# Chapter 6

# Discussion and future perspectives

In this dissertation I have explored the dynamics of OSN diversity, as measured by olfactory receptor expression, and the effects of genetic and environmental factors on its regulation. For this, I utilised an RNAseq-based approach that allowed me to study the complete OR repertoire. I have shown that the transcriptional profiles obtained via RNAseq from whole tissue extracts are accurate and highly reproducible, and outperform other technologies available. Further, the high-throughput and unbiased character of the technique allowed the generation of a comprehensive catalogue of the transcripts present in the olfactory system, both known and novel; and the generation of full-length gene models for hundreds of OR and VR genes. The combination of RNAseq with FACS and single-cell technologies resulted in a precise characterisation of the molecular profile of the OSN transcriptome. Moreover, it allowed the discovery of novel subdivisions of mature OSNs. Importantly, the study of single OSNs permitted me to assess the wide-held belief that OR expression is monogenic and monoallelic, hereby directly proven.

From the data, I can conclude several things. First, expression levels of OR genes in WOM samples are an accurate reflection of the number of OSNs in the MOE that express particular receptors. Thus, the transcriptional profiles inform on the proportions of the different OSN types found in the neuroepithelium. Second, such diversity of OSN types is stereotypical in animals of the same genetic background, irrespective of sex and (largely) of age. Third, the presence of genetic variation results in high divergence of the relative proportions of different OSN types, with most being susceptible to altered abundance based on their genomic context. Fourth, the final distribution of OSN diversity is controlled by genetic elements that act in *cis*, and is not affected by sustained alterations of the olfactory environment. And fifth, the persistent but interleaved presentation of olfactory stimuli alters the abundance of a subset of OSN types,

in a time-dependent, odour-specific and reversible manner (Figure 6.1).

## 6.1 Understanding the mouse olfactory system by RNAseq.

High-throughput RNA-sequencing has greatly advanced and developed in the last few years. With the sequencing costs dropping, RNAseq has substituted the use of microarrays and has become routine for transcriptional profiling. I have exploited the strengths of RNAseq to characterise and better understand the transcriptional dynamics of the mouse olfactory system. In this dissertation, I have demonstrated the accuracy and reproducibility of the expression estimates obtained with this methodology, and its superiority when compared to other techniques such as a greater dynamic range of expression values and better correlation with qRT-PCR expression estimates.

During the course of my PhD, several groups published results from similar experiments to mine[329–331]. Shiao et al. performed RNAseq in WOM samples of male and female BALB/c mice and concluded that males have overall higher expression of OR genes[329]. But they failed to notice that males also have higher expression of all the canonical markers of mature OSNs and, therefore, it is likely that the observed differentials are only a product of varying proportions of OSNs in their whole tissue samples. Additionally, they sequenced only one sample of each sex (that was the pool of three individuals), which makes it very difficult to test for differential expression with confidence, since any observed differences could be the result of technical variation with no biological relevance. Indeed, both my data and that of Kanageswaran et al. [330] failed to identify any convincing differential OR expression by sex.

A similar situation occurs in the experimental design of Kanageswaran et al.[330]; they sequenced several biological replicates of MOE samples from CD1 and B6 mice, but only a single replicate of FAC-sorted OSNs from OMP-GFP animals. Thus, the comparisons of the transcriptomes of the OSNs versus the whole tissue are underpowered and are also flawed since the genetic background was not controlled. Finally, Shum et al. presented WOM RNAseq data of two adult B6 females and reconstructed OR gene models using a strategy similar to mine; however, their sequencing depth was much lower. An analysis of their reconstructed gene models led to the proposal that the presence of introns leads to higher expression levels[331]. However, the authors did not recognise that the genes they classified as *intronless* were so because they did not have enough

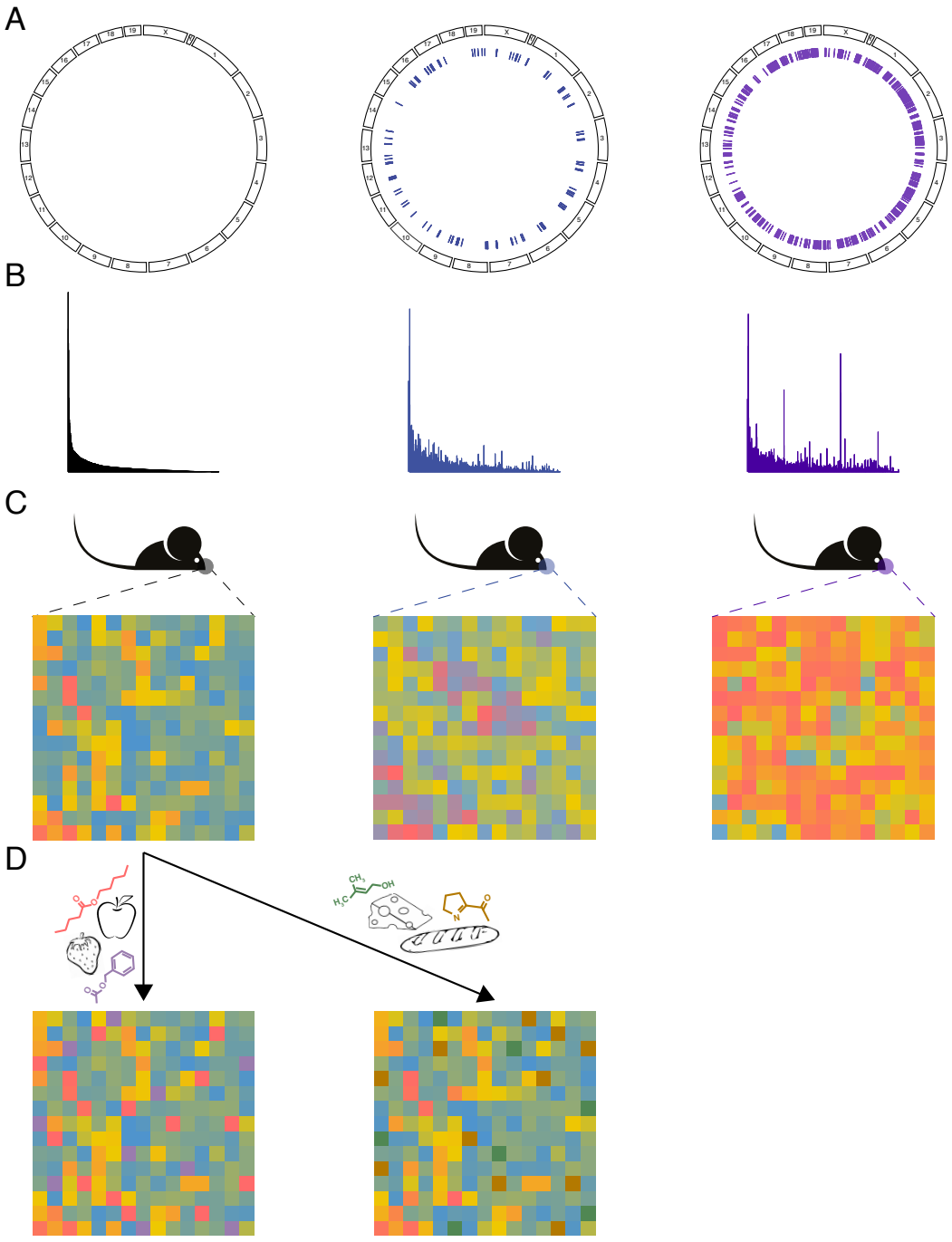


**Figure 6.1** – **Genetic and environmental regulation of OSN diversity.** Model on the impact of genetic and environmental factors on OR expression and OSN type abundance. **A)** Representation of the mouse genome, where all chromosomes are arranged in a circle. The chromosome number is indicated. As an inner circle is a depiction of genetic variation events, each as a vertical line. **B)** The first genome (left) produces a particular distribution of expression for all the OR genes, ordered by decreasing abundance order. For the second (centre) and third (right) genomes, OR genes are ordered the same as in the first. The presence of genetic variation results in OR expression profiles that are different from each other. Greater amounts of variation (right) result in a more divergent profile. **C)** The unequal expression levels for different OR genes results in unequal numbers of OSNs expressing such receptors. Each square represents a different OSN and the colour indicates the particular OR gene it expresses. Thus, each mouse with a unique genome has a unique pattern of OSN diversity in its nose. **D)** The proportion of each OSN type is amenable to modification upon olfactory stimulation. On the left is represented a country mouse that feeds on fruit and seeds; constant exposure to odorants from fruits results in the enrichment of OSNs that express ORs that recognise such odorants (pink and purple). On the right is a city mouse that instead feeds on cheese and bread and therefore has more OSNs that express ORs activated by the molecules in these foods (green and brown).

depth to reconstruct their full gene models. Indeed, from all the OR genes proposed to be intronless, 91.9% have introns in the reconstructed models I presented. Thus, the low expression levels were the cause of the lack of introns in their models, and not the other way around.

An important advantage of RNAseq is that it does not depend on the genome annotation; this has allowed me to discover novel genes that, to date, remain unannotated in the mouse genome. Several hundred loci showed evidence of expression of multi-exonic structures, many of which contained protein features and domains. The surprisingly high number of such novel genes indicates that despite the high quality of the sequence and annotation of the mouse genome, the use of specialised tissues that are often not considered in gene annotation pipelines are a source of additional information. It is likely that many of the putative genes I have defined are specifically expressed in the olfactory system and serve specialised functions; examples are the two lipocalin genes (*Lcn16* adn *Lcn17*) that were validated as true genes, and are odorant binding proteins. These two novel genes can be found in the same orientation in the rat genome, but synteny is disrupted in the primate lineage and there are no orthologues present in primates or humans.

RNAseq combined with other experimental strategies, such as cell sorting based on expression of particular marker genes, provide a powerful strategy to deconstruct complex tissues such as the MOE. In collaboration, I utilised this approach to characterise the transcriptional profile of the OSNs only, excluding the other cell types present in WOM samples. By differential expression analysis, I defined a large list of genes that are specific to the neurones and provide information on their molecular processes and pathways. These will be useful as a reference for future studies, since discriminating neuronal processes from supporting functions is very informative on the nature of the phenomena being studied.

Perhaps more interesting was the finding that the expression of *Omp*, the canonical marker that defines mature OSNs, is not expressed in a continuum but, instead, segregates into two discrete populations. The analysis of the genes that differentiate these two subpopulations revealed that while both are mature OSNs, the $GFP^{low}$ cells are slightly less mature than the $GFP^{high}$ cells. A recent study characterised the temporal expression of OR genes, *Adcy3* and *Omp* in differentiating precursor cells into mature OSNs. Indeed, *Omp* was found to be the last gene to be activated, after both OR genes and *Adcy3* had been turned on[312]. Therefore, it is conceivable that the difference between OSNs expressing either low or high levels of *Omp* coincides with a discrete functional

event in the final maturation of the neurones. For example, it could be that the cells from the $GFP^{high}$ population have successfully established the negative feedback that ensures OR singular expression; the release of the unfolded protein response might be necessary to achieve high *Omp* expression. Alternatively, it could be that the distinction reflects the successful innervation of a glomerulus in the MOB. In either case, this finding likely marks an event in the functional maturation of OSNs and deserves further study.

Probably the most important advantage of RNAseq as a technology, is that it provides a comprehensive and unbiased profile of all the genes expressed in a particular sample. As such, it is a very attractive strategy to study the basic questions of OR gene expression regulation. Ever since the discovery of the OR genes and the study of their expression in the MOE, it has been assumed that they are expressed in a monogenic and monoallelic fashion in every OSN[66, 67, 69, 72]. However, all the evidence supporting monogenic expression of ORs stems from testing coexpression of a few combinations of two receptors but the full OR repertoire has never been tested[70, 182]. Therefore, the study of the transcriptome of single OSNs provides an unbiased method to account for *all* the different OR genes that are transcribed in a particular neurone. By sequencing 21 different individual OSNs, I was able to identify abundant OR expression of one OR gene in 19 of these. Additional receptor genes showed evidence of expression, but at very low levels. Indeed, taking together all the sequencing data supporting OR gene expression, in each OSN over 98.1% was concentrated on a single OR, with the remaining scattered across a few other receptors. It is not clear whether these low-abundance OR genes are biologically meaningful or whether they represent leaky transcription that has no impact on the sensing capabilities of the OSN. It has been shown that genes that are expressed at very low levels do not correlate with protein expression[305]; therefore, it is possible that only the abundant OR translates into protein. Additionally, low expression levels of some OR genes could be observed in other various cell types, which suggests that their expression is not related to olfactory function. However, until proteomic techniques match the sensitivity of the transcriptomic methodologies, this will remain unresolved. To date, studies of the membrane proteome of the cilia of OSNs have been able to identify only a few dozen OR genes, due to their low expression in WOM preparations[332].

Based on these data, the 'one neurone - one receptor' rule of OR expression is supported, for the first time, on a scale that accounts for each and every receptor annotated in the mouse genome. However, this is based only on 19 OSNs, which is far from representative sampling of the diversity of neurones present in the MOE. Thus, I cannot

rule out that some OSNs indeed express several receptors at high levels, but were not captured in the limited sample analysed here. Furthermore, the stringent QC criteria applied selected against the inclusion of neurones that contained two or more OR genes expressed at high levels, because some were clear cases of carry-over from other samples; therefore, it was not possible to ascertain the validity of the observed expression. In order to explore further the possibility that some OSNs express two ORs, collaborations are in place to obtain OSNs from mouse lines that express a specific receptor gene along a reporter fluorescent protein. By sequencing the population of neurones that express a particular OR gene, we will be able to gain insight into the levels of expression of other receptors. If the data indicates that coexpression of receptors is likely, a single-cell approach could then be used to definitely prove that this occurs within a single OSN.

The OMP-GFP animals used for the single-cell RNAseq experiments are in a mixed B6×129P2 genetic background. Using the SNPs present in the exons of the abundant OR genes, it was possible to infer the allele expressed in each OSN. By examining the sequencing data directly, I was able to confirm that OR expression is monoallelic, and that this is extremely tightly regulated. Over 99.7% of all the sequencing reads covering variable positions supported the expression of one of the alleles; the remaining reads presented any of the other three nucleotides, suggesting that these contain sequencing errors and low quality base calls, rather than being the transcription of the other allele. Therefore, the paradigm of expression of only one allele of the chosen OR gene holds true, at least based on the data from the small subset of OSNs tested.

Lastly, the combination of powerful technologies such as single-cell capture and preparation, along with RNAseq, allow for the discovery of minor populations of cells that are undetectable when bulk RNAseq is performed. In this way I was able to identify two OSNs that do not express any OR genes at high levels. Instead, they have very abundant expression of *Gucy1b2*, a soluble guanylyl cyclase, and *Trpc2*, the cation channel that is fundamental in signal transduction in VSNs. During the analysis of these data, a paper reported the existence of two different subpopulations of OSNs in the MOE that are positive for *Trpc2*[131]. Further characterisation of one of these –the type B cells– resulted in the identification of expression of *Gucy1b2* and suggested the lack of expression of chemoreceptors[322]. By profiling the complete transcriptome of two of these cells, I was able to identify over 50 genes that are not expressed in the canonical –OR-expressing– OSNs and, therefore, constitute the molecular fingerprint of this novel neurone type. Additionally, I confirmed that no known chemoreceptor is expressed at abundant levels. The *Trpc2*$^+$ OSNs have recently been shown to innervate glomeruli in

the MOB, which suggests that they might indeed be chemosensory cells relying olfactory information[131, 322]. Many of the components involved in the signalling pathway used to generate action potentials are present in these cells, and *Gucy1b2* is expressed at levels similar to OR genes; thus, it is possible that this guanylyl cyclase might be taking on the role of the receptor protein, but further studies are necessary to test this.

## 6.2 Almost all OR genes are expressed in the MOE.

The MOE contains a variety of OSN subpopulations, each defined by the particular receptor they express. The great majority express OR genes, but minor subpopulations also express TAARs or GC-D, and possibly several other subdivisions are yet to be characterised. In this dissertation I have presented, for the first time, the complete expression profile of the receptor repertoire in mice, with particular focus on the OR genes. Since the majority of the OR genes have been defined by computational methods, they lack evidence supporting their role in chemo-signalling. The most basic requirement for an OR to be implicated in olfaction, is that it is expressed in a sensory tissue such as the MOE. Therefore, the evidence that nearly the complete repertoire of putatively functional OR genes are indeed expressed in the WOM samples supports their involvement in transducing olfactory information. This is particularly relevant since extra-olfactory functions have been reported for some OR genes, that are not only expressed elsewhere[333], but have actually been shown to act in processes such as sperm chemotaxis[297, 298], muscle cell regeneration and migration[334] and serotonin release in gut cells[335].

Between 10 and 20 OR genes that are annotated as functional receptors lack expression data in any of the WOM samples. However, taking together the expression profiles of the different strains, only five lack expression in all the samples profiled. Thus, some of the receptors that are not expressed in a particular strain, might represent pseudogenes as a result of functional variation, or might be expressed at such low values that they were not detected. From the five OR genes with no unique counts in any sample, two are identical copies of each other (*Olfr247*), located ~8 kb apart; for these, all sequencing reads are multimapped and therefore it is impossible to know if one or the two genes are expressed. From the other three (*Olfr891*, *Olfr952* and *Olfr1061*), *Olfr952* has some multireads mapped and therefore could be expressed, while the other two do not. These could be cryptic pseudogenes, might be expressed at a different age or be present in an extremely low number of OSNs.

An important contribution of my work has been the construction of full-length gene models for a large number of OR genes. This has greatly increased the amount of information recovered for each receptor and, also, the amount of sequence that is unique. This has allowed me to estimate expression levels with more accuracy, since OR genes with close paralogs are systematically underestimated unless the more divergent non-coding regions of the genes are considered. But the additional sequence will also be very helpful for studies that are based on methodologies that rely on hybridisation approaches, such as qRT-PCR, NanoString nCounter, arrays and *in situ* hybridisation, one of the most popular for the study of OR expression. Genes that before were inaccessible are now available for study with probes specific enough to differentiate them from other receptor genes.

## 6.3 The MOE is a mosaic of OSN types.

The regulation of OR expression in OSNs is only partially understood. Achieving monogenic expression relies on a basal state of generalised repression of the whole OR gene repertoire; as the OSN reaches maturity, a single OR allele is activated[203, 205]. This is inefficient enough so that typically only one event can occur before a negative feedback mechanism ensures the process is shut down[209, 211]. However, how a particular receptor is chosen is still an open question. Very often, it is described as a *random* or *stochastic* process[202, 210], implying that any OSN can chose any of the 1,250 OR genes, and that any receptor has the same probability of being chosen. Contrary to this, each OSN has a restricted subset of the repertoire available for expression, depending on its location on the epithelium. Particular OR genes are expressed in restricted zones of the MOE[18, 66, 67, 78] and, therefore, only the OSNs located within those regions can choose them. Furthermore, it has been shown that different ORs are expressed in varying numbers of OSNs, with some being much more abundant than others[62]. This is at least partly influenced by the activity of enhancer elements and the number and organisation of transcription factor binding sites in the receptors' promoters[186, 188]. Thus, *random choice* is an unfortunate choice of words.

By profiling the entirety of the OR gene repertoire I have demonstrated that the expression levels of different receptor genes are highly variable, spanning at least four orders of magnitude. Importantly, I have also shown that the RNAseq expression estimates correlate with the number of OSNs expressing a particular OR gene; therefore, a highly expressed OR implies a high number of OSNs in the MOE expressing such a re-

ceptor. With this in mind, the unequal expression levels observed from the RNAseq data represent unequal proportions of each OSN type. This disparate distribution could be the result of two mechanisms, that might act alone or in combination: 1) differences in the frequency with which a particular OR gene is chosen, or 2) variation in the life-span of each OSN type such that, with time, those that live longer become more abundant in the overall population. It is unlikely that the second mechanism acts alone, given that unequal expression of ORs are already observed in newborn animals; in these, presumably not enough time has passed to allow survival dynamics to impact the proportion of each OSN type.

Based on the sequencing of several biological replicates from both male and female mice, I have determined that each individual expression pattern for the OR repertoire is exactly the same, as long as the genome remains unchanged (there are no OR genes in the Y chromosome). The rank correlation between different individuals is almost perfect (median rho = 0.98, p-value < 2.2e-16) which indicates that each receptor has equivalent values in the distribution, and the proportions of the different OSN types are preserved. Therefore, the contribution of each OSN type to the MOE's neural population is *determined* by the genetic architecture of the animal. The fact that males and females are indistinguishable, indicates that the OSN repertoire in the MOE is not influenced by the physiology or hormonal balance of the organism; nor is it altered by the differences in the olfactory environment produced by each sex. Furthermore, it is virtually unchanged at different ages within the controlled lab environment. The OR expression profiles of B6 animals of 10 or 24 weeks of age are equivalent and as highly correlated as between animals of the same age. This is consistent with a study of ageing female B6 mice, from 2 to 31 months of age, where the expression of 531 OR genes was assessed to find that only 4.3% of these were significantly differentially expressed[304]. But also, OSN abundance is unaffected by social and behavioural differences; for example, both sexes establish social hierarchies when group-housed[336], but no differences were evident between the different cage-mates. Thus, even though social interaction and behaviour are highly driven by olfactory cues, regulation of these processes is not achieved by differences in the receptor repertoire expressed in each animal. Instead, recent data suggests that internal state does alter olfactory perception, but by post-transcriptional mechanisms. Rather of changing the expression of the receptors, their activation is blocked by the influence of cycling hormones[337].

However, the distribution of OSN types observed in different strains of mice is highly variable (Figure 6.1B). Whereas B6, 129, CAST or OMP-GFP animals are all highly

correlated with those of the same strain, the differences between genetic backgrounds are large. More than 65% of the ORs were significantly DE between at least a pair of strains. This indicates that the great concordance observed between biological replicates is not the result of intrinsic stability or tight regulation of the expression levels of the OR genes. The different strains analysed in this dissertation are all inbred laboratory mouse strains. All the classical laboratory strains were derived from a small pool of founders from the *Mus musculus* (*M. m.*) *domesticus*, *M. m. musculus* and *M. m. castaneus* subspecies[321, 338]. Therefore, their genomes are a combination of regions from different genetic origin. The classic strains are mostly of *M. m. domesticus* origin (86 to 96% of the genome) with only small contributions from the other two subspecies[338]. Therefore, the genomes from B6 and 129 animals are closely related; these contain only around 4.4 million SNPs and 800 thousand small indels[321]. As a comparison, any two humans differ, on average, at around two to three million basepairs, considering SNPs only[339]. In contrast, CAST animals are a wild-derived strain, that pertains to the *M. m. castaneus* subdivision. As such, it is a lot more divergent from the inbred classic laboratory strains, and contains more than four times the amount of variation[321]. Despite the disparate divergence of these strains, the OR expression levels for the whole repertoire are remarkably dissimilar between all, with up to 50% of all receptor genes significantly differentially expressed. This indicates that genetic variation has a very significant effect on the regulation of the final distribution of the different OSN types (Figure 6.1A-B).

Analysis of the distribution of genetic variation in the mouse genome has revealed that OR genes have slightly more variation than the average gene[271] and they tend to be enriched in regions of copy number variation both in humans[255] and mice[340]. These two characteristics reflect the evolutionary dynamics of the OR gene family. The olfactory system has evolved to discriminate a large catalog of molecules by diversifying the repertoire of receptors available for detection. The ability to sense a larger number of odorants increases the amount of information an animal can gather from their surroundings, and provides a reservoir of detectors to adapt to novel environments. As such, diversification of the OR repertoire should be beneficial. Analysis of introgression events between different mouse species lends support for this hypothesis. Introgression events between the house mouse (*M. m. domesticus*) and the Algerian mouse (*Mus spretus*) are common in the wild, but most hybridisation events tend to be removed by drift and selection[341]. An analysis of the genomes from diverse wild-caught mice identified some regions where hybridisation occurred between the two species. Interestingly, the intro-

gressed regions that have prevailed in the genomes of several individuals, are enriched in OR genes, suggesting that their maintenance is beneficial to the animals[341].

Based on the above, genetic variation to diversify the OR repertoire is beneficial. Several examples have shown that single amino acid changes are able to shift the binding specificity of an OR[224] and, often, they result in the innervation of separate glomeruli[85]. In contrast to inbred mice, wild animals will possess up to twice as many alleles for the OR gene catalogue. Thus, individuals heterozygous at some OR loci will likely possess increased detection and information processing capabilities. Furthermore, genomic variation will also affect the non-coding portions of the OR genes, their promoters and regulatory elements. Alteration of transcription factor binding sites has been shown to impact the final number of OSNs that express the affected allele[188]. Thus, the accumulation of non-coding genetic variation will have an effect on the proportion of the neuronal population that is taken by each particular allele (Figure 6.1C). Hence, the combination of coding variation that alters the detection properties of the receptors, along with non-coding variants that modify the number of OSNs that express each OR, will ultimately produce a unique repertoire of ORs with a specific OSN distribution, which in turn will impact olfactory sensing.

In support of this, analysis of the 1000 Genomes Project data has revealed that any two individuals differ in around 30% of their OR genes, either by possessing differing sets of segregating pseudogenes or by coding variation that has an impact on the response profile of the receptor to its ligands[241]. Further, several examples exist on the effects of genetic variation in OR genes and differences in perception[241, 266–268, 270]. A recent study tested the variability in human perception of a set of odorants, based on several descriptors. While the gross perception based on pleasantness was very similar between individuals, it was highly specific when detailed descriptors were used. Furthermore, the perceptual profiles were highly variable between individuals; so much so, that the authors proposed that with enough odorants and descriptors, it would be possible to create an *olfactory fingerprint* for every person. What's more, the similarity of two perceptual fingerprints was correlated with the similarity of their HLA profiles (human leukocyte antigen system, analogous to the major histocompatibility complex in animals), suggesting that the olfactory fingerprint might be capturing genetic information[342]. Thus, these data suggest that human perception is indeed highly variable, as is the OR profile of each individual.

## 6.4 Plastic control of OSN diversity.

Little information is available on how differences in the number of OSNs expressing a particular OR gene affect detection and/or perception of a particular stimulus. An interesting study generated a mouse with a "monoclonal nose", where over 95% of all the OSNs expressed the M71 receptor. Correspondingly, the rest of the OR repertoire was dramatically reduced. The small number of neurones expressing other receptor types still innervated particular glomeruli, though these were co-innervated by M71 axons. Not surprisingly, EOG recordings upon exposure to acetophenone (the ligand of M71) were greatly increased, and responses to other odorants were diminished. Similarly, acetophenone elicited widespread glomerular activation while other odorants did not elicit detectable responses. Despite the low number of OSNs expressing most receptors, the mice were able to detect and discriminate between different odorants and even between enantiomer pairs. However, their ability to differentiate mixtures of enantiomers was greatly impaired[343]. Thus, these data suggests that a low number of OSNs expressing a given receptor are sufficient to bind odorants and transmit the information, which can be used for olfactory learning tasks. However, the discrimination capacity is greatly weakened, perhaps because the glomerular activation is not strong enough to allow differentiation between similar patterns. Unfortunately, the authors did not test the detection threshold of these animals to common odorants. Therefore, it is unclear whether these animals also have reduced sensitivity; it could be that they are able to detect a ligand only when it is present at high concentrations.

It is tempting to speculate that animals with varying proportions of each receptor type will have different capabilities to detect and discriminate differing sets of odorants. If so, the ability to tune the proportion of OSNs devoted to the recognition of important odorants would be greatly beneficial, especially if the starting abundance dictated by the genetic background is low. Several studies have shown that odorant stimulation increases the life-span of the OSNs that are activated and, with time, these OSN types become enriched in the MOE[286, 288, 292]. Consistently, I have found that the intermittent exposure of animals to either a cocktail of four different odorants, or subsets of these, results in the differential expression of specific OR genes (Figure 6.1D). Interestingly, no changes could be detected when the odorants were present 24 hours a day. Presumably, an odorant that is always part of an animal's environment is non-informative and thus

it would not be advantageous to devote any more OSNs to its continuous detection. At the OSN level, adaptation stops OSNs from responding to sustained stimulation. Based on my results, it seems likely that constant stimulation also blocks enhanced survival.

In the acutely exposed animals the changes in OR expression observed increased with time, which argues in favour of a survival-mediated mechanism. Moreover, it is difficult to imagine a plausible mechanism through which an odorant can influence the choice of its cognate OR during neurogenesis. However, in all experiments, a subset of genes were also consistently downregulated, suggesting a decrease in OSN number. Though unexpected at first glance, the adult MOE maintains a fine balance of the total number of OSNs[22]; therefore, to increase the frequency of some OSN types it may be necessary to decrease others. If all OR genes were to decrease equally, it is likely that the changes would be small enough not to be detected by expression profiling. However, a more parsimonious scenario is one where the receptors expressed in overlapping regions with those that increase frequency are the ones affected, while the rest of the repertoire remains unchanged. This has been observed in a mouse where the coding sequence of *MOR28* was removed; only ORs expressed in the same zone were able to populate the OSNs initially devoted to express MOR28[207]. Thus, it is not unreasonable to assume that the increase in a particular OSN type would take the space of OSNs from the same region. Still, it is likely that many ORs are expressed in the region overlapping that of an activated OR; hence the downregulation of each should be small and it might require longer times to reach a differential that is detectable by RNAseq. Indeed, the number of downregulated ORs is much larger (44.4% of the total) in the animals that were exposed for 24 weeks, than in those that were exposed for only 10 weeks (6.7%, 25% and 33.3% for the groups exposed to R-carvone, heptanal or both).

Recently, a study reported that short-term exposure (5 hours) of animals to particular odorants results in the downregulation of the ORs that respond to them, at the mRNA level; for some OR genes, these changes could be observed as soon as 30 minutes after the start of the exposure[344], but no shorter times were tested. Therefore, it is not clear whether this could be occurring in the acutely exposed animals. The authors performed a comprehensive analysis by RNAseq of animals exposed to acetophenone and generated a list of downregulated OR genes. These were further shown to colocalise with phosphorylated ribosomal protein S6 (pS6), which is a marker of activated OSNs[344]. In parallel, an independent group used the presence of pS6 to capture the OSNs that were activated upon stimulation, and then performed RNAseq to identify the ORs expressed; these were then validated in a heterologous cell system[345]. Together, the lists of

receptors identified by these groups, overlap with four different OR genes that were significantly DE in the animals exposed to the mix of four odorants (24 weeks), that contained acetophenone. Interestingly, all four receptors were downregulated in the RNAseq data.

Based on the above, it seems that at least some of the downregulated OR genes are activated by acetophenone. Thus, it could be that only neurones expressing some OR genes are able to modulate their life-span or that the increase of some OSN types is not big enough to be detected by RNAseq; instead, the temporary downregulation from exposure events close to the time of tissue collection could be identified as overall downregulation. Further, it could also be possible that different mechanisms operate depending on the affinity of each receptor for a given ligand. To better understand the dynamics of the changes observed, I have established a collaboration with Casey Trimmer and Joel Mainland (Monell Chemical Senses Center) to test some of the DE ORs in an *in vitro* response assay in heterologous cells. Preliminary results indicate that three out of five DE ORs tested indeed respond to the mix of four odorants used as stimulus (data not shown), but more systematic and thorough tests are being carried out at present.

Conflicting data is available on the effect of odorant exposure on OR expression and OSN number. Whereas several studies have concluded that OSN activation leads to increased life-span which, with time, should increase OSN number[286, 288, 292], others have proposed that olfactory stimulation results in a reduced number of the activated OSNs[293, 294]. In some cases, these changes have been shown to be specific to a particular OR, whereas other receptors remain unchanged[294]. Thus, while the analysis of particular OR-ligand pairs reveal interesting phenomena, the observations cannot be generalised. In this respect, my data provides the first comprehensive study of the response of the complete OR repertoire to a particular olfactory stimulation paradigm.

## 6.5 Functional impact of differences in OSN number.

As mentioned previously, it is not clear what is the functional consequence of altering the number of OSNs that express a particular OR. One hypothesis is that a greater number of detectors would result in enhanced sensitivity towards the odorants that are recognised with high affinity. To directly test this, I have created a transgenic mouse

line –in a B6 genetic background– where the CDS of *Olfr1507* (the most abundant OR in this strain) has been replaced by that of *Olfr2*. For this I utilised CRISPR-Cas9 technology. I created a vector for homologous recombination (HR) that contained the coding sequence of *Olfr2*, flanked by 1kb homology arms matching the *Olfr1507* locus. The vector was microinjected into B6 embryos, along with two guideRNA molecules that produce double-strand cuts in the intended site of HR. These embryos were then allowed to develop to term in foster mothers.

*Olfr1507* is expressed 35 times more abundantly than *Olfr2*. Since the abundance of a particular OR gene is controlled by the genetic architecture in *cis*, I expect to greatly increase the expression –and therefore the number of OSNs– of *Olfr2* in these animals. The response profile of *Olfr2* (better known as *I7*) has been very well characterised and, thus, these transgenic animals will provide an opportunity to assess the impact of increasing the cell number of a given OR on odour detection.