Chapter 7

Polymorphism of human MHC-linked ORs.

7.1. Introduction

Within the human species, both the ability to sense various chemicals and how these chemicals are perceived varies widely. The chemical androsterone, for example, cannot be smelt by some individuals, whilst for those who can smell it, the smell is either considered to be similar to urine or sandalwood. Although this difference in perception may be caused by non-genetic factors (for example, a traumatic head injury or an infection of the nasal mucosa can produce damage leading to anosmias), studies of differences in the perception of androsterone and other odorants between fraternal and identical twins have indicated that there is a genetic component of this observed variation (Wysocki and Beauchamp, 1984, Wysocki *et al.*, 1989, Gross-Isseroff *et al.*, 1992). This genetic component of differences in olfactory perception could be located at a number of steps within the olfactory pathway: for example, odorant binding proteins appear to be involved in facilitating the transfer of odorants across the mucus layer to olfactory receptors (Tegoni *et al.*, 2000), so allelic differences in these genes may play a role in differing perceptions of odorants. Similarly, allelic variations in G-proteins (Rana *et al.*, 2001) which are coupled to olfactory receptors and trigger an increase in cyclic AMP when an odorant binds to an olfactory receptor may also be important in individual's differing olfactory abilities.

Within the olfactory system as it is currently understood, however, the large number of olfactory receptor genes within the human genome compared to the number of odorant binding proteins or G-proteins implicated in the olfactory system suggests that the majority of differences in smelling ability caused by genetic factors can be predicted to be due to genetic variations within the

repertoire of olfactory receptor genes. This genetic variation within the olfactory receptor family appears to be able to take 2 forms: firstly, there may be variation within the sequence of olfactory receptors (Gilad *et al.*, 2000, Sharon *et al.*, 2000), or secondly, there may be variation in the number of olfactory receptor genes at certain chromosomal sites within the genome (Trask *et al.*, 1998, Linardopoulou *et al.*, 2001).

In order to assess genetic variation that could potentially have a role in some anosmias, two MHC-linked olfactory receptor genes (hs6M1-17 and hs6M1-20) were resequenced in 10 different cell lines. These cell lines were derived from different donors, representing different HLA haplotypes and different ethnic origins. Eight of the 10 cell lines were HLA homozygous, whereas two (BM19.7, BM28.7) were HLA hemizygous (Ziegler *et al.*, 1985, Volz *et al.*, 1992). Other MHC-linked ORs were resequenced by collaborators in Berlin (Anke Ehlers and Armin Volz) and in Cambridge (Simon Forbes).

7.2. Alleles of hs6M1-17

For hs6M1-17, sequence variations were observed at 10 positions within the gene (Table 7.1). Of these 10 variations, all but 2 are predicted to affect the protein sequence of the OR. The most notable variation is at amino acid position 55 where one cell line (BM19.7 (East African)) has a stop codon as opposed to the functional CAG codons found at this position in the other 10 DNA samples surveyed (including DNA from the Human Genome Project). This stop codon, which can be predicted to make the OR gene non-functional, may explain why there have been a number of other changes within this allele.

Allele	1	2	3	4	5
DNA 163 AA 55	CAG Glu	CAG Glu	CAG Glu	CAG Glu	TAG stop
DNA 183 AA 61	TTC Phe	TTC Phe	TTC Phe	TTC Phe	TTT Phe
DNA 265 AA 89	CGC Arg	CGC Arg	CGC Arg	CGC Arg	AGC Ser
DNA 361 AA 121	CGC Arg	CGC Arg	CGC Arg	CGC Arg	TGC Cys
DNA 412 AA 138	CGG Arg	CGG Arg	CGG Arg	CGG Arg	TGG Try
DNA 478 AA 160	CCT Pro	CCT Pro	CCT Pro	TCT Ser	TCT Ser
DNA 521 AA 174	CCG Pro	CCG Pro	CAG Glu	CCG Pro	CCG Pro
DNA 736 AA 246	GTG Val	ATG Met	ATG Met	ATG Met	GTG Val
DNA 762 AA 254	GCA Ala	GCA Ala	GCA Ala	GCA Ala	GCC Ala
DNA 929 AA 310	ATG Met	ATG Met	ATG Met	ATG Met	AGG Arg
Cell lines	BM 28.7 LG2 Genomic	SA H2LCL WT51 YAR	KR3598 OLGA	AMAI	BM19.7

Table 7.1: Polymorphisms observed in hs6M1-17. The differences in the DNA sequences and the amino acid found in the OR protein of the 5 alleles are listed, alongside cell lines found to carry a specific allele.

In addition to the change at amino acid (AA) 55, 5 changes that are only found within this allele can be observed (AA 89: Arg → Ser, AA 61 Phe → Phe (synonymous mutation), AA 121 Arg → Cys, AA 138 Arg → Try, AA 310 Met → Arg). In contrast to the variation that has been generated within this allele, comparing the other 4 alleles only 3 nonsynonymous mutations can be observed (AA 160 Pro → Ser, AA 174 Pro → Glu, AA position 246 Val → Met). The high variation in the allele containing the stop codon compared to the other alleles of hs6M1-17

suggests that selectional forces conserving the structure of OR genes acted less strongly to maintain the DNA sequence of the BM19.7 allele after the mutation at DNA position 163 rendered the allele non-functional. This idea, that the selectional forces acting to conserve the amino acids of the OR protein have been relaxed since the stop codon mutation partially mediates against the hypothesis that OR pseudogenes or fragments or these pseudogenes play a functional role within the human genome (Chapter 6).

7.3. Alleles of hs6M1-20

Resequencing hs6M1-20 in the 10 individuals led to 6 different combinations of alleles being found (Table 7.2). 8 substitutions at the DNA level were observed. These consisted of 1 silent mutation at amino acid position 255 and 7 changes predicted to code for different amino acids. 3 of these changes appear to be nonpolar amino acid for nonpolar amino acid (Val → Phe, Phe → Leu, Val → Ile), whilst the other 4 changes can be predicted to have more of an impact upon the protein structure (Leu → Pro, Phe → Ser, Leu → Arg, Ser→ Cys).

In 3 samples, this gene appears to be heterozygous. The DNA sample from population KR3598, for example, has 2 alleles that differ at DNA position 362. The difference between the alleles in the SA and OLGA cell lines is even greater: they differ at 5 and 4 positions respectively.

Allelic combination	1	2	3	4	5	6
DNA 139 AA 47	GTC Val	GTC Val	TTC Phe	TTC Phe	GTC Val	GTC Val
DNA 167 AA 56	CTT Leu	CTT Leu	CCT Pro	CCT Pro	CTT/CCT Leu/Pro	CTT/CCT Leu/Pro
DNA 311 AA 104	TTC Phe	TCC Ser	TTC Phe	TTC Phe	TTC Phe	TTC Phe
DNA 339 AA 113	TTC Phe	TTC Phe	TTG Leu	TTG Leu	TTC/TTG Phe/Leu	TTC/TTG Phe/Leu
DNA 359 AA 120	CTC Leu	CTC Leu	CGC Arg	CGC Arg	CTC/CGC Leu/Arg	CTC/CGC Leu/Arg
DNA 362 AA 121	TCT Ser	TCT Ser	TGT Cys	TGT/TCT Cys/Ser	TGT/TCT Cys/Ser	TGT/TCT Cys/Ser
DNA 475 AA 159	GTA Val	GTA Val	ATA Ile	ATA Ile	GTA/ATA Val/Ile	ATA Ile
DNA 765 AA 255	CTT Leu	CTT Leu	CTC Leu	CTC Leu	CTC Leu	CTC Leu
Cell lines	BM28.7 LG2 AMAI Genomic	WT51	BM19.7 H2LCL YAR	KR3598	SA	OLGA

Table 7.2: Polymorphisms observed in hs6M1-20. The differences in the sequences of the 5 alleles are listed, alongside cell lines found to carry a specific allele.

7.3. Alleles of other MHC-linked OR genes.

The polymorphisms observed in hs6M1-17 and hs6M1-20 were compared against other polymorphisms in MHC-linked OR genes (data generated by Anke Ehlers and Armin Volz, Berlin, and Simon Forbes, Cambridge, summarised in table 7.3 (Ehlers *et al.*, 2000, Ziegler *et al.*, 2000)). In all 52 point mutations were detected. On the nucleotide level, the majority of these changes are transitions (C → T, A → G) rather than transversions (C → A, C → G, G → T, T → A). Contrary to what might be expected, however, these point mutations appear to be largely

equally distributed throughout codon positions: in fact there are slightly more mutations that alter the first nucleotide of a codon than the other 2 nucleotides within a codon. The apparent lack of a selective pressure producing more mutations in the third and second nucleotide positions than in the first nucleotide position means that the majority of nucleotide mutations produce nonsynonymous changes within the OR protein.

Hs6M1-4 is similar to hs6M1-17 in having a functional and a non-functional allele found within different cell lines. (Non-functional alleles are both disrupted by a stop codon.) In contrast to hs6M1-17, however, the non-functional allele has not amassed a number of mutations that are not found in the other alleles of this gene. This suggests the selective pressure is stronger on hs6M1-4. Reasons for this include the idea that this could be a more recent mutation that has not been around for a long enough period of time to accumulate the same number of mutations as hs6M1-17. Alternatively, the pseudogene allele of hs6M1-4P could be functional outside the olfactory system whereas the pseudogene allele of hs6M1-17 could be totally non-functional.

Although, not included in table 7.3, functional and non-functional alleles were observed in hs6M1-19P. At this locus in addition to the pseudogene form found in the genomic sequence, there appears to be a functional form without the 16 base pair deletion that renders hs6M1-19 a pseudogene.

Across the cluster, the number of alleles of MHC-linked OR genes ranges from 2 (hs6M1-1, hs6M1-10, hs6M1-18) up to as many as 7 (hs6M1-17, hs6M1-20), although the average is 3-4. It is interesting to note that members of the same subfamily appear to contain a similar number of point mutations, for example, hs6M1-1 and hs6M1-10 both have 1 point mutation, whilst hs6M1-12 and hs6M1-16 have 4 and 3 respectively. This may reflect similar evolutionary pressures

acting on members of a subfamily, although these events occur at different amino acid positions in the 2 proteins in both cases.

Name	No. of point mutations	No. of alleles	Position in consensus seq.	AA change	DNA change	Codon position
hs6M1-1	1	2	105	Leu → Leu	**A → G**	3
hs6M1-3	4	4	108	Ala → Thr	**G → A**	1
			221	Gln → Arg	**A → G**	2
			223	Val → Ile	**G → A**	1
			256	Ile → Met	**A → G**	3
hs6M1-6	6	3	71	Tyr → His	**T → C**	1
			108	Ala → Thr	**G → A**	1
			117	Ser → Ser	**G → A**	3
			143	Val → Ala	**T → C**	2
			211	Leu → Leu	C → G	1
			215	Ala → Thr	**G → A**	3
hs6M1-10	1	2	232	Gln → Arg	**A → G**	2
hs6M1-12	5	4	19	Pro → Pro	**A → G**	3
			29	Phe → Leu	**T → C**	1
			37	Leu → Leu	**A → G**	3
			48	Ala → Val	**C → T**	2
			78	Gln → Gln	**A → G**	3
hs6M1-15	3	3	79	Met → Val	**A → G**	1
			277	Thr → Thr	**C → T**	3
			294	Asp → Asn	**G → A**	1
hs6M1-16	3	3	62	Ser → Ser	**C → T**	3
			63	Asn → Asp	**A → G**	1
			208	Pro → Pro	**C → T**	3
hs6M1-18	1	2	162	Ala → Thr	**G → A**	1

Name	No. of point mutations	No. of alleles	Position in consensus seq.	AA change	DNA change	Codon position
hs6M1-20	8	7	47	Val → Phe	G → T	1
			56	Leu → Pro	**T → C**	2
			104	Phe → Ser	C → G	2
			113	Phe → Leu	C → G	3
			120	Leu → Arg	T → G	2
			121	Ser → Cys	C → G	2
			159	Val → Ile	**G → A**	1
			254	Leu → Leu	**T → C**	3
hs6M1-21	4	4	21	Leu → Trp	T → G	2
			104	Phe → Phe	**C → T**	3
			231	Gly → Arg	**G → A**	1
			236	Phe → Phe	**T → C**	3
hs6M1-4	6	5	11	Ile → Leu	A → C	1
			81	Val → Val	**G → C**	3
			96	Thr → Thr	A → G	3
			179	*Val → Ala*	***C → T***	*2*
			193	*Gln → Stop*	***C → T***	*1*
			203	Ile → Ile	T → A	3
hs6M1-17	10	7	*54*	*Gln → Stop*	***C → T***	*1*
			60	*Phe → Phe*	***T → C***	*3*
			88	*Arg → Ser*	*C → A*	*1*
			120	*Arg → Cys*	***C → T***	*1*
			137	*Arg → Trp*	***C → T***	*1*
			159	Pro → Ser	**C → T**	1
			173	Pro → Gln	C → A	2
			245	Val → Met	**G → A**	1
			253	*Ala → Ala*	*A → C*	*3*
			309	*Met → Arg*	*T → G*	*2*

Table 7.3: Summary of all polymorphisms found within MHC-linked OR genes. Transitions are indicated in bold, whilst italics are used to indicate where changes are only observed in the non-functional allele of the 2 loci, hs6M1-4 and hs6M1-17.

There are, however, 4 pairs of olfactory receptor genes where polymorphisms are found at the same position relative to the consensus protein sequence (see Chapter 4). At consensus sequence position 104, for example, hs6M1-20 and hs6M1-21 both show polymorphisms (although in –21 it is silent, whilst in –20 a phenylalanine becomes a serine). Other shared positions for polymorphisms include 108, where both hs6M1-3 and hs6M1-6 have a G to A transition which changes an alanine into a threonine. Hs6M1-17 and hs6M1-20 have 2 positions in common where polymorphisms are located, although the changes (at position 120, Leu $\rightarrow$ Arg and Arg $\rightarrow$ Cys; at position 159, Val $\rightarrow$ Ile and Pro $\rightarrow$ Ser) involve different nucleotide changes.

The distribution of the polymorphic amino acid sites is shown in Figure 7.1. From this it can be seen that the largest number of polymorphisms are found within the first half of the olfactory receptor protein consensus sequence. 5 regions show a significant amount of polymorphism: cytoplasmic region 1 , transmembrane region 2, extracellular region 2 and transmembrane region 3 (which are all located next to each other), and cytoplasmic region 3. These results are surprising in the light of the conservation profile of the MHC-linked OR proteins (Chapter 4): transmembrane region 2 was found to be highly conserved in this conservation profile, and so it might be expected to show a lower percentage of polymorphic sites.

7.4. Single nucleotide polymorphism (SNP) large scale analysis

Data about single nucleotide polymorphisms (SNPs) were extracted from the Ensembl database (Chapter 2) and mapped onto the detailed plots of the major and minor human MHC-linked OR clusters. This revealed a total of 561 SNPs within the major MHC-linked OR cluster (561 per 718800 bp = density of 1 SNP per 1281 bp) and 207 SNPs within the minor MHC-linked OR cluster (207 per 200000 bp = density of 1 SNP per 966 bp). These figures are

Figure 7.1a

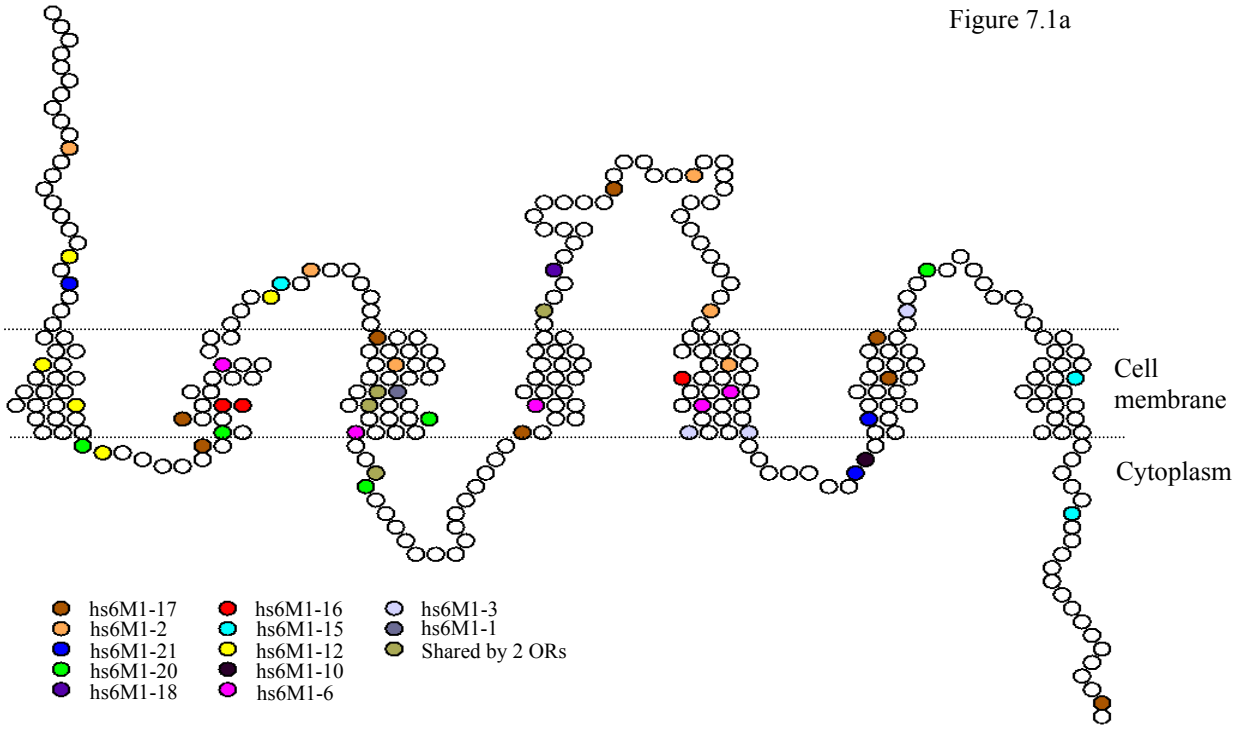

Figure7.1b

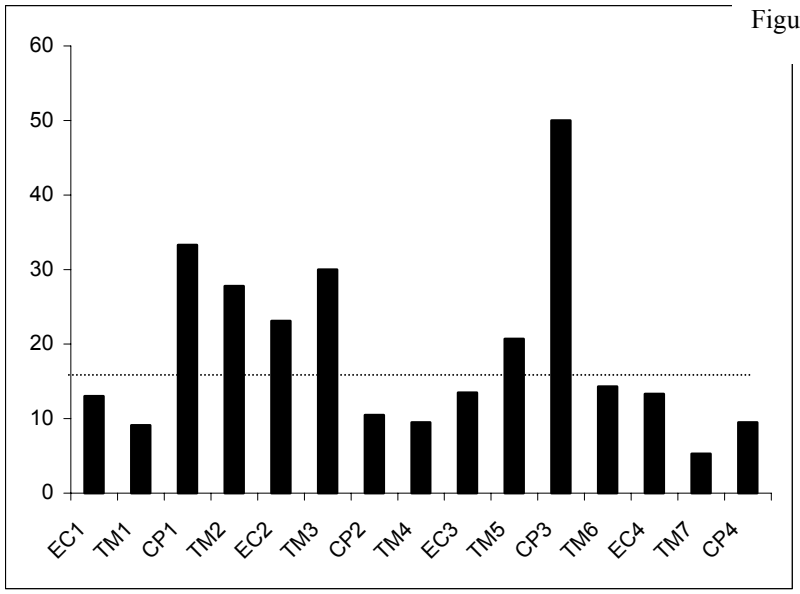

Figure 7.1a: Polymorphisms found in human MHC-linked ORs. The positions of the polymorphisms are displayed with reference to the position of the polymorphism in the consensus protein sequence. The polymorphisms found in different genes are indicated by different coloured residues. The light green colour indicates the 4 positions where 2 ORs both have a polymorphism (details in text).
Figure 7.1b: Percentage of polymorphisms per regions of the consensus MHC-linked OR protein. This shows the number of polymorphic residues with respect to the number of other amino acid residues within the defined section of the protein. The dotted line shows the average figure found across the protein of 16.9%.

slightly higher then the average reported figure of 1 SNP per 1910 bp for the human genome (Sachidanandam *et al.*, 2001), but this difference is likely to reflect the steady accumulation of SNP data since publication of this paper rather than a higher rate of SNPs per base pairs in these regions of the genome. This estimate of the number of SNPs in the human genome was, in any case, fairly conservative, since other studies have suggested the figure may be higher (1 SNP per 721 bp, generalised from 2 Mb of sequence tagged sites (Wang *et al.*, 1998), 1 SNP per 100-300 bp (dbSNP database, *http://www.ncbi.nlm.nih.gov/SNP/.*))

The position of SNPs within the region is shown in Figure 7.2. The distribution within the major OR cluster is striking in its inequality: SNPs appear to be concentrated in a 260 Kb region located at the centromeric end of the cluster, with far fewer SNPs located where the bulk of the olfactory receptor genes are found. In the minor cluster, the distinction in SNP frequency is less pronounced, but there do seem to be fewer SNP in the middle of the cluster. Whether these differences in the distribution of SNPs are significant is debatable, especially since the coverage of SNPs in the public database may represent a partial rather than a complete picture of SNPs within the human genome. With regard to the major cluster, however, it is interesting that there is much higher number of SNPs in the region nearest the MHC, one of the regions within the human genome with the highest amount of variation between individuals (Horton *et al.*, 1998).

The vast majority of these SNPs are located within non-coding sequence. In the major MHC-linked OR cluster, 507 (90.4%) are predicted to exist outside gene loci (both OR genes and other genes within the region), whilst a higher percentage of 95.4% are associated with pseudogene loci and non-coding sequence. The true percentage of SNPs not implicated with affecting coding sequence is likely to be between these 2 figures, since some of the olfactory genes seem to be coding in some individuals and pseudogenes in other individuals.

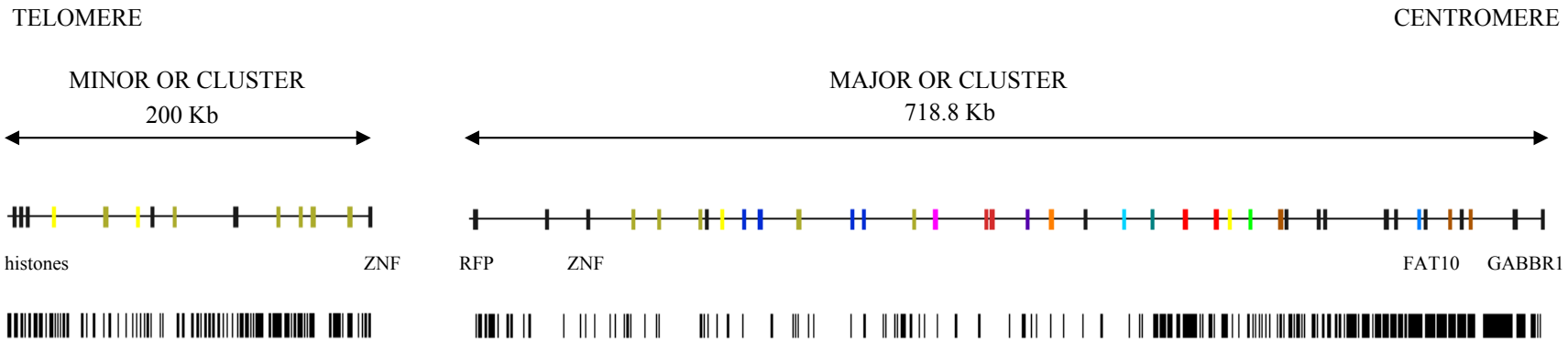

Figure 7.2: Distribution of single nucleotide polymorphisms (SNPs) across the major and minor MHC-linked OR clusters (extracted from the Ensembl database). The major and minor OR clusters were both analysed: OR genes are coloured according to their subfamily designation, or where they do not belong to a subfamily they are coloured pale green. The majority of genes are not labelled: for gene names refer to Figure 4.6. The track underneath the gene name shows the SNP distribution across the region. Within the major OR cluster, this distribution appears to be more dense towards the centromeric end of the cluster. This centromeric end is also the nearest end to the MHC, a well-characterised variable region within the human genome.

Within the minor cluster, 94.7% (196) of the SNPs are found outside gene loci (both OR genes and other genes within the region), the percentage rises to 97.6% (202) assuming pseudogenes are non-coding. Across the two clusters, 32 of the 768 SNPs are associated with olfactory receptor loci (both genes and pseudogenes). These SNPs were analysed further to see if they added to the allelic diversity already described or to see whether these SNPs can be confirmed by the data already described.

7.5. SNPs of MHC-linked ORs

Analysing the SNP data (Table 7.4) associated with the MHC-linked ORs suggested that the resequencing of many of the MHC-linked alleles had found the large majority of polymorphic sites within the OR genes. Generally, the resequencing strategy found more polymorphic sites than the SNP genome-wide approach, and SNPs detected were found at the same sites as those that had already been identified. For example, in hs6M1-4P, resequencing found 6 point mutations whereas the number of SNPs that were identified was 3. These 3 SNPs had already been uncovered by the resequencing approach.

Hs6M1-17 was the exception to this rule: 3 extra point mutations were identified using the SNP data, meaning this OR gene contains at least 13 point mutations. This high number of point mutations could be attributed to the pseudogene status of hs6M1-17 in some haplotypes. However, it appears that hs6M1-17 has a higher number of point mutations than many of the pseudogenes, as the 8 pseudogenes with identified SNPs have on average 1-2 point mutations. Taking hs6M1-17 as the model for how successfully the SNP data manages to identify point mutations, (assuming 13 mutations, 6 of which were identified by the SNP analysis), the number of point mutations per pseudogene can be estimated as 2-4 suggesting hs6M1-17 has a much higher mutation rate than some OR genes within the cluster.

Name	No. of point mutations	Position in consensus seq.	AA change	DNA change	Codon position
hs6M1-8P	1	130	Tyr → Tyr	T → C	3
hs6M1-35	2	253	Ile → Asn	T→ A	2
		305	Arg → Arg	A → G	3
hs6M1-13P	1	257	Tyr → Tyr	C → T	3
hs6M1-14	4	107	Ser → Ser	C → G	3
		149	Ser → Ser	T → C	3
		220	Ala → Ala	C → G	3
		225	Cys → Cys	C → T	3
hs6M1-32	1	144	Ala → Ala	T → C	3
hs6M1-22P	1	124	Ile → Ile	A → T	3
hs6M1-2P	1	40	Asn → Asn	C → T	3
hs6M1-29P	4	55	Asn → Thr	A → C	2
		117	Ala → Val	C → T	2
		224	Val → Glu	T → A	1
		228	Ser → Leu	C → T	2
hs6M1-31P	1	127	Pro → Ser	C → T	1
hs6M1-30P	1	230	Ala → Thr	G → A	1
hs6M1-19P	1	252	Pro → Arg	C → G	2
hs6M1-17	+3	85	Phe → Leu	T → C	1
		188	Phe → Leu	C → A	3
		227	Pro → Pro	A → G	3

Table 7.4: SNPs in MHC-linked ORs found by searching the public databases.

This high mutation rate of hs6M1-17 suggests selective pressures are more relaxed on hs6M1-17 than any other OR identified in this analysis. One tentative explanation for this could be that, as the non-functional allele of hs6M1-17 begins to propagate throughout the population, the protein is no longer expressed and so the selective pressure on hs6M1-17 is lost. This allows the number of point mutations within different haplotypes to increase dramatically, producing the situation

that can currently be observed The contrast between hs6M1-17 and pseudogenes that appear to be pseudogenic in all haplotypes can be explained by hypothesizing that these 'pseudogenes' have been recruited for other purposes within the genome: for example, they may be expressed as a 5 transmembrane domain protein, or they may form other genomic structures, for example, CpG islands, like one of the OR genes from the chromosome 17 cluster (Glusman *et al.*, 2000), or nuclear matrix attachment regions (Gimelbrant and McClintock, 1997). Alternatively, as has been suggested for loci within the MHC, it may be that these pseudogenes are maintained as they are involved in generating new alleles through gene conversion (Haino *et al.*, 1994, The MHC Sequencing Consortium, 1999).

The SNP data also provide tentative support for the non-pseudogenic status of hs6M1-14. 4 point mutations were observed in this gene, but these are all present in the third nucleotide of the codons producing the 4 amino acids and these mutations produce no changes to the predicted amino acid that will be translated. This higher rate of codon conservation in the first and second coding positions suggests some form of selective pressure is acting upon this locus. This apparent selective pressure, alongside the high conservation of this locus compared to the mouse OR gene, mm17M1-6 (Chapter 5) appears to imply that this gene may have a functional role in spite of its lack of open reading frame.

A point mutation was also observed in hs6M1-19P. This suggests at least 3 alleles of hs6M1-19P are present within the human species: 2 non-functional alleles, and 1 functional allele observed in the resequencing study. The hs6M1-19P data remain the only data that suggests insertions and deletions may also be present within the MHC-linked olfactory receptor cluster: SNP data do not include these type of mutation events.

7.6. Conclusions

In total, 73 point mutations (52 from resequencing OR genes, 21 from SNP survey) have been described. This produces an average value of 2.2 point mutations per olfactory receptor locus, although there are genes for which no mutations have been reported which brings this average value down. This value can, however, be compared with the figure of 1.7 point mutations per olfactory receptor gene within the chromosome 17 cluster (26 point mutations identified in 15 olfactory receptor genes) (Sharon *et al.*, 2000). The higher frequency within the MHC-linked cluster may reflect the proximity to the MHC, where class I and class II alleles are characterized by an extremely high number of alleles (Bodmer *et al.*, 1999). Proximity to the MHC has been proposed to explain the high variability of the GABBR1 locus (Peters *et al.*, 1998).

The functional implications of these polymorphisms, with the exception of the non-functional alleles caused by stop codons (hs6M1-4 and hs6M1-17) or deletions (hs6M1-19), are difficult to assess. 46 amino acid changes are nonsynonymous in a variety of positions across the protein, variability is not just restricted to the position where the ligand is thought to interact with the protein. The functional importance of these amino acid variations cannot be predicted. Within the transmembrane domains implicated in ligand binding, however, variations can be predicted to be likely to have a very large effect, since even conservative amino acid changes (such as Val $\rightarrow$ Ile in a mouse OR, transmembrane domain 5) result in different preferences for odorant binding (octanal $\rightarrow$ heptanal in the mouse OR) (Krautwurst *et al.*, 1998).

Differences in the ability to sense different odorants within the human species are clearly present and it is likely that a large amount of this heterogeneity is caused by variations in olfactory receptor gene repertoires. It is clear, however, that compared to some gene families, such as those

involved in immune defence and those involved in development, the olfactory receptor gene family is likely to be under a lower amount of selective pressure. This is because, although the olfactory system will contribute to survival chances, mutations in the immune defence system or developmental processes are likely to have a larger effect on an organisms' survival chances (Trask *et al.*, 1998, Gilad *et al.*, 2000, Sharon *et al.*, 2000). In the light of these lower selectional pressures, it is both possible to imagine non-functional olfactory receptor alleles spreading quickly across the population, and it is also possible to imagine that an olfactory receptor gene could amass a large number of alleles that would produce a number of proteins with no significant difference to an organisms' survival or mate-finding chances. Considering these 2 possibilities, therefore, it appears that the number of polymorphisms within the MHC-linked OR cluster is fairly low. This could be explained by the recruitment of olfactory receptor genes into other biological systems or it could imply that the mutational rate within olfactory receptor gene clusters is lower than that found in other regions of the genome. The mutational rate within OR clusters may be lower than that found within other areas of the genome because, rather than diversity being generated by a high number of alleles within a moderate number of genes, it may be that diversity is generated by a high number of genes.