

5. Diskussion

Ziel dieser Arbeit war es, die HLA-gekoppelten OR Gene auf Polymorphismen hin zu untersuchen, um einem bestimmten OR Haplotyp einen bestimmten HLA-Haplotyp zuzuordnen. Die Ergebnisse könnten dazu beitragen, individuelle Geruchspräferenzen für HLA-differente Partner zu erklären oder um Ursachen für ungeklärte wiederholte Aborte zu finden. Des Weiteren sollte die Expression von HLA-gekoppelten OR Genen in anderen Geweben untersucht werden. Die Resultate hier deuten an, daß diese Rezeptoren noch andere Funktionen besitzen als nur die des Riechens. Schließlich sollte noch die Exon/Intron Struktur dieser Gene analysiert werden. Die genomische Organisation einiger OR Gene weist auf potentielle Promotorregionen hin.

5.1. Polymorphismusanalyse

Allelische Variabilität bei humanen M-OR Genen wurde zwar vermutet (Mombaerts, 1999b), konnte aber erst vor kurzem definitiv gezeigt werden (Gilad et al., 2000; Ehlers et al., 2000; Sharon et al., 2000; Ziegler et al., 2000a). Die Variabilität der OR Gene, die in dieser Arbeit demonstriert wird, ist daher eine der ersten systematischen Studien, die belegt, daß Polymorphismen bei OR Genen eher die Regel als die Ausnahme sind. Eine noch größere Allelfrequenz pro Gen findet man bei HLA-Klasse I und -Klasse II Genen. Von den HLA-A und -B Genen z.B. sind jeweils mehr als 100 verschiedene Allele beschrieben worden (Bodmer et al., 1999). Auch für das GABBR1 Gen, das ebenfalls in der 6p21.3-p22.1 Region lokalisiert ist (s. Abb. 4), wurden Polymorphismen gefunden (Peters et al., 1998). Weiterhin wurde die Vermutung, daß es OR Gene gibt, die sowohl funktionell als auch nicht funktionell sein können (Mombaerts, 1999b), in dieser Arbeit bestätigt. Eine unterschiedliche Anzahl funktionsfähiger OR Gene in verschiedenen Individuen haben Trask und Mitarbeiter nachgewiesen (Trask et al., 1998a), die jedoch auf der Anwesenheit zusätzlicher OR Gene bei manchen Individuen beruht. Trask und Mitarbeiter haben eine 36 kb große Region beschrieben, die drei OR Gene enthält, von denen eines potentiell funktionstüchtig ist. Duplikate mit hoher Ähnlichkeit dieser Region treten in vielen subtelomeren chromosomalen Bereichen auf. Bei allen 45 untersuchten Individuen ist dieses Segment auf 3q, 15q und 19p gefunden worden. Sieben bis elf Kopien dieser Region wurden bei den einzelnen Individuen gefunden.

5.1.1. OR-Pseudogene

Die Haupt- und Neben-OR-Cluster, die von HLA-F ca. 135 kb bzw. 2000 kb entfernt sind, enthalten 34 M-OR Gene (s. Abb. 4) (Younger et al., 2001). Von diesen OR Genen sind zumindest in einigen Haplotypen 17 (50%) potentiell funktionstüchtig, während es sich bei

den restlichen 50% vermutlich um Pseudogene handelt. Dieser Prozentsatz an Pseudogenen scheint in Abhängigkeit vom betrachteten OR-Cluster sehr unterschiedlich zu sein. Im Gegensatz zu den Ergebnissen bei den HLA-gekoppelten OR Clustern, haben Glusman und Mitarbeiter in der chromosomalen Region 17p13.3 einen Pseudogen-Anteil nachgewiesen, der weniger als ein Drittel beträgt (Glusman et al., 2000; Gilad et al., 2000). Aufgrund der bisher insgesamt bekannten OR Gene wird in einer anderen Studie vermutet, daß die Anzahl der Pseudogene im humanen Genom ca. 53% beträgt (Fuchs et al., 2001), was mit den hier dargestellten Ergebnissen sehr gut übereinstimmt. Betrachtet man jedoch die zwei HLA-gekoppelten OR-Cluster getrennt, so findet man bereits relativ große Unterschiede: Im Haupt-Cluster sind 14/25 (56%) der OR Gene potentiell funktionstüchtig, im Neben-Cluster jedoch nur 3/9 (33%). Dabei ist jedoch zu bedenken, daß für die Pseudogene weitere Allele existieren könnten, die einen ORF haben. Außerdem liegen für *hs6M1-14P* und *-24P* bereits EST-Daten vor, die belegen, daß diese beiden Gene zumindest transkribiert werden (Younger et al., 2001).

Für die komplett sequenzierten Pseudogene (*hs6M1-2P*, *-7P* und *13P*) konnten zwei bis drei Allele definiert werden. Auch wenn die Anzahl der analysierten Pseudogene nicht sehr hoch ist, weist doch alles darauf hin, daß sich die Häufigkeit der Polymorphismen der Pseudogene und der funktionellen Gene nicht wesentlich unterscheidet (s. Abschnitt 5.1.2): Sieben von 13 analysierten Genen mit ORF weisen ebenfalls zwei bis drei Allele auf. Im Kontrast dazu weisen die Pseudogene auf Chromosom 17 mehr Polymorphismen auf als die funktionellen Gene (Gilad et al., 2000). Bei einer Analyse von sieben intakten OR Genen und fünf Pseudogenen in 50 Individuen wurden 3,8 +/- 1,4 Haplotypen für die Pseudogene und nur 2,3 +/- 0,4 Haplotypen für die funktionellen Gene gefunden.

Besonders interessant sind die Gene *hs6M1-4P*, *-17* und *-19P*, da sie in einigen Zelllinien einen ORF besitzen, in anderen jedoch nicht. Bei *hs6M1-17* tritt das inaktive Allel (*hs6M1-17*05*) nur bei einer Zelllinie auf, während die inaktiven Allele bei *hs6M1-4P* (*-4P*01*) und *-19P* (*-19P*01*) sehr viel häufiger vorkommen (s. Tab. 15). In beiden Fällen besitzen mehr als die Hälfte der untersuchten Chromosomen 6 ein nicht funktionelles Allel. Das Allel *hs6M1-19P*01* ist vermutlich aus dem potentiell funktionstüchtigen Allel *hs6M1-19P*02* durch eine Deletion entstanden, da es unwahrscheinlich ist, daß sich 16 bp genau in die korrekte Position von einem nicht funktionellen Allel einfügen, um ein funktionelles Allel entstehen zu lassen. Damit stellt das Gen *hs6M1-19P* ein gutes Beispiel für „Pseudogenisierung“ innerhalb der menschlichen Spezies dar. Bisher wurde - bezüglich der OR Gene - "Pseudogenisierung" nur zwischen verschiedenen Spezies beschrieben (Rouquier et al., 1998a). Es wird deshalb angenommen, daß die Evolution der OR Gene sich in den olfaktorischen Fähigkeiten widerspiegelt, wobei Säugetiere mit ausgeprägtem Geruchsinne, wie z.B. Nagetiere, nur wenige Pseudogene aufweisen, während Säuger mit

einem weniger gut ausgeprägten Geruchssinn, wie z.B. Menschen, viele Pseudogene besitzen. Andererseits könnte es aber auch sein, daß die Pseudogene doch eine Funktion besitzen und deshalb durch positiven Selektionsdruck entstanden sind. Die Pseudogene könnten zum Beispiel durch Genkonversion die Variabilität des OR Gen Repertoires vergrößern. Eine derartige Funktion erfüllen die Pseudogene im Immunsystem von Vögeln, wo die Antikörpervielfalt auf diese Weise entsteht: Hühner z.B. besitzen nur ein funktionelles V_L -Gen. Die 25 V_L -Pseudogene dienen als Donorsequenzen für die Genkonversionen, die zu Nukleotid-Deletionen oder -Insertionen führen. 97% dieser Ereignisse führen letztendlich zu einem Gen mit offenem Leseraster (Sayegh et al., 1999). Für diese Hypothese spricht, daß bereits zahlreiche Genkonversions-Ereignisse bei OR Genen nachgewiesen wurden (Sharon et al., 1999; Sharon et al., 2000). Auch in dieser Arbeit wurden in zwei Fällen Mutationen nachgewiesen, die bei verschiedenen OR Genen an korrespondierenden Positionen auftreten: bei *hs6M1-3* und *-6* sowie bei *hs6M1-20* und *-21* (s. Abb. 7). Es kann also auch hier davon ausgegangen werden, daß Genkonversion stattgefunden hat. Dieser Mechanismus führt damit auf somatischer Ebene zu einer größeren Rezeptorvielfalt. Gegen diese Hypothese spricht, daß die Donorsequenzen sowohl in den Studien von Sharon (Sharon et al., 1999; Sharon et al., 2000) als auch in dieser Arbeit funktionelle Gene sind. Allerdings ist hierzu anzumerken, daß zumindest in dieser Arbeit nur drei Pseudogene komplett sequenziert wurden, so daß es durchaus sein kann, daß nur aufgrund der geringen Datenmenge noch keine Pseudogene als Donorsequenzen nachgewiesen werden konnten.

5.1.2. Aminosäuresubstitutionen innerhalb der ORs

Alle analysierten OR Gene wiesen in der analysierten Stichprobe mehrere Allele auf. Bis auf eine Ausnahme (*hs6M1-1*) gilt dies auch für die resultierenden Proteine. Insgesamt wurden bei den 13 HLA-gekoppelten OR Genen mit ORF 52 Einzelbasenaustausche gefunden, von denen 36 zu einer veränderten Aminosäuresequenz führen. Ganz anders sehen die Ergebnisse von Sharon und Mitarbeitern aus, die OR Gene auf Chromosom 17 analysierten. In dieser Studie wurden 14 OR Gene in 15 Individuen bzw. in einem Pool von 200 Individuen untersucht, wobei 26 Substitutionen gefunden wurden, von denen 21 zu einem Aminosäure Austausch führten (Sharon et al., 2000). Die HLA-gekoppelten OR Gene scheinen damit ca. doppelt so viele Polymorphismen aufzuweisen wie die OR Gene auf Chromosom 17, obwohl sogar noch fünf Individuen weniger untersucht wurden. Das hier analysierte Gen auf Chromosom 17 (*hs17M1-20*), das dem OR17-24 in der Sharon-Studie entspricht, weist ebenfalls zahlreiche Mutationen auf, während bei Sharon und Mitarbeitern nur eine einzige Mutation gefunden wurde. Eine mögliche Erklärung für diese Diskrepanz wäre, daß die untersuchten Individuen in dieser Arbeit eine größere Genvarianz aufweisen. In beiden Studien findet man bei *hs17M1-20* bzw. OR17-24 eine identische Mutation in EC3 an Aminosäureposition 193, was dafür spricht, daß diese Substitution weit verbreitet ist (in den

hier dargestellten Ergebnissen wurde ein Verhältnis von 8/13 gefunden) und somit evolutionär relativ alt sein sollte. Ein weiterer Unterschied besteht in der Heterozygotie der OR Gene. Während in dieser Arbeit insbesondere drei Zelllinien (KR3598, SA und OLGA) heterozygot für zahlreiche HLA-gekoppelte OR sind, beschreibt die Sharon-Studie kein einziges für einen OR heterozygoten Individuum. Das mag aber, wie oben bereits erwähnt, an der Auswahl der analysierten Individuen liegen. Unterschiede in der Methodik sind eine weitere Möglichkeit die Differenzen zu erklären. Sharon und Mitarbeiter haben sowohl direkt aus der PCR als auch klonierte DNA sequenziert. Während man bei der Sequenzierung von PCR-Produkten bei heterozygoten Positionen zwei Banden in verschiedenen Spuren findet, kann man unterschiedliche Allele in Klonen nur durch eine Analyse von zahlreichen Klonen nachweisen. Leider geht aus der Sharon-Studie nicht hervor, welche Methode für welche Gene verwendet wurde. Dies wäre besonders für das *hs17M1-20* Gen interessant gewesen, das bei Aminosäureposition 193 in dieser Arbeit in sieben Zelllinien heterozygot ist, während die identische Substitution bei Sharon anscheinend überall homozygot auftritt. Außerdem deklarieren Sharon und Mitarbeiter OR17-24 bzw. *hs17M1-20* als Pseudogen, obwohl die hier erhaltene Sequenz einen kompletten ORF aufweist. Auch Zozulya und Mitarbeiter kommen zu dem Ergebnis, daß zumindest die in der NCBI Datenbank verfügbaren Sequenzen auf einen funktionellen Status von OR17-24 hinweisen (Zozulya et al., 2001).

Mit Ausnahme von EC4 und TM7 sind in allen Domänen Polymorphismen in der Aminosäuresequenz gefunden worden. Diese Mutationen könnten die Interaktion mit extrazellulären Liganden oder mit intrazellulären Proteinen, die z.B. an der Signal-Transduktions-Kaskade beteiligt sind, beeinflussen. Durch den Vergleich von Sequenzen von 197 ORs verschiedener Spezies wurde herausgefunden, daß die Transmembran (TM) Domänen 3, 4, 5 und 6 die größte Sequenzdiversität zeigen (Pilpel and Lancet, 1999). Auch die extrazellulären Regionen EC2 und besonders EC4 (entspricht EC1 und EC3 in deren Nomenklatur) scheinen sehr variabel zu sein. Bei den hier analysierten Zelllinien wurde in EC4 keine einzige Substitution gefunden (s. Abb. 6), und in TM7 wurde nur eine einzige stumme Mutation (*hs6M1-15*) gefunden (s. Tab. 15). Auch die beiden in Tab. 14 aufgeführten Pseudogene (*hs6M1-2P* und *-7P*) weisen weder in EC4 noch in TM7 eine Mutation auf. Es wird vermutet, daß die Austausch in den TM Domänen die Liganden Affinität beeinflussen (Buck and Axel, 1991; Pilpel and Lancet, 1999). Selbst ein so konservativer Austausch wie die Aminosäuresubstitution Val159Ile in TM4 von *hs6M1-20* könnte die Ligandenspezifität verändern, was bei einem vergleichbaren Austausch in der TM5 Domäne eines murinen OR gezeigt werden konnte: Die Präferenz dieses OR wechselte durch die Substitution von Octanal zu Heptanal (Krautwurst et al., 1998). Dagegen ist der Arg138Trp Polymorphismus in *hs6M1-17* ein Beispiel für eine drastische Substitution. Es bleibt zu vermuten, daß auch dies Auswirkungen auf die Funktion des Rezeptors nach sich

zieht.

M-OR Gene beinhalten intrazelluläre Sequenzmotive, die hoch konserviert sind und vermutlich bei der G-Protein Bindung eine Rolle spielen (Hedin et al., 1993; Pilpel et al., 1998; Pilpel and Lancet, 1999; Skoufos, 1999; Mombaerts, 1999b; Mombaerts, 1999c; Zozulya et al., 2001). Beispiele dafür sind die ORIC1, ORIC2 und ORIC3 Regionen in den cytoplasmatischen Domänen CP1, 2 und 3 (Pilpel and Lancet, 1999). Die Gene *hs6M1-17* und *-20* zeigen in CP2 Mutationen, während *hs6M1-3*, *-20* und *-21* in CP3 Polymorphismen aufweisen (s. Tab. 15). Diese Substitutionen könnten dazu führen, daß die Signalkaskade gestört ist. Neben den intrazellulären Domänen beinhalten aber auch die TM1, TM2 und TM7 Domänen hoch konservierte Regionen (Zozulya et al., 2001). Um z.B. degenerierte Oligonukleotide herzustellen, werden diese routinemäßig in die Domänen TM2 und TM7 gelegt. Sowohl das Alignment von 347 OR in der Studie von Zozulya als auch das Alignment der 15 HLA-gekoppelten OR Gene mit ORF in Abb. 2 zeigt eine besonders konservierte Region in TM2 mit dem Motiv PMYFFL.

Alle G-gekoppelten Rezeptoren besitzen ein hochkonserviertes Cystein in EC2 und ein weiteres in EC3 (bei Zozulya EC1 und EC2) (Zozulya et al., 2001). Es gibt Hinweise dafür, daß diese beiden Cysteine eine intra- oder inter-molekulare Disulfidbrücke bilden, was in einigen Fällen wichtig für die Ligandenerkennung ist (Le Gouill et al., 1997; Zeng and Wess, 1999; Blanpain et al., 1999). Die hier dargestellten Ergebnisse (s. Abb. 2) und die Resultate von Zozulya und Mitarbeitern weisen ein hochkonserviertes Cystein am Ende von EC2 (bzw. EC1 bei Zozulya) und drei in EC3 auf. Dies ist eher ungewöhnlich für G-gekoppelte Rezeptoren und nur wenige, wie z.B. Chemokine oder P2Y1-Rezeptoren, weisen zusätzliche disulfidbrückenbildende Cysteine im aminoterminalen Ende und in EC4 auf (bzw. EC3 in anderen Veröffentlichungen) (Blanpain et al., 1999; Hoffmann et al., 1999). Es wird spekuliert, daß diese vier hochkonservierten Cysteine zwei Disulfidbrücken auf der extrazellulären Oberfläche der OR ausbilden.

Mit Hilfe der großen Anzahl olfaktorischer Rezeptoren läßt sich erklären, warum Säugetiere Tausende von Gerüchen differenzieren können. Die Variabilität dieser Rezeptoren (Ehlers et al., 2000; Sharon et al., 2000) bietet sich an, um individuell unterschiedliche Geruchswahrnehmungen, die im Extremfall in spezifischen Anosmien gipfeln können, zu erklären. Leider gibt es bis jetzt noch keine Ergebnisse bezüglich der Ligandenspezifität der HLA-gekoppelten OR, so daß es etwas voreilig wäre, über die funktionellen Konsequenzen der einzelnen Aminosäure Austausch zu spekulieren. Man kann aber darüber diskutieren, ob die Individuen, die homozygot für ein nicht funktionelles Allel von z.B. *hs6M1-4P* (eines der drei Gene mit funktionellen und nicht funktionellen Allelen) sind, Symptome einer spezifischen Anosmie aufweisen. Derartige olfaktorische Defizite bieten die Möglichkeit, Verhaltensweisen mit genetischen Polymorphismen zu korrelieren (Mombaerts, 1999b).

Personen, die heterozygot für ein aktives und ein inaktives Allel sind, könnten eine andere Geruchsschwelle besitzen als Personen, die zwei aktive Allele besitzen. Vermutlich ist aber die Situation insgesamt komplizierter, da es Hinweise darauf gibt, daß Geruchsspezifitäten verschiedener OR überlappen können. Ein fehlendes olfaktorisches Signal, verursacht durch die Abwesenheit des entsprechenden funktionellen Rezeptors, könnte durch Signale anderer Rezeptoren mit überlappenden Spezifitäten kompensiert werden (Malnic et al., 1999; Tsuboi et al., 1999).

5.1.3. Definition von Haplotypen

Obwohl der Polymorphismusgrad von OR verglichen mit dem von HLA-Molekülen offenbar limitiert ist, war es möglich mit den 18 untersuchten Chromosomen 6 dreizehn verschiedene Haplotypen für HLA-gekoppelte OR Gene zu definieren (s. Tab. 16), obwohl nur die dreizehn analysierten Gene mit ORF für diese Kalkulation verwendet wurden. Die Anzahl der OR Haplotypen ist also größer als die Anzahl der HLA-Klasse I Haplotypen bei den hier verwendeten Zelllinien. Das liegt zum einem daran, daß im Vergleich zu 13 OR Genen nur drei HLA-Klasse I Gene analysiert wurden. Es ist also durchaus denkbar, daß einige der Zelllinien zwar einen identischen Satz an HLA-Klasse I Genen haben, aber genau betrachtet für diese Region trotzdem nicht homozygot sind. Entsprechend waren vier der HLA-Klasse I homozygoten Zelllinien im Bereich der beiden OR-Cluster nicht mehr homozygot.

Die Zelllinien H2LCL und YAR besitzen für alle untersuchten Genorte identische OR Allele, mit der einzigen Ausnahme, daß H2LCL für das Gen *hs19M1-4* im Gegensatz zu YAR heterozygot ist (s. Tab. 15). YAR ist jüdischer Herkunft, während die Herkunft von H2LCL nicht mit Sicherheit angegeben werden kann (obwohl in Tab. 1 eine kaukasische Herkunft angenommen wird). Wie in Abschnitt 4.1.6 beschrieben, besitzen auch noch andere Zelllinien kleinere Blöcke identischer Allele. Es wäre interessant zu erfahren, ob diese mehr oder weniger langen gemeinsamen Allelblöcke generell in der Bevölkerung auftreten. Die CEPH Familien wären für eine derartige Untersuchung gut geeignet, da alle Mitglieder bereits HLA-typisiert sind (Ziegler, persönliche Mitteilung.). Diese Familien wurden auch verwendet, um das Kopplungsungleichgewicht zwischen HLA-A und HFE zu untersuchen (Malfroy et al., 1997). Die enge Nachbarschaft polymorpher OR und MHC-Klasse I Gene, die zusammen mit einer reduzierten Rekombinationsrate zu solchen „ausgeweiteten“ Haplotypen führt, könnte bei der Ausbildung von Geruchspräferenzen bei der Partnerwahl beteiligt sein (Wedekind et al., 1995; Wedekind and Furi, 1997; Penn and Potts, 1998a). Ein Zusammenhang zwischen Produkten von MHC und OR-ähnlichen Genen wurde bereits 1976 von Yamazaki und Kollegen vermutet (Yamazaki et al., 1976). Produkte des MHC beeinflussen den individualspezifischen Geruch bei Inzucht-Mäusen und -Ratten (Yamazaki et al., 1979; Singh et al., 1987; Brown et al., 1989; Penn and Potts, 1998a; Penn and Potts, 1998c; Yamazaki et al., 1999). Der MHC-abhängige Körpergeruch spielt eine wichtige Rolle

bei der Partnerwahl (Yamazaki et al., 1976; Potts et al., 1991; Penn and Potts, 1998b), wobei MHC unähnliche Paarungspartner bevorzugt werden. Bei Mäusen beeinflusst bereits eine Punktmutation im H-2K Gen (als H-2 bezeichnet man den MHC bei Mäusen) den Fortpflanzungserfolg (Yamazaki et al., 1986). Auch bei Menschen wurde eine Präferenz für HLA-unähnliche Partner nachgewiesen: Körpergerüche werden als "attraktiver" empfunden, je weniger HLA-Klasse I Antigene der Geruchsauscheider und der Geruchsempfänger teilen (Wedekind et al., 1995; Wedekind and Furi, 1997). Des weiteren scheint der Mensch für sich selbst Parfums auszuwählen, die in irgendeiner Form seinen eigenen Körpergeruch verstärken und somit die "Immungenetik" dieser Person stärker zum Ausdruck bringen. In einer Verhaltensstudie mit 137 weiblichen und männlichen Probanden wurde nachgewiesen, daß eine signifikante Korrelation zwischen dem MHC und dem Duftstoff besteht, nach dem man gerne riechen möchte (Milinski and Wedekind, 2001).

Ein funktioneller Zusammenhang zwischen HLA-Komplex und HLA-gekoppelten OR Genen könnte entstanden sein, um Inzucht zu vermeiden oder um die Selektion eines MHC-divergenten Partners zu begünstigen. Die resultierenden MHC-heterozygoten Nachkommen wären somit besser gegen Krankheiten geschützt (Potts and Wakeland, 1993; Beauchamp and Yamazaki, 1997; Penn and Potts, 1998a). Um eine derartige Interaktion zwischen HLA und OR Genen schlüssig zu belegen, sind jedoch noch weitere Arbeiten notwendig (s. Abschnitt 5.5).

5.2. Vorkommen von HLA-gekoppelten OR Genen bei verschiedenen Säugetieren

Es wird allgemein angenommen, daß die zahlreichen OR Gene sowohl durch Duplikation einzelner Gene sowie ganzer Cluster entstanden sind (Glusman et al., 1996; Sullivan et al., 1996; Trask et al., 1998b). Die resultierenden Gene, vor allem wenn sie evolutionsbiologisch noch jung sind, zeigen erwartungsgemäß eine extrem hohe Sequenzhomologie zu ihren Vorläufern. Die HLA-gekoppelten OR Gene *hs6M1-3*, *-4P*, *-5P* und *-6* z.B. gehören zu dieser Gruppe. Ihre Nukleotidsequenzen weisen eine Ähnlichkeit von ca. 90% auf. Aus diesem Grunde war es nicht möglich, Sonden für die Expressionsanalysen herzustellen, die jeweils nur eines dieser Gene erkannten. Dies ist auch auf den Southern Blots mit verschiedenen humanen und tierischen genomischen DNAs zu sehen (s. Abb. 11). Bei den EcoRI verdauten humanen Zelllinien wurden z.B. mit der jeweiligen *hs6M1-3*, *-4P* und *-6* Sonde immer vier Banden gefunden (s. Abb. 10 und Abb. 11), was mit großer Wahrscheinlichkeit die Kreuzreaktion der jeweiligen Sonde mit den anderen drei Genen widerspiegelt. Die uns nahe verwandten Schimpansen zeigen ebenfalls vier Kopien, wobei die Restriktionsfragmente identische Größen haben (s. Abb. 11). Es handelt sich deshalb vermutlich um Orthologe. Des weiteren kann man darüber spekulieren, ob z.B. das

hs6M1-4P Gen, das beim Menschen sowohl funktionelle als auch nicht funktionelle Allele aufweist, beim Schimpansen noch bei allen Individuen funktionstüchtig ist. Die etwa gleich große Häufigkeit der beiden Allele beim Menschen (s. Tab. 15) deutet jedoch darauf hin, daß die Mutation schon relativ alt ist, weshalb eine ähnliche Situation bei Schimpansen zu erwarten ist. Identische Bandenmuster bei Mensch und Schimpanse findet man übrigens auch für die sequenzähnlichen Gene *hs6M1-16*, *-12* und/oder *-13P*. Orthologe nicht HLA-gekoppelter menschlicher OR Gene im Schimpansen-Genom sind bereits nachgewiesen worden. Sharon und Mitarbeiter haben von 15 OR Genen aus der humanen Region 17p13 Orthologe in der syntenen Region 19p15 des Schimpansen gefunden (Sharon et al., 1999). Mit allen verwendeten Sonden wurden auch Kopien im Pavian-Genom gefunden. Bei *hs6M1-3* und *-6* ist die Anzahl der Banden zwar jeweils identisch mit der von Mensch und Schimpanse, aber die Größe der Restriktionsfragmente zeigt Unterschiede. Dieser Größenunterschied könnte aber auf die hohe DNA Konzentration in dieser Spur zurückzuführen sein, die einen Einfluß auf die Laufgeschwindigkeit hat.

Die grüne Meerkatze weist einen geringeren Verwandtschaftsgrad zum Menschen auf: Hier zeigen nur 10 von 15 verwendeten Sonden positive Signale.

Beim Schwein werden z.T. viele positive Banden mit einer Sonde – wie z.B. bei *hs6M1-3* – gefunden. Diese sind jedoch alle sehr schwach, d.h. die Sequenzhomologie ist im Vergleich zu den Affen deutlich reduziert. Daraus folgt vermutlich, daß im Laufe der Evolution die Vorläufer der HLA-gekoppelten OR Gene beim Schwein viel häufiger dupliziert wurden als bei Primaten. Bei Rind und Ratte sind die Ähnlichkeiten der HLA-gekoppelten OR Gene im Vergleich zum Schwein noch weiter reduziert: Es sind nur noch sehr wenige positive Banden zu sehen, die außerdem sehr schwach sind.

Von den *hs6M1-14P* und *-16* Genen wurden bereits funktionelle Orthologe im Genom der Maus gefunden (Younger et al., 2001). Diese Gene wurden auf Chromosom 17 gefunden, in einer Region, die synten zur Region 6p21.3-22 des Menschen ist. Auch von OR Genen auf Chromosom 17 wurden Orthologe in der syntenen Region auf Chromosom 11 der Maus gefunden (Lapidot et al., 2001).

Die hier beschriebenen Experimente resultieren lediglich in Informationen über die Existenz homologer Sequenzen in anderen Spezies; eine Aussage über die Funktionsfähigkeit der entsprechenden Gene ist aber nicht möglich. Eine Sequenzanalyse von OR Genen bei Hominiden (N=183), Altweltaffen (N=41), Neuweltaffen (N=49), Lemuren (N=35), Nagern (N=33) und Fischen (N=3) hat gezeigt, daß der Prozentsatz an Pseudogenen bei den Hominiden mit 50% am größten ist, während bei Nagern und Fischen – zumindest in dieser Studie - kein einziges Pseudogen gefunden wurde (Rouquier et al., 2000). Ein MHC-gekoppeltes OR Pseudogen bei der Maus (*mm17M1-5P*) ist aber von Younger et. al beschrieben worden (Younger et al., 2001). Anzumerken ist außerdem, daß die Anzahl der

analysierten Fisch OR Gene in der Rouquier Studie sehr gering war (Rouquier et al., 2000). Innerhalb der Hominiden ist der Pseudogenanteil beim Menschen am größten (53%) (Fuchs et al., 2001) und beim Orang Utan am geringsten (39%) (Rouquier et al., 2000). Dies entspricht auch den Ergebnissen von Sharon und Mitarbeitern (Sharon et al., 1999), die den OR Cluster auf Chromosom 17p13 analysierten. Innerhalb dieses Clusters wurde beim Orang Utan kein einziges Pseudogen gefunden, während im entsprechenden humanen Cluster sechs von insgesamt fünfzehn OR Genen keinen ORF aufweisen. Auch hier wurde ein Gen (OR17-31) beschrieben, das bei einigen Individuen einen ORF aufweist, bei anderen aber nicht. Da Mensch und Orang Utan vor ca. 8 Mio Jahren die letzten gemeinsamen Vorfahren hatten (Kumar and Hedges, 1998), haben viele der OR Gene im humanen Genom ihre Funktionsfähigkeit in einer evolutionsbiologisch relativ kurzen Zeit verloren. Es bleibt zu vermuten, daß die geistige Weiterentwicklung der Hominiden die Relevanz des Geruchssinns reduziert hat, z.B. weil die Qualität von Nahrungsmitteln auch über andere Sinneseindrücke eingeschätzt werden konnte. Ob die Pseudogene tatsächlich lediglich dazu führen, daß Menschen weniger Geruchsqualitäten auseinanderhalten können, oder ob sie vielleicht doch einen biologischen Sinn erfüllen, wurde bereits in Abschnitt 5.1.1 diskutiert.

5.3. Expression von HLA-gekoppelten OR Genen in verschiedenen Geweben

Bereits seit 1992 ist es bekannt, daß OR Gene nicht nur im olfaktorischen Epithel, sondern auch in anderen Geweben exprimiert werden (Parmentier et al., 1992; Vanderhaeghen et al., 1993; Vanderhaeghen et al., 1997a; Dreyer, 1998). Auch im Rahmen dieser Arbeit konnte gezeigt werden, daß Transkripte der analysierten HLA-gekoppelten ORs nicht nur im Riechepithel, sondern auch in zahlreichen anderen Organen vorkommen. Es ist deshalb wahrscheinlich, daß die olfaktorischen Rezeptoren neben der Erkennung von Gerüchen in der Atemluft auch andere Funktionen erfüllen. J. Dreyer hat eine Hypothese aufgestellt, nach der olfaktorische Rezeptoren als Zelloberflächenrezeptoren innerhalb des Organismus dienen könnten (Dreyer, 1998). Des weiteren wurde eine Beteiligung an der Embryo- und Organogenese in Erwägung gezogen. Prinzipiell könnten ORs dafür besonders geeignet sein, weil sie erstens zu einer sehr umfangreichen Multigenfamilie gehören und zahlreiche hochspezifische Zelloberflächenrezeptoren kodieren könnten, und zweitens sind sie fast überall im Organismus zu finden. Unterstützt wird diese Hypothese zusätzlich dadurch, daß OR Transkripte in der humanen embryonalen Niere gefunden wurden (Wetzel et al., 1999). Im embryonalen Rattenherz wurde ebenfalls ein OR nachgewiesen (Drutel et al., 1995). Auch die Ergebnisse hier zeigen, daß vier der sechs analysierten Gene in fötalen Geweben (Herz, Niere, Leber, Milz (nur z.T.), Thymus (nur z.T.) und Lunge) vorkommen. Bei einigen Säugetieren wurden OR Transkripte im Hoden nachgewiesen (Hund, Mensch, Maus und

Ratte) (Parmentier et al., 1992; Vanderhaeghen et al., 1993; Vanderhaeghen et al., 1997b). Sowohl M-ORs als auch V1-ORs wurden bei Mäusen in heranreifenden Spermatozyten gefunden, wo sie wahrscheinlich co-exprimiert werden (Tatsura et al., 2001). Es wird spekuliert, daß sie eine Rolle bei der Spermien/Eizell-Interaktion oder bei der Spermatogenese spielen (Vanderhaeghen et al., 1993; Vanderhaeghen et al., 1997a; Branscomb et al., 2000). Innerhalb der Säugetiere sollen 10% der OR's im Hoden exprimiert werden (Branscomb et al., 2000). Im Rahmen dieser Arbeit konnte die Expression von zahlreichen HLA-gekoppelten OR Genen im Hoden ebenfalls demonstriert werden (mRNA-Blot-Ergebnisse: Abschnitt 4.2.2; RACE-Ergebnisse: Abschnitt 4.3.2.3, s. Abb. 15 und Abb. 16). Die Autoradiographie der RACE-Produkte aus der Hoden-cDNA-Bibliothek demonstriert anschaulich, daß mit Ausnahme von *hs6M1-4P* alle analysierten HLA-gekoppelten OR Gene im Hoden exprimiert werden (s. Abb. 16), wobei der funktionelle Status von *hs6M1-4P*, wie in Abschnitt 4.1.4 bereits erwähnt, vom jeweiligen Haplotyp abhängig ist. Bis auf *hs6M1-18* war mit Hilfe des mRNA-Blots die Expression im Hoden für alle analysierten Gene nachzuweisen (s. Tab. 17). Eine mögliche Erklärung, warum *hs6M1-18* dort nicht detektiert werden konnte, ist, daß die Hybridisierung an nicht amplifizierte RNA deutlich weniger sensitiv ist. Die Diskrepanzen zwischen den Ergebnissen des mRNA-Blots und der RACE können zusätzlich zu den Sensitivitätsunterschieden auch daran liegen, daß die RNA von unterschiedlichen Individuen stammt, die ein unterschiedliches Expressionsmuster aufweisen könnten.

Im Hoden wird ein breites Spektrum HLA-gekoppelter OR Gene exprimiert, in Niere und Lunge aber nur ganz bestimmte OR Gene (s. Abb. 15). Umgekehrt werden einige HLA-gekoppelte OR Gene nur in wenigen spezifischen Organen exprimiert (z.B. *hs6M1-16* bzw. *-12* und *-18*), während andere in zahlreichen Geweben zu finden sind (s. Tab. 17). Man kann darüber spekulieren, ob je nach Funktion des Organs ein anderes Expressionsmuster von OR Genen auftritt. Für den Hoden könnte es zum Beispiel essentiell sein, daß in diesem Gewebe besonders viele OR Gene exprimiert werden, da sie als Vermittler bei der Chemotaxis zwischen Spermien und Eizellen dienen könnten (Vanderhaeghen et al., 1993; Vanderhaeghen et al., 1997a). Für andere Organfunktionen könnte es dagegen wichtiger sein, gezielt einzelne OR zu exprimieren; vielleicht weil spezifische Liganden benötigt werden, um eine ganz bestimmte Signalkaskade auszulösen.

Die bisher erwähnten Expressionsanalysen gaben zwar Aufschluß darüber, ob ein bestimmtes HLA-gekoppeltes OR Gen in einem bestimmten Organ exprimiert wird, aber sie vermitteln keine Information darüber, in welchen Zellen des Organs die Transkripte vorkommen. Dies sollte mit Hilfe der *In-situ*-Hybridisierung analysiert werden. Für die Expression von OR Genen im Hoden ist bereits bekannt, daß OR Transkripte in Spermatozyten und Spermatozyten vorkommen (Parmentier et al., 1992; Vanderhaeghen et al., 1993). Leider konnte mit Hilfe der *In-situ*-Hybridisierung keine zellspezifische Expression

von HLA-gekoppelten OR Genen nachgewiesen werden. Weder die *hs6M1-16* noch die *hs6M1-10* Sonde resultierte in einer positiven Anfärbung im Hodengewebe. Auch die Hybridisierung von *hs6M1-10* mit Nieren- und Lebergewebe war negativ. Diese Ergebnisse stimmen nicht mit den Resultaten der mRNA-Blot Experimente überein. Eine mögliche Erklärung dafür ist, daß die Zellen, die eigentlich diese Gene exprimieren, in dem nur einige µm dünnen Schnitt nicht vorhanden waren. Da die Schnitte jedoch sowohl zentrale wie auch Randbereiche beinhalteten, ist diese Erklärung nicht besonders wahrscheinlich. Wie bereits oben angedeutet, ist es denkbar, daß die Expression der Gene von Individuum zu Individuum unterschiedlich ist. Während sowohl der mRNA-Blot als auch die cDNA-Bibliothek jeweils mRNA von zahlreichen Individuen enthielt (s. Abschnitt 4.2.2), wurden bei der *In-situ*-Hybridisierung maximal sieben verschiedene Individuen getestet (s. Abschnitt 4.2.3). Die wahrscheinlichste Erklärung ist jedoch, daß die Expression so gering ist, daß die Konzentration der Transkripte in den Schnitten unterhalb der Nachweisgrenze lag. Dafür spricht außerdem, daß die mRNA-Blots z.T. bis zu drei Wochen exponiert werden mußten, bis sich ein positives Ergebnis zeigte.

Außer von Neuronen aus dem olfaktorischen Epithel (Malnic et al., 1999; Touhara et al., 1999), gibt es leider noch keine Funktionsanalysen bezüglich der OR Expression in verschiedenen Körperorganen. Um die Funktion der OR's in diesen Geweben wirklich verstehen zu können, müßten zunächst einmal die Liganden für die dort exprimierten Rezeptoren gefunden werden. Mit großer Wahrscheinlichkeit sind sie gänzlich anderer Art als die Liganden, die bevorzugt mit ORs im MOE interagieren. Im MOE werden vermutlich größtenteils flüchtige Substanzen, die aus der Umgebung stammen, gebunden. Dagegen ist anzunehmen, daß die exprimierten ORs in anderen Körperorganen mit löslichen Substanzen aus den Körperflüssigkeiten (Blut oder Lymphe) oder aber mit Oberflächenmolekülen von anderen Zellen interagieren. Diese Überlegungen bieten einen guten Ansatzpunkt für weitere Arbeiten.

5.4. Genomische Organisation der HLA-gekoppelten OR Gene

Ziel dieses Abschnittes war es, die genomische Struktur der humanen HLA-gekoppelten OR Gene zu charakterisieren, um mehr über die Regulation ihrer Expression zu erfahren, und um Zusammenhänge zwischen der genomischen Struktur und der Funktion aufzudecken. Bisher ist von der Struktur der OR Gene bekannt, daß die kodierende Region keine Introns beinhaltet und ca. 1 kb umfaßt (Buck and Axel, 1991; Ben-Arie et al., 1994). Eine Analyse von humanen und murinen OR Genen hat erst vor kurzem ergeben, daß die Gene stromaufwärts ein oder mehrere nicht kodierende Exons besitzen. Eine Distanz von 1,3 bis 11,1 kb zwischen dem ersten 5' nicht kodierenden Exon und dem kodierenden Exon wurde bereits beschrieben (Glusman et al., 1996; Asai et al., 1996; Walensky et al., 1998; Qasba

and Reed, 1998; Sosinsky et al., 2000). Auch im Rahmen dieser Arbeit wurden nicht kodierende Exons gefunden, die jedoch sowohl 5' wie auch 3' lagen und einen maximalen Abstand von ~100 kb vom kodierenden Exon aufwiesen. Des Weiteren wurde in extremen Maße alternatives Spleißing gefunden.

5.4.1. Differentielle Polyadenylierung von HLA-gekoppelten OR Genen

Von *hs6M1-16* und *-21* wurden zahlreiche Hoden-Transkripte gefunden, die eine Polyadenylierung innerhalb der kodierenden Region aufwiesen (s. Abb. 19 und Abb. 20). Allerdings weist keines dieser Transkripte die üblichen Polyadenylierungssignale auf. Die Polyadenylierung findet an verschiedenen Stellen statt, wobei in mehreren unabhängigen Klonen immer wieder die gleiche Position gefunden wurde. Da gleichlange Transkripte in zwei unabhängigen Bibliotheken gefunden wurden, ist es sehr unwahrscheinlich, daß es sich um Artefakte handelt. Bisher ist ein derartiges Phänomen im Tierreich noch nicht beschrieben worden. Es gibt allerdings Untersuchungen mit einem Kristallprotein-Gen aus *Bacillus thuringiensis*, das in Kartoffel- und Tabakpflanzen überführt wurde (Diehn et al., 1998; Haffani et al., 2000). In dieser Studie stellte sich bei einer anschließenden Expressionsanalyse heraus, daß es nur vorzeitig polyadenylierte mRNA Transkripte gab, die ebenfalls immer an ganz bestimmten Positionen endeten. Auch hier konnte das übliche Polyadenylierungssignal AAUAAA nicht gefunden werden. Allerdings handelt es sich bei den Pflanzen um vorzeitig polyadenylierte Fremd-DNA, während es sich bei den olfaktorischen Genen um zelleigene Gene handelt.

Die vorzeitige Polyadenylierung von Transkripten könnte einen Mechanismus darstellen, um die Expression eines Gens zu kontrollieren. Von dem potentiell funktionstüchtigen Gen *hs6M1-21* wurden z.B. nur kurze Transkripte gefunden, aber nie die komplette kodierende Region bzw. eine 3' UTR. Es ist daher anzunehmen, daß es von diesem Gen im Hoden keinen funktionellen Rezeptor gibt. Die Regulation der OR Gene ist noch völlig ungeklärt. Es wurde aber bereits gezeigt, daß im MOE von Vertebraten ein Neuron meist nur eines der 1000 OR Gene exprimiert (Ngai et al., 1993; Ressler et al., 1993; Barth et al., 1997; Malnic et al., 1999). Ob dies in anderen Organen auch der Fall ist, bleibt noch zu erforschen. Zu erwähnen ist in diesem Kontext auch noch, daß andere Proteine mit sieben Transmembran-Regionen nur an der Oberfläche exprimiert werden, wenn die cytoplasmatische Domäne 1 vorhanden ist und die richtigen Motive aufweist (Schülein et al., 1996; Younger et al., 2001), d.h. eine vorzeitige Polyadenylierung könnte die dafür notwendigen Sequenzbereiche eliminieren.

Es wurde bereits erwähnt, daß sowohl die Transkripte aus *Bacillus thuringiensis* in der Kartoffel bzw. Tabakpflanze als auch die vorzeitig polyadenylierten humanen HLA-gekoppelten OR Transkripte keine Polyadenylierungssignale aufwiesen. Eine Analyse von

157775 humanen polyadenylierten EST's hat ergeben, daß nur 73% dieser Sequenzabschnitte das übliche AAUAAA- oder das seltenere AUUAAA-Signal aufweisen (Beaudoing et al., 2000). Bei zahlreichen EST's wurden Varianten der üblichen AAUAAA Sequenz gefunden, die sich nur in einer Base unterscheiden. Interessanterweise zeigen die Sequenzen, die zwei oder mehr Polyadenylierungsstellen besitzen, in dem Polyadenylierungssignal proximal zur kodierenden Region eine größere Varianz als in dem Signal, das am weitesten im 3' Bereich lokalisiert ist. Letzteres tendiert eher dazu, die übliche AAUAAA Sequenz zu besitzen (Beaudoing et al., 2000). Insofern ist es nicht ungewöhnlich, daß bei den in dieser Arbeit gefundenen Transkripten keine Polyadenylierungssignale gefunden wurden, da die Poly-A Bereiche weit innerhalb der kodierenden Region lagen, und sie somit eine größere Variabilität in diesen Bereichen zeigen könnten als die üblichen Signale. Außerdem wäre ein allgemein übliches Signal nicht sinnvoll, da es vermutlich grundsätzlich zu einem Abbruch bei der Transkription der kodierenden Sequenz führen würde und somit dieses Gen nie funktionsfähig wäre. Man kann also darüber spekulieren, ob in der kodierenden Sequenz der OR Gene durch einzelne Basenaustausche abgeschwächte Polyadenylierungssignale vorhanden sind, die je nach Organ oder Entwicklungsstadium zu einer vorzeitigen Polyadenylierung führen und das Gen damit ausschalten.

5.4.2. Genomische Organisation der 5' UTR von HLA-gekoppelten OR Genen

Im Rahmen dieser Arbeit wurden Transkripte von 13 OR Genen gefunden (insgesamt wurden 19 Gene analysiert). Zwölf dieser cDNAs wurden im Hoden nachgewiesen (s. Abschnitt 4.3.2.3.). Bei fünf HLA-gekoppelten Genen weisen diese Transkripte zusätzliche Exons im 5' Bereich auf: *hs6M1-12*, *-16*, *-18*, *-21* und *-27*. Alle Exons in der 5' UTR sind nicht kodierend. Bis auf *hs6M1-12* liegen diese Gene in einem gemeinsamen genomischen Bereich, der ca. 110 kb umfaßt und ungefähr 340 bis 450 kb von HLA-F entfernt liegt (s. Abb. 17 und Abb. 4). Die Anzahl der zusätzlichen nicht kodierenden Exons schwankt bei den jeweiligen Genen zwischen eins und vier. Das bedeutet, daß diese Gene alternativ gespleißt werden. Alle HLA-gekoppelten OR Gene, von denen zusätzliche 5' UTR Exons gefunden wurden, weisen zwei bis drei verschiedene Spleißvarianten auf (s. Abb. 17). Bei *hs6M1-16* wurde außerdem noch Spleißing innerhalb eines 5' Exons nachgewiesen (s. Abschnitt 4.3.2.3). Alternatives Spleißing in der 5' UTR scheint relativ häufig bei HLA-gekoppelten OR Genen vorzukommen, was man deutlich anhand der Autoradiographie der 5' RACE-cDNA-Produkte erkennen kann (s. Abb. 15). Besonders im Hoden zeigen bis auf eine Ausnahme (*hs6M1-6*) alle analysierten Gene mehrere unterschiedlich große Banden, die darauf schließen lassen, daß es von jedem Gen unterschiedliche Spleißvarianten gibt. In Lunge und Niere dagegen wurden auch Transkripte gefunden, die nur eine konkrete Größe aufwiesen (s. Abb. 15). Im Hoden werden gegenüber anderen Organen generell besonders viele Gene

und diese meist in verschiedenen Spleißvarianten exprimiert (Walker et al., 1999). Es wird vermutet, daß der Hoden diese vielen Proteine benötigt, um die zahlreichen Reifungsschritte bei der Spermatogenese zu regulieren (Walker et al., 1999). Eine detaillierte Hypothese gibt es jedoch dafür noch nicht.

Alternatives Spleißing in der 5' UTR von OR Genen ist bereits gezeigt worden. Asai und Mitarbeiter haben Transkripte mit unterschiedlichen 5' Exons in verschiedenen Geweben der Maus nachgewiesen (Asai et al., 1996), während Sosinsky und Mitarbeiter verschieden gespleißte Isoformen im olfaktorischen Epithel des Menschen gefunden haben (Sosinsky et al., 2000). Auch das Spleißen innerhalb eines Exons ist bereits beschrieben worden (Sosinsky et al., 2000). Im Vergleich zu den bisher veröffentlichten Ergebnissen (Glusman et al., 1996; Asai et al., 1996; Walensky et al., 1998; Qasba and Reed, 1998; Sosinsky et al., 2000) ist allerdings die maximale Distanz zwischen dem kodierenden Exon und dem ersten 5' nicht kodierenden Exon in der hier vorliegenden Untersuchung fast um das zehnfache größer. Für das *hs6M1-21* Gen z.B. wurde eine Distanz von ca. 100 kb ermittelt (s. Abb. 17). Alternatives Spleißing in der 5' UTR ist wahrscheinlich ein genereller Mechanismus, der der Zelle erlaubt, unterschiedlich regulierte Promotoren für dasselbe Gen zu verwenden. Eine EST Analyse einer menschlichen Gendatenbank (TIGR) hat ergeben, daß ein Drittel aller Gene alternatives Spleißing zeigt; wobei 80% der Gene in der 5' UTR gespleißt sind (Mironov et al., 1999).

Es war bisher nicht bekannt, daß verschiedene OR Gene identische 5' Exons besitzen können, wie es bei *hs6M1-21*, *-27* und *-18* der Fall ist (s. Abb. 17). Alle drei Gene besitzen ein gemeinsames erstes 5' Exon, das bei *hs6M1-21* ~100 kb, bei *hs6M1-27* ~80 kb und bei *hs6M1-18* ~30 kb stromaufwärts von der kodierenden Region liegt. *hs6M1-21* und *-27* weisen noch weitere identische Exons auf, die zwischen der kodierenden Region und dem ersten 5' nicht kodierenden Exon liegen (s. Abb. 17). Identische Exons von verschiedenen Genen wurden im humanen Genom auch bei den beiden Zinkfinger Genen PEG3 und ZIM2 nachgewiesen (Kim et al., 2000). ZIM2 besteht aus 11 Exons, von denen die ersten sieben stromaufwärts von PEG3 liegen. Diese sieben Exons sind identisch mit den ersten sieben Exons von PEG3. Das Phänomen, daß verschiedene Gene ein identisches erstes 5' Exon besitzen, deutet daraufhin, daß sie einen gemeinsamen Transkriptionsstartpunkt besitzen.

In der Nähe dieser drei überlappenden Gene (*hs6M1-21*, *-27* und *-18*) liegt auch *hs6M1-16*, das aber eine entgegengesetzte Transkriptionsrichtung aufweist. Auch dieses Gen weist zusätzliche Exons auf, die allerdings nur eine maximale Distanz von ~5 kb zum kodierenden Exon aufweisen. Das erste 5' Exon von *hs6M1-16* liegt nur 80 bp entfernt vom gemeinsamen ersten 5' Exon der anderen drei Gene (s. Abb. 17). Damit ist die Region, in der sich der Promotor vermutlich befindet, stark eingegrenzt. Da 80 bp kaum Platz bieten für zwei Promotoren, kann man darüber spekulieren, ob an dieser Stelle ein bidirektionaler

Promotor lokalisiert sein könnte, der sowohl die drei überlappenden Gene in die eine Transkriptionsrichtung als auch *hs6M1-16* in die entgegengesetzte Transkriptionsrichtung kontrolliert. Es gibt zahlreiche Beispiele für bidirektionale Promotoren im menschlichen Genom (Li and Seetharam, 1998; Orii et al., 1999; Bennett et al., 1999; Lee and Song, 2000; Dong et al., 2000), aber es wäre eine absolut neue Erkenntnis, wenn bidirektionale Promotoren mehr als zwei Gene kontrollieren könnten.

Die genomische Struktur von *hs6M1-21*, *-27* und *-18* besitzt Ähnlichkeiten zum Operon Modell von Jacob und Monod bei Prokaryonten (Jacob and Monod, 1961). Auch hier werden mehrere Strukturgene durch nur einen Promotor kontrolliert. Allerdings werden bei diesem Modell alle Strukturgene gleichzeitig in einem einzigen mRNA Molekül hintereinander transkribiert. Bei den drei überlappenden HLA-gekoppelten OR Genen ist die Situation vermutlich sehr viel komplexer. Eine hintereinandergeschaltete Transkription erscheint allerdings unwahrscheinlich, weil bei einer Transkription von *hs6M1-21* gleichzeitig die Gene *hs6M1-27* und *-18* herausgespleißt werden müßten. Die Entscheidung, welches der drei Gene exprimiert wird, kann in diesem Falle nicht auf Transkriptionsebene erfolgen, da eine Repression oder Induktion des gemeinsamen Promotors immer alle drei Gene gleichzeitig betreffen würde. Die Regulation wird vermutlich durch einen nachgeschalteten Prozeß erfolgen. Man kann darüber spekulieren, ob zunächst ein riesiges 100 kb umfassendes mRNA-Primärtranskript entsteht, das noch alle drei Gene *hs6M1-18*, *-21* und *-27* sowie *hs6M1-17*, *-19P* und *-20* (aber Antisinnstrang) enthält. Die Entscheidung, welches der drei Gene schließlich exprimiert wird, könnte dann über alternatives Spleißing erfolgen, bei dem jeweils die kodierende Region der anderen beiden Gene herausgespleißt würde, oder aber, wie bereits in Abschnitt 5.4.1 beschrieben, durch vorzeitige Polyadenylierung. Um die Regulation der Expression in diesem OR Gen Cluster verstehen zu können, müßte man auch die Frage beantworten, ob alle drei Gene in einer Zelle exprimiert werden oder ob vielleicht nur ein Gen transkribiert wird, wie es bei den meisten Neuronen im MOE der Fall ist.

5.4.3. Genomische Organisation der 3' UTR und der kodierenden Region von HLA-gekoppelten OR Genen

Von *hs6M1-21* konnten keine Transkripte gefunden werden, die die gesamte kodierende Region einschließlich der 3' UTR umfassen (s. Abschnitt 4.3.3). Es wurden aber zahlreiche kurze Transkripte von diesem Gen gefunden, die nur den vorderen Teil der kodierenden Region beinhalten. Daher ist es wahrscheinlich, daß von diesem Gen lediglich verstümmelte mRNA-Fragmente existieren. In Abschnitt 5.4.1 wurde bereits diskutiert, daß die frühzeitige Polyadenylierung von Genen einen Regulationsmechanismus darstellen könnte, um Gene auszuschalten. Es wäre also denkbar, daß *hs6M1-21* im Hoden nicht benötigt wird.

Von *hs6M1-16*, *-18* und *-27* dagegen konnten cDNA Klone, die die vollständige kodierende

Region einschließlich der 3' UTR enthielten, identifiziert werden. Während bei *hs6M1-18* und *-27* keine Introns in der 3' UTR nachgewiesen wurden, ist dieser Bereich bei *hs6M1-16* gespleißt (s. Abschnitt 4.3.3. und Abb. 24) und weist drei verschiedene Spleißvarianten mit maximal zwei zusätzlichen Exons auf. Generell sind Introns in der 3' UTR selten. Eine Studie mit EST Daten aus dem "TIGR Human Gene Index" ergab, daß bei den Genen, die alternatives Spleißing zeigen, nur 20% in der 3' UTR Introns aufweisen (Mironov et al., 1999). Erstmals wird mit den Ergebnissen hier bewiesen, daß ORs - oder G-gekoppelte Proteine im allgemeinen – einen gespleißten 3' flankierenden Bereich besitzen können.

Ebenfalls zum ersten Mal wurde nachgewiesen, daß OR Gene von Säugetieren innerhalb der kodierenden Region gespleißt sein können (s. Abschnitt 4.3.3., Abb. 23 und Abb. 24). Derartige Phänomene für ORs waren bisher nur bei *Caenorhabditis elegans* und *Drosophila melanogaster* bekannt (Dryer, 2000). Von allen höheren Tieren wie z.B. Neunaugen, Knochenfischen und Säugetieren sind bislang nur intronlose kodierende OR Sequenzen veröffentlicht (Dryer, 2000). Von *hs6M1-27* wurden zwei verschiedene Isoformen bezüglich der kodierenden Sequenz gefunden. Die eine Isoform beinhaltet die komplette kodierende Region, während die andere Form in dieser Region durch ein 272 bp großes Intron unterbrochen ist (s. Abb. 23). Die Sequenz nach dem Intron ist nicht mehr *in frame*, so daß ein eventuell resultierendes Protein maximal drei TM Domänen besitzen könnte und damit vermutlich nicht funktionsfähig ist.

Von *hs6M1-16* wurden fünf verschiedene Spleißvarianten gefunden (s. Abb. 24). Zwei davon beinhalten die komplette kodierende Region, unterscheiden sich aber in der 3' UTR. Zwei andere Transkripte, die sich ebenfalls nur in der 3' UTR unterscheiden, beginnen in der kodierenden Region exakt erst an Aminosäure Position 79, während das fünfte Transkript stromaufwärts von dieser Aminosäure (79) 13 weitere Basenpaare besitzt. Das Codon 79 (s.o.) kodiert für ein Methionin und könnte bei den Transkripten, bei denen die ersten 234 bzw. 221 bp der kodierenden Region fehlen, als alternatives Start-Codon für einen OR ohne die ersten zwei TM Domänen dienen. Interessanterweise würde dieser um die ersten 78 Aminosäuren verkürzte OR wieder mit einer extrazellulären Domäne beginnen. Anhand mutierter Chemokin-Rezeptoren wurde bereits nachgewiesen, daß G-gekoppelte 7-TM Rezeptoren auch funktionstüchtig sein können, wenn ihnen die ersten beiden TM Domänen fehlen (Ling et al., 1999). Bei 61% der HLA-gekoppelten OR's wurde ein Methionin an AS Position 79 nachgewiesen (Younger et al., 2001). Neun dieser Gene (*hs6M1-2P*, *-7P*, *-8P*, *-9P*, *-15*, *-16*, *-21*, *-22P*, *-24P*) besitzen funktionelle Spleißstellen an den entsprechenden Sequenzabschnitten, die eine Expression von diesem alternativen Start-Codon aus ermöglichen würden. Es wird daher vermutet, daß auch weitere ORs dieses Phänomen aufweisen könnten. Durch die Deletion der ersten beiden Domänen würden z.B. auch die Rastermutationen von *hs6M1-7P* und *-22P* eliminiert werden, womit diese beiden

Pseudogene als funktionstüchtige Proteine exprimiert werden könnten (Younger et al., 2001).

Aus diesen Ergebnissen folgt direkt die Frage: Erfüllen diese Spleißing-Ereignisse einen biologischen Sinn? Im Falle von *hs6M1-16* könnte man spekulieren, daß das potentiell funktionstüchtige, aus fünf TM Domänen bestehende Protein eine andere Ligandenaffinität aufweisen könnte als das 7 TM Protein, wodurch die Rezeptorvielfalt noch weiter vergrößert werden würde. Beim *hs6M1-27* Gen dagegen sieht die Situation gänzlich anders aus. Hier führt das gespleißte Produkt zu einem aus drei TM Domänen bestehenden Rezeptor, dem die letzten vier TM Domänen fehlen. Es ist nicht anzunehmen, daß dieses Produkt funktionstüchtig ist, da vermutlich die Bindungstasche für Liganden fehlt (Buck and Axel, 1991). Man kann darüber spekulieren, ob dies ein weiterer Mechanismus ist, um die Expression von Genen zu kontrollieren, wie es z.B. in Abschnitt 5.4.1 und 5.4.2 diskutiert wurde. Eine weitere Möglichkeit wäre jedoch, daß sich die verstümmelten Proteine mit anderen verstümmelten Rezeptor-Fragmenten verbinden. Auf diese Weise könnte ein Mechanismus entwickelt worden sein, der die Rezeptorvielfalt enorm vergrößert, da immer andere Kombinationen möglich sind. In diesem Falle würden auch die zahlreichen Pseudogene im humanen Genom eine völlig neue Bedeutung bekommen. Pseudogene wie z.B. *hs6M1-4P*, bei dem ein Stop-Codon in der dritten extrazellulären Domäne in einigen Allelen zu verkürzten Transkripten führt, könnten durch ein Rearrangement mit anderen Rezeptor-Bruchstücken wieder funktionsfähig werden. Eine Parallele zu dieser Hypothese könnte man vielleicht zur Bildung von Antikörpern ziehen. Im Genom von Maus und Mensch gibt es zahlreiche unterschiedliche Gensegmente für die variablen Regionen eines Antikörpers. Durch zufällige somatische Rekombination werden die verschiedenen V-, D-, und J-Gensegmente in den einzelnen Lymphozyten jeweils anders zu einer vollständigen V-Domäne zusammengesetzt, wodurch eine geradezu unglaubliche Variabilität bei den Produkten dieser Gene entsteht. Im Gegensatz zu der oben beschriebenen Hypothese handelt es sich hier allerdings um eine Art Spleißereignis auf DNA Ebene (Smith et al., 1971; Honjo et al., 1974; Tonegawa et al., 1974; Janeway and Travers, 1997). Rearrangements dieser Art konnten auf genomischer Ebene für OR Gene bisher nicht nachgewiesen werden.

Die Hypothese, daß es eine Kombinatorik bei Proteinfragmenten gibt, ist analog zur Kombinatorik bei der Bildung der Antikörper auf DNA Ebene. In beiden Fällen würde aus einer relativ kleinen Anzahl von Genen eine große Anzahl an Proteinen erzeugt werden. Die neuesten Erkenntnisse des Humanen Genom Projektes ergeben, daß die Anzahl der Gene viel geringer ist als bisher angenommen (Baltimore, 2001). Daher könnte ein einzelnes Gen für zahlreiche verschiedene Proteine kodieren. In diesem Kontext wäre es interessant, wenn verschiedene Kombinationen von OR-Protein-Fragmenten zu einer Vielzahl von funktionellen OR-Proteinen führen.

5.5. Ausblick

Die Ergebnisse dieser Arbeit zeigen, daß alle analysierten HLA-gekoppelten OR Gene polymorph sind. Somit ist die Grundvoraussetzung für eine individuelle Geruchspräferenz bei der HLA-abhängigen Partnerwahl gegeben. Es sind allerdings noch mehr Arbeiten notwendig, um einen funktionellen Zusammenhang zwischen HLA-Komplex und HLA-gekoppelten OR Genen zu beweisen. Zunächst müßte eine größere Anzahl Individuen untersucht werden, um eine statistisch signifikante Aussage darüber zu erhalten, ob eine gehäufte Kopplung zwischen bestimmten HLA- und OR-Haplotypen besteht. Eine häufig vorkommende HLA/OR-Haplotyp Kombination könnte im Laufe der Evolution entstanden sein, weil sie dazu beigetragen hat, daß man in Abhängigkeit von seinem eigenen HLA-Typ eine Geruchspräferenz für einen HLA-unähnlichen Partner ausbildet. Dies führt dann letztendlich zu einem selektiven Vorteil bei der Fortpflanzung. In diesem Kontext sind in Zusammenarbeit mit dem Fertility Center Berlin unter der Leitung von Prof. Dr. Kentenich bereits Blutproben von Paaren gesammelt worden, die sich einer künstlichen Befruchtung unterzogen hatten. In einer Kooperation des Instituts für Immunogenetik und dem Department of Pathology in Cambridge soll von diesen Paaren eine HLA- und OR-Haplotyp Bestimmung erfolgen, um eine Korrelation zwischen HLA/OR-Typ und einer nicht erfolgreichen Befruchtung oder eines erfolgten Abortes nachzuweisen. Paare, die nicht erfolgreich bei der Fortpflanzung waren, könnten eine HLA/OR-Haplotyp Kombination aufweisen, die zu einem selektiven Nachteil bei der Fortpflanzung geführt hat. Die Ergebnisse dieser Analysen könnten vielleicht bei *in vitro* Befruchtungen berücksichtigt werden.

In dieser Arbeit wurde nachgewiesen, daß HLA-gekoppelte OR Gene in vielen Organen exprimiert werden. Es wäre äußerst interessant zu erfahren, welche Funktionen sie dort erfüllen. Um diese Frage zu beantworten, müßte man aber zunächst die Liganden dieser Rezeptoren identifizieren. Bisher hat man nur Liganden für OR in der Nase nachgewiesen; welcher Art diese Moleküle in anderen Geweben sein könnten, ist bisher noch völlig offen. Bei der Vielzahl der Organe, in denen diese Rezeptoren exprimiert werden, und bei der großen Anzahl an OR im humanen Genom, wird aber die vollständige Beantwortung dieser Frage noch einige Jahre in Anspruch nehmen.

Die Tatsache, daß verschiedene OR Gene gemeinsame erste 5' Exons besitzen, läßt vermuten, daß sie über einen gemeinsamen Promotor reguliert werden. Um dies eindeutig zu beweisen, sind allerdings noch weitere Experimente notwendig. Erstens müßten *in silico* Promotor Analysen erfolgen, um festzustellen, ob vielleicht bekannte Promotorsequenzen in dieser Region auftreten. Zweitens könnte man die Expression dieser Gene in eukaryotischen Zelllinien testen. Dafür müßte allerdings ein Konstrukt erstellt werden, das die entsprechende Region – die kodierenden Bereiche der drei Gene einschließlich der 5' UTR mit dem ersten gemeinsamen Exon – enthält. Da diese Region aber ca. 100 kb umfaßt, ist die Erstellung

eines derartigen Konstruktes nicht mit den üblichen Methoden (PCR, Restriktionsverdau) möglich. Man müßte alternative Methoden finden, um ein derartiges Konstrukt zu erstellen. Sollte die Expression mit Hilfe eines derartigen riesigen Konstruktes gelingen, wäre es z.B. interessant zu erfahren, ob die drei überlappenden Gene (*hs6M1-21*, *-27* und *-18*) gemeinsam in einer Zelle exprimiert werden, wie es bereits für andere Gene in einzelnen olfaktorischen Neuronen nachgewiesen wurde (Rawson et al., 2000), oder ob nur jeweils einer dieser Rezeptoren präsent ist. Diese Frage könnte man dann mit Hilfe von Einzelzellen PCR beantworten.