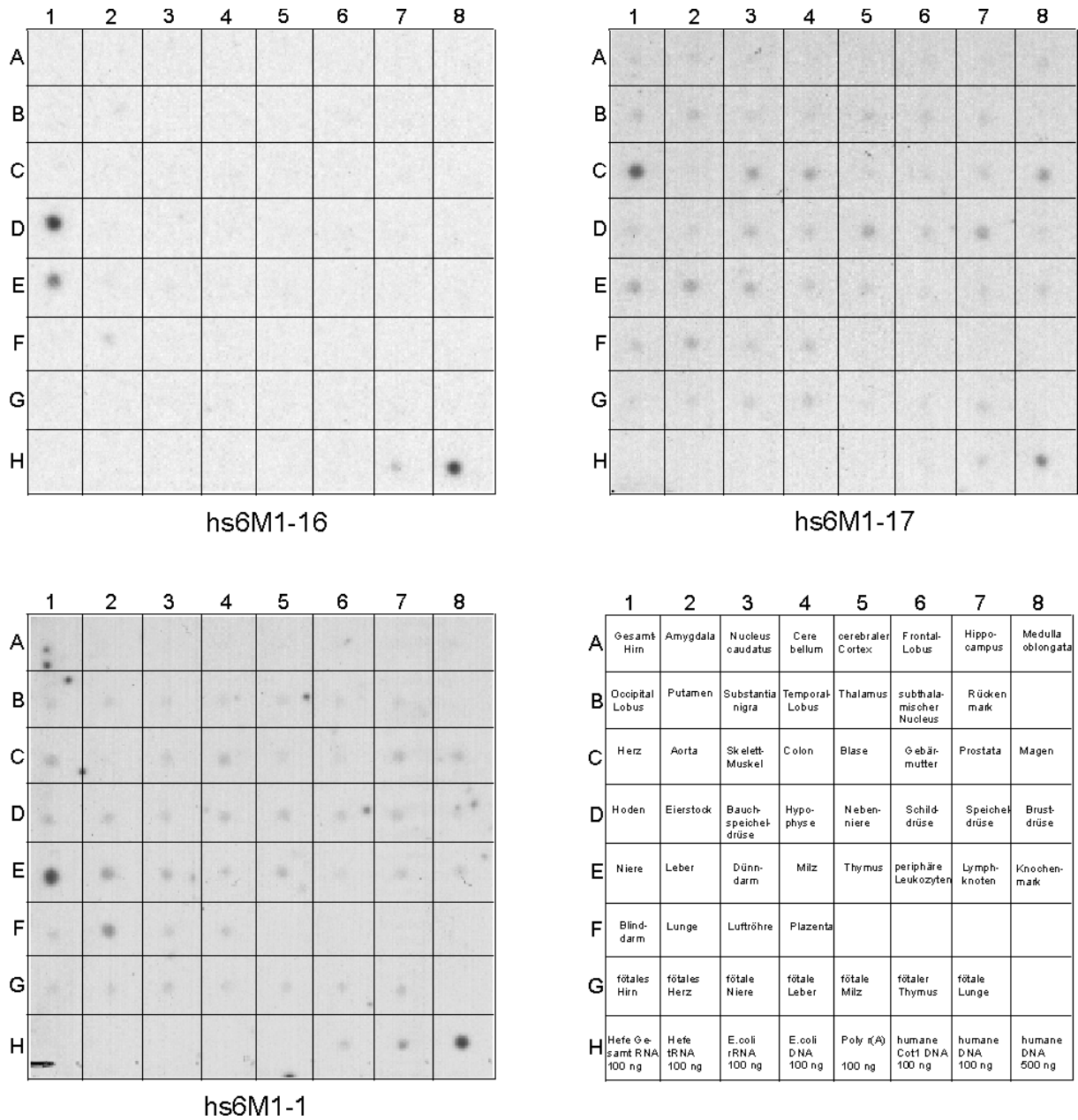


Abb. 12: Drei exemplarisch ausgewählte Autoradiographien von einem Dot-Blot mit humaner mRNA (Clontech). Die Sonden, die die OR Gene *hs6M1-17* und *-1* detektieren, sind spezifisch für diese Gene (s. Abb. 10), während die *hs6M1-16* Sonde vermutlich noch mit *hs6M1-12* kreuzreagiert. *hs6M1-16(-12)* wird nur in drei Geweben exprimiert (Hoden, Niere, Lunge), während die anderen beiden Gene in sehr vielen Geweben transkribiert werden.

Tab. 17: Autoradiographie-Ergebnisse eines Dot-Blots mit humaner mRNA (Clontech). Die Sonden für die *hs6M1-16* und *-6* Gene detektieren vermutlich noch die in den Klammern angegebenen OR Transkripte. Die Anzahl der + soll die Stärke des Signals demonstrieren.

	hs6M1-							hs6M1-					
	16 (12)	10	6 (3/ 4P)	1	17	18		16 (12)	10	6 (3/ 4P)	1	17	18
Gesamt-Hirn			+		+		Speichel-drüse		+	+++	+	++	
Amygdala					+		Brust-drüse			+	+		
Nukleus caudatus							Niere	+++	++++	++++	++++	++	+++
Cerebellum							Leber		+++	+++	+++	++	+
cerebraler Cortex							Dünn-darm		++	++	++	++	
Frontal Lobus			+				Milz		+	+	++	+	
Hippocampus			+				Thymus		+	+	++	+	
Medulla oblongata			+				periphere Leukozyten				+	+	
Occipital-Lobus		+	++	+	+		Lymphknoten		+	+	++	+	
Putamen		+	++	+	+		Knochenmark		+	+	++	+	
Substantia nigra		+	++	+			Blinddarm			+	+	+	
Temporal-Lobus		+	++	+			Lunge	+	+++	+++	+++	++	++
Thalamus		+	+++	+	+		Luftröhre		+	+	+	+	
Subthalam. Nukleus		+	++	+	+		Plazenta		++	+	+	+	
Rückenmark		+	++	+	+		fötales Hirn			+			
Herz		++	++++	++	++++		fötales Herz			+	+	+	
Aorta		+	+				fötale Niere		+	+++	+	+	
Skelettmuskel		+	+++	+	++		fötale Leber		++	++	+	+	
Colon		+	+++	++	++		fötale Milz		+	+	+		
Blase		+	+	+			fötales Thymus		+	+	+		
Gebärmutter			+				fötale Lunge		+	+	+	+	
Prostata		+	++	++	+		Hefe Gesamt RNA 100 ng						
Magen		+	+++	++	++		Hefe tRNA 100 ng						
Hoden	++++	+	+	++	+		E. coli rRNA 100 ng		++				
Eierstock		+	+	+	+		E. coli DNA 100 ng		++				
Bauchspeicheldrüse		+	+	+	+		Poly r(A) 100 ng						
Hypophyse		+	++	++	+		humane cot1 DNA 100 ng				++		+
Nebenniere		++	+++	+	++		humane DNA 100 ng	+	+	++	+++	+	++
Schilddrüse		+	++	+	+		humane DNA 500 ng	+++	++++	++++	++++	++	+++

### 4.2.3. *In-situ*-Hybridisierung

Die Ergebnisse von Abschnitt 4.2.2 geben nur Aufschluß über die Transkription von HLA-gekoppelten OR Genen in ganzen Geweben. Da Organe jedoch aus einer Vielzahl von verschiedenen Zelltypen bestehen, war es jetzt interessant zu erfahren, ob die Gene nur in eindeutig spezifizierbaren Zelltypen exprimiert werden. Deshalb wurde die Methode der *In-situ*-Hybridisierung an Gewebeschnitten gewählt (s. Abschnitt 3.19). Für die Analyse wurden zwei Gene ausgewählt, die stark positive Signale auf dem mRNA-Blot gezeigt hatten (s. Abschnitt 4.2.2): *hs6M1-16 (-12)* und *hs6M1-10*. Mittels PCR wurden 500 bis 700 bp große Sequenzfragmente von der jeweiligen kodierenden Region amplifiziert (Primer und Bedingungen: *hs6M1-16* (s. Tab. 2 und Tab. 13): FAT11A/ FATB (55°C); *hs6M1-10* (s. Tab. 2): OL-AE 800f / OL-AE 1525r (52°C). Die PCR Produkte wurden in die EcoRV Schnittstelle von pZErO<sup>TM</sup>-2 (Invitrogen) hineinkloniert (s. Abschnitt 3.16 und 3.17). Nachdem die Richtung des klonierten Fragmentes durch Sequenzieren ermittelt worden war, konnte der Promotor identifiziert werden, der den anti-Sinnstrang erzeugen würde. Mit einem Restriktionsenzym, das 5' des OR Gens schnitt, wurde das Plasmid linearisiert. Mit Hilfe der entsprechenden RNA-Polymerase und <sup>35</sup>S-UTP wurde der radioaktive anti-Sinn-Strang, d.h. die Ribosonde<sup>2</sup>, erzeugt (s. Abschnitt 3.19.1). Die *hs6M1-16* Ribosonde wurde mit Hodengewebe hybridisiert und die *hs6M1-10* Ribosonde mit Nieren-, Hoden- und Lebergewebe (s. Abschnitt 3.19.3). Bei der Niere handelte es sich um Normalgewebe von sieben verschiedenen Individuen, bei der Leber um ein Adenocarcinom und bei dem Hoden um Normalgewebe sowie um einen Keimzelltumor. Die Expositionszeiten betragen 3-40 Tage.

Auch nach langen Expositionszeiten (40 Tage) zeigte keine der beiden Ribosonden positive Anfärbungen in den drei verschiedenen Geweben, was darauf hindeutet, daß die OR Gene nur sehr schwach exprimiert sind.

## 4.3. Genomische Organisation von ausgewählten HLA-gekoppelten OR Genen

### 4.3.1. Suche von HLA-gekoppelten OR Transkripten in verschiedenen cDNA Bibliotheken mit Hilfe der radioaktiven Hybridisierung

Um das 5' und 3' Ende von HLA-gekoppelten OR Genen zu finden, wurde eine Hoden cDNA Bibliothek (Normalgewebe), die aus 28000 Klonen bestand, und eine Lungen cDNA Bibliothek (Plattenepithelcarcinom), die aus 55000 Klonen bestand, durchsucht. Die Bibliotheken bestanden aus cDNA Klonen, die auf eine Filtermembran aufgebracht worden

---

<sup>2</sup> Die Transkription der *hs6M1-16 (-12)* Ribosonde sowie die anschließende Hybridisierung wurden von Kollegen im Labor von Dr.Hermann Herbst im Universitätskrankenhaus Eppendorf in Hamburg durchgeführt.

waren (Ressourcen-Zentrum, Berlin). Von ausgewählten HLA-gekoppelten OR Genen mit ORF (*hs6M1-1, -3, -4P, -6, -10* und *-16*) wurden per PCR Sonden generiert, die 500-700 bp der 5' kodierenden Region beinhalteten (s. Tab. 2). Die PCR Produkte wurden auf ein PA-Gel aufgetragen und anschließend aus diesem isoliert und gereinigt (s. Abschnitt 3.7.1 und 3.8.1). Jede Sonde wurde getrennt mit  $\alpha$ [<sup>32</sup>P]-dCTP radioaktiv markiert (s. Abschnitt 3.13.1.1), um den korrekten Einbau bei jeder einzelnen Sonde kontrollieren zu können. Bei der anschließenden Hybridisierung (s. Abschnitt 3.13.1.2) wurden alle Sonden in einem gemeinsamen Gemisch auf die verschiedenen Blots gegeben. Nach einer Expositionszeit von zwei Tagen wurde ein einziger positiver Klon in der Hoden cDNA Bibliothek gefunden; die Lungen cDNA Bibliothek zeigte kein positives Signal.

Der positive Klon aus der Hoden Bibliothek wurde kultiviert. Nach der Isolierung des Plasmids wurde die Größe der cDNA bestimmt, indem die HindIII/EcoRI-verdaute DNA über ein Agarose Gel aufgetrennt wurde. Die cDNA war ~3,35 kb groß. Da 3,35 kb nicht auf einmal sequenziert werden konnten, wurde "Primer-Walking" durchgeführt: Die cDNA wurde zunächst mit den im Vektor enthaltenen Sequenzier-Primern T7 und M13for von beiden Seiten ansequenziert. Anhand der erhaltenen Sequenzdaten konnten neue Primer für die jeweiligen Enden festgelegt werden. Diese besaßen neben der cDNA-spezifischen Sequenz einen Überhang, der komplementär zu entsprechend vorhandenen Sequenzier-Primern war (s. Tab. 10 und Tab. 11). Nach erfolgter PCR konnte somit, ohne zu klonieren, sequenziert werden. Mit Erhalt der neuen Sequenzdaten konnten die nächsten Primer hergestellt werden. Nachdem die komplette Sequenz ermittelt worden war, folgte eine BLAST-Analyse. Die cDNA entsprach dem *hs6M1-32* Gen, das im Neben-Cluster distal von RFP liegt. Die cDNA ist kolinear zur genomischen Sequenz, d.h. es sind keine Introns herausgespleißt worden. Sie beginnt ~2250 bp stromaufwärts vom Start-Codon und endet ~140 bp stromabwärts vom Stop-Codon.

Um noch mehr Transkripte von HLA-gekoppelten OR Genen im Hoden und anderen Geweben zu finden, sollte nun eine Adapter-ligierte cDNA Bibliothek verwendet werden, bei der mit Hilfe von RACE (Rapid Amplification of cDNA Ends) die Transkripte isoliert werden sollten (s. Abschnitt 4.3.2).

### 4.3.2. Suche von HLA-gekoppelten OR Transkripten in verschiedenen cDNA Bibliotheken mit Hilfe von RACE (Rapid Amplification of cDNA Ends)

Für die RACE Methode wurde das Marathon-Ready™ cDNA Kit (Clontech) verwendet, das aus einer Adapter-ligierten cDNA Bibliothek bestand.

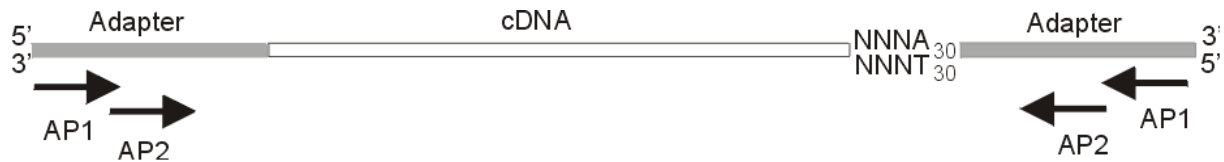


Abb. 13: Schematische Darstellung der Marathon-Ready™ cDNA Bibliothek (Clontech). Die gekaufte cDNA enthielt bereits die ligierten Adapter Sequenzen.

Im Rahmen dieser Arbeit wurden drei verschiedene Marathon-Bibliotheken verwendet: Hoden, Niere und Lunge. Analysiert wurden die 12 HLA-gekoppelten OR Gene *hs6M1-1*, *-3*, *-6*, *-10*, *-12*, *-15*, *-16*, *-17*, *-18*, *-20*, *-21* und *-27* sowie die 4 Pseudogene *hs6M1-4P*, *-13P*, *-14P* und *-19P*, von denen vermutet wurde, daß sie ebenfalls transkribiert werden könnten. Weiterhin wurden die drei OR Gene auf den Chromosomen 7, 17 und 19 (*hs7M1-2*, *hs17M1-20* und *hs19M1-4*) näher analysiert.

Für die RACE-PCR wurden genspezifische Primer (GSP) festgelegt (s. Abschnitt 2.7.2), die als Gegenprimer zu den jeweiligen Adaptorprimern dienten. In Analogie zu den mitgelieferten Adapterprimern (AP1 und AP2) wurde ein äußerer (GSP1) und ein innerer (GSP2) Primer (s. Abb. 14 und Abb. 18) entworfen. Mit den äußeren Oligonukleotiden wurde zuerst eine "Pool"-PCR durchgeführt, die alle vorhandenen äußeren genspezifischen Primer (GSP1) sowie den AP1 Primer beinhaltete (s. Tab. 3 u. Tab. 5). Die Schmelztemperatur der verschiedenen GSP1 durfte daher nicht zu sehr variieren. Mit Aliquots der "Pool-PCR" wurden unter Verwendung der inneren Oligonukleotide (AP2 und GSP2) genspezifische PCRs durchgeführt (s. Tab. 4 u. Tab. 6). Durch die zwei verschachtelten PCRs wurde die Sensitivität bei guter Spezifität der Reaktion erhöht. Eine weitere Verbesserung der Spezifität wurde durch eine Eigenschaft der cDNA Bibliothek erreicht: Das 3' Ende beider Stränge war verkürzt, d.h. die zum Adapterprimer (AP1) komplementäre Sequenz war nicht vorhanden. Deshalb wurde die AP1 Bindungsstelle erst generiert, nachdem der genspezifische Primer verlängert wurde. So sollte eine unspezifische Anlagerung des AP1 verhindert werden.

Um eine Kontamination der Bibliotheken sowie der PCR-Ansätze zu verhindern, wurden die Pipettierschritte in einem Raum durchgeführt, der weit entfernt von dem Labor lag, in dem normalerweise mit PCR-Produkten gearbeitet wurde. Die benötigten Primer-Lösungen, PCR-Puffer, Nukleotide und Taq wurden nur in diesem Raum verwendet. Bei der "inneren"-PCR wurde die Zugabe des "Pool-PCR"-Produktes – als letzten Schritt – wiederum in einem anderen Raum ausgeführt.



#### 4.3.2.1. 5' RACE: Primerdesign, PCR und nachfolgende Analyse

Die Primer für die 5' RACE wurden zunächst in die erste extrazelluläre Domäne der jeweiligen OR's lokalisiert. Nachdem aber bekannt wurde, daß zwei mutierte 7-TM-Chemokin-Rezeptoren auch funktionsfähig waren, wenn ihnen die ersten beiden Domänen fehlten (Ling et al., 1999), wurden alle späteren Oligonukleotide in die dritte TM Domäne gelegt. Dadurch sollte ein evt. mehr stromabwärts gelegener Transkriptions-Startpunkt für die ORs berücksichtigt werden. Die Primer für die Gene *hs6M1-1*, *-3*, *-4P*, *-6*, *-10*, *-12*, *-15*, *-16*, *-17*, *-18*, *-20*, *-21*, *hs7M1-2*, *hs17M1-20* und *hs19M1-4* lagen innerhalb der ersten 15-120 bp stromabwärts vom Start-Codon, während die Primer für die Gene *-13P*, *-14P*, *-19P* und *-27* ca. 270 bis 300 bp stromabwärts vom Start-Codon lagen (s. Abb. 14, Tab. 3 und Tab. 4).

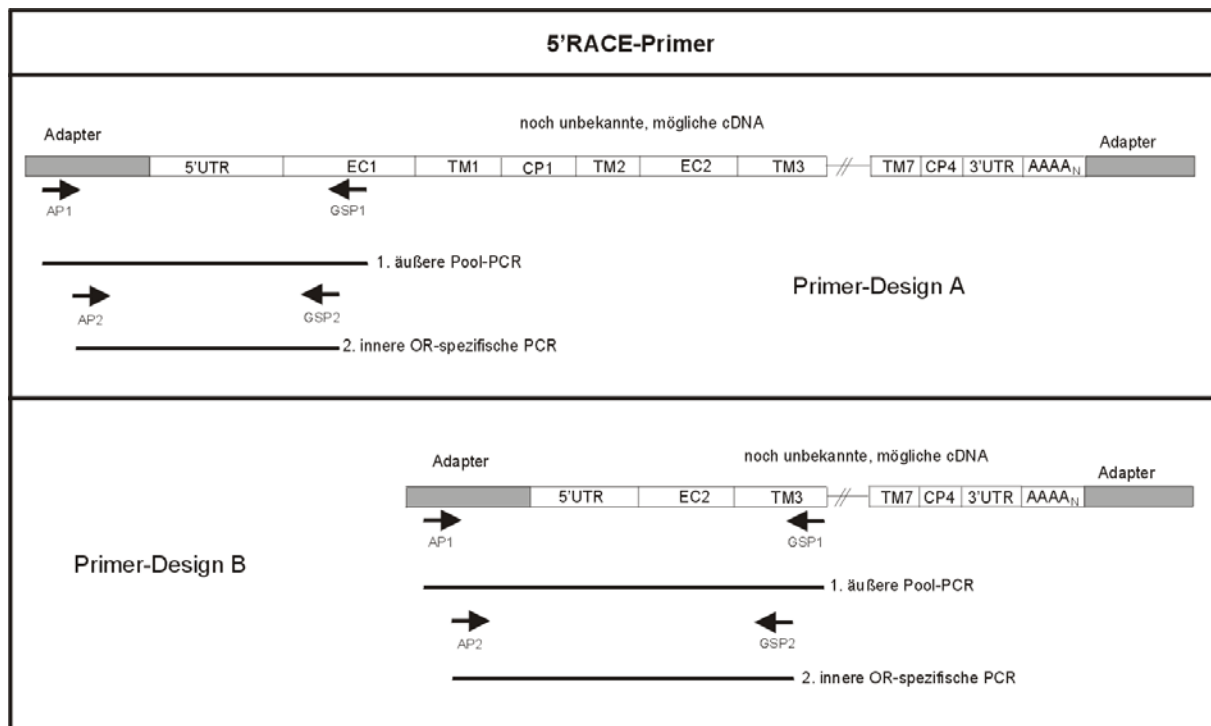


Abb. 14: Schematische Darstellung des Primer-Designs für die 5' RACE (s. Tab. 3 und Tab. 4). Primer Design A: Die zuerst generierten genspezifischen Primer (GSP) wurden direkt stromabwärts vom Start-Codon lokalisiert. Primer Design B: Hier wurden die genspezifischen Primer in die dritte TM Domäne gelegt. In der äußeren "Pool-PCR" wurden alle GSP1s in einem PCR-Ansatz vereinigt. In der darauffolgenden 2. inneren PCR wurde jeweils nur ein GSP2-Primer verwendet.

Nach erfolgter "Pool"- und innerer PCR wurden die Amplifikationsprodukte auf ein PA-Gel aufgetragen und photographisch erfaßt. Die PCR-Produkte, die sichtbare Banden ergaben, wurden in TOPO-TA kloniert, um sie später zu sequenzieren. Wenn mehrere Banden zu sehen waren, wurden diese z.T. aus dem Gel isoliert und gereinigt, um jede Bande separat zu klonieren. Die anderen PCR-Produkte wurden direkt kloniert (s. Abschnitt 3.17.2 und 3.14.2.2). Nach den ersten Sequenz-Analysen stellte sich allerdings heraus, daß viele der

rekombinierten Klone keine OR cDNA enthielten. Um die Sequenzierung dieser uninteressanten Fragmente zu vermeiden, wurden die Klone mit zwei verschiedenen Methoden selektioniert: Erstens wurden die Gele mit den PCR-Produkten auf eine Membran transferiert, um sie anschließend mit entsprechenden OR-spezifischen Sonden zu hybridisieren. Dadurch sollte sichergestellt werden, daß die Amplifikationsprodukte tatsächlich OR cDNA Transkripte enthielten. PCR-Produkte, die bei dieser Methode ein negatives Ergebnis zeigten, wurden verworfen. Die zugehörigen Ergebnisse werden im Abschnitt 4.3.2.2 erläutert. Zweitens wurde nach der Klonierung eine "selektive" PCR mit den rekombinierten Klonen durchgeführt. Bei dieser Amplifikation wurden die inneren Primer verwendet, die zuvor für die genspezifische PCR benutzt worden waren. Nur die Klone, die bei dieser PCR ein positives Produkt ergaben, wurden für weitere Sequenzanalysen eingesetzt.

#### 4.3.2.2. Analyse der RACE-Produkte

Wie in Abschnitt 4.3.2.1 bereits erwähnt, sollten die RACE-Produkte vor der Klonierung zunächst auf ihre OR Spezifität hin getestet werden. Der linke Teil von Abb. 15 zeigt Ethidiumbromid gefärbte Gele mit den aufgetrennten PCR Produkten der 5' RACE. Rechts sind Autoradiographien von Southern-Hybridisierungen mit OR-spezifischen Sonden dargestellt. Die Sonden wurden per PCR hergestellt und umfassten die ersten 100-300 bp der kodierenden Region sowie 100-200 bp der 5' UTR. Für diese PCR wurden die jeweiligen 5' UTR Primer von der Polymorphismusanalyse (s. Tab. 2) und als Gegenprimer die genspezifischen Primer von der 5' RACE verwendet (s. Tab. 3). Es wurden nur OR Gen Sonden des Chromosom 6 Clusters verwendet, welche jeweils einzeln mit  $^{32}\text{P}$  markiert worden waren, dann aber als Mischung verwendet wurden.

Die Intensität der Banden auf den PCR-Gelen (links) korreliert nicht immer mit der Bandengröße auf den Blots (rechts). Besonders deutlich ist dies bei der Hoden-Bibliothek zu erkennen. So zeigt z.B. *hs6M1-15* ein relativ starkes Signal bei ca. 220 bp in der Autoradiographie, während auf dem PCR-Gel nur eine schwache Bande bei ca. 700 bp zu sehen ist. *hs6M1-27* zeigt zwar eine Übereinstimmung bei den drei eng zusammenliegenden Banden (bei ca. 400-500 bp), aber auf dem Blot ist noch eine stark positive Bande bei ca. 50 bp zu sehen, die auf dem Gel nicht erscheint. Nur bei *hs6M1-17* und *-20* scheinen die Banden relativ gut übereinzustimmen. Die Diskrepanz zwischen diesen beiden Darstellungen (PCR-Gel und Autoradiographie) macht deutlich, wie wichtig es ist, die RACE-Produkte zunächst auf ihre OR-Spezifität hin zu testen. Bei der kurz exponierten Hoden-Autoradiographie (s. Abb. 15) sind nicht bei allen HLA-gekoppelten OR Genen Signale zu sehen. Betrachtet man aber eine Autoradiographie, die 24 h exponiert wurde (Abb. 16), kann man erkennen, daß von fast allen getesteten Genen cDNA-Transkripte im Hoden vorkommen. Wie oben bereits erwähnt, wurden keine Sonden für *hs7M1-2*, *hs17M1-20* und

*hs19M1-4* eingesetzt. *hs6M1-4P* ist in vielen Haplotypen ein Pseudogen.

Im Gegensatz zum Hoden werden in Niere und Lunge nicht alle getesteten Gene exprimiert. In der Niere wurden z.B. nur bei *hs6M1-27*, *-1* und *hs7M1-2* eindeutige PCR-Produkte erhalten. Die Gene *hs6M1-27* und *-1* zeigen auch in der Autoradiographie Banden entsprechender Größe: *hs6M1-27* bei ~250 bp und *hs6M1-1* bei ~200 bp. Die 200 bp Bande von *hs6M1-1* erzeugt ein besonders kräftiges Signal in der Autoradiographie. Zwei schwächere Banden erscheinen bei ca. 500 bp, die auf dem Gel nicht zu erkennen sind. In der Autoradiographie sind außerdem noch positive Banden für die folgenden Gene zu erkennen: *hs6M1-12*, *-16* und *-6*. Die Banden von *hs6M1-16* und *-6* sind allerdings sehr schwach. Von allen anderen Genen scheinen in der Nieren-Bibliothek keine cDNA-Transkripte vorhanden zu sein. Bei einer Expositionszeit von 4 h (nicht dargestellt) zeigt sich nur bei *hs6M1-1* ein stark positives Signal, die Gene *hs6M1-27* und *-12* weisen nur schwache Banden auf und von *hs6M1-6* und *-16* sind keine positiven Signale zu sehen.

Im Lungengewebe scheint eine ähnliche Situation vorzuliegen. Eindeutige PCR-Produkte findet man für die folgenden Gene: *hs6M1-1*; *-10*, *-12*, *-15*, *-16*, *-17* und *-27*, sowie für *hs19M1-4*. In der Autoradiographie zeigen aber nur *hs6M1-1*, *-6*, *-10* und *-12* ein positives Signal, wobei die Banden die gleiche Größe wie die auf dem PCR-Gel haben. Die schwach positive Bande bei dem *hs19M1-4* Gen scheint eine Kreuzreaktion mit einer der verwendeten Sonden darzustellen. Nach einer Expositionszeit von 24 h findet man noch schwach positive Banden bei *hs6M1-27* und *-18* (nicht dargestellt).

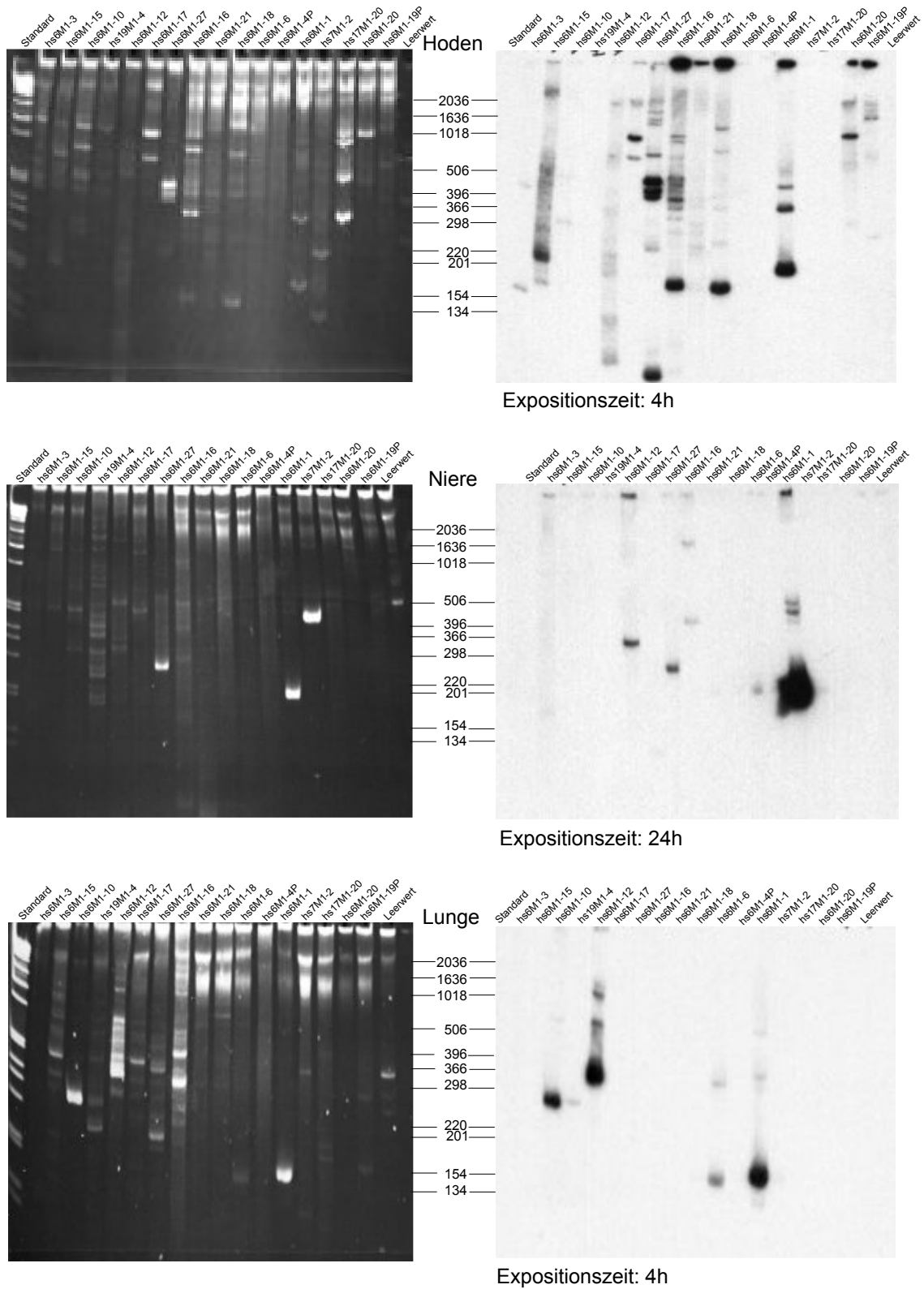


Abb. 15: Links: 5' RACE PCR-Produkte aus der Hoden-, Nieren- und Lungen-cDNA-Bibliothek. Aufgetrennt über ein PA Gel. Rechts: Die entsprechenden Autoradiographien nach Hybridisierung mit Sonden von HLA-gekoppelten OR Genen (*hs7M1-2*, *hs17M1-20* und *hs19M1-4* Sonden wurden nicht verwendet).

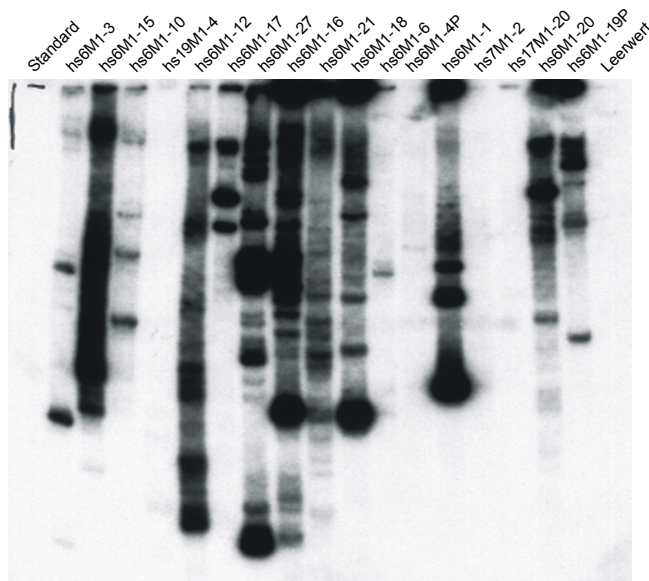


Abb. 16: Autoradiographie eines Blots mit 5' RACE Produkten aus der Hoden-cDNA-Bibliothek. Hybridisierung mit Sonden von HLA-gekoppelten OR Genen. Bis auf die Expositionszeit (hier 24 h) ist diese Autoradiographie identisch mit der entsprechenden Autoradiographie in Abb. 15 (4 h Expositionszeit).

Bei der Autoradiographie des Hoden-Blots sind bis auf *hs6M1-6* immer mehrere Banden für ein bestimmtes Gen zu sehen. Im Gegensatz dazu findet man für einige Gene in Niere und Lunge lediglich eine Transkriptgröße: *hs6M1-27* und *-12* in der Niere und *hs6M1-10* in der Lunge. Einige dieser einzeln vorkommenden Transkripte entsprechen vermutlich gleichgroßen Banden im Hoden. So ist z.B. eine ~250 bp Bande von *hs6M1-27* in der Niere auch im Hoden zu sehen, obwohl sie dort relativ schwach ausgeprägt ist. Dagegen sind anderen Banden von *hs6M1-27* sehr prominent im Hoden. Es scheint, daß im Hoden viele unterschiedlich gespleißte Transkripte eines Gens vorkommen, während in Niere und Lunge - zumindest von einem Teil der OR Gene - nur eine Spleißvariante existiert.

#### 4.3.2.3. Genomische Organisation der 5' UTR HLA-gekoppelter OR Gene

Wie oben bereits erwähnt, wurden die RACE-PCR Produkte kloniert und sequenziert. Mit Hilfe des BLAST-Programms (Altschul et al., 1990) konnten die cDNA-Sequenzdaten mit bereits veröffentlichten genomischen Sequenzen der entsprechenden PAC's und BAC's, die diese Gene beinhalten, verglichen werden. Dadurch war es möglich, die Größe und Lage der Exons und Introns zu ermitteln. Nur Klone, die die entsprechenden GT/AG Spleißstellen (die seltene Spleißstelle AC/AT wurde nicht gefunden) in der intronischen Sequenz aufwiesen, wurden in den Ergebnissen dieser Arbeit verwertet. Alle anderen Klone, ohne entsprechende Spleißstellen, wurden als Artefakt angesehen.

Mit der Lungen- und Nieren Bibliothek wurde nur die 5' RACE durchgeführt. In der Lunge sind für die folgenden Gene cDNA-Transkripte gefunden worden: *hs6M1-1*, *-10* und *-12*. Gemäß Abb. 15 sind dies die Gene, die am stärksten in der Lunge exprimiert werden. *hs6M1-6* Transkripte konnten nicht gefunden werden. Dieses Gen scheint aber auch wesentlich schwächer exprimiert zu werden als die anderen drei Gene. Die gefundenen Transkripte sind kollinear mit der genomischen Sequenz, d.h. es wurden keine Introns

gefunden. Bei *hs6M1-1* ist die 5' UTR ~70 bp lang, bei *hs6M1-10* ~275 bp und bei *hs6M1-12* ~320 bp.

In der Niere wurde bislang nur ein Gen (*hs7M1-2*) gefunden, das am 5' Ende ebenfalls intronlos ist und eine 5' UTR von ~330 bp aufweist. Obwohl die OR Gene *hs6M1-12*, -27 und -1 in Abb. 15 relativ kräftige Banden zeigen, gelang es nicht, Transkripte für diese Gene in der Niere zu finden.

Die meisten OR-Transkripte sind im Hoden vorhanden. Die 5' UTR folgender OR Gene wurden im Hoden nachgewiesen: *hs6M1-3*, -6, -10, -12, -16, -17, -18, -20, -21, -27, *hs17M1-20* und *hs7M1-2*. Die 5' UTR der Gene *hs6M1-3*, -6, -10, -17, -20 und *hs17M1-20* war kolinear zur genomischen DNA. Im Gegensatz dazu enthielt die 5' UTR der Gene *hs6M1-12*, -16, -18, -21, -27 und *hs7M1-2* auf genomischer Ebene zusätzliche Introns. Das 5' Exon des am weitesten zentromer gelegenen HLA-gekoppelten OR Gens *hs6M1-12* (s. Abb. 4) liegt z.B. ~1,5 kb stromaufwärts vom Start-Codon. Vier unabhängige Klone zeigten diese Isoform. Außerdem wurden im Hoden von *hs6M1-12* auch die Transkripte ohne 5' Intron gefunden, die bereits mit Lungen cDNA gefunden worden waren

Die Gene *hs6M1-21*, -27, -18 und -16 liegen in einem Bereich, der ca. 110 kb umfaßt (s. Abb. 17). Von diesen ist *hs6M1-16* das am weitesten zentromer gelegene Gen. Seine transkriptionelle Orientierung ist umgekehrt zu *hs6M1-21*, -27 und -18. Bei *hs6M1-21*, das von dieser Gruppe am weitesten telomer gelegen ist, wurden Exons gefunden, die z.T. über 100 kb stromaufwärts vom Start-Codon liegen. Von diesem Gen wurden drei alternativ gespleißte Transkripte gefunden. Transkript A (n=1) (s. Abb. 17) besitzt nur ein zusätzliches Exon, das ~100 kb stromaufwärts vom Start-Codon liegt; Transkript B (n=3) weist neben diesem 100 kb entfernten Exon 1 ein zweites zusätzliches Exon (Exon 3) auf, das ca. 50 kb stromaufwärts vom Start liegt; und Transkript C (n=1) besitzt, zusätzlich zu Exon 1, das ~98 kb stromaufwärts vom Start gelegene Exon 2.

Die kodierende Region von *hs6M1-27* liegt ~20 kb zentromer von der kodierenden Region des *hs6M1-21* Gens. Von diesem Gen wurden zwei alternativ gespleißte cDNAs gefunden. Transkript D (n=3) weist drei zusätzliche Exons in der 5' UTR auf. Das am weitesten stromaufwärts gelegene Exon (~80 kb vom Start-Codon entfernt) ist identisch mit Exon 1 von *hs6M1-21*. Das nächst folgende Exon liegt ca. 78 kb stromaufwärts und ist identisch mit Exon 2 von *hs6M1-21*. Das 3. Exon des *hs6M1-27*-Transkripts D liegt ~30 kb stromaufwärts vom Start-Codon und ist wiederum identisch mit Exon 3 von *hs6M1-21*. Das zweite Transkript von *hs6M1-27* (Transkript E (n=2)) besitzt nur ein zusätzliches Exon, das ca. 75 kb stromaufwärts vom Start-Codon liegt. Dieses Exon ist mit keinem der bereits beschriebenen Exons identisch.

Auf *hs6M1-27* folgt in zentromerer Richtung *hs6M1-20*. Dieses Gen besitzt eine andere Transkriptionsrichtung; und es konnten keine zusätzlichen Exons nachgewiesen werden

(Transkript F (n=5)). Auf *hs6M1-20* folgt zentromer *hs6M1-19P* in der gleichen Orientierung. Von diesem Gen – einem potentiellen Pseudogen - konnten keine Transkripte in der Hoden-cDNA-Bibliothek gefunden werden.

Ungefähr 9,5 kb zentromer von *hs6M1-19P* liegt *hs6M1-18*. Von diesem Gen, mit derselben Transkriptionsrichtung wie *hs6M1-21* und *-27*, wurden drei alternativ gespleißte Transkripte gefunden (G, H, I). Transkript G (n=1) weist vier zusätzliche Exons auf. Drei davon liegen innerhalb der ersten 4 kb stromaufwärts vom Start-Codon, jeweils ~1kb voneinander getrennt. Das am weitesten 5' gelegene Exon liegt dagegen ca. 30 kb vom Start entfernt. Dieses Exon ist identisch mit Exon 1 von *hs6M1-21* und *-27*. Transkript H von *hs6M1-18* weist nur dieses gerade erwähnte Exon 1 auf. Transkript I besitzt ebenfalls nur ein zusätzliches Exon, das aber nur ~0,6 kb vom Start entfernt liegt.

12 kb zentromer von *hs6M1-18* liegt *hs6M1-17*. Dieses Gen weist wieder dieselbe Transkriptionsrichtung wie *hs6M1-20* und *-19P* auf. Von diesem Gen wurde ein Transkript (J) (n=1) gefunden, das keine zusätzlichen Exons besitzt.

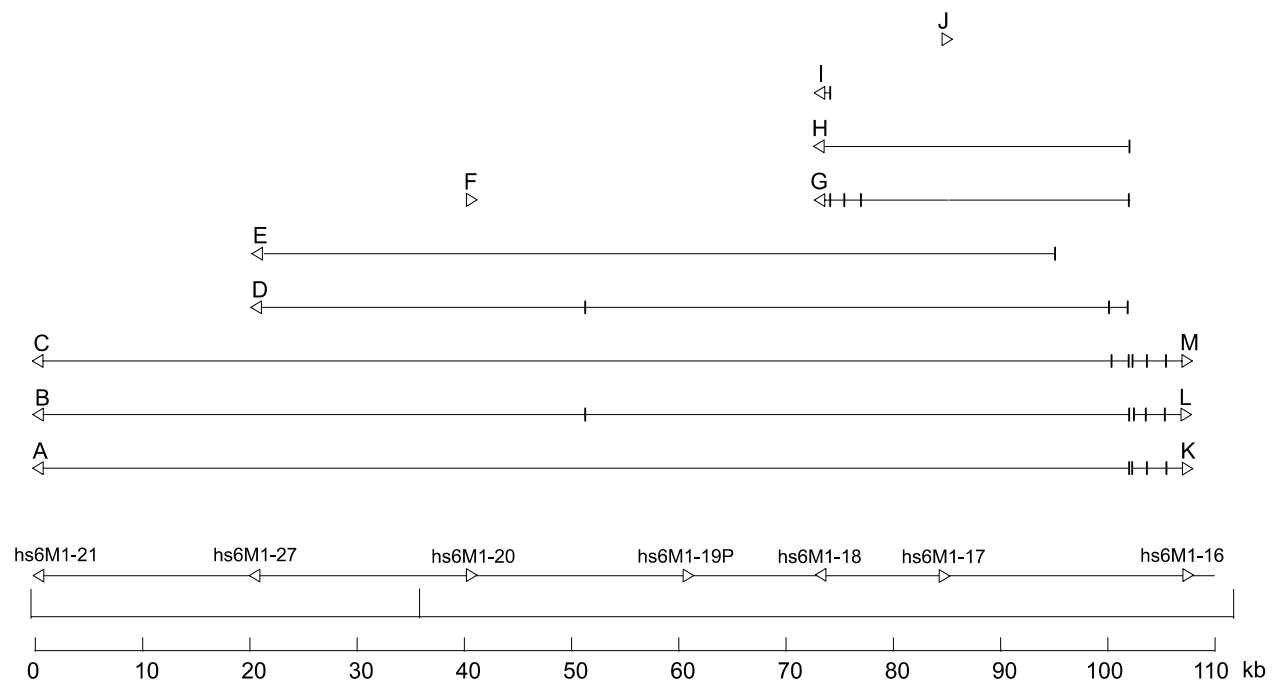


Abb. 17: Schematische Darstellung der 5' UTR von sechs HLA-gekoppelten OR Genen, die im Hoden exprimiert werden (von *hs6M1-19P* wurden keine Transkripte gefunden). Die weißen Pfeile geben die Position der kodierenden Region sowie die Transkriptionsrichtung der OR's an. Intronische Sequenzen sind als horizontale Linien und Exons als vertikale Striche dargestellt.

*hs6M1-16* liegt ~20 kb zentromer von *hs6M1-17*. Von der 5' UTR Region dieses Gens wurden drei alternativ gespleißte Klone gefunden (K, L, M). Das erste Exon der Transkripte K und M ist identisch und liegt ~4,5 kb stromaufwärts vom Start-Codon. Nur 50 bp stromabwärts dieses Exons beginnt das erste Exon von Transkript L. Bei allen drei Transkripten folgt ein Exon, das ca. 3,2 kb stromaufwärts vom Start liegt, wobei das 3' Ende dieses zweiten Exons bei allen drei Transkripten identisch, das 5' Ende bei Transkript K aber 11 bp kürzer ist, woraus folgt, daß der Spleißprozess innerhalb eines Exons stattfand. Dies wird durch das Vorhandensein einer Spleiß-Akzeptor-Stelle unterstützt. Das nächst folgende Exon beginnt ~1kb stromaufwärts vom Start und ist bei allen drei cDNAs identisch.

Erwähnenswert ist, daß der Abstand zwischen dem am weitesten stromaufwärts gelegenen gemeinsamen Exon von *hs6M1-21*, *-27* und *-18* und dem am weitesten stromaufwärts gelegenen Exon von *hs6M1-16* nur 80 bp beträgt. Damit wird die Promotorregion für diese Gene stark eingegrenzt.

#### 4.3.2.4. 3' RACE: Primerdesign, PCR und nachfolgende Analyse

Die 3' RACE wurde nur bei den Genen durchgeführt, die bei der 5' RACE positive Ergebnisse geliefert hatten: *hs6M1-1*, *-3*, *-6*, *-10*, *-12*, *-16*, *-18*, *-20*, *-21*, *-27*, *hs7M1-2* und *hs17M1-20*. Als Ausgangsmaterial wurde nur die Hoden-cDNA-Bibliothek verwendet. Die Primer für die 3' RACE (s. Tab. 5 und Tab. 6) wurden unter Berücksichtigung der bereits vorhandenen cDNA-Daten generiert, d.h. es wurde darauf geachtet, daß sie nicht in einem Intron lokalisiert waren. Um die gesamte kodierende Region bei der PCR zu erfassen, wurden die Primer wenige Basenpaare stromaufwärts vom Start-Codon positioniert (s. Abb. 18).

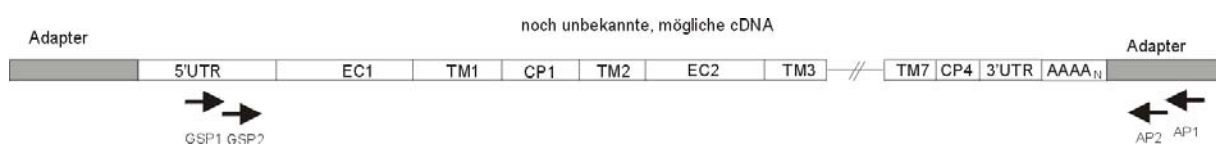


Abb. 18: Position der genspezifischen 3' RACE Primer (s. Tab. 5 und Tab. 6)

Nach erfolgter "Pool"- und innerer PCR wurden die RACE-Produkte, wie bereits in Abschnitt 4.3.2.1 für die 5' RACE beschrieben, auf ihre OR Spezifität hin getestet: Die Proben wurden auf ein Gel aufgetragen, das anschließend auf eine Membran transferiert und danach mit OR spezifischen Sonden hybridisiert wurde. Die genomischen OR-Sonden beinhalteten diesmal die vordere Hälfte der jeweiligen kodierenden Region und wurden mit Hilfe der Primer aus Tab. 2 hergestellt. Zur Markierung wurde in diesem Falle die nichtradioaktive ECL-Methode gewählt, bei der die Sonden mit Meerrettich-Peroxidase markiert wurden (s. Abschnitt 3.13.2). Ergab die ECL ein positives Ergebnis wurden die RACE-Produkte in TOPO-TA kloniert. Die rekombinierten Klone wurden, entsprechend der 5' RACE, einer positiven Selektion



unterzogen: Es wurde eine PCR durchgeführt mit Primern, die zur Herstellung der zu klonierenden 3' Race Produkte gedient hatten. Die bei dieser PCR gefundenen positiven Klone wurden sequenziert und mit Hilfe von BLAST analysiert.

#### 4.3.2.5. Ergebnisse der 3' RACE und Länge der gefundenen Transkripte

Mit den in Abb. 18 dargestellten 3' RACE-Primern wurden nur bei den folgenden Genen Poly-A-Schwänze gefunden: *hs6M1-16*, *-18* und *-21*. Die Analyse der anderen Gene (*hs6M1-1*, *-3*, *-6*, *-10*, *-12*, *-20*, *-27*, *hs7M1-2* und *hs17M1-20*) brachte keine positiven Ergebnisse, wobei allerdings anzumerken ist, daß von diesen Genen wesentlich weniger rekombinierte Klone getestet wurden. Obwohl die Transkripte polyadenyliert waren, enthielt keines der gefundenen Transkripte die komplette kodierende Region. Zunächst wurde an einen Artefakt der Clontech-Bibliothek gedacht. Um diese Möglichkeit auszuschließen, wurde Hoden-cDNA mit dem Superscript Kit (Gibco BRL) selbst hergestellt (s. Abschnitt 3.12). Die dafür benötigte Hoden-mRNA wurde entweder aus Hodengewebe isoliert (s. Abschnitt 3.3), das von einer Prostata-OP stammte, oder gekauft (Clontech). Die cDNA Synthese wurde mit Hilfe geankerter poly-dT-Primer (Oligo-dT-SMART, s. Tab. 13) durchgeführt. Anschließend wurde mit Primern, die spezifisch für vier verschiedene Haushaltsgene waren, die Qualität der cDNA getestet (s. Tab. 12).

Der Einfachheit halber wird nachfolgend die cDNA aus selbst isolierter mRNA als AE-cDNA und die, aus der bei Clontech gekauften mRNA hergestellte, als CT-cDNA bezeichnet. In diesen cDNAs wurde erneut nach OR-Transkripten gesucht. Für die RACE-PCR wurden wiederum die beiden ineinander verschachtelten OR-spezifischen Oligonukleotide verwendet, sowie Primer, die spezifisch für den Oligo-dT-SMART Primer waren (SMART-TAG1 und -TAG2). Die nachfolgenden Schritte von der Klonierung bis zur Sequenzierung wurden wie oben bereits beschrieben durchgeführt.

Bis auf eine Ausnahme (*hs6M1-18*) wurden auch hier wiederum nur kurze Transkripte gefunden. Von dem *hs6M1-18* Gen wurden nur zwei Transkripte identifiziert. Beide stammten aus der CT-Bibliothek. Das eine Transkript war nur 85 bp lang, während das zweite gefundene Transkript ausnahmsweise die gesamte kodierende Region umfaßte. Die 3' UTR dieses Transkriptes beinhaltete ~2000 bp und war nicht gespleißt. Von *hs6M1-16* und *-21* wurden mehrere Transkripte gefunden.

Bei *hs6M1-21* waren alle Transkripte kürzer als die kodierende Region. Elf Transkripte wurden aus der Marathon-Bibliothek und fünf aus der CT-cDNA isoliert. Abb. 19 zeigt ein Alignment dieser Sequenzen. Bei *hs6M1-21* scheint es drei "Hot Spots" bezüglich der Polyadenylierung zu geben. Der erste "Hot Spot" beginnt bei der Basenpaar-Position 87 der kodierenden Region. Vier cDNA-Transkripte aus der Marathon-Bibliothek und drei Transkripte aus der CT-cDNA sind hier polyadenyliert. Die letzten beiden Basen vor dem Poly-A variierten z.T. bei den einzelnen Transkripten. Dies war jedoch nicht weiter

verwunderlich, da die reverse Transkription mit geankerten Oligo-dT-Primern, die am 3' Ende zwei zusätzliche degenerierte Basen besaßen, durchgeführt worden war. Der zweite Adenylierungs "Hot Spot" liegt bei Position 143, wobei die drei entsprechenden cDNAs alle aus der Marathon-Bibliothek stammten. Eine weiteres Transkript ist 3 bp länger als diese drei, was sich jedoch nicht mit den degenerierten Primern erklären läßt. Der dritte Adenylierungs "Hot Spot" liegt bei Position 282. Drei cDNAs aus der Marathon-Bibliothek und ein Transkript aus der CT-cDNA weisen diese Isoform auf.

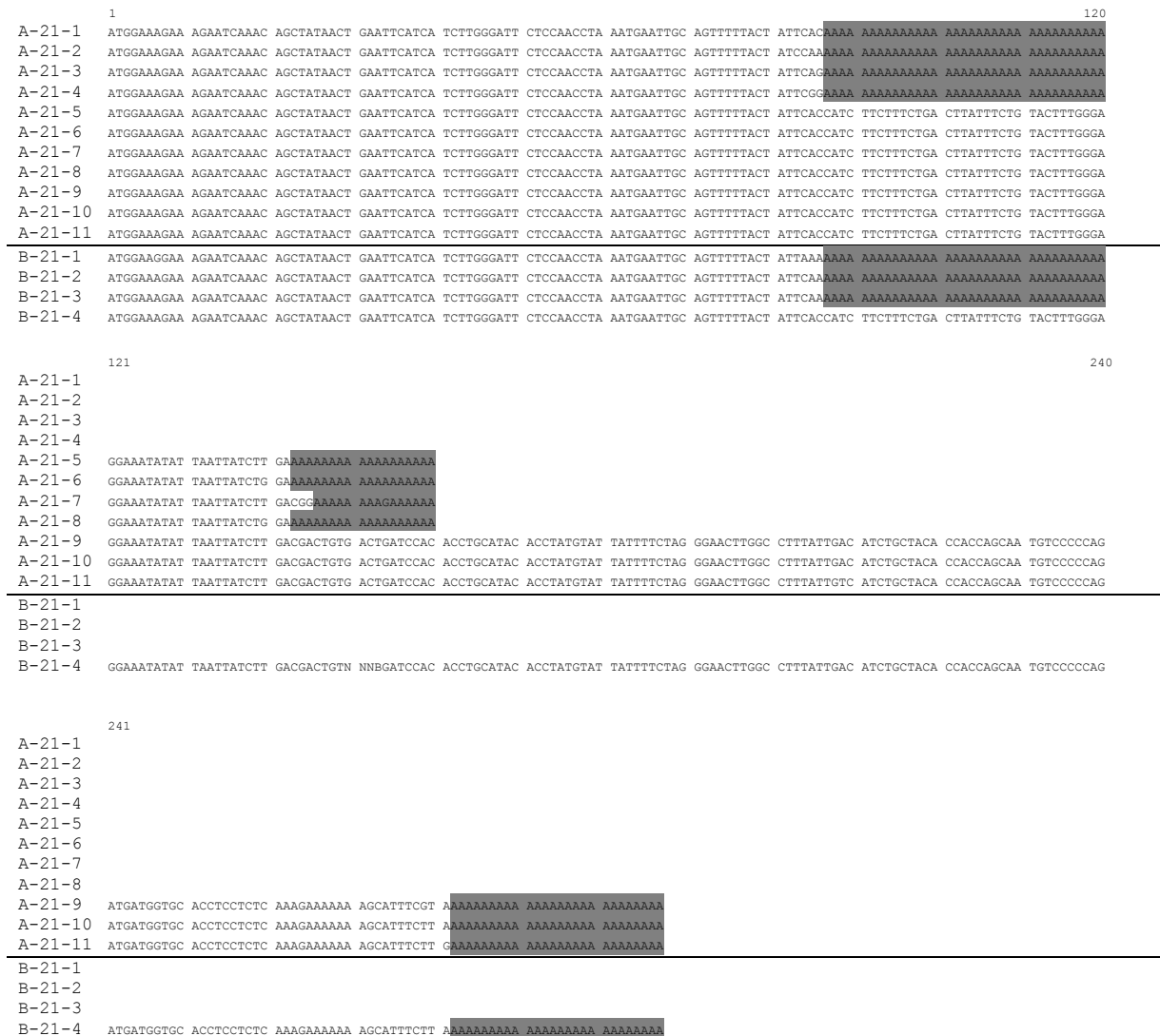


Abb. 19: Alignment der gefundenen *hs6M1-21* Transkripte. Die gefundenen Transkripte beinhalten maximal 281 bp (kodierende Region: 966 bp). Die Polyadenylierung beginnt bei drei verschiedenen "Hot Spots". Oberhalb der horizontalen Linie sind die Transkripte aus der Marathon-Bibliothek aufgeführt; unterhalb der Linie sind Transkripte aus der CT-cDNA (hergestellt aus gekaufter Hoden mRNA, Clontech) gezeigt. Zwei der drei "Hot Spots" sind in beiden Bibliotheken identisch.

Ein ähnliches Phänomen wurde auch bei *hs6M1-16* gefunden (s. Abb. 20), obwohl die Situation hier noch etwas komplexer ist. Insgesamt wurden acht Transkripte gefunden. Sieben davon stammen aus der CT-cDNA und eines aus der Marathon-Bibliothek. Das Transkript aus der Marathon-Bibliothek ist bei Position 189 polyadenyliert und somit identisch mit fünf weiteren Transkripten aus der CT-cDNA. Zwei Transkripte aus der CT-cDNA weisen allerdings an Position 190 ein Guanin statt des Adenins auf. Da dieses Nukleotid nicht mit der PAC Sequenz übereinstimmt, und es sich somit nicht um ein verlängertes Transkript handeln kann, ist anzunehmen, daß es sich hierbei um einen Sequenzierartefakt handelt. Eine andere Isoform endet bei Position 71.

Schließlich wurde noch ein Transkript von *hs6M1-16* identifiziert, das 80 Basenpaare in der 5' UTR - im Anschluß an den 5' Primer - aufwies, dem jedoch die ersten 234 bp der kodierenden Region fehlten. Das vermeintliche Start-Codon dieser Isoform liegt deshalb in der zweiten extrazellulären Domäne, was dazu führt, daß ein daraus abgeleitetes Protein erst an einer Position beginnen würde, bei der Proteine der anderen erwähnten Transkripte bereits zu Ende wären. Selbst diese Isoform endet bei Position 596 und enthält deshalb ebenfalls kein Stop-Codon.

Bemerkenswert ist noch, daß in den kodierenden Regionen der OR Gene keine Polyadenylierungssignale gefunden wurden, die Anlaß für die kurzen Transkripte geben könnten.

Da in zwei voneinander unabhängig hergestellten Bibliotheken Transkripte gefunden wurden, die an unerwarteten, aber identischen Positionen polyadenyliert sind, ist es sehr unwahrscheinlich, daß diese Resultate von Artefakten bei der cDNA Synthese herrühren.

	1	120
hs6M1-16	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
8S071908	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
7S122107	-----	
7S092814	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
7S122101	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
7S122104	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
7S122109	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
7S122108	ATGGTTAACG AAAGCTCCCG -ATGGGCTTC CTCCTTCTGG GCTTCTCTGA ACACCCAGCA CTGAAAAGGA CTCTCTTTGT GGTGTCTTC ACTTCCTACC TCTTGACCCT GGTGGGCAAC	
7S122113	ATGGTTAACG AAAGCTCCCG CATGGGCTTC CTCCTTCTGG GCTTCTCTGA GCACCCAGCA CTGAAAAGGA CAAAAAAAAA AAAAAAAAAA	
	121	240
hs6M1-16	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGACC TCTCTTCTT GGACCTCTGC TTTACCACAA GTTGTGTCCT CCAGATGCTG	
8S071908	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
7S122107	-----	
7S092814	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
7S122101	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
7S122104	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
7S122109	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
7S122108	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
7S122113	ACACTCATCA TCCTGCTGTC TGTACTGTAC CCCAGGCTCC ACTCTCCAAT GTACTTTTTC CTCTCTGAAA AAAAAAAAAA AAAAAAAAAA	
	241	360
hs6M1-16	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
8S071908	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S122107	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S092814	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S122101	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S122104	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S122109	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S122108	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
7S122113	GTCAACCTCT GGGGCCAAA GAAGACCATC AGCTTCTGG GATGCTCTGT CCAGCTCTTC ATCTTCTGT CCTGGGGAC CACTGAGTGC ATCTCTCTGA CAGTATGTC CTTTGACCCT	
	361	480
hs6M1-16	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
8S071908	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S122107	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S092814	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S122101	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S122104	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S122109	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S122108	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
7S122113	TACGTGGCTG TCTGCCAGCC CCTCCACTAT GCCACCATCA TCCACCCCG CCTGTGCTGG CAGCTGGCAT CTGTGGCCTG GGTATGAGT CTGGTTCAAT CGATAGTCCA GACACCATCC	
	481	600
hs6M1-16	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
8S071908	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S122107	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S092814	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S122101	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S122104	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S122109	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S122108	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
7S122113	ACCCTCCACT TGCCCTTCTG TCCCCACCAG CAGATAGATG ACTTTTTATG TGAGGTCCCA TCTCTGATTC GACTCTCCTG TGGAGATACC TCCTACAATG AAATCCAGTT GGCTGTGTCC	
	640	
hs6M1-16	AGTGTCACTC TCGTGGTTGT GCCTCTCAGC CTCATCCTTG	
8S071908	AGTGTCACTC TCGTGGTTGT GCCTCTCAGC CTCATCCTTG	
7S122107	AAAAAAAAAA AAAAAAAAAA	
7S092814	AAAAAAAAAA AAAAAAAAAA	
7S122101	AAAAAAAAAA AAAAAAAAAA	
7S122104	AAAAAAAAAA AAAAAAAAAA	
7S122109	AAAAAAAAAA AAAAAAAAAA	
7S122108	AAAAAAAAAA AAAAAAAAAA	
7S122113	AAAAAAAAAA AAAAAAAAAA	

Abb. 20: Alignment der gefundenen *hs6M1-16* Transkripte, die nicht die gesamte kodierende Region umfassen. Die Polyadenylierung beginnt bei drei verschiedenen Positionen. Oberhalb der horizontalen Linie ist das einzige in der Marathon-Bibliothek gefundene Transkript aufgeführt; unterhalb der Linie sind Transkripte aus der CT-cDNA (hergestellt aus gekaufter Hoden mRNA, Clontech) aufgeführt. Jeweils die erste Zeile zeigt die genomische *hs6M1-16* Sequenz des PAC's. In beiden Bibliotheken wurden Transkripte gefunden, deren Polyadenylierung bei Position 189 beginnt. Einem Transkript aus der CT-Bibliothek fehlen die ersten 234 bp der kodierenden Region, dargestellt durch Minuszeichen (7S122107).

### 4.3.3. Analyse der 3' UTR und der kodierenden Region

Wie in Abschnitt 4.3.2.5 beschrieben, wurden mit den genspezifischen Primern, die direkt stromaufwärts vom Start lokalisiert sind (s. Abb. 18), nur kurze Transkripte gefunden, die nicht die gesamte kodierende Region umfaßten. Um die Chance zu erhöhen, doch noch die 3' UTR einschließlich Poly-A dieser Gene zu finden, wurden zwei weitere ineinander verschachtelte Primer entwickelt, die in den letzten beiden Domänen TM7 und CP4 lokalisiert waren (s. Abb. 21, Tab. 7).

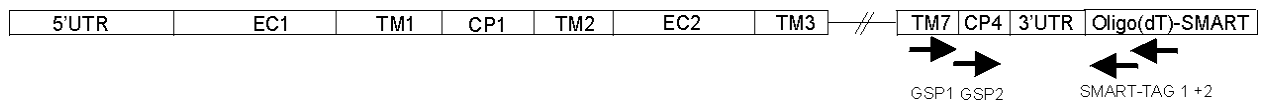


Abb. 21: Schematische Darstellung der Primerposition, um die 3' UTR zu finden.

Diese Primer wurden nur für die Gene *hs6M1-21*, *-27*, *-18* und *-16* hergestellt, weil diese Gene aufgrund ihrer überlappenden Exons besonders interessant erschienen (s. Abschnitt 4.3.2.3). Als cDNA Bibliothek wurde in diesem Falle die CT-cDNA (hergestellt aus gekaufter mRNA, Clontech) verwendet. Daher wurden die SMART-Oligonukleotide als Gegenprimer verwendet.

Von *hs6M1-21* konnten trotz zahlreicher Versuche mit veränderten PCR-Parametern keine Transkripte in der 3' UTR gefunden werden (s. Tab. 7). Von den OR Genen *hs6M1-16*, *-18* und *-27* konnte jedoch die 3' UTR kloniert werden. Von *hs6M1-27* wurden zwei Isoformen gefunden. Beide Transkripte weisen in der 3' UTR keine zusätzlichen Exons auf, sind aber unterschiedlich lang (s. Abb. 23). Die eine cDNA besitzt eine ca. 900 bp lange 3' UTR, während die andere eine 44 bp lange 3' UTR aufweist. Die längere der beiden cDNAs besitzt ein Polyadenylierungssignal 24 bp vor dem Poly-A. Von *hs6M1-18* wurde nur ein Transkript mit 845 bp in der 3' UTR gefunden, das kolinear zur genomischen Sequenz ist. Ein Polyadenylierungssignal wurde 17 bp stromaufwärts von Poly-A gefunden. Im Gegensatz dazu besitzt *hs6M1-16* eine gespleißte 3' UTR, von der drei Isoformen existieren (s. Abb. 24). Transkript D und E besitzen ein identisches Exon, das ca. 1 kb stromabwärts vom Stop-Codon liegt. Zusätzlich zu diesem Exon weist Transkript D noch ein weiteres Exon auf, das ca. 700 bp vom Stop entfernt liegt. Transkript F besitzt nur ein Exon das ca. 1,5 kb stromabwärts vom Stop-Codon liegt. Nur in diesem am weitesten stromabwärts gelegenen Exon wurde ein Polyadenylierungssignal gefunden - 23 bp stromaufwärts von Poly-A.

Da die 3' UTR Sequenzen der drei Gene (*hs6M1-16*, *-18* und *-27*) nun bekannt waren, konnten spezifische Primer in dieser Region generiert werden. Somit war jetzt eine PCR mit ausschließlich spezifischen Primern möglich (s. Tab. 8 und Abb. 22). Diese PCR sollte dazu dienen, die Amplifikation der gesamten bisher noch nicht gefundene kodierenden Region zu erleichtern.

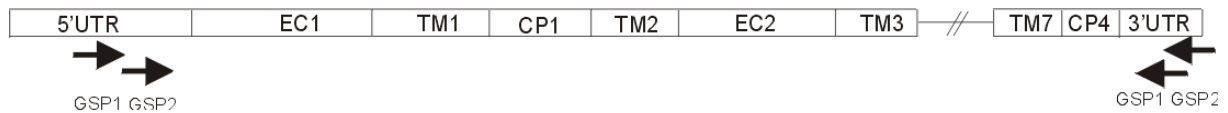


Abb. 22: Schematische Darstellung der spezifischen Primer, die zur Amplifikation der kodierenden Region von HLA-gekoppelten OR Genen verwendet wurden.

Auf diese Weise wurden für *hs6M1-18* mehrere Klone isoliert, die eine komplette kodierende Region umfaßten.

Von *hs6M1-27* wurden zwei verschiedene Isoformen der kodierenden Region gefunden. Die eine Isoform besitzt die vollständige kodierende Sequenz, während die kodierende Region der anderen Form nach 411 bp durch ein Intron unterbrochen wird, das 272 bp lang ist (s. Abb. 23). Die nachfolgende cDNA aus dem kodierenden Bereich ist nicht mehr *in frame*.

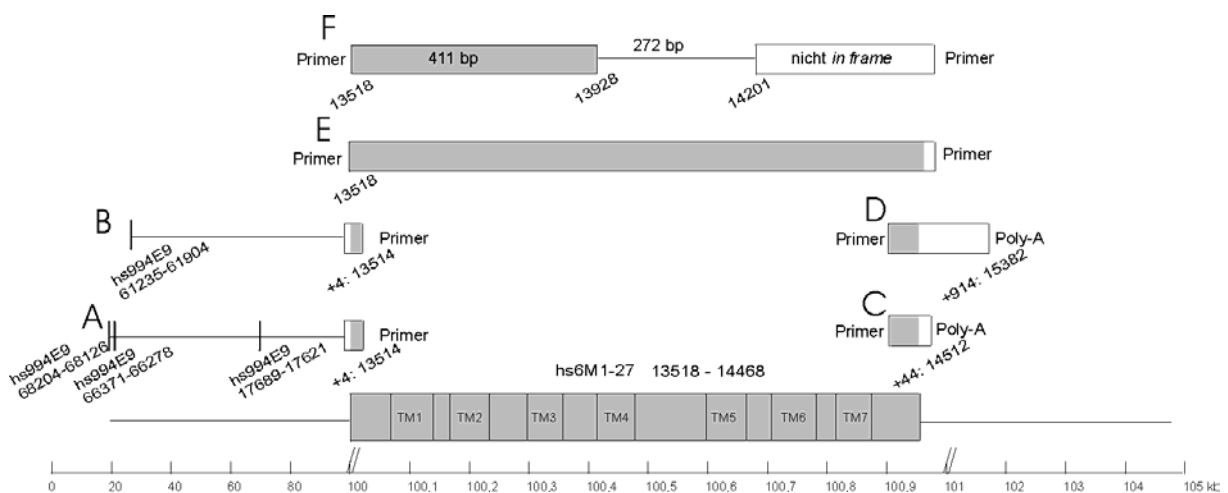


Abb. 23: Gefundene Transkripte von *hs6M1-27*. Die Zahlen geben die Position auf dem entsprechenden PAC oder BAC an. Die kodierende Region von *hs6M1-27* befindet sich auf dem BAC ba150A6, während die ersten 5' Exons auf dem PAC dJ994E9 gefunden wurden.

Bei *hs6M1-16* wurde der 3' UTR Primer in das gemeinsame Exon von Transkript D und E gelegt (s. Abb. 24) – ca. 1 kb stromabwärts vom Stop-Codon. Es wurden vier alternativ gespleißte Transkripte gefunden. Ein weiteres fünftes Transkript wurde mit den Primern gefunden, wie sie in Abb. 18 beschrieben worden sind. Transkript G (n=5) (s. Abb. 24) besitzt die komplette kodierende Region, die sich in der 3' UTR ungespleißt fortsetzt und schließlich durch den Primer (~1 kb stromaufwärts vom Stop-Codon) begrenzt wird. Das 5' Ende dieser cDNA wird durch den anderen Primer beendet, der ca. 100 bp stromaufwärts

vom Start-Codon liegt. Transkript H (n=3) ähnelt Transkript G, ist aber in der 3' UTR gespleißt. Dieses Exon ist identisch mit dem bereits zuvor gefundenen Exon von Transkript D und E. Transkript I (n=3) und J (n=6) sehen gänzlich anders aus. Beide Transkripte starten in der 5' UTR mit dem entsprechenden Primer. Nach 80 bp wurde jedoch ein 253 bp langes Intron herausgespleißt. Dadurch fehlen diesen beiden Transkripten die ersten 234 Basen der kodierenden Region. Erstaunlicherweise beginnt das folgende Exon mit einem Start-Codon, das in der zweiten extrazellulären Domäne liegt. Der restliche Anteil dieser beiden Transkripte entspricht den beiden zuvor beschriebenen cDNAs, d.h. es folgt eine ungespleißte oder eine gespleißte 3' UTR. Transkript K (n=1) ähnelt I und J. Diesem Transkript fehlen allerdings nur 221 bp des vorderen Teils der kodierenden Region. Es ist somit stromaufwärts von dem ATG in der zweiten extrazellulären Domäne 13 bp länger als I und J. Auch hier wurden die korrekten GT/AG Spleißstellen gefunden. In der 3' UTR ist diese cDNA nicht gespleißt. Nach dem Stop-Codon folgen noch 92 weitere Basen bis sich Poly-A anschließt.

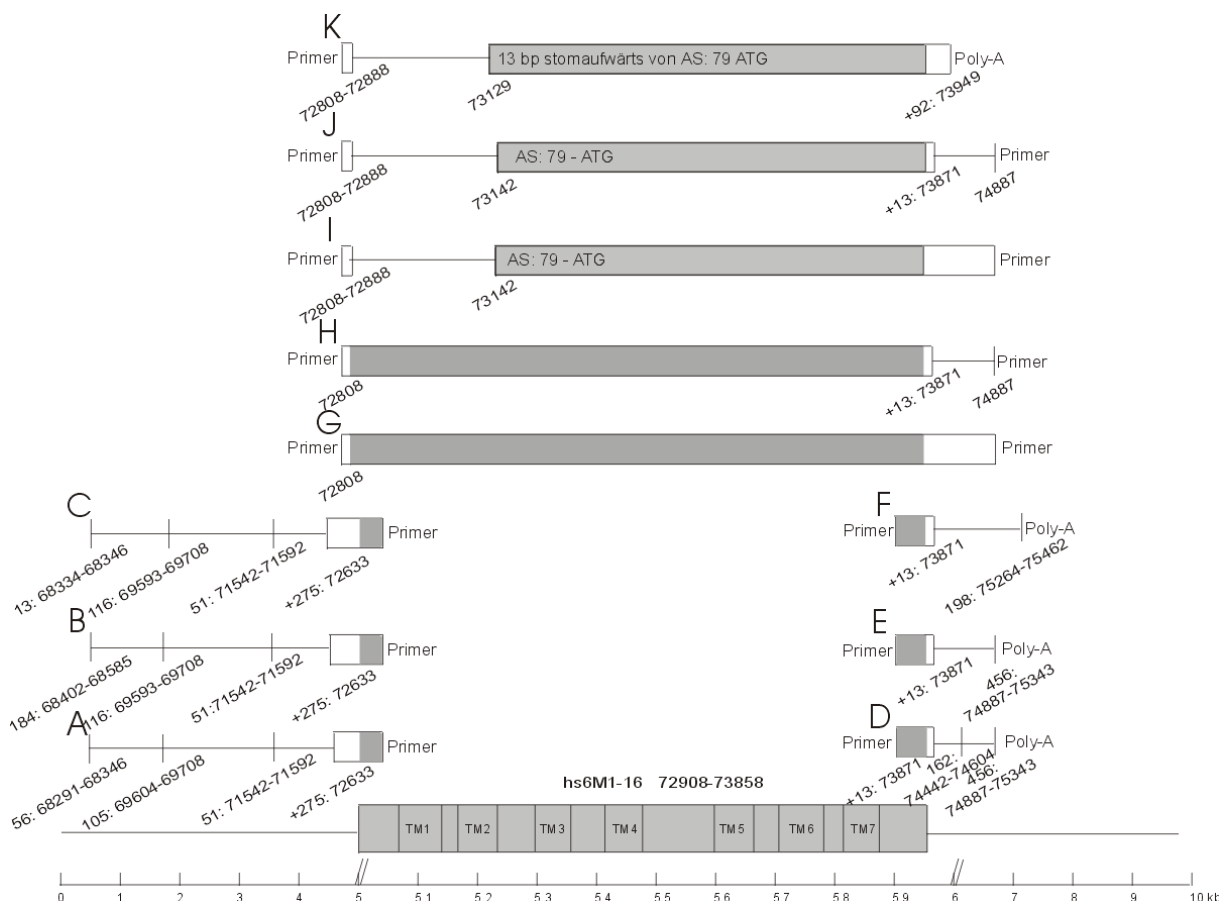


Abb. 24: Schematische Darstellung der verschiedenen Transkripte von *hs6M1-16*. Die Zahlen geben die Anzahl der Basenpaare in einem Exon an sowie die Position auf dem entsprechenden PAC (dJ994E9).