

Single-cell transcriptional profiles and spatial patterning of the mammalian olfactory epithelium

IAN TIETJEN¹, JASON RIHEL² and CATHERINE G. DULAC^{*,2}

¹Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine, Boston, MA, USA and

²Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA

ABSTRACT In order to gain insights into the regulatory control of neuronal diversity in the mammalian olfactory system, we have identified the transcriptional profile of individual olfactory neurons. A single cell microarray strategy was performed to search for candidate genes involved in the molecular specification of dorso-ventral zones of olfactory receptor (OR) expression. Several transcripts were identified that display differential expression in distinct OR zones, including a novel family of genes, the Lozenge-like (Lzl) genes which share sequence consensus motifs with Lozenge, a transcription factor involved in the patterning of the *Drosophila* olfactory and visual systems.

KEY WORDS: *olfactory, development, microarray, single cell*

Foreword

How is the enormous tissue and cell diversity of most multicellular organisms generated during embryonic development? The groundbreaking work of Nicole Le Douarin and her colleagues at the Institut d'Embryologie in Nogent-sur-Marne uncovered specific mechanisms of neural crest development into the peripheral nervous system and other non-neural derivatives, and offered to the scientific community a set of sophisticated tools – both intellectual and technological- to use in other experimental systems. The ability to manipulate individual neurons and precursors and more importantly, the development of experimental designs to uncover the developmental potentials of single neuronal precursors at the cellular, molecular or transcriptional levels, were themes largely developed in Nogent in the mid-80s and early 90s (Baroffio *et al.*, 1988; 1991). The study presented here, aimed at identifying the transcriptional profile of individual olfactory neurons and at gaining access into the transcriptional control of neuronal diversity, has benefited from the rich scientific legacy of research in the Nogent Institute.

Introduction

In the highly complex mammalian brain as well as in nervous systems of simpler organisms, like that of the nematode *Caenorhabditis elegans*, hundreds to thousands of different classes of neurons can be distinguished based on morphological, physiological and molecular criteria.

Genetic analysis in invertebrates has successfully identified cascades of signaling events involved in the differentiation of specific neuronal cell types; for example the emergence of distinct photoreceptor cells in the *Drosophila* eye (reviewed in (Wolff *et al.*, 1997) or the specification of chemosensory neurons in *C. elegans* (Lanjuin and Sengupta, 2002, Sagasti *et al.*, 1999, Sarafi-Reinach *et al.*, 2001, Sarafi-Reinach and Sengupta, 2000, Troemel *et al.*, 1999). In turn, the search for vertebrate homologs of *Drosophila* and *C. elegans* transcriptional regulators has led to significant insights into the molecular control of neurogenesis in higher organisms (reviewed in (Livesey and Cepko, 2001, Tanabe and Jessell, 1996). However, the much larger neuronal heterogeneity found in the vertebrate brain, together with the lack of direct experimental approaches to identify specific transcripts expressed by rare neuronal subpopulations, have considerably limited the scope of these searches (reviewed in Cao and Dulac, 2001).

In the olfactory epithelium of rodents, mature olfactory sensory neurons (OSNs) each express a single olfactory receptor gene from a large family of approximately one thousand genes. The olfactory epithelium appear loosely organized into large domains of receptor expression such that all neurons expressing the same

Abbreviations used in this paper: ION, immature olfactory neuron; Lzl, lozenge-like gene; MOE, mammalian olfactory epithelium; OR, olfactory receptor; OSN, olfactory sensory neuron; VPC, vomeronasal precursor cell; VSN, vomeronasal sensory neuron.

*Address correspondence to: Dr. Catherine G. Dulac. Howard Hughes Medical Institute, Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. e-mail: dulac@fas.harvard.edu

receptor transcript appear randomly dispersed within one of four broad zones segregated along a dorso-ventral axis (Buck and Axel, 1991, Chess, 1994, Malnic *et al.*, 1999, Ressler *et al.*, 1993, Vassar, 1993). Specific questions related to olfactory neurogenesis such as the regulation of olfactory commitment and the patterning of the epithelium into zones of receptor expression have so far been largely impaired by the extremely large cellular heterogeneity of the olfactory epithelium.

In order to gain direct access into the signaling events involved in olfactory neurogenesis, we have developed a technique to extensively characterize the transcriptional profiles of individual olfactory neurons and progenitor cells (Tietjen *et al.*, 2003). Here, we describe a similar single-cell microarray strategy aimed at analyzing the molecular specification of the dorso-ventral zones of olfactory receptor expression. This search led to the identification of several transcripts with differential expression in distinct OR zones, including a novel gene family called Lozenge-like (Lzl), which display sequence similarity with *Drosophila* Lozenge, a transcriptional regulator involved in the patterning of the fly olfactory and visual systems.

Results

General strategy

Olfactory neurogenesis was analyzed at the single-cell level by comparing the transcriptional profiles of mature olfactory sensory neurons (OSNs) and immature olfactory neurons (IONs). IONs are undergoing their differentiation process and display immature neuronal and olfactory characteristics. Because the respective ratios of these cell-types change as the animal develops, OSNs were obtained from adult olfactory epithelium (Tietjen *et al.*, 2003), while IONs were obtained from olfactory epithelium of newborn pups. Individual cells were picked at random from dissociated mouse olfactory epithelia at the respective ages and single-cell cDNA was obtained as previously described (Dulac and Axel, 1995). Southern blot analysis was then performed to ensure that single-cell cDNA samples contained strong expression of ubiquitous transcripts such as alpha-tubulin and GAPDH, in addition to verifying the represented cell-type. OSNs were characterized by strong expression of olfactory marker protein (OMP) and olfactory receptor (OR), while IONs were identified by the additional expression of developmental markers such as PCNA and GAP43 (data not shown). Neither OSNs nor IONs were found to express the transcription factors Mash1, Ngn1, or NeuroD, which are markers of earlier olfactory progenitor and precursor cells. Furthermore, *in situ* hybridization with the specific olfactory receptor probe isolated by degenerate PCR from each OSN- and ION-cDNA enabled us to unambiguously identify the epithelial zone from which each neuron originated (Figure 1A). In total, cDNAs from 9/45 adult and 10/19 neonatal olfactory epithelium cells were kept for further analysis and hybridized to Affymetrix Mu11K high-density oligonucleotide arrays.

Transcriptional profiles of single neurons and sensory precursors

During the course of our study, amplified single-cell cDNAs were prepared from a vast collection of sensory neurons including OSNs, sensory neurons of the adult VNO (VSNs) and photoreceptors of the retina. In addition, olfactory progenitor cells

(OPCs), which give rise to OSNs, were obtained from E15 olfactory epithelium (Tietjen *et al.*, 2003) and vomeronasal precursor cells (VPCs) were isolated from E16 VNO. OPCs were identified via Southern blot by their strong expression of Mash1, Ki67 and Cdc2 and by lack of OMP or olfactory receptor expression. VSNs were identified by the strong expression of OMP, while VPCs were identified by the expression of Mash1 or Ngn1.

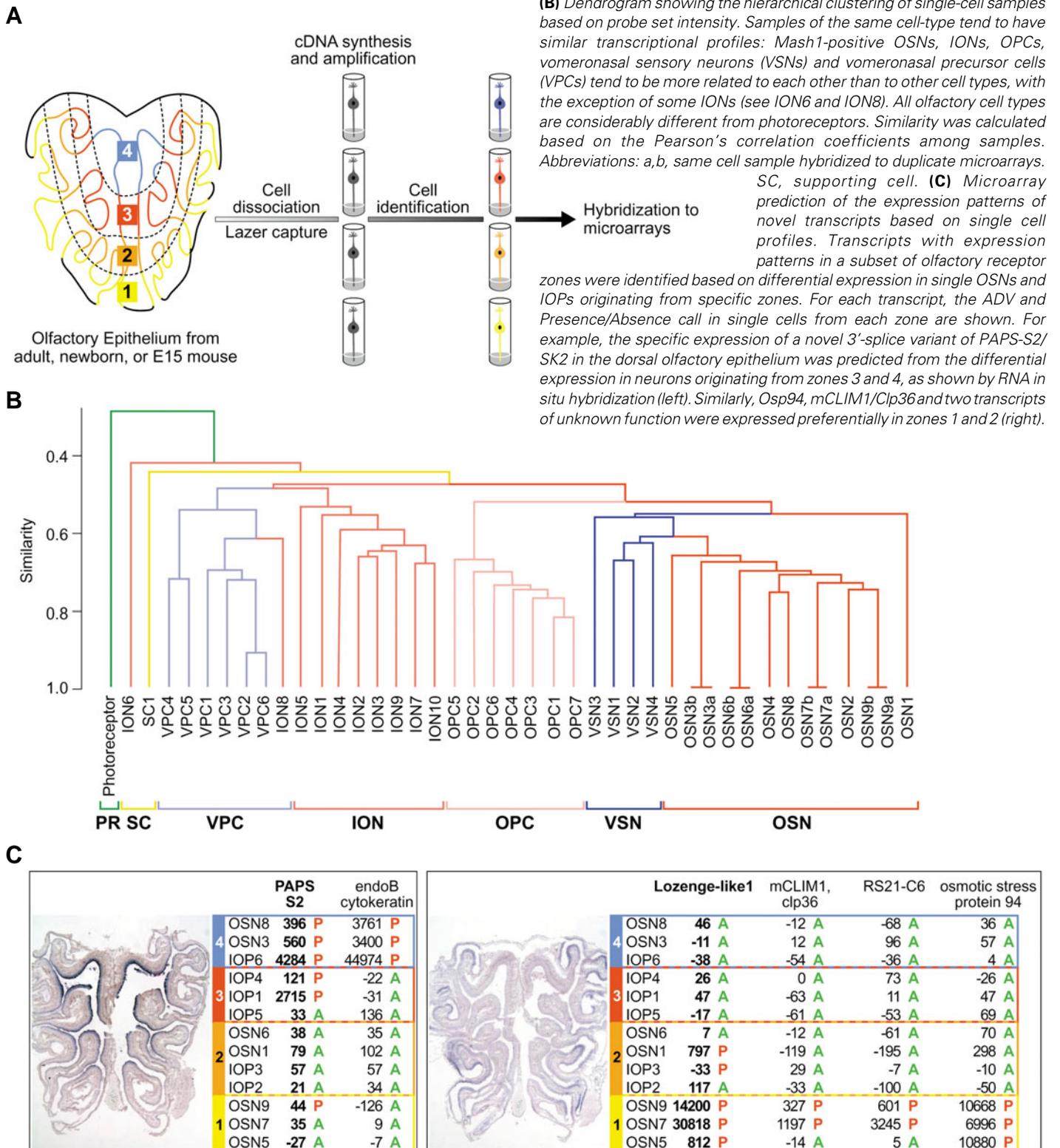
A broad comparison of gene expression across all single cell cDNAs (Figure 1B) confirms the tight clustering of profiles of cells of a given type, for example of all mature VNO neurons or of all olfactory progenitors, while distinct sensory types such as photoreceptor show considerable divergence. The hierarchical clustering of the cell profiles (Figure 1B) indicates that VNO and olfactory neurons, as well as VNO and olfactory precursors, share more resemblance to each other than each neuron type to its respective precursor. This clearly emphasizes the similarity between mature chemosensory neurons together with the tremendous changes in gene expression occurring during neuronal development. In contrast, the absence of co-clustering of the single cell profiles generated from neurons belonging to the same zone (OSN3, OSN8 and ION6 in zone 4, ION1, ION4 and ION5 in zone 3, OSN1, OSN6, ION2 and ION3 in zone 2 and OSN5, OSN7 and OSN9 in zone 1) (Figure 1B) suggests that, at least in the data collected from the 11K murine array, zone-specific transcriptional differences are likely to be relatively minor. However, despite the large number of developmental differences among different cell-types, the extreme sensitivity of single-cell transcriptional profiling still allowed us to identify several zone-specific transcriptional differences, as shown below.

Specific transcription in olfactory zones and the lozenge-like family

We reasoned that subtle changes in gene expression in mature olfactory neurons from different MOE zones could be identified by comparing the transcriptional profiles of single OSNs. Transcriptional profiles of OSNs originating from adult MOE (OSNs 1-9) have been described previously (Tietjen *et al.*, 2003). Using the same single-cell microarray method, we also obtained the transcriptional profiles of several OSNs from neonatal MOE (ION 1-7). The quality of cDNA samples representing OSNs and IONs was determined based on the strong expression of alpha-tubulin and GAPDH by Southern blot and the detection of a single olfactory receptor sequence by degenerate PCR. Following microarray hybridization, however, the average pairwise correlation coefficients of gene expression were considerably lower for neonatal IONs than for adult OSNs (0.53 +/- 0.01 versus 0.68 +/- 0.05), indicating a greater level of transcript heterogeneity among neonatal cells. Indeed, Southern blot and microarray analyses indicate that neonatal IONs have variable expression of markers of cell division (e.g., PCNA), axon growth (e.g., GAP43) and other developmentally-regulated events (e.g., RGS11, Sox11; data not shown), suggesting that these cells are still undergoing various steps of neuronal maturation.

To determine spatial differences in MOE neuronal gene expression, we first identified the original zone of each olfactory neuron using its olfactory receptor sequence as a probe for RNA *in situ* hybridization. Each olfactory receptor sequence gave robust expression in an apparently random subset of OSNs confined to one zone (data not shown). Based on 3-4 cells attributed to each zone, we then

Fig. 1. Monitoring transcription in single olfactory neurons. (A) Strategy. The olfactory epithelium of rodent is subdivided into four broad zones (colored numbers and lines), each characterized by the expression of a subset of olfactory receptor genes. In addition, each epithelium comprises cells at various stages of neuronal development. Olfactory epithelia dissected from adult, newborn, or e15 mice were dissociated. Individual olfactory neurons and olfactory precursors were picked at random and seeded into individual PCR tubes before undergoing cell lysis, cDNA synthesis and PCR amplification. The zone of origin and the developmental stage of each cell were retrospectively determined by PCR amplification of olfactory receptor genes present in each single cell cDNA and by Southern blot hybridization with developmental- and olfactory-specific markers. cDNA samples representing neurons or precursors of interest were then hybridized to Affymetrix Mu11K GeneChips. **(B)** Dendrogram showing the hierarchical clustering of single-cell samples based on probe set intensity. Samples of the same cell-type tend to have similar transcriptional profiles: Mash1-positive OSNs, IONs, OPCs, vomeronasal sensory neurons (VSNs) and vomeronasal precursor cells (VPCs) tend to be more related to each other than to other cell types, with the exception of some IONs (see ION6 and ION8). All olfactory cell types are considerably different from photoreceptors. Similarity was calculated based on the Pearson's correlation coefficients among samples. Abbreviations: a,b, same cell sample hybridized to duplicate microarrays. SC, supporting cell. **(C)** Microarray prediction of the expression patterns of novel transcripts based on single cell profiles. Transcripts with expression patterns in a subset of olfactory receptor zones were identified based on differential expression in single OSNs and IOPs originating from specific zones. For each transcript, the ADV and Presence/Absence call in single cells from each zone are shown. For example, the specific expression of a novel 3'-splice variant of PAPS-S2/SK2 in the dorsal olfactory epithelium was predicted from the differential expression in neurons originating from zones 3 and 4, as shown by RNA in situ hybridization (left). Similarly, *Osp94*, *mCLIM1*/*Clp36* and two transcripts of unknown function were expressed preferentially in zones 1 and 2 (right).



searched the microarray data for transcripts predicted to be enriched in the ventral-most (Zone 1) or dorsal-most (Zone 4) zone of the MOE (Figure 1A). As zone-specific transcripts are likely to represent a very low percentage of the total transcripts expressed by neonatal and adult OSNs, while developmental and other cell-intrinsic differences of OSNs are likely to represent a much higher percentage (Figure 1B), we observed that a large portion of candidate zone-specific transcripts represented false-positive signals. However, using a standard hierarchical clustering approach (Eisen *et al.*, 1998; see Experimental Procedures), two factors were successfully identified as preferentially expressed in the most dorsal zones of the adult MOE, while four transcripts were identified in the most ventral zones (Figure 1C). Similar results were also obtained using self-organizing map and confidence interval approaches (Tamayo *et al.*, 1999; Tietjen *et al.*, 2003; data not shown). Predictions made from the arrays were confirmed by *in situ* hybridization. Two transcripts were identified that display preferential expression in the most dorsal zones of the olfactory epithelium. One encodes a novel 3' splice variant of the sulfate-activating enzyme PAPS synthase2/SK2 and is expressed in a gradient with zone 4 showing the highest expression level (Figure 1C). The function of this gene in the adult olfactory epithelium is unknown but one may suspect that this enzyme activity could be related with differential glucoconjugate synthesis (Kurima *et al.*, 1998).

Conversely, several transcripts, including one encoding an alpha-interferon induced protein of unknown function, were found to be strongly and differentially expressed in the ventral zones of the olfactory epithelium (Figure 1C). This latter transcript captured our attention because it encodes a protein which displays distant but noticeable similarity with the runt domain-containing protein Lozenge, a transcription factor involved in the pre-patterning of the *Drosophila* eye and olfactory organs and in the regulation of the proneural gene *amos* in olfactory sensillae (Canon and Banerjee, 2000, Flores *et al.*, 1998, Goulding *et al.*, 2000). A 29% sequence similarity to Lozenge was found over a stretch of 147 amino acids encompassing part of the conserved DNA-binding and transactivating portions of the Lozenge runt domain (data not shown). Database searches and screening of a mouse olfactory cDNA library revealed a total of three mouse genes predicted to encode proteins that share a common motif similar to part of and extending downstream of the Lozenge runt domain. These genes are thus named Lozengelike 1-3 (Lzl1-3, Figure 2A). Similar searches also revealed numerous splice variants of Lzl1 and Lzl2 in mouse, including some that lack the conserved Lzl domain (Figure 2B). While most Lzl1 and Lzl2 sequences are predicted to contain one Lzl motif, Lzl3 is predicted to contain two motifs (Figure 2A). Lzl1-3 were all identified within a ~23 kb genomic region of mouse Chromosome 12E (data not shown), suggesting that they arose from recent gene duplication events. Although divergent Lzl sequences were identified in human, rat and zebrafish databases (data not shown), database searches revealed no obvious similarities to other genes or other known protein motifs.

Strikingly, two members of the Lzl family display zone-restricted expression in adult MOE. Lzl1a, the transcript originally identified by microarray, is specifically expressed in Zones 1 and 2, as determined by expression of olfactory receptor sequences with established zone identities on adjacent sections (Figure 2C). All Lzl1 splice variants investigated gave identical expression patterns ($n=5$, data not shown). In contrast, Lzl2a is detected in a subset of basal cells and underlying lamina propria found exclusively in the dorsal MOE (Figure 2C). By

in situ hybridization, we estimate the onset of detectable expression to the first postnatal week. Multiple splicing variants of lozenge-like1 were identified in the olfactory cDNA library while the expression of lozenge-like3 was not detected in this tissue. In addition, lozenge-like1 and 2 were found by Northern blot analysis in a variety of tissues including adult lung, liver, spleen and heart, indicating that their expression is not olfactory specific (data not shown).

Discussion

Neurons represent only a minor cellular component of sensory organs and, within each sensory epithelium, small subpopulations of receptor cells are specialized morphologically and molecularly for the detection of a restricted range of environmental cues. Subsets of photoreceptor cells in the retina specialize in the detection of photons of specific wavelengths and hair cells of the inner ear each develop a specific configuration of stereocilia and ionic currents to provide a narrow tuning to sound frequency. Similarly, the mammalian olfactory epithelium contains a highly heterogeneous and constantly renewing population of neurons and neuronal precursors. In the mouse, each mature olfactory sensory neuron (OSN) expresses a unique olfactory receptor gene from a large family of approximately one thousand genes and all neurons expressing the same receptor transcript are randomly dispersed within one of four broad zones of the olfactory epithelium (Buck and Axel, 1991, Chess, 1994, Malnic *et al.*, 1999, Ressler *et al.*, 1993, Vassar, 1993).

What are the transcriptional regulators underlying such sensory diversity? Genetic analysis of the mouse retina, inner ear and olfactory epithelium has pointed to the essential role played by basic helix-loop-helix (bHLH) containing transcription factors related to the *Drosophila* proteins *achaete-scute* and *atonal* in controlling the broad sensory identity of neuronal progenitors (reviewed in (Bertrand *et al.*, 2002). Progenitors of the inner ear express the *atonal* homolog *Math1* and mice with targeted deletion of *Math1* fail to form mature inner hair cells (Bermingham *et al.*, 1999). Similarly, mature olfactory sensory neurons do not develop in mice with a targeted deletion of the *achaete-scute* homolog *Mash1* (Guillemot *et al.*, 1993). Expression of *Mash1* in early olfactory progenitor cells (OPCs) controls expression of the bHLH containing transcription factors *Ngn1* and *NeuroD*, which in turn regulate olfactory differentiation (Cau *et al.*, 1997). However, within a given modality, the identity of transcriptional regulators by sensory progenitors to achieve sensory diversity, for example to choose a specific olfactory receptor, an opsin type or a ion channel isoform, has remained largely elusive, primarily because progenitor cells represent such a rare and diverse population in developing sensory epithelia. More than a decade after the discovery of the large family encoding olfactory receptors (Buck and Axel, 1991), the mechanism by which a single olfactory neuron chooses the expression of a unique receptor gene is still a mystery. Various hypotheses have been formulated that invoke the expression of complex combinations of transcription factors or the existence of somatic recombination or chromatin remodeling to position a regulatory element upstream of the chosen receptor gene. Regardless of the exact mechanism for single receptor choice, the mere existence of a dorso-ventral patterning of receptor expression into four zones of the olfactory epithelium implies mechanism of spatial signaling within the epithelium.

How can regulators of the zones of receptor expression be identified? In a first approach we aimed at investigating differences

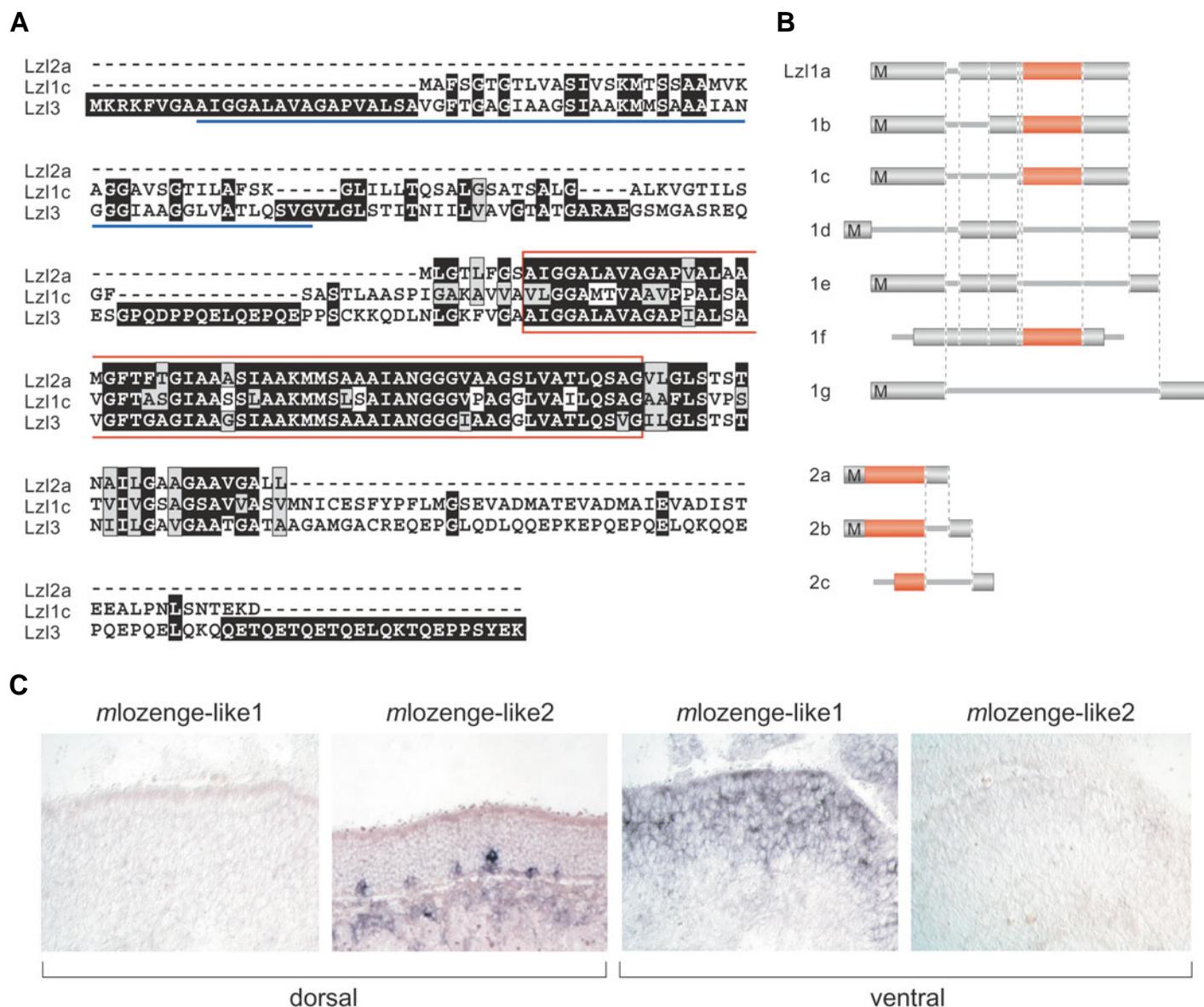


Fig. 2. The Lozengelike family. (A) Amino acid sequence alignment of *Lzl1-3*. The conserved region with similarity to the *Drosophila* *Lozenge* transcription factor is underlined in red. *Lzl3* also has a second conserved region, underlined in blue. **(B)** Schematic of the alignment of *Lzl1* and *Lzl2* splice variants. Overlapping amino acid sequences are shown in grey. The conserved *Lzl* domain is shown in blue. The putative start Methionine, when known, is indicated (M). **(C)** Expression patterns of *Lozengelike1a* and *2a* in the dorsal (zone 4) and ventral (zone 1) portions of adult MOE. *Lozengelike1a* is expressed throughout the ventral MOE, while *Lozengelike2a* is found only in basal cells and underlying lamina propria of the dorsal MOE.

in gene expression between neurons belonging to different zones of the adult epithelium. The mature olfactory epithelium is a particularly striking example of a composite neuronal structure in which precursors at different developmental stages as well as a thousand neuronal types grouped into four broad zones of the epithelium are interspersed. The extremely convoluted structure of the olfactory turbinates in rodents preclude any attempt to investigate molecular differences between the OR zones by surgical dissection of the epithelium. However, olfactory receptor identification from each OSN cDNA allowed us in turn to determine the OR zone of origin for each neuron and thus to uncover transcriptional differences between olfactory neurons belonging to distinct OR zones. Despite

the prediction that zonal differences would represent a minor portion of the cell profiles, the extreme sensitivity and cellular resolution provided by the single cell profiles enabled us to readily identify genes with differential expression between neurons belonging to different OR zones. Our search identified several closely related genes, called here *lozenge-like 1 to 3*, that share common motifs within a domain of the *lozenge* protein that comprises parts of and extends 3' to the runt domain (Canon and Banerjee, 2000, Flores *et al.*, 1998, Goulding *et al.*, 2000). The function of these genes is unknown, but significant differences within the canonical runt domain motifs make it unlikely that *lozenge-like* proteins play a role as genuine runt domain-containing transcription

factors. Strikingly, two members of this family display restricted expression in zones of the olfactory epithelium: lozenge-like 1 is specifically expressed in the ventral epithelium, while lozenge-like 2 is detected in a subset of basal cells found exclusively in the dorsal olfactory epithelium. Further studies will be required to ascribe a function of the Lzl family and determine whether it shares any functional property of sensory patterning with its closest homolog in the *Drosophila* genome Lozenge.

The large-scale analysis of transcriptional profiles obtained from individual olfactory neurons and olfactory progenitors has provided a unique snapshot of the different regulatory networks concurring or competing within a single cell at a specific developmental stage to control cell proliferation and olfactory specification. More generally, the single cell profiling approach presented here offers an unprecedented level of sensitivity and cellular resolution that will be of great use to uncover the molecular basis of neuronal identity in the brain.

Materials and Methods

The protocols for isolation of dissociated olfactory neurons and subsequent single cell RT-PCR is covered extensively in Dulac and Axel (1995) and Dulac (1998). Following Southern blot analysis of the single cell samples with a number of cell specific probes, the samples are re-amplified by PCR to generate enough starting material and then biotinylated for hybridization to microarrays.

Reamplification

For each reaction, the following solution is made on ice: 80 μ l of ultrapure H₂O, 10 μ l of 10X PCR buffer II, 10 μ l of 10X MgCl₂, 0.2 μ l each of 100 mM dATP, dCTP, dGTP, dTTP, 1 μ l of Amplitaq (Applied Biosystems) and 5 μ l of AL-1 primer (AL-1 sequence: 5' ATT GGA TCC AGG CCG CTC TGG ACA AAA TAT GAA TTC 3'). Next, 0.75 μ l of single cell cDNA is added. Each sample is run in triplicate and each set of reactions includes a negative control with no cDNA. On a thermocycler, the following PCR conditions are run: 94°C for 1 min 30 seconds, 42°C for 2 minutes and 72°C for 3 minutes. This program is repeated for 30 cycles. The samples are then held to 72°C for 20 minutes, followed by 4°C until the samples are ready to process. Each tube is then purified separately using a Qiagen PCR purification kit. Samples are eluted from the column with 30 μ l of warm EB buffer. The triplicate reactions are then pooled for further analysis. The total amount of cDNA following re-amplification and purification should be between 20-50 μ g. Southern blots are used to check the re-amplified cDNA and ensure that the re-amplification worked properly.

Microarray hybridization

The samples are first briefly digested with DNase in a reaction containing 15 μ g of purified single cell cDNA, 6 μ l of 10X One-Phor-All buffer (Amersham Pharmacia), 3.75 μ l of 0.2 U/ μ l RQ1 DNase (diluted in 10mM Tris, pH 7.5; Promega) and EB buffer (Qiagen kit) to a final volume of 60 μ l. The samples are incubated at 37°C for 13 minutes and then heat inactivated at 99°C for 15 minutes. After a brief spin in a centrifuge to collect the sample, 1.5 μ l of Biotin-N₆-ddATP (1.0 mM; Perkin Elmer) and 2.25 μ l of terminal transferase (15U/ μ l; Invitrogen) are added to the samples. The samples are incubated at 37°C for 1.5 hours and then heat inactivated at 65°C for 15 minutes. The samples are then stored on ice for 5 minutes.

The final cocktail that is used for hybridization to GeneChips follows the standard Affymetrix hybridization protocols. First, the B2 oligo and 20X eukaryotic controls are thawed at 65°C for 5 minutes. The following reagents are then added to the labeled cDNA samples: 5 μ l of B2 control (Affymetrix), 15 μ l of 20X eukaryotic hybridization controls (Affymetrix), 3 μ l of herring sperm DNA (10mg/mL), 3 μ l of Acetylated BSA (50mg/mL), 150 μ l of 2X chip hybridization buffer (Affymetrix) and H₂O for a final volume of 300 μ L. This

hybridization cocktail is heated to 9°C for five minutes, 45°C for five minutes and spun at 13,000 rpm on a tabletop centrifuge for five minutes. The samples are now ready for hybridization.

During preparation of the hybridization cocktail, the Genechips are warmed to room temperature. The arrays are then wet with 1X hybridization buffer and incubated at 45°C for 10 minutes while rotating at 50 rpm. Finally, the 1X buffer is removed and 200 μ l of single cell cDNA hybridization cocktail is added to each standard array. Hybridization occurs overnight at 45°C for 16 hours while rotating at 50 rpm.

Subsequent washing and staining of the Genechips follows the Affymetrix protocols for standard arrays as detailed in the Affymetrix technical manual.

Analysis

For any given single cell sample, approximately 20-25% of the probes are called «Present» by the Affymetrix Genechip software v3.2 (using the following parameters: positive/negative minimum, 2.2; positive/negative maximum, 3.0; positive ratio minimum, 0.24; positive ratio maximum 0.33; average log ratio minimum, 0.80; average log ratio maximum, 1.2). Replicate single cells (i.e. of the same cell type) typically have correlation coefficients in the range of 0.60-0.70. To effectively search for differentially expressed genes with such high within-sample variability, a number of strategies are employed. First, multiple single cell samples are collected for each cell type of interest. Typically, at least four or five examples are needed to average out cell-cell variability. Second, when using hierarchical clustering, self-organized maps, or other data analysis software to visualize differences in gene expression, the stringencies of the parameters are lowered to account for the increased variability. For example, genes expressed highly in only a fraction of a specific cell type could be considered candidate genes, even if they fall below statistical significance using standard methods (e.g. Student's t-test). A third way to deal with cell-cell variability is to use data from single cultured cells, which have high correlation coefficients (in the .9 range), to generate a baseline of cell-to-cell noise. After plotting the best-fit regression curve for the cultured cell data with mADV > 250, another curve is deduced that encompasses 95% of the data points. This curve represents a statistically significant confidence interval for future data analysis (see Tietjen *et al.* 2003 for a more detailed discussion).

With these considerations, a given transcript was considered to have significant differences between two cell types based on three criteria. First, the mADV of the probe set, inclusive of all single cell samples, should be equal to or greater than 250. Second, the variability in probe set expression level, as defined by the ratio of mADVs between the two cell types, should exceed the fold-change threshold derived from the confidence interval described above. Finally, the corresponding probe set should display differences in ADV levels between the two types of samples that are considered statistically significant by the Student's unpaired t-test ($p < 0.05$). This is a stringent set of criteria to reduce the number of false positives. However, it is important to also consider candidate genes that fall below these thresholds, especially genes expressed in only subsets of cell types.

Clustering was performed using the Agglomerative Hierarchical Clustering feature of the XLstat Pro 6.1 software for PC. Clustering was based on the Pearson's correlation coefficient and the aggregation used the weighted-pair group average.

Isolation of Olfactory Receptor Sequences from single OSN samples

Olfactory receptor sequences were isolated from single OSN and ION samples by degenerate PCR essentially as described in Malnic *et al.* (1999). In no case was more than one olfactory receptor sequence isolated per single-cell sample.

Identification of OSN transcripts with zone-restricted expression

Candidate transcripts in OSNs with expression segregated according to OR zones boundaries were identified by hierarchical clustering using Gene Cluster and TreeView v. 1.45 software. Average linkage clustering on normalized genes was performed for all transcripts in which at least 3 cells had an ADV \geq 500. This approach predicted numerous zone-

restricted transcripts that were predicted to have strong and significant expression exclusively in Zone 1 or Zone 4, of which only a few (~20%) were confirmed by *in situ* hybridization as predicted. Transcripts (Corresponding probe set) confirmed by *in situ* hybridization include PAPS S2 (aa242579_s_at), endoB cytokeatin (m11686_s_at), Lzl1a (aa711625_s_at) and osmotic stress protein 94 (u23921_s_at). To identify additional zone-restricted transcripts in OSNs, a second round of hierarchical clustering was performed on all transcripts in which at least 1 cell had an ADV \geq 250. Two additional transcripts were identified that fit into a highly similar gene expression cluster with Lzl1a and were also confirmed by *in situ* hybridization. These transcripts (Corresponding probe set) are mCLIM1/clp36 (aa509478_s_at) and RS21-C6 (aa516966_s_at).

Lzl family analysis

Lzl1a and Lzl2a fragments were initially obtained by PCR from mouse whole MOE cDNA using specific primers. To obtain full-length Lzl sequences, a Lambda-ZAP cDNA library generated from adult MOE was probed with radiolabeled Lzl1a or Lzl2a PCR products according to manufacturer's instructions (Stratagene). This strategy identified full-length Lzl1a, Lzl2a and additional Lzl1 isoforms. BLAST searches were performed on NCBI (<http://www.ncbi.nlm.nih.gov>) and TIGR sequence databases (<http://www.tigr.org>) to identify additional Lzl1 isoforms, Lzl2 isoforms and Lzl3. Amino acid motifs were analyzed using SMART (<http://smart.embl-heidelberg.de>). Genomic analysis was performed using the UCSC mouse genome browser (<http://genome.ucsc.edu/>).

In situ hybridization analysis

RNA *in situ* hybridization was performed as described (Schaeren-Weimers and Gerlin-Most, 1993). MOEs were dissected from mice in age ranging from newborn to adult. Tissues were either fixed in 4% paraformaldehyde or freshly embedded in Tissue-Tek OCT compound (Sakura).

Acknowledgements:

This work was in part supported by the Howard Hughes Medical Institute and a Wellcome Trust Functional Genomics Programme Grant (066790/E/02/Z).

References

- BERMINGHAM, N.A., HASSAN, B.A., PRICE, S.D., VOLLRATH, M.A., BEN-ARIE, N., EATOCK, R.A., BELLEN, H.J., LYSAKOWSKI, A. and ZOGHBI, H.Y. (1999). Math1: An essential gene for the generation of inner ear hair cells. *Science* 284: 1837-41.
- BERTRAND, N., CASTRO, D.S. and GUILLEMOT, F. (2002). Proneural genes and the specification of neural cell types. *Nat Rev Neurosci* 3: 517-30.
- BUCK, L. and AXEL, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell* 65: 175-187.
- CANON, J. and BANERJEE, U. (2000). Runt and lozenge function in *Drosophila* development. *Semin Cell Dev Biol* 11: 327-36.
- CAO, Y. and DULAC, C. (2001). Profiling brain transcription: Neurons learn a lesson from yeast. *Curr Opin Neurobiol* 11: 615-20.
- CAU, E., GRADWOHL, G., FODE, C. and GUILLEMOT, F. (1997). *Mash1* activated a cascade of bhlh regulators in olfactory neuron progenitors. *Development* 124: 1611-1621.
- CHESS, A., SIMON, I., CEDAR, H., AXEL, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. *Cell* 78: 823-834.
- FLORES, G.V., DAGA, A., KALHOR, H.R. and BANERJEE, U. (1998). Lozenge is expressed in pluripotent precursor cells and patterns multiple cell types in the *Drosophila* eye through the control of cell-specific transcription factors. *Development* 125: 3681-7.
- GOULDING, S.E., ZUR LAGE, P. and JARMAN, A.P. (2000). Amos, a proneural gene for *Drosophila* olfactory sense organs that is regulated by lozenge. *Neuron* 25: 69-78.
- GUILLEMOT, F., LO, L.C., JOHNSON, J.E., AUERBACH, A. and ANDERSON, D.J. and JOYNER, A.L. (1993). Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* 75: 463-76.
- KURIMA, K., WARMAN, M.L., KRISHNAN, S., DOMOWICZ, M., KRUEGER, R.C., JR., DEYRUP, A. and SCHWARTZ, N.B. (1998). A member of a family of sulfate-activating enzymes causes murine brachymorphism. *Proc Natl Acad Sci USA* 95: 8681-5.
- LANJUIN, A. and SENGUPTA, P. (2002). Regulation of chemosensory receptor expression and sensory signaling by the kin-29 ser/thr kinase. *Neuron* 33: 369-81.
- LIVESEY, F.J. and CEPKO, C.L. (2001). Vertebrate neural cell-fate determination: Lessons from the retina. *Nat Rev Neurosci* 2: 109-18.
- MALNIC, B., HIRONO, J., SATO, T. and BUCK, L.B. (1999). Combinatorial receptor codes for odors. *Cell* 96: 713-23.
- RESSLER, K.J., SULLIVAN, S.L. and BUCK, L.B. (1993). A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* 73: 597-609.
- SAGASTI, A., HOBERT, O., TROEMEL, E.R., RUVKUN, G. and BARGMANN, C.I. (1999). Alternative olfactory neuron fates are specified by the lim homeobox gene *lim-4*. *Genes & Development* 13: 1794-1806.
- SARAFI-REINACH, T.R., MELKMAN, T., HOBERT, O. and SENGUPTA, P. (2001). The lin-11 lim homeobox gene specifies olfactory and chemosensory neuron fates in c. *Elegans*. *Development* 128: 3269-81.
- SARAFI-REINACH, T.R. and SENGUPTA, P. (2000). The forkhead domain gene *unc-130* generates chemosensory neuron diversity in c. *Elegans*. *Genes Dev* 14: 2472-85.
- TANABE, Y. and JESSELL, T.M. (1996). Diversity and pattern in the developing spinal cord. *Science* 274: 1115-23.
- TROEMEL, E.R., SAGASTI, A. and BARGMANN, C.I. (1999). Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in c. *Elegans*. *Cell* 99: 387-98.
- VASSAR, R., NGAI, J., AXEL, R. (1993). Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* 74: 309-318.
- WOLFF, T., MARTIN, K., RUBIN, G. and ZIPURSKY, S.L. (1997). The development of the *drosophila* visual system. In *Molecular and cellular approaches to neural development*, (ed. COWAN, W. M. JESSELL, T. and ZIPURSKY, S. L.). Oxford University Press, New York, pp.474-508.