Chapter 9

MHC-linked vomeronasal receptor (VR) genes

9.1. Introduction

In addition to the major olfactory epithelium (MOE), where olfactory receptor (OR) genes are expressed, mammals also have another anatomically distinct organ where odorants are perceived. This secondary organ is known as the vomeronasal organ and it is located at the base of the nasal septum connected to the nasal cavity by a small duct (Bargmann, 1997, Keverne, 1999). In rodents, removing the VNO interferes with the detection of pheromones which are defined as chemicals that convey information about reproductive and social status between members of the same species (Wysocki and Lepri, 1991). Changing the ability of the VNO to detect pheromones generally produces changes in mating and aggressive behaviour in rodents.

In contrast to the major olfactory epithelium, where a large amount is known about the pathway from olfactory receptor to olfactory bulb, very little is known about how pheromones are detected in the VNO. Three families of pheromone receptors have been identified, as is the case for ORs, these families all code for proteins belonging to the 7 transmembrane G-protein coupled receptor superfamily. These families, known as V1Rs (Dulac and Axel, 1995), V2Rs (Herrada and Dulac, 1997, Matsunami and Buck, 1997) and V3Rs (Pantages and Dulac, 2000), are all expressed within the rat or mouse VNO and contain about 100-150 members. The V2Rs differ from the other 2 families in that they possess a large extracellular N-terminal domain, similar to that found in extracellular calcium-sensing receptors and metabotropic glutamate receptors.

The existence of functional VR genes within the human species is controversial. Seven different human V1R sequences were identified using PCR and library screening with rodent V1R

sequences but this approach failed to produce any functional human V1Rs (Giorgi *et al.*, 2000). The lack of functional human V1Rs is supported by anatomical evidence that suggests although there is a foetal VNO in humans, the adult version is an atrophied, obsolete organ (Tirindelli *et al.*, 1998). On the other side of the controversy, one group reported finding a functional V1R-type pheromone receptor gene expressed in the olfactory mucosa (Rodriguez *et al.*, 2000) suggesting pheromones could be perceived through the main olfactory system as they are in rabbit (Hudson and Distel, 1986) and pig (Rodriguez *et al.*, 2000). This could explain the observation of proposed pheromone-regulated behaviour, such as the synchronization of menstrual cycles among women living together (McClintock, 1971, Stern and McClintock, 1998). There have also been a number of studies that reported finding a structurally intact VNO (Garcia-Velasco and Mondragon, 1991, Moran *et al.*, 1991, Stensaas *et al.*, 1991).

9.2. Identification of VR genes in the human extended MHC class I region

Five pheromone receptor loci of the V1R-type were identified in the human extended MHC. These are found within a genomic region of 662765 bp, separated by a large number of other genes, including a cluster of histone genes (Figure 3.1). In contrast to the olfactory receptor genes located centromeric of this sequence, the VR genes are all pseudogenic, and a large distance separates the loci: hs6V1-5P and hs6V1-1P are separated from the 3 other VR genes by 371 Kb and 250 Kb respectively (Figure 3.1). They are generally dissimilar on the protein level, with shared protein identities typically below 30%. There is, however, evidence that hs6V1-3P and hs6V1-4P may be related as they share a protein identity of 63.9%. hs6V1-2P, the other pheromone receptor located within the core group of 3 is also similar to hs6V1-3P and hs6V1-4P: it has a shared protein identity of greater than 40%. The lack of conservation within these pheromone receptors is not surprising as these genes are all pseudogenic, disrupted by frameshifts

and stop codons. The pseudogenic properties of these genes is something that appears to be shared by the majority of the VR type 1 genes within the human genome.

9.3. Identification of VR genes in the human genome

In addition to the 5 MHC-linked VR genes, a further 46 VR genes were identified in the human genome, using a similar method to that used to create the human OR database. This total represents a first attempt to identify human V1R genes; in contrast to the human ORs, it cannot be regarded as a comprehensive identification. Nevertheless, an estimate of the total number of V1R genes within the rat genome suggested a total number between 30 and 100, so the 51 VR loci identified in this project seems likely to represent between 50 to 100% of the human VR type 1 repertoire. In the light of this the number of functional V1R genes within the human genome appears to be very small. Of the 51 VR genes identified, only one, located on chromosome 19 (hs19V1-5), is predicted to produce a functional protein. The majority of the V1 pheromone receptor genes are characterised by pseudogenic reading frames, usually containing several stops and frameshifts.

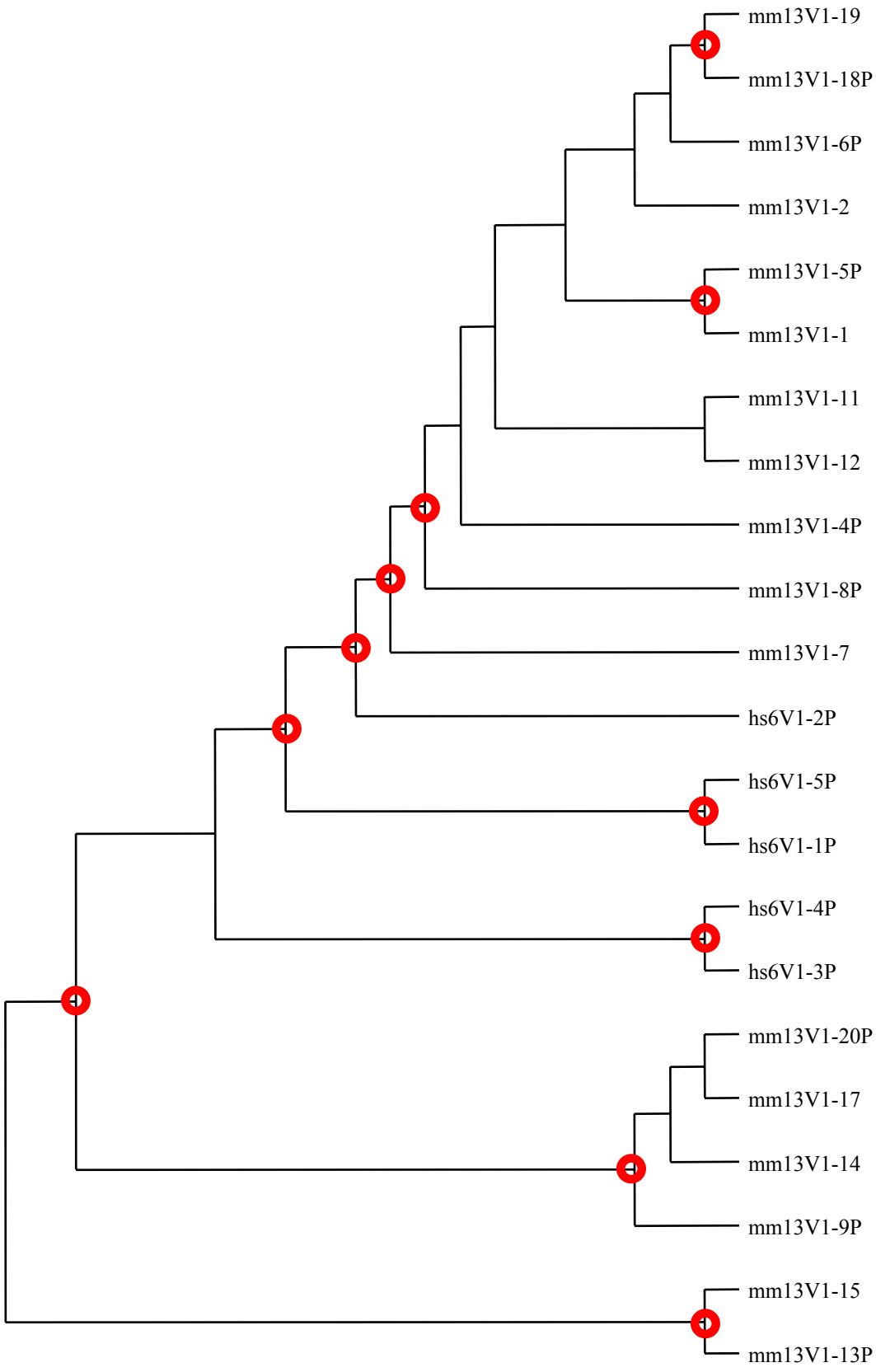
In order to amass this degree of pseudogenicity, therefore, the decline of the VR gene family either started several million years ago, or these genes may have rapidly mutated as the VNO became less important in human evolution. The majority of the loci remaining in the human genome today are found in clusters (36 out of 51 = 70.5%) which suggests the functional VR gene family may have been arranged in clusters, similar to the OR clusters. As with the OR family, the majority of VR genes appear to be more closely related to other VR genes located on the same chromosome than to other VR genes within the human genome, suggesting local duplications may have been the major pathway through which this gene family proliferated.

The proximity of the MHC-linked VR genes to the MHC-linked OR genes suggested that there may be a relationship between clusters of olfactory receptor genes and clusters of pheromone receptor genes. This relationship between VR genes and OR genes can also be observed on other chromosomes, notably chromosome 1 (VR genes: 1q44, OR genes: 1q43), chromosome 3 (VR genes: 3p25.1-2, OR genes: 3p25.3), and chromosome 15 (both VR and OR genes: 15q11.2). This association between VR genes and OR genes could be the remnants of a functionally important association or it could reflect the fact that the 2 gene families both relied on local duplications to provide new members. Clusters of OR genes and VR genes may have ended up in similar genomic locations owing to the ability of these chromosomal regions to promote local duplications.

9.4. Identification of VR genes in the syntenic mouse region (mouse chromosome 13)

Using sequences obtained from the human region, the Ensembl mouse draft sequence was searched for potential orthologs of these human VR genes. With the exception of hs6V1-1P, the most closely related sequences to these genes were found on mouse chromosome 13, localising to a position downstream of a cluster of histone genes and the mouse *Hfe* locus. The mouse VR genes, therefore, appear to have been conserved in a similar position to their human orthologs. Potential orthologs for hs6V1-1P were also localised to this area, but they also appear to be localised to mouse chromosome 7. In contrast to the OR genes, the VR genes are too pseudogenic to accurately predict orthologous relationships, although from the phylogenetic tree (Figure 9.1), one subgroup of VRs are associated with hs6V1-2P, and there are at least 2 additional subgroups not associated with human orthologs.

Figure 9.1 (next page): Phylogenetic tree (maximum parsimony) showing the relationships between the VR genes in the human MHC extended class I and VR genes on mouse chromsome 13. 142 sites were used and 500 bootstrap replicates were performed. The red rings at branch points indicate where bootstrap values are over 70%.


mm13V1-19
mm13V1-18P
mm13V1-6P
mm13V1-2
mm13V1-5P
mm13V1-1
mm13V1-11
mm13V1-12
mm13V1-4P
mm13V1-8P
mm13V1-7
hs6V1-2P
hs6V1-5P
hs6V1-1P
hs6V1-4P
hs6V1-3P
mm13V1-20P
mm13V1-17
mm13V1-14
mm13V1-9P
mm13V1-15
mm13V1-13P

20 mouse VR genes were identified in a segment of sequence approximately 700 Kb in length. 9 of these genes were considered to be functional, with another 11 pseudogenes owing to frameshifts, stop codons and a lack of a starting methionine. The ratio of genes to pseudogenes, therefore, is 0.81 which resembles the gene to pseudogene ratio of the ORs in the human genome (0.8) rather than the gene to pseudogene ratio of the ORs in the mouse genome (3.6). This seems to suggest that the VR family in mouse has undergone or is undergoing the same type of contraction that has been observed for the OR family within the human genome, where the reduction in the fraction of functional OR genes within the OR repertoire has been considered to be due to the reduced functional importance of the sense of smell (Sharon *et al.*, 1999, Rouquier *et al.*, 2000). This is a higher amount of pseudogenicity than that observed by Del Punta *et al.* (2000). If this gene to pseudogene ratio is consistent throughout the mouse genome (and there is no reason that the MHC-linked VR genes cannot be considered representative of this repertoire), it does appear that a higher number of V1R genes than might be predicted are pseudogenes, suggesting a possible decline in the importance of these genes in detecting pheromones. Alternatively, it may be that the poor quality of the draft mouse sequence accounts for a number of these genes lacking open reading frames, although the identification of several complete open reading frames in OR genes identified from the mouse genome sequence seems to suggest this is not the case.

9.5. Conclusions

5 VR pseudogenes were identified within the human extended MHC class I region. These were compared to other VR genes within the human genome, and to VR genes from the mouse syntenic region. The MHC-linked VR genes are similar to other VR genes within the human genome in that the majority of these genes are disrupted by stop codons and frameshifts. There was limited evidence suggesting that some chromosomes have clusters of VR genes and OR

genes located closely together: this could represent the fact that the 2 gene families proliferate or are regulated within the same genomic environment, or it could be due to an ancient relationship between the 2 families.

The mouse MHC-linked VR genes could suggest the 2 families of genes have a shared evolutionary fate. The high number of pseudogenes within the mouse VR repertoire appears to suggest that the VNO could be in the process of being made redundant or at least downsized from the structure it once was. It is tempting to speculate that, as the OR gene family grew in size and function, the VR family was downsized. This suggests the VNO was the primitive chemosensory organ, something that may be supported by observations that a VNO is present in most amphibia, reptiles and nonprimate mammals but absent in birds and adult catarrhine monkeys and apes (Stoddart, 1980). This primitive VNO-non-primitive MOE distinction is, however, too simplistic for three reasons. Firstly, there are a number of other gene families that are VNO receptors and these may not be in decline in the mouse genome. Secondly, the VNO may be variably present in species such as catarrhine primates (Old World monkeys, apes and humans (Smith *et al.*, 2001)), it cannot necessarily be regarded as continually declining throughout evolution. Thirdly, the basis for this suggestion would require a VNO-type structure and VR genes to be more functionally important than a MOE and OR genes within a species that had evolved earlier than mouse or human. There is currently little evidence for VR genes in a species such as *Danio rerio* (zebrafish), although data from the zebrafish genome project (The Sanger Centre in collaboration with the zebrafish community, Zebrafish Workshop 2000) may serve to refute or add evidence to this hypothesis. Primitive or not, it is clear however that similar selective pressures would have acted on both these gene families, and their fate is likely to have been linked in some way.