Neuron, Vol. 38, 161–175, April 24, 2003, Copyright ©2003 by Cell Press

Single-Cell Transcriptional Analysis of Neuronal Progenitors

Neurotechnique

Ian Tietjen,[1,4] Jason M. Rihel,[1,4] Yanxiang Cao,[2,4] Georgy Koentges,[1,3] Lisa Zakhary,[1] and Catherine Dulac[1,*]
[1]Howard Hughes Medical Institute
Department of Molecular and Cellular Biology
Harvard University
Cambridge, Massachusetts 02138
[2]Department of Applied Research
Affymetrix Inc.
3380 Central Expressway
Santa Clara, California 95051
[3]Functional Genomics
Wolfson Institute of Biomedical Research
University College London
The Cruciform Building
Gower Street
London WC1E 6AU
United Kingdom

Summary

The extraordinary cellular heterogeneity of the mammalian nervous system has largely hindered the molecular analysis of neuronal identity and diversity. In order to uncover mechanisms involved in neuronal differentiation and diversification, we have monitored the expression profiles of individual neurons and progenitor cells collected from dissociated tissue or captured from intact slices. We demonstrate that this technique provides a sensitive and reproducible representation of the single-cell transcriptome. In the olfactory system, hundreds of transcriptional differences were identified between olfactory progenitors and mature sensory neurons, enabling us to define the large variety of signaling pathways expressed by individual progenitors at a precise developmental stage. Finally, we show that regional differences in gene expression can be predicted from transcriptional analysis of single neuronal precursors isolated by laser capture from defined areas of the developing brain.

Introduction

In the highly complex mammalian brain, thousands of different classes of neurons can be distinguished based on morphological, physiological, and molecular criteria. Developmental neuroscience must tackle the formidable challenge of uncovering the numerous and intricate cascades of signaling events and regulatory networks underlying the establishment of neuronal diversity and specificity. Genetic approaches in invertebrates have successfully identified discrete classes of genes involved in the emergence of specific neuronal cell types (Sagasti et al., 1999; Wolff et al., 1997). In turn, the identification of vertebrate homologs of *Drosophila* and *C. elegans* transcriptional regulators has led to significant insights into the transcriptional control of neurogenesis in higher organisms (Bertrand et al., 2002; Tanabe and Jessell, 1996). However, these "vertical" approaches can only attempt to reconstruct in vertebrates the regulatory hierarchies that have already been discovered in other species. In turn, parallel approaches of expression profiling, as afforded by microarray analysis, have provided new guiding principles to illuminate the fine structure of regulatory networks (Panda et al., 2002). However, the extraordinary cellular and molecular complexity of the mammalian nervous system has considerably limited the scope of present microarray approaches (Cao and Dulac, 2001), as highly specific transcripts within rare precursor populations are rendered undetectable by homogenizing whole brain areas for RNA isolation.

We have developed an experimental protocol in which single-cell cDNA synthesis (Dulac and Axel, 1995; Tanabe et al., 1998; Yamagata et al., 2002) is combined with Gene chip analysis and laser capture-mediated cell isolation. To test the accuracy and reliability of this procedure, we have focused our analysis on the mammalian olfactory system, which includes the highly heterogeneous population of olfactory sensory neurons (OSN) in the main olfactory epithelium (MOE) and the progenitors of OSN target cells, the mitral cells of the olfactory bulb. Mature OSNs each express a unique olfactory receptor gene from a large family of a thousand genes. The MOE is partitioned into four broad zones of receptor expression such that, within each zone, neurons expressing distinct receptors are randomly interspersed (Buck, 2000). Despite advances on the essential role played by basic helix-loop-helix (bHLH) containing transcription factors Mash1, Ngn1, and NeuroD in ensuring olfactory differentiation (Cau et al., 1997, 2002; Guillemot et al., 1993), the identity of transcriptional regulators involved in further sensory diversity, including the choice of a given olfactory receptor or the formation of the MOE zones, has remained largely elusive, primarily because progenitor cells represent such a rare and diverse population.

We show here that a genome-wide transcriptional analysis can be performed successfully at the single-cell level, whether the cells are isolated from dissociated tissue or laser captured from intact brain slices. The unprecedented sensitivity and cellular resolution of these large-scale, single-cell molecular fingerprints have led us to uncover multiple regulatory pathways underlying olfactory neuronal diversity.

Results

General Strategy

Olfactory neurogenesis was analyzed at the single-cell level by comparing the transcriptional profiles of mature olfactory sensory neurons (OSNs) and olfactory progenitor cells (OPCs) (Figure 1). OPCs are actively dividing neuroepithelial cells committed to the olfactory lineage, which ultimately give rise to fully mature OSNs (Murray

*Correspondence: dulac@fas.harvard.edu
[4]These authors contributed equally to this work.

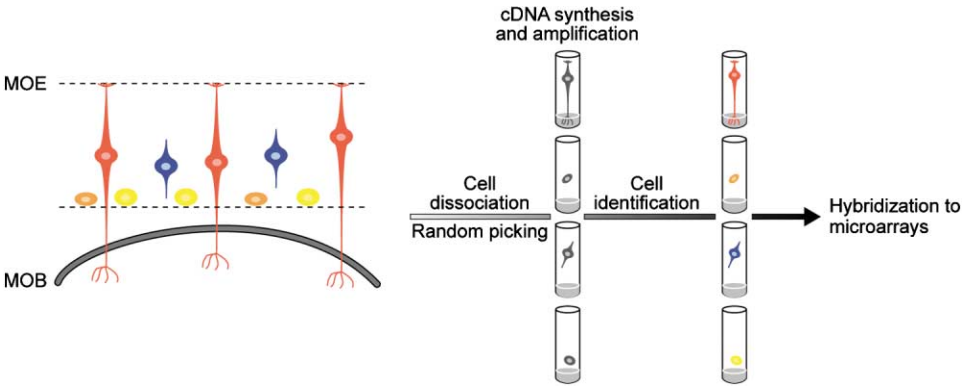


Figure 1. Monitoring Transcription in Single Olfactory Neurons and Precursors

The main olfactory epithelium (MOE) of mammals continuously regenerates OSNs and comprises neuronal precursors at various stages of development (cells with different shapes and colors). Individual OSNs or OPCs were picked at random from dissociated MOEs and seeded into individual PCR tubes before undergoing cell lysis, first-strand cDNA synthesis, and PCR amplification. The identity and the developmental stage of each cell were retrospectively determined by PCR Southern blot analysis of single-cell cDNAs to identify presence of developmental- and olfactory-specific markers. cDNA samples representing OSNs or OPCs of interest were then hybridized to Affymetrix Mu11K GeneChip probe arrays. MOB, main olfactory bulb.

and Calof, 1999). The respective ratios of these three cell types change as the animal develops, such that OPCs are abundant in the embryonic epithelium but are very rare in the adult. This prompted us to harvest OSNs and OPCs from adult and embryonic day 15 (E15) MOE, respectively, to increase the probability of successfully isolating multiple representatives of each cell type.

Individual cells were picked at random under the microscope from dissociated MOEs and seeded into lysis buffer-containing PCR tubes as previously described (Dulac and Axel, 1995; see Experimental Procedures). Subsequent reverse transcription and PCR amplification of the whole-cell transcriptome typically generated 10 to 20 μg of single-cell cDNA. The cell type represented by each single-cell cDNA was identified retroactively by Southern blot analysis and PCR amplification aimed at detecting various markers (Figure 1 and see Experimental Procedures). OSNs were characterized by the robust expression of olfactory marker protein (OMP) and olfactory receptor (OR) together with the lack of Mash1. In turn, OPCs were identified by the robust expression of Mash1, by the cell division markers cdc2 (Riabowol et al., 1989) and Ki67 (Scholzen and Gerdes, 2000), and by the absence of OMP and OR.

A stringent selection was performed to eliminate single-cell cDNAs failing to display robust expression of ubiquitous transcripts such as α-tubulin and GAPDH as well as cell type-specific markers. In total, cDNAs from 9/45 adult and 7/45 E15 MOE cells were kept for further analysis. To obtain sufficient amounts of cDNA for microarray hybridization from each sample, an additional round of PCR amplification was performed (see Experimental Procedures). Samples were then rechecked for the robust expression of all appropriate markers by Southern blot and PCR amplification. Ten micrograms of each single-cell cDNA sample was then labeled and hybridized to Affymetrix Mu11K high-density oligonucleotide arrays containing 13027 probe sets (see Experimental Procedures). Using the Affymetrix GeneChip analysis software, the expression level of each probe set was quantified by both the average difference value (ADV) and the "Present/Absent" absolute call algorithm (Lockhart et al., 1996; see Supplemental Data set S1 at http://www.neuron.org/cgi/content/full/38/2/161/DC1).

General Characteristics of the Single-Cell Profiles

The hybridization of amplified single-cell cDNAs to microarrays is an entirely unvalidated method; therefore, prior to any analysis of biological significance of the single-cell profiles, we investigated their general characteristics, including the reproducibility, complexity, and fidelity of the data obtained.

Reproducibility of Microarray Hybridization

Duplicate hybridizations of the same OSN cDNA to distinct batches of microarrays yielded nearly identical profiles with an average correlation coefficient of 0.97 (Figure 2A, left; Table 1). In contrast, comparison of the transcriptional profiles of two individual OSNs revealed larger variation with a lower correlation coefficient (Figure 2A, second from left; Table 1), presumably reflecting biologically significant differences in gene expression.

Representation of Transcriptome Complexity

Is the complexity of the single-cell cDNA sufficient to provide useful transcriptional information profiles? Approximately 20%–25% of the total number of probe sets in OSNs and OPCs predicted a "Present" signal by the Affymetrix GeneChip analysis software algorithms (2657 ± 347 and 3158 ± 1046 probe sets, respectively). Although these numbers are somewhat lower than those obtained from whole tissue such as heart, liver, and MOE (3248, 4437, and 3612, respectively), these values are comparable, suggesting that a large number of genes are detected in amplified single-cell cDNA samples.

Is the distribution of rare to abundant genes comparable in single-cell cDNA and whole-tissue RNA? On average, approximately 38.4% of all Present probe sets per OSN sample detected transcripts at a very low abundance level, as defined by a mean ADV (mADV) of 250 or lower (Figure 2B). In contrast, 6.1% of all Present probe sets per OSN detected transcripts with an mADV $>$10,000, suggesting that few transcripts were

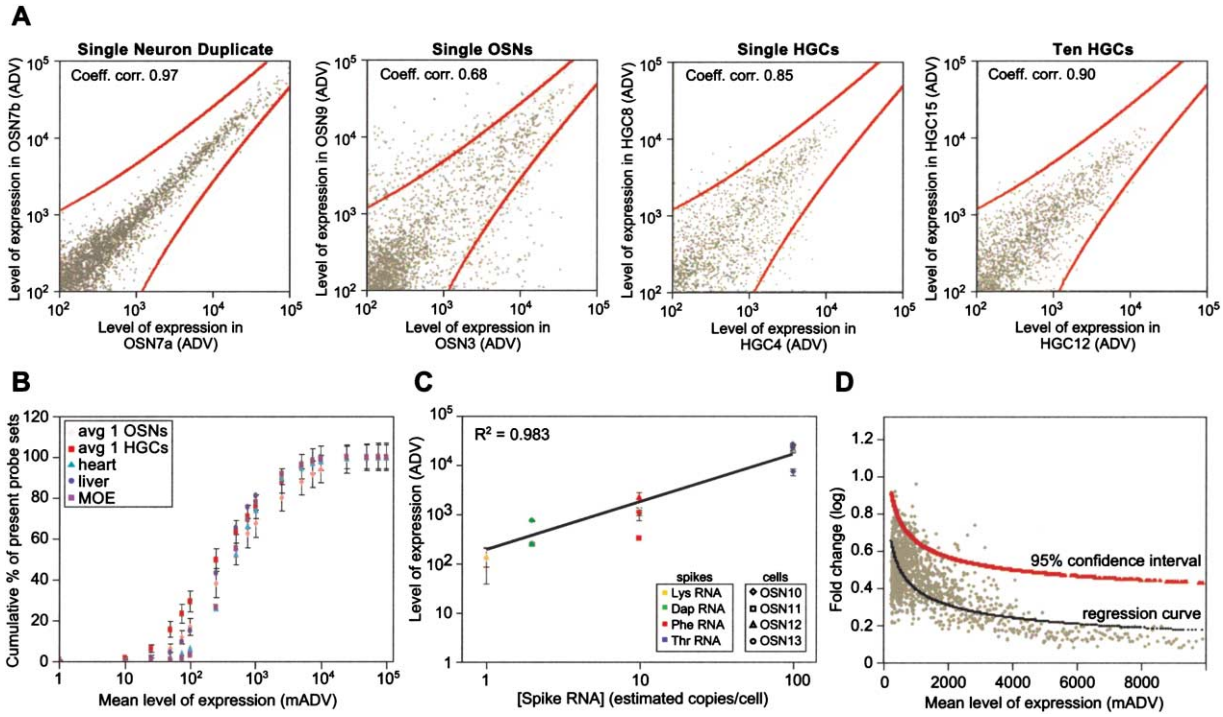


Figure 2. Accuracy of Transcript Amplification in Single-Cell cDNA Samples

(A) Reproducibility of the single-cell cDNA profiles. Scatter plots of data obtained from two repeat hybridizations of the same olfactory neuron cDNA (OSN7, left), from hybridization of cDNAs from two olfactory neurons (OSN3 and OSN9, second from left), from two single HGCs (HGC4 and HGC8, second from right), and from two samples of ten HGCs (HGC12 and HGC15, right). Each data point represents an individual probe set whose hybridization intensity is described by the ADV. As seen on the left panel, the duplicate hybridization to microarrays is highly reproducible (correlation coefficient = 0.97), and "identical" cells (single HGC correlation coefficient = 0.85) or groups of "identical" cells (ten HGC correlation coefficient = 0.90) provide similar profiles. In contrast, cDNA samples from different OSNs have considerably more variability (correlation coefficient = 0.68). Red curves denote the equation describing the 95% confidence interval in (D). The large majority of data points in all examples lies between the two red curves and denote transcripts with lower variability. However, a subset of data points predicts differentially expressed transcripts in one of two OSNs.

(B) Distribution of transcripts according to expression level. For each cell or tissue type, data points represent the cumulative percentage of probe sets predicted to be "Present" at a given mean average difference value (mADV) or lower (mean ± SEM). Comparison of data obtained from single olfactory sensory neurons (1 OSNs), single human glioblastoma cells (1 HGCs), unamplified heart, liver, and MOE RNA indicates that the single-cell cDNA amplification does not dramatically distort the overall distribution of rare (ADV $<10^2$), moderate (ADV = 10^2 to 10^4), and high (ADV > 10^4) abundant transcripts.

(C) Detection of transcripts at known copy number in single cells. Control bacterial poly(A)-tailed RNAs of known concentrations were "spiked" into the single-cell lysis buffer. Individual cells were seeded into the spiked lysis buffer, followed by cDNA synthesis and amplification. A total of four RNA spikes (dots of different colors), corresponding to roughly 1, 2, 10, and 100 RNA copies per cell, were spiked into four single OSN samples (dots of different shapes). Each data point represents duplicate microarray hybridizations (mean ± SEM). Spike transcripts added at ~100 copies/cell were detected in four of four experiments with a mADV of 19,959 ± 4,652, while spikes added at ~10 copies/cell were consistently detected with a mADV of 1,367 ± 451. Thus, the mADV of spikes at these concentrations appears proportional to the estimated spike copy number in the original single cell. Moreover, spiked transcripts originally present at ~1 and ~2 copies per cell were each amplified and detected by the array in two of four experiments, with a mADV in these experiments of 189 ± 54 and 525 ± 145, respectively.

(D) Variability in gene expression as a function of mADV. The average variability in gene expression (log fold change) was scored for all pairwise combinations of single HGCs as a function of the mADV (see Experimental Procedures). The regression curve that best fits the data at a mADV ≥ 250 approximates the level of variability that can be expected at a given mADV (black curve). A curve representing the 95% confidence interval is then determined (red curve). When comparing two cell samples or groups of cell samples, a probe set with a log fold change greater than that defined by the 95% confidence interval equation may be considered to have significant variability in expression.

detected at high abundance. Comparison with data obtained with unamplified tissue samples (Figure 2B) suggests that the relative proportion of transcripts expressed at low, moderate, and high levels has not been significantly altered in the single-cell cDNA synthesis and amplification.

Hybridization Efficiency of Single-Cell cDNA

Single-cell cDNA represents only the 3′-most 600 bp of mRNA transcripts. Is the short single-cell cDNA template hybridizing efficiently to microarray probe sets? Numerous transcripts expected to be present in OSNs or OPCs were successfully detected by the microarrays including, but not limited to, constitutive neuronal and OSN-specific markers (see Supplemental Table S1 and Data set S1 at http://www.neuron.org/cgi/content/full/38/2/161/DC1). In addition, the single-cell profiles for a large number of transcripts including the diagnostic probes described above, OCAM, RGS11, and five expressed sequence tags of unknown function, were verified independently by Southern blot. Microarray probe sets and Southern blot analyses of these markers in single-cell cDNA samples were consistently in agreement (76 of 89

Table 1. Statistical Analysis of Single-Cell Transcriptional Profiles

Cell Type		Pairwise Corr. Coeff. (Mean ± SD)
OSN (repeat)	n = 4	0.97 ± 0.01
Single HGC	n = 6	0.86 ± 0.03
Groups of 10 HGCs	n = 5	0.92 ± 0.04
10 pg dilution	n = 6	0.84 ± 0.03
100 pg dilution	n = 7	0.95 ± 0.02
OSNs	n = 9	0.68 ± 0.05
OPCs	n = 7	0.71 ± 0.05
OSNs versus OPCs		0.57 ± 0.04
MOBCs	n = 5	0.58 ± 0.02
AOBCs	n = 5	0.60 ± 0.02
MOBCs versus AOBCs		0.54 ± 0.03
MOE 1 versus 2	n = 2	0.88
MOE versus heart		0.42

Pairwise correlation coefficients of gene expression levels (mean ± standard deviation) are shown for various cell and tissue samples, as measured by the Average Difference Value (ADV) for each microarray probe set.

trials, 85.4%, not shown). Disagreements usually resulted from the 5′ localization of probe sets that could not accommodate the 3′ biased single-cell cDNA. This was the case for OMP which, although strongly detected by Southern blot, was rarely detected by microarray because the OMP probe sets are located 2 kb upstream of the transcript 3′ end (probe set information is available at http://www.netaffx.com). We also noted that Mash1 consistently failed to appear on microarray data obtained from Mash1-positive samples, including single-cell cDNAs (our data) and a variety of embryonic neural tissues (not shown), suggesting an unidentified problem with this particular probe set. Thus, it appears from our initial set of controls that single-cell cDNA templates contain complex transcriptional information that can be successfully and reproducibly analyzed by hybridization to microarrays.

Proportionality and Reproducibility of cDNA Amplification

A critical issue for the validity of our experimental strategy is to determine whether the respective ratios of transcripts in the amplified single-cell cDNA represent the original single-cell transcriptome. Although the nature of the original single-cell RNA is inaccessible, we reasoned that the degree of confidence in our procedure could be directly evaluated by the reproducibility of data obtained from multiple samples containing low (i.e., cellular) amounts of identical starting material.

"Spiking" Experiments

To provide a first measure of the accuracy of single-cell PCR amplification, we spiked four distinct poly(A)-tailed bacterial RNAs, each diluted to a known concentration, into the lysis buffer of each tube (see Experimental Procedures). A single cell was then added, and the regular cDNA synthesis and amplification procedure was performed. The experiment was performed four times. Subsequent analysis of control probe sets specific for the 3′ region of each bacterial spike indicated that the ADVs were roughly proportional to the estimated copy number of each spike mRNA (Figure 2C). Moreover, bacterial spikes at concentrations as low as ~1 and ~2 copies per cell were detected in two out of four experiments. The variable detection of bacterial spikes at these dilutions could result from either the stochastic absence of these spikes in two of four experiments, or from the lower limits of detection. In either case, these data illustrate the ability of the single-cell cDNA synthesis and amplification procedure to conserve the respective ratios of transcripts between the original single-cell RNA and the amplified single-cell cDNA. Moreover, they illustrate the ability to consistently amplify and detect low to extremely low levels of transcription in the single-cell profiles.

Genome-Wide Transcriptional Similarity among Identical Cells

A transcriptome-wide measure for the accuracy of the single-cell cDNA synthesis was performed by comparing the entire profile of individual cells and groups of cells obtained from a homogeneous population. A T98G human glioblastoma cell line (HGC) was synchronized at Go-phase by serum starvation. Individual cells (HGC1 to 10) and suspensions of ~10 cells (HGC11 to 15) were isolated, and amplified cDNAs were hybridized to Affymetrix HuGeneFL microarrays containing 7070 probe sets (see Experimental Procedures and Supplemental Data set S5 at http://www.neuron.org/cgi/content/full/38/2/161/DC1). The relative number and distribution of probe sets predicted to be Present were comparable to those of olfactory samples (Figure 2B and data not shown). Strikingly, correlation coefficients of gene expression in pairwise comparison of profiles of single HGCs and of groups of 10 HGCs, with an average of 0.86 ± 0.03 and 0.92 ± 0.04, respectively (Figure 2A, right and second from right; Table 1), equaled those of whole MOE profiles and approached those obtained by duplicate hybridization of the same single-cell cDNA to distinct microarrays, demonstrating that "identical" cells indeed generate similar transcriptional profiles. Additionally, correlation coefficients of gene expression obtained by pairwise comparison of 10 HGCs were similar to that of 1 HGCs (Figure 2A, Table 1), suggesting that a 10-fold increase in the amount of starting material does not significantly increase the quality of the procedure.

Amplification from RNA versus Single-Cell Sources

To test further the fidelity of single-cell PCR amplification, total RNA was prepared from the human glioblastoma cell culture, and cDNA synthesis and amplification was performed from dilutions containing 10 pg and 100

pg total RNA. These dilutions correspond to the range of total RNA hypothesized to exist in single mammalian cells. Hybridization of cDNA derived from 10 pg total RNA (~0.3–0.5 pg mRNA, n = 6) and 100 pg total RNA (~3–5 pg mRNA, n = 7) to HuGeneFL microarrays yielded high correlation coefficients of gene expression (0.84 ± 0.03 and 0.95 ± 0.02, respectively), remarkably similar to those obtained from single HGC and groups of 10 HGCs (Table 1; see Supplemental Data set S6).

Thus, these experiments clearly demonstrate that, although some distortion of the original single-cell transcript distribution cannot be excluded, the procedure appears extremely reproducible from sample to sample, such that a high degree of confidence can be granted to the comparison of single-cell profiles.

Quantitative Model of Single-Cell Variability

The transcriptional profile of single HGCs also offered a unique opportunity to quantify the cell-to-cell noise among "identical" cells and thus enabled us to distinguish biologically significant transcriptional differences from random inaccuracies in the reverse transcription and PCR amplification or from the intrinsic variability of transcription in single cells. For this purpose, fold changes in transcript expression level among pairwise combinations of HGCs were scored in log scale and plotted as a function of the mADV of each transcript (Figure 2D and see Experimental Procedures). As expected, more variability is observed for genes expressed at lower levels, as summarized by the best-fit regression curve of the data for mADV ≥ 250 (dark line in Figure 2D). From the spike mRNA experiments, a mADV of 250 roughly corresponds to 1–2 transcript copies per cell, allowing us to choose this number as a natural cutoff limit. Another curve can then be deduced that encompasses 95% of the data points, thus representing a statistically significant prediction or confidence interval within which a particular fold change can be considered above the background cell-to-cell noise (red line in Figure 2D).

This led us to establish a stringent set of criteria for the successful identification of transcriptional differences between two cell populations. We considered a given transcript 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, the equivalent of 1–2 copies per cell. 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 as defined by the confidence interval described above (Figure 2D and Experimental Procedures). 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 ($p < 0.05$).

OPC- and OSN-Specific Transcription

We then searched for differences in gene expression between single OSNs and OPCs. Using the three criteria established above, 458 probe sets were predicted to be enriched in one of the two cell types. Among those, 209 probe sets representing 187 individual genes were preferentially expressed in OSNs, and 249 probe sets representing 197 individual genes were predicted to be enriched in OPCs. A complete list of transcripts is provided in Supplemental Table S1 and Data set S2 at http://www.neuron.org/cgi/content/full/38/2/161/DC1.

The expression of 40 transcripts predicted in OSNs or OPCs, including randomly selected genes of unknown function, transcriptional regulators, and signaling molecules, was systematically verified by RNA in situ hybridization. Of these transcripts, 36 out of 40 (90%) matched the respective patterns predicted by microarray, while 4 gave an ambiguous or no signal, demonstrating a very tight concordance between predicted and actual expression patterns of the sampled transcripts. We also expanded our investigation to a small subset of transcripts that were predicted preferentially in OPCs but did not meet 1–2 of the criteria described above. Within this pool, 6 of 9 transcripts (66%) matched the patterns predicted by microarray. A complete summary of the expression patterns of these transcripts is shown in Supplemental Table S3, and representative examples are shown in Figures 3A and 3B and Supplemental Figure S1. Direct comparison of OPC-specific transcription with profiles defined for other early neuronal progenitors revealed that 42% of the OPC-enriched transcripts are also enriched in neurospheres (data not shown), suggesting a high degree of similarity between the two populations of olfactory and brain-derived rapidly dividing neuronal progenitors, while only 3% are common between OPCs, embryonic, neuronal, and hematopoietic stem cells (Ramalho-Santos et al., 2002)

Transcripts with predicted enrichment in OSNs were found to have expression in the neuroepithelial layer of the mature and E15 MOE, but not in the basal layer, where immature progenitors are thought to reside. In contrast, most transcripts with preferential expression in OPCs (19 out of 23) were found in the precursor cells of the E15 MOE and basal layer of the mature MOE. An additional four transcripts were detected in precursor cells of the E15 MOE but not in the mature MOE. This observation suggests that although there are some differences between OPCs originating from E15 and from mature MOE, the transcriptional profile of the two progenitors is mostly similar.

Although all the OSNs and OPCs chosen for analysis in our study share expression of key marker proteins, such as OMP and Mash1, respectively, it remains possible that they could be subdivided into different cellular classes. This possibility seems especially critical for the Mash-positive OPCs, which may represent separate developmental lineages or different developmental stages. To test this hypothesis, we used Genecluster 2.0 to perform class predictions based on the generation of self-organized maps (SOMs) (Golub et al., 1999). When a 2-cluster SOM is generated, the OSNs and OPCs segregate perfectly along the OSN-OPC distinction; moreover, crossvalidation studies (see Golub et al., 1999) demonstrate these two categories are readily distinguished (not shown). When 3 × 1 and 4 × 1 cluster SOMs are generated, the algorithm does segregate the OSNs and OPCs into further subdivisions, but these subcategories perform very poorly on crossvalidation. This analysis supports the idea that our Mash-positive samples reflect a common cell type at a similar develop-

A

probe set	gene name	p value	mADV OSNs (±sem)	mADV OPCs (±sem)	logFC/ 95% CI ratio
OPC-enriched					
aa250191_s_at	Hes6	0.0004	-8 ±62	10839 ±3887	8.10
m69293_rc_at	Id2	0.0248	96 ±44	3093 ±2113	2.38
m60523_s_at	Id3	0.0205	435 ±342	1777 ±813	0.90
m64292_s_at	Tis21	0.0238	127 ±201	1142 ±669	1.26
u81603_s_at	Eya2	<0.0001	-60 ±12	49 ±26	0.63
x80339_s_at	Six1	0.0415	60 ±68	430 ±290	0.94
u52951_s_at	Enx1	0.0029	-7 ±17	4799 ±2244	6.29
u78103_s_at	Eed	0.0964	123 ±104	1616 ±1649	1.57
d83596_s_at	Etf	0.0031	33 ±55	854 ±381	1.73
Msa.22493.0_s_at	Rbtn1	0.0045	-254 ±131	753 ±481	3.04
y08361_s_at	RIL-pending	0.0307	-157 ±125	10679 ±8076	8.04
aa606494_s_at	Unknown 1 (5730427N09Rik)	0.0009	184 ±189	3834 ±1415	2.18
w12941_s_at	Unknown 2 (1110004C05Rik)	0.0273	-74 ±40	2673 ±1985	5.18
aa409089_rc_s_at	Unknown 3 (2410018C03Rik)	0.0036	-65 ±58	2471 ±1220	5.05
OSN-enriched					
aa270796_s_at	Oda10	<0.0001	2568 ±416	43 ±37	2.83
aa546047_s_at	Unknown 4 (AA546047)	0.0004	8598 ±2648	53 ±95	4.46
Msa.2916.0_s_at	Axcam	0.0416	1000 ±711	-10 ±20	3.94
d50311_s_at	Mef2b	0.0428	946 ±741	-100 ±47	3.81

B

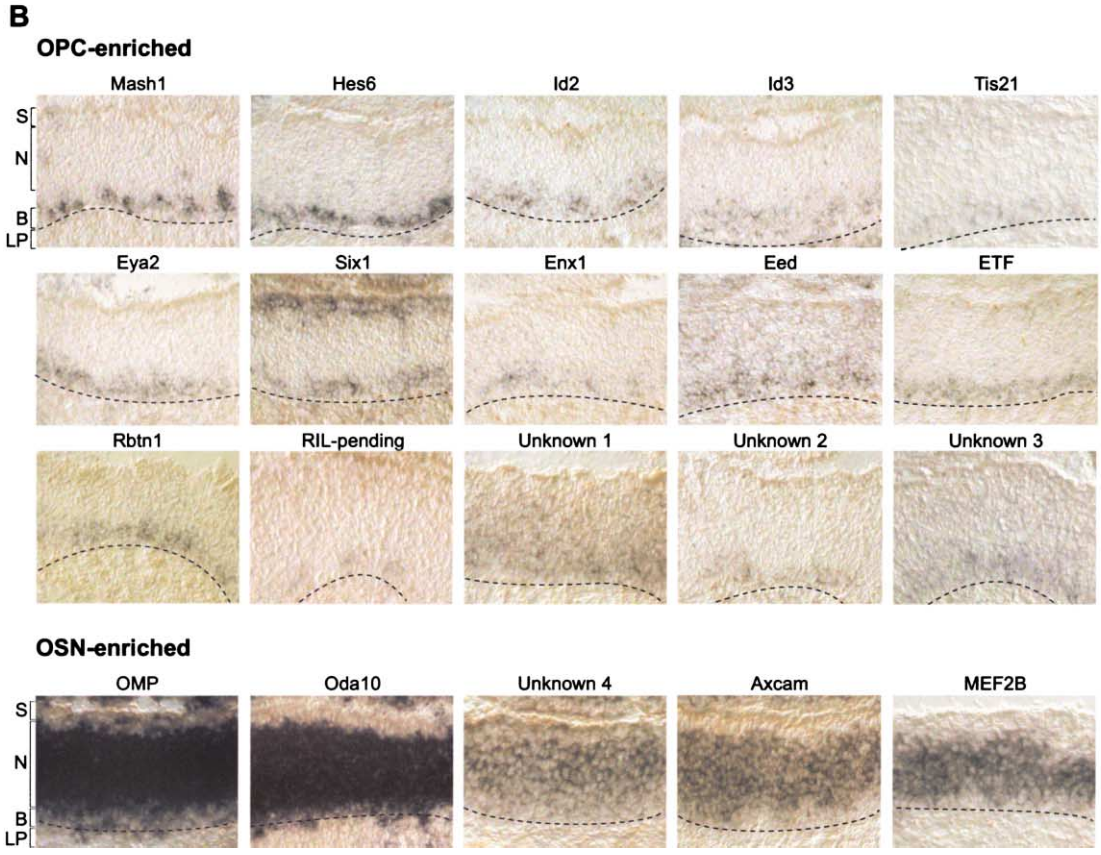


Figure 3. Gene Expression in OSNs and OPCs

(A and B) Microarray (A) and in situ hybridization (B) data from selected transcripts with predicted differential expression by OSNs and OPCs. (A) For each transcript, the corresponding microarray probe set, gene name, and values for the three criteria for predicted differential expression (p value, mADV [for OSNs and OPCs], and log FC/95 Confidence Interval ratio) are shown. When comparing the profiles of OSNs versus OPCs,

mental stage. Due to the low abundance of most transcripts identified in our study, definitive demonstration of the coexpression of Mash1 and OPC-specific transcripts will await further improvement in the sensitivity of transcript detection by double in situ hybridization.

The Transcriptional Profile of Mash1-Positive Olfactory Progenitors

The transient expression of the bHLH-containing transcription factor Mash1 by olfactory progenitors corresponds to a critical step in their progression to a neuronal fate. In the early embryonic MOE, Mash1 is expressed by a population of rapidly dividing apical neuroepithelial cells, which then translocate to form a basal pool of dividing Mash1-positive neuronal progenitors (Cau et al., 1997, 2000, 2002). These OPCs fail to appear in the Mash1 mutant, which consequently does not produce any substantial number of OSNs (Guillemot et al., 1993). Mash1 has also been shown to control Notch signaling components and transcription factors essential for olfactory differentiation of basal cells, leading to the hypothesis that it functions as a proneural gene during olfactory development. Therefore, the ability to monitor the transcriptional profile of single progenitors provides a unique opportunity to uncover early Mash1-dependent and -independent signaling events that are essential for olfactory neurogenesis.

OPC-specific transcripts displayed distinct types of expression patterns in the E15 MOE. A subset of transcripts, including Hes6, Tis21, Enx1, ETF, and Eed, were found predominantly or exclusively in basal neuronal progenitors, while others, notably Pax6, Id2, Eya2, and Six1, were equally expressed by both apical and basal progenitors (Figure 4; see Supplemental Figure S1 at http://www.neuron.org/cgi/content/full/38/2/161/DC1). This result is in agreement with the existence of distinct pools of basal and apical Mash1-positive progenitors and further indicates that the transcriptional programs of basal and apical progenitors have already significantly diverged. Moreover, the presence of transcripts with exclusive basal expression in all OPCs of our collection enables us to further and more precisely identify these Mash1-positive cells as basal neuronal progenitors.

Among the OPC-specific transcripts, we identified members of the Notch pathway. RBP, a member of the transcriptional complex controlled by Notch signaling, and Hes6, which inhibits Hes1 transcription (Bae et al., 2000; Koyano-Nakagawa et al., 2000) both appear upregulated in Mash1-positive cells. In contrast, none of the OPCs were found to express Hes1, shown to be present in apical Mash1-positive precursors (Cau et al., 2002). Expression of other members of the Notch pathway, Hes5, Serrate1, and Serrate2, which have been identified in the developing MOE (Cau et al., 2002), could not be investigated through the Affymetrix Murine11K probe arrays. Hes5, as assessed by Southern Blot analysis, does not appear to be present at significant levels in OPC cDNAs.

OPCs also express multiple factors involved in both the positive and negative regulations of cell proliferation, suggesting that this cell type is in a highly dynamic state. Three members of the Id protein family, Id1, Id2, and Id3, had strong and specific expression in OPCs, hinting at a potential role in preventing neuroblasts from cell cycle withdrawal further neuronal differentiation (Norton et al., 1998). However, the simultaneous detection of the antiproliferative gene Tis21 indicates that OPCs also receive signals to leave the cell cycle. Indeed, Tis21 expression has been shown in neuroepithelial cells which, at their last division, will generate postmitotic neurons (Iacopetti et al., 1999).

Interestingly, the vast majority of differentially expressed transcripts appear exclusively expressed in basal neuronal progenitors, consistent with the idea that the process of neurogenesis and patterning is initiated in these cells. We identified two interacting Polycomb-group genes, Enx1 and Eed (Figure 3A), which suggest a role in chromatin remodeling in olfactory specification (Beuchle et al., 2001; Schumacher and Magnuson, 1997). In addition, we identified the embryonic TEA domain-containing transcription factor ETF (Guss et al., 2001; Yasunami et al., 1995) and a LIM domain-containing nuclear protein expressed in specific areas of the developing brain, Rhombotin1 (Rbtn1) (Figure 3A; Hinks et al., 1997). ETF appears transiently expressed by Mash1-positive precursors of the basal embryonic epithelium, while Rbtn1, although clearly excluded from apical precursors, is more widely expressed in basal precursors as well as in immature olfactory neurons, suggesting a role throughout olfactory differentiation. Finally, we identified three genes, Eya2, Six1, and Pax6, which belong to a regulatory network involved in eye development (Figure 3A). Eya2 and Six1 are homologs of *Drosophila eyes absent (eya)* and *sine occulis (so)*, respectively. *Eya* and *so* interact physically, synergize to induce ectopic eyes, and function within a transcriptional network that also includes *eyeless/Pax6* in order to control cell proliferation, patterning, and neuronal specification within the developing eye (Heberlein and Treisman, 2000; Pignoni et al., 1997). The expression of three members of the same genetic network in the developing MOE may indicate that the function of the complex is conserved in olfactory neurogenesis.

Mash1-Dependent and -Independent Transcriptional Control of Olfactory Neurogenesis

The availability of the Mash1$^{-/-}$ knockout (Guillemot et al., 1993), which lacks the open reading frame of the

a logFC/95% confidence interval value greater than 1 indicates that the average variability of OSNs versus OPCs is significantly greater than the threshold defined by the 95% confidence interval equation in Figure 2D. A subset of transcripts shown (Id3, Eya2, Six1, Eed, Rbtn1) was investigated by RNA in situ hybridization but did not meet 1–2 of the three criteria.

(B) In situ hybridization shows that transcripts predicted to be enriched in OPCs (top three rows) are preferentially expressed in the globose basal layer of P21 olfactory epithelium, which corresponds to the location of Mash1-positive OPCs (top left panel). In contrast, transcripts predicted to be enriched in OSNs (lower row) are preferentially expressed in the neuronal layer of P21 MOE, identified by the expression of OMP (left panel). Dotted lines indicate the basal layer/lamina propria border. S, Sustentacular cell layer; N, neuronal layer; B, globose basal cell layer; LP, lamina propia.

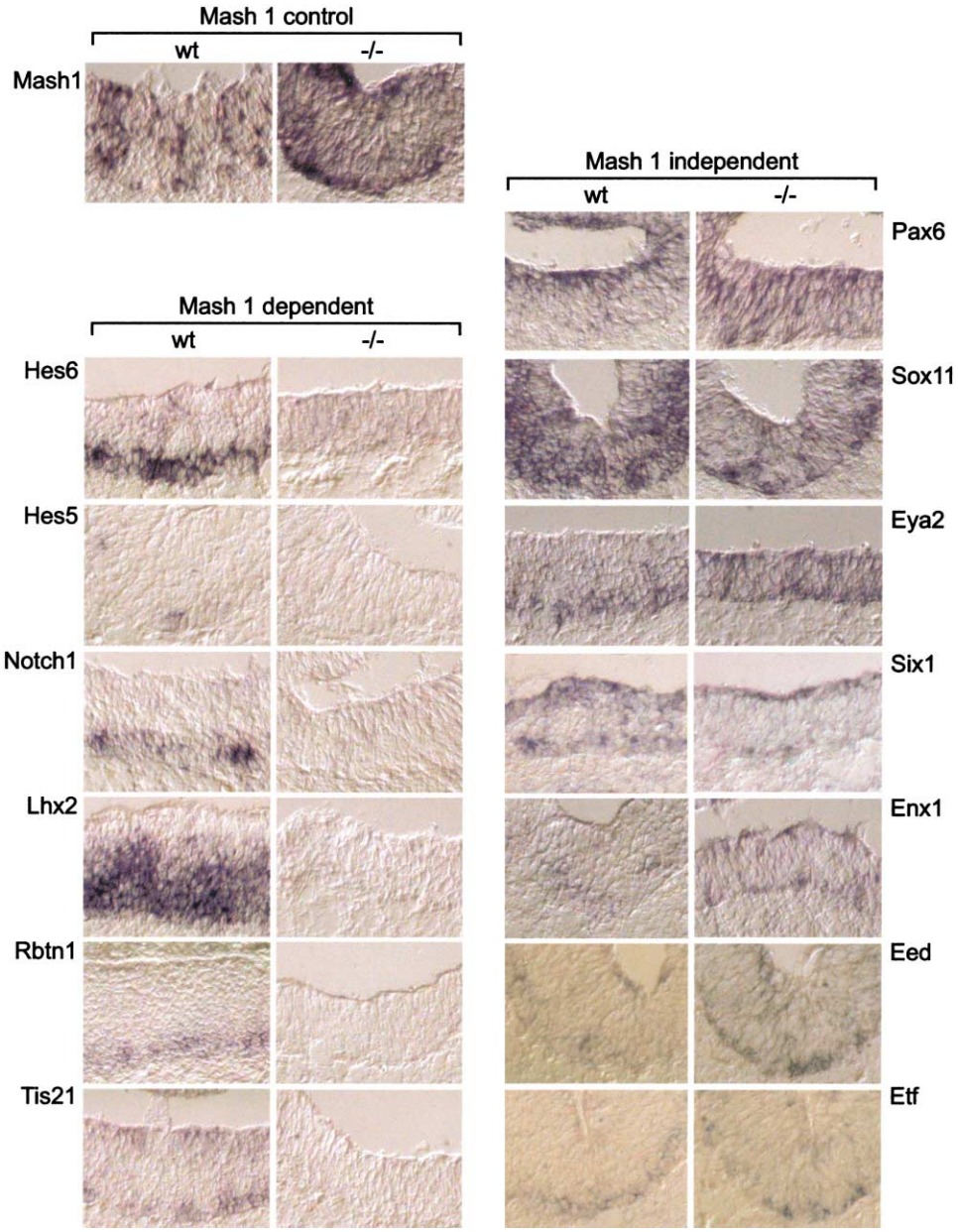


Figure 4. Mash1-Dependent and -Independent Pathways in Olfactory Development

Expression patterns of OPC-enriched transcripts were investigated by RNA in situ hybridization on wild-type (wt, left of each column) and Mash1-deficient (−/−, right of each column) E15 MOE. Top left: Mash1 transcript expression is abundant in both wt and Mash1$^{-/-}$ MOE. Bottom left: the expression of several OPC-enriched transcripts is dependent on Mash1. Expression of these transcripts is significantly reduced (Hes6, Notch1, Lhx2) or virtually eliminated (Hes5, Rbtn1, Tis21) in Mash1 mutants. Right: several OPC-enriched transcripts associated with neurogenesis are still expressed in the absence of Mash1 function.

Mash1 gene, has allowed us to perform partial epistasis analysis between Mash1 and other transcriptional regulators as identified in OPCs on arrays. The expression of genes related to cell proliferation or Notch signaling, mainly Hes5, Hes6, Notch1, and Tis21, was abolished or strongly reduced in the mutant E15 MOE (Figure 4). Control in situ hybridization with a full-length Mash1 probe reveals the presence of numerous Mash1-positive progenitors in the E15 mutant epithelium (Figure 4), suggesting that, in these cells, Mash1 controls expression of Hes5, Hes6, Notch1, and Tis21, thereby promoting neurogenesis and exit of cell cycle.

In contrast, the expression of a large subset of OPC-specific transcripts involved in neurogenesis and neuronal specification was maintained in the Mash1$^{-/-}$ epithelium. The extent of expression of these genes throughout the entire MOE well exceeds that of the very small subset of OSNs known to survive in the ventrolateral region of the Mash1$^{-/-}$ MOE. The expression of Rbtn1 and Lhx2, another transcription factor expressed

in the embryonic MOE and potentially associated with Rbtn1 function (Cau et al., 2002; Porter et al., 1997), is abolished in Mash1$^{-/-}$ MOE (Figure 4). However, the two interacting Polycomb-group genes (Enx1 and Eed), as well as the transcription factors ETF, Sox11, an early olfactory marker (C.D., unpublished), and the three members of the retinal development signaling network (Six1, Eya2, and Pax6), were consistently detected in the Mash1 mutant epithelium, suggesting that although the ultimate survival of neuronal precursors relies on Mash1 function, some aspects of neurogenesis and neuronal specification might be readily initiated in the absence of Mash1 expression.

Transcriptional Patterning of the Embryonic MOE

The expression of olfactory receptor genes is segregated along a dorso-ventral axis such that all neurons expressing a given OR reside within one of four distinct zones of the olfactory epithelium. Although some transcripts have been shown to display gradients of gene expression along the dorso-ventral axis of the MOE (Norlin et al., 2001), no systematic approach has been performed to investigate how this patterning is achieved.

Although olfactory receptor (OR) transcripts are not detectable in OPCs, in situ hybridization on the E15 MOE detects OR expression in more advanced precursors, which are already segregated into four distinct epithelial zones (Buck, 2000). We performed systematic in situ hybridization on serial sections of the MOE with RNA probes corresponding to ORs and to OPC-specific transcription factors. This set of experiments revealed that a subset of the genes identified with the single-cell profiling study display a dramatic restriction of their expression to a subset of the MOE zones both in adult and in the embryo. Indeed, while Sox 11, Hes6, and Mash1 (Figure 5, left column), as well as Six1 and Lhx2 (not shown), are widely expressed along the whole epithelium, Eya2, Rbtn1, Id2, and Id3 are present in progenitors of one or two ventral zones exclusively (Figure 5, middle and right columns). Thus, it appears that restricted expression of transcriptional regulators in distinct zones of the MOE can be detected in early progenitors that have yet to express a given OR. These genes in turn represent good candidates to participate in the combinatorial coding of a dorso-ventral MOE patterning axis, a process that ultimately results in the zonal organization of OR expression.

Remarkably, the zonal expression of Eya2 (Figure 5A, right) and Id2 (not shown) appear unaffected in the Mash1 mutant, suggesting that the control of the MOE dorso-ventral patterning is at least partially independent from Mash1-dependent neurogenesis.

Transcriptional Profile of Neuronal Progenitors Captured from Brain Slices

Patterning events in the embryo rely on the spatial segregation of molecularly distinct subpopulations, and neurons and neuronal precursors of specific types are often closely intermingled within a defined structure. Laser capture microscopy on tissue slices offers the unique opportunity to isolate specific cell types from precise tissue locations.

In order to successfully isolate single precursors with only minor contamination by adjacent cells, we restricted the thickness of tissue sections to 5–7 μm, corresponding to a single-cell layer, and narrowed to its minimum the diameter of the laser beam (see Experimental Procedures). Tissue sections were treated with hematoxylin stain, enabling us a thorough visual inspection of the captured sample under the microscope to assess the presence of a single nucleus (see Figure 6A, insert). In a pilot experiment, we investigated the ability to successfully generate amplified single-cell cDNA from laser-captured samples. Single neurons were laser captured from frozen sections of vomeronasal organ (VNO), and single-cell cDNAs generated according to our standard protocol were checked extensively by Southern blot analysis for a large variety of ubiquitous and VNO-specific markers. Results consistently indicated that the quality of single-cell cDNAs obtained from laser-captured samples was indistinguishable from that of hand-picked VNO neurons from our lab collection (not shown).

Based on these results, we directly aimed at discovering genes differentially expressed between two targets of olfactory neuronal projections, the nascent precursor cells from the main olfactory bulb (MOBCs) and from the accessory olfactory bulb (AOBCs). For this purpose, we laser captured cells from parasagittal sections at E15.5, a stage at which only mitral cell progenitors have become postmitotic and have migrated into the future bulb (Hinds, 1968). To provide a spatial guide for the laser dissection of AOB versus MOB precursors, we performed in situ hybridization on adjacent sections with Id2, a marker that distinguishes the two embryonic bulb structures at E15.5 (Figure 6A, top panels, and Experimental Procedures). Following laser capture, single cells were dissected from the plastic caps (Figure 6A, third panel) and placed into individual PCR tubes for single-cell cDNA amplification (see Experimental Procedures). Confirmation of the identity of each cell was obtained by Southern blot with Id2 (Figure 6A, bottom panel). Of 36 amplified cDNAs, 10 samples (5 MOBCs and 5 AOBCs) were labeled and hybridized to Murine11K Affymetrix microarrays as described (see Supplemental Data set S3 at http://www.neuron.org/cgi/content/full/38/2/161/DC1). Systematic inspection of data quality was performed, confirming that single-cell cDNA obtained by laser microdissection was similar in quality to that obtained from live cells in suspension. The average number of genes expressed in MOBCs (2985 ± 329) and in AOBCs (3092 ± 368) was similar to that obtained from live cells. Moreover, the complexity of low, middle, and high abundance transcripts was also comparable (data not shown). In addition, comparisons between samples give correlation coefficients comparable to, although slightly lower than, live OPCs picked from the MOE (Table 1). This lower correlation coefficient between early bulb precursors is expected because these cells are present at different maturation stages and have not been selected for the expression of a defined developmental marker as were Mash1-positive OPCs.

To find genes differentially expressed between MOBCs and AOBCs, we applied the three criteria established previously. 131 probes representing 125 genes were specifically and highly upregulated in MOBCs, while 20 probes representing 18 genes were specifically

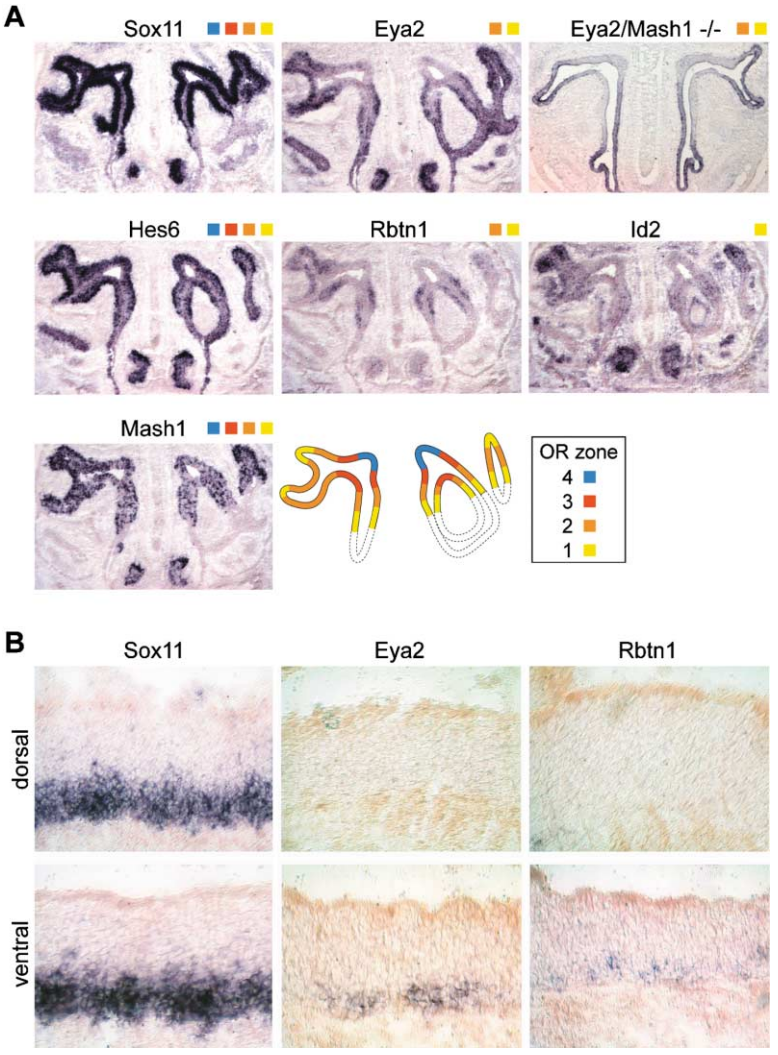


Figure 5. OPC-Enriched Transcription Factors in Olfactory-Receptor Zones

(A) RNA in situ hybridization patterns on adjacent, coronal sections of E15 MOE. Sox11, Hes6, and Mash1 are strongly expressed throughout the embryonic MOE. In contrast, Eya2, Rbtn1, Id2, and Id3 (not shown) are preferentially expressed in ventral zones. The expression of Eya2 in Mash1-deficient MOE (Eya2/Mash1$^{-/-}$) is similar to the zone-restricted pattern in wild-type tissue. A schematic of the four zones in e15 MOE is shown at lower right.

(B) The zone-restricted expression of transcripts is maintained in p21 MOE. Representative examples of Sox11, Eya2, and Rbtn1 expression in dorsal (zone 4) and ventral (zone 1) regions of MOE are shown, as determined by olfactory receptor expression on adjacent sections (not shown).

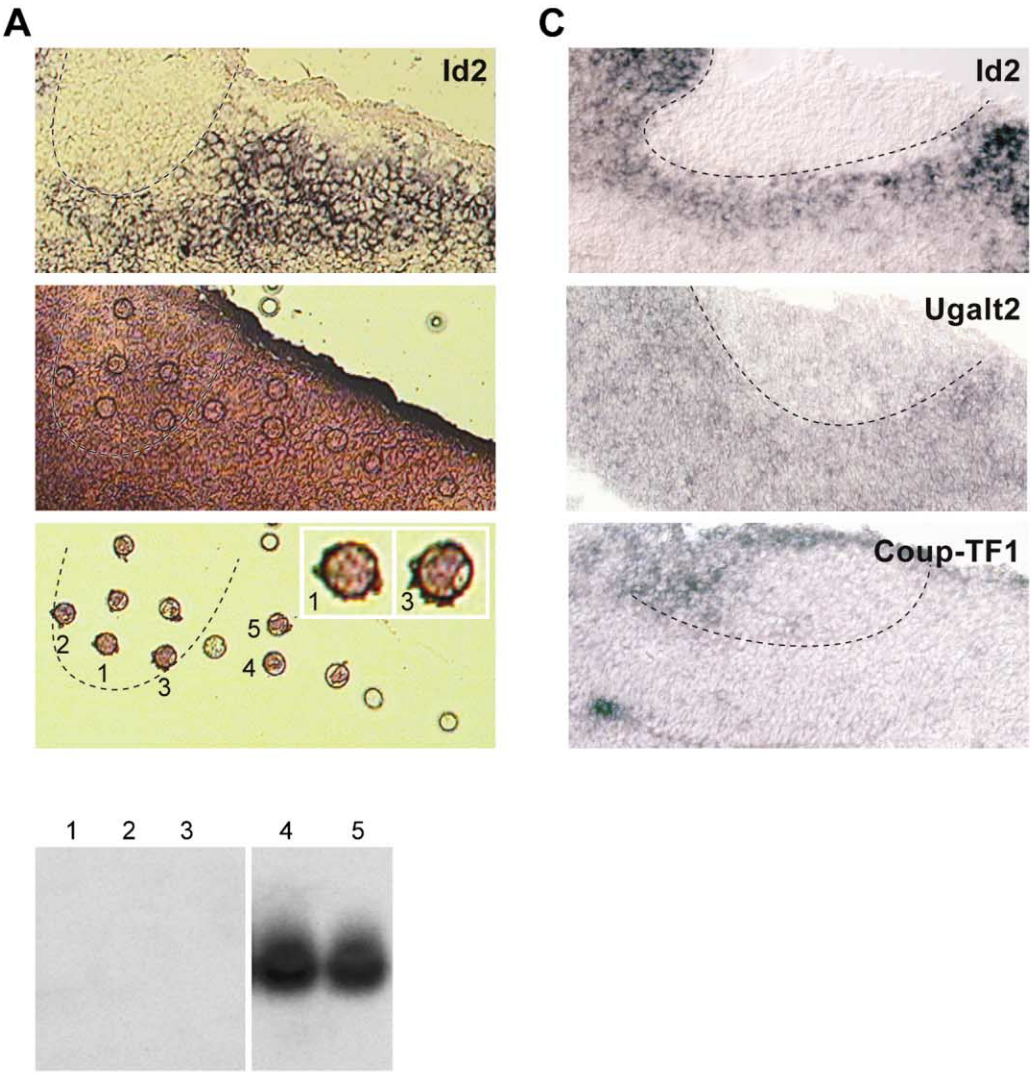
upregulated in AOBCs (see Supplemental Table S2 and Data set S4 at http://www.neuron.org/cgi/content/full/38/2/161/DC1). A significant number of these enriched genes have known functions in transcriptional regulation and repression, cell cycle control, translational control, the regulation of the cytoskeleton, and cell adhesion, thus offering promising leads toward the dissection of transcriptional networks specifying distinct mitral cell populations.

The expression pattern of several of these transcripts was tested by in situ hybridization on E15.5 olfactory bulbs (Figure 6C), confirming the very tight concordance between array and expression data already observed with hand-picked cells: 11/13 (85%) transcripts gave the expected expression patterns as predicted by the array (see Supplemental Table S3 at http://www.neuron.org/cgi/content/full/38/2/161/DC1). Two of these genes (aa611744 and COUP-TF1) selectively stained only the posterior half of the AOB (Figure 6C and data not shown), attesting to the cellular heterogeneity of mitral cell precursors at that stage.

Discussion

We have demonstrated here the ability to monitor the transcriptional profile of individual cells. Although others have reported on the representation of cDNA amplified from dilute starting material (Iscove et al., 2002) or the ability to detect amplified material on nylon filters (Klein et al., 2002), the hybridization of amplified single-cell cDNAs to microarrays for the purpose of monitoring large-scale changes in gene expression is an unvalidated method. We have documented the sensitivity, reliability, and fidelity of the single-cell cDNA amplification, resulting in a faithful representation of the original single-cell transcriptome and therefore the ability to predict meaningful differences between specific and rare cell types.

We have established useful criteria that permit the successful identification of transcriptional differences between individual cells, while reducing false calls resulting from random variability in gene expression or PCR distortions. Similar results were obtained using the



B

probe set	gene name	p value	mADV MOBCs (±sem)	mADV AOBCs (±sem)	logFC/95% CI ratio
MOBC-enriched					
m69293_rc_at	Id2	0.1016	15364 ±12117	341 ±146	3.67
d87990_s_at	Ugalt2	0.0354	2318 ±1149	174 ±63	1.73
AOBC-enriched					
x74134_s_at	Coup-TF1	0.0856	554 ±294	7587 ±5222	2.21

Figure 6. Laser Capture of Single Cells and Gene Expression in Embryonic Olfactory Bulb

(A) Laser capture microscopy of single MOBCs and AOBCs from a section of E15.5 olfactory bulb. Top: Id-2 is expressed in the developing MOB, but not in the AOB. Second from top: adjacent section following laser capture of single cells. Dark circles indicate areas where laser pulses melted the plastic cap. Second from bottom: cell samples adhering to the cap. Close inspection of captured samples (insert) demonstrate the presence of single nuclear structures. Bottom: Southern blot on single-cell cDNAs from captured cells confirms expression of Id-2 in MOBCs (cells 4–5) but not AOBCs (cells 1–3).

(B and C) Microarray (B) and in situ hybridization (C) data of selected transcripts with predicted differential expression in the MOB (Id2, Ugalt2) and AOB (coup-tf1). Details are the same as Figure 3.

self-organizing map and class predictor strategies of Golub et al. (1999) to identify those genes that behave as strong "predictors" of OSN versus OPC cells. One difference, however, is that the 95% confidence interval provides an experimentally defined approach to vary the fold change required for selection as a function of

the mADV, instead of a simple uniform fold change for all mADV, thereby providing a more stringent filter for gene selection. We also found that one could reduce the stringency of transcript selection, by using only a subset of the established criteria, and still obtain meaningful information, although with a higher percentage of error. In particular, it appears from our data that a lower stringency filter permits identification of transcripts with expression by a cell subset only, for example, coup-TF1 in the posterior AOB.

Some limitations of the presented strategy have emerged during our study. First, although the vast majority of microarray probe sets appear to function normally, probe sets that do not correspond to the 3′-most portions of the transcript do not successfully detect single-cell cDNA. In addition, in the absence of cell-specific markers for rare or poorly characterized cell types, the ability to identify and isolate one's favorite neuronal cell type and precursor might present some challenge, a problem further enhanced by the fact that, due to cell-to-cell variability, the analysis of at least 4–5 samples per cell type is required to eliminate false positives from cell type comparisons. This issue is likely to be partially alleviated by enhanced cell identification afforded by the laser capture microscopy.

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.

In order to test the utility of our new experimental startegy, we focused our attention on unraveling Mash1-dependent and -independent epistatic interactions. The persistent expression of OPC-specific transcripts associated with neurogenesis in basal cells throughout the entire epithelium of the Mash1$^{-/-}$ mutant well exceeds the presence of a small subset of surviving OSNs already described in the mutant MOE. Thus, our data indicates that some aspects of neurogenesis still occur in the absence of Mash1. Our identification of an Eya2, Six1, and Pax6 transcriptional network in OPCs of wild-type and Mash1 mutant documents how networks that control the expression of bHLH factors are changed and redeployed under Mash1-dependent and -independent conditions in different developmental systems (Heanue et al., 1999; Ohto et al., 1999).

In conclusion, the level of reproducibility, sensitivity, and accuracy of our single-cell approach provides unrivalled cellular resolution for large-scale transcriptional analysis that has immediate application for the study of olfactory neurogenesis and can be more vastly applied to investigate the molecular nature of neuronal diversity throughout the developing, the adult, and the diseased brain.

Experimental Procedures

Single OSN and OPC cDNA Preparation and Analysis

Single-cell cDNA synthesis and amplification was performed for all samples according to Dulac and Axel (1995). Single OSNs and OPCs were obtained from c57Bl/6J mice. OSNs and OPCs were obtained from dissected MOEs from adult and embryonic day 15 (E15) embryos, respectively. Small pieces of tissue were dissociated for 10 min at 37°C in phosphate-buffered saline (PBS) (without Ca^{2+} and Mg^{2+}), 0.025% trypsin, 0.75 mM EDTA. After gentle trituration of the tissues in Dulbecco's modified Eagle's medium plus 10% fetal calf serum, cells were collected by centrifugation and resuspended in ice-cold PBS. The cell suspension was observed on a Leitz inverted microscope. OSNs from adult MOE preparations were identified as bipolar neurons with an axonal process and a dendrite terminating in an olfactory knob. Cells from E15 preparations were picked independent of cell morphology. Isolated cells were picked with a mouth pipette fitted with a beveled microcapillary. Single cells were seeded into thin-walled PCR tubes (Perkin-Elmer) containing 4 μl of ice-cold cell lysis buffer (50 mM Tris-HCl [pH 8.3], 75 mM KCl, 3 mM $MgCl_2$, 0.5% NP-40, containing 80 ng/ml pd(T)19-24 [Pharmacia], 5 U/ml Prime RNase inhibitor [5′-3′ Incorporated], 324 U/ml RNAguard [Pharmacia], and 10 μM each of dATP, dCTP, dGTP, and dTTP). Lysis was subsequently performed at 65°C for 1 min. First-strand cDNA synthesis was then initiated by adding 50 U of MMLV and 0.5 U of AMV reverse transcriptases (Invitrogen) followed by incubation at 37°C for 15 min. Samples were heat inactivated at 65°C for 10 min, and poly(A) was added to the first-strand cDNA product by adding an equal volume of 200 mM potassium cacodylate (pH 7.2), 4 mM $CoCl_2$, 0.4 mM DTT, 200 μM dATP containing 10 U of terminal transferase (Roche Diagnostics) at 37°C for 15 min. Samples were heat inactivated at 65°C for 10 min, and the contents of each tube was brought to 100 μl with a solution made of 1× PCR buffer II (Applied Biosystems), 2.5 mM $MgCl_2$, 100 μg/ml bovine serum albumin, 0.05% Triton X-100 and containing 1 mM of dATP, dCTP, dGTP, dTTP, 10 U of AmpliTaq polymerase (Applied Biosystems), and 5 μg of the PCR primer AL1. The *AL1* sequence is 5′-ATT GGA TCC AGG CCG CTC TGG ACA AAA TAT GAA TTC $(T)_{24}$-3′. PCR amplification was then performed according to the following schedule: 94°C 1 min, 42°C 2 min, and 72°C 6 min with 10 s extension per cycle for 25 cycles. An additional 5 U of Taq polymerase was added before performing 25 more cycles of PCR without the 10 s extension per cycle. In this manner, 10–20 μg of PCR-amplified cDNA was synthesized from RNA of individual neurons.

Five microliter aliquots of each single-cell cDNA were checked for the presence of ubiquitous and cell type-specific markers by Southern blot hybridization (see below). All single-cell cDNA considered for further study showed strong expression of α-tubulin and GAPDH. OPCs were identified by the strong expression of Mash1, cdc2, and Ki67 (Guillemot et al., 1993; Riabowol et al., 1989; Scholzen and Gerdes, 2000). OSNs were identified by the strong expression of olfactory marker protein (OMP).

In addition, the olfactory receptor sequence was identified from single OSNs using degenerate PCR essentially as previously described (Malnic et al., 1999). Olfactory receptor sequences were confirmed by DNA sequencing. In no case was >1 OR sequence detected per single OSN sample.

Single OPC and OSN samples that met the above criteria were subsequently reamplified by PCR. 2.25 μl of each single-cell cDNA was added to 300 μl of 1× PCR buffer II, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 5 U AmpliTaq, and 5 μg AL1 primer. Samples were then PCR amplified with 30 cycles of 94°C 1.5 min, 42°C 2 min, 72°C 3 min. Approximately 50 μg of sample was obtained in this manner. 100 μl aliquots were added to PCR purification columns (Qiagen) and rechecked for the presence and intensity of control markers by Southern blot. OSN samples were also rechecked for the presence of olfactory receptor sequence by degenerate PCR. Samples with robust expression of all diagnostic markers were then hybridized to microarrays (see below).

"Spiked" Single OSN Preparation and Analysis

Poly(A)-tailed *B. subtilus* Lys, Dap, Phe, and Thr RNA were prepared as described (Lockhart et al., 1996) and "spiked" into a subset of OSN samples. Spike RNAs were added into the single-cell lysis buffer to the following final concentrations: Lys, 10^{-6} pg/sample; Dap, 2×10^{-6} pg/sample; Phe, 10^{-5} pg/sample; Thr, 10^{-4} pg/sample. Assuming 1 μg of poly(A)-RNA with an average length of 2 kb = 1.52 pmol, this translates into ~1, 2, 10, and 100 copies/sample, respectively. Samples were then amplified by single-cell PCR and analyzed as described above.

HGC cDNA Preparation and Analysis

A T98G human glioblastoma cell line (CR1-1690) was serum starved 3 days, trypsinized 5 min, and resuspended in PBS without Ca^{2+} or Mg^{2+}. 88% of all cells were determined to be in G1/G0 phase by FACS analysis (data not shown). Resuspended cells were counted with a hemocytometer and diluted to a concentration of 20 cells/μl. 1HGC samples were picked as described above. 10HGC samples were obtained by adding 0.5 μl of the cell suspension directly to the single-cell lysis buffer. cDNA was obtained as described above and checked for the strong expression of α-tubulin and GAPDH by Southern blot. Samples with robust expression of these markers were directly hybridized to microarrays (see below).

Diluted HGC RNA Preparation and Analysis

Total RNA was extracted with RNAzol (Tel-Test) from a T98 human glioblastoma cell line according to manufacturer's specifications. The total RNA (~2000 ng/μl) was then serially diluted to a concentration of 200 pg/μl and 20 pg/μl. 0.5 μl of each dilution was directly added to the single-cell lysis buffer. cDNA was obtained as described above and checked for the strong expression of α-tubulin and GAPDH by Southern blot. Samples with robust expression of these markers were directly hybridized to microarrays (see below).

Laser Capture Microdissection and Sample Preparation

E15.5 mouse olfactory and accessory olfactory bulbs were dissected and freshly embedded in Tissue-Tek OCT compound (Sakura). 5–7 μm parasaggital cryosections of E15.5 bulbs were placed on permafrost-coated glass slides and taken through the following staining and dehydration procedure: 70% EtOH 30 s, H_2O 5 s, Mayer's Hematoxylin (Sigma) 1 min, H_2O 5 s, Bluing Reagent (Shandon) 1 min, 70% EtOH 10 s, 95% EtOH 10 s, Eosin Y (Shandon) 20 s, 95% EtOH 30 s ×2, 100% EtOH 1 min ×2, xylene 5 min ×2, air dry 15 min. Single-cell nuclei were then laser microdissected using a PixCell II Laser Capture Microdissection microscope (Arcturus). Single cells were removed from the Capsure caps (Arcturus) using a microneedle and placed into PCR tubes. Lysis buffer was added directly to the laser-captured sample, and single-cell cDNA was generated as described above. Samples were checked by Southern blot for the strong expression of α-tubulin and GAPDH, and Id2 expression in MOBCs and absence in AOBCs was confirmed. Samples were reamplified as described above, rechecked for the presence and intensity of control markers, and hybridized to microarrays (see below).

Microarray Hybridization

Ten micrograms of single-cell cDNA was digested with 1U RQ1 DNase (Promega) in 1× One-Phor All buffer (Pharmacia) in 80 μl total volume at 37°C for 13 min, followed by 99°C for 15 min. DNA fragments were end-labeled with 25 μM Biotin-N6 ddATP (NEN) and 45 U terminal transferase (GIBCO-BRL) at 37°C for 1.5 hr, 65°C for 5 min, then on ice for 5 min.

DNA fragments were then hybridized to HuGeneFL or Mu11K GeneChip probe arrays (Affymetrix) (Mody et al., 2001). Microarrays were prewarmed on a rotisserie at 60 rpm, 45°C, 15 min with 1× MES (0.1 M Mes, 1.0 M NaCl, 0.01% Triton X-100 [pH 6.7]), 0.5 mg/ml acetylated BSA (GIBCO-BRL), and 0.5 mg/ml herring sperm DNA (Promega). Immediately before hybridization, single-cell cDNA was warmed to 99°C for 5 min and cooled in a water bath at 45°C. cDNA was then added to microarrays and incubated in a rotisserie at 60 rpm, 45°C for 16 hr.

Following hybridization, arrays were washed with 6× SSPE-T (0.9 M NaCl, 60 mM NaH_2PO_4, 6 mM EDTA, 0.01% Triton X-100 [pH 7.6]) at 22°C on a fluidics station (Affymetrix) for 10 × 2 cycles, followed by incubation with 0.1× MES buffer on a rotisserie at 60 rpm, 45°C, 30 min. Arrays were again washed as above. Hybridized samples were then stained with a streptavidin-phycoerythrin conjugate (Molecular Probes) at 40°C for 15 min and washed as above. To enhance the signals, arrays were further stained with anti-streptavidin antibody (Vector) at 40°C for 30 min, washed as above, stained with a streptavidin-phycoerythrin conjugate at 40°C for 15 min, and washed as above. Arrays were scanned at a resolution of 3 μm, using a specifically designed confocal scanner (Affymetrix).

Microarray Data Analysis

Microarray image data were analyzed using Affymetrix GeneChip System v3.2 software with the following changes to Data Analysis 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. These values lower the stringency of the Absolute call to compensate for the lack of full-length transcripts in single-cell cDNA.

All subsequent analysis was performed using Microsoft Excel. To obtain the 95% confidence interval equation quantifying significant 1HGC gene expression variability, the mean ADVs (mADVs) and absolute values of the log fold change (logFC) were calculated for each probe set in pairwise comparisons of the 1HGC microarray data (6 samples for 15 total pairwise comparisons). All pairwise comparisons were independently ranked from highest to lowest by mADV, and the corresponding logFC and mADV were then averaged across all ranked, pairwise comparisons. This ranked logFC was plotted against the ranked mADV, thereby providing a measure of FC variability at a given mADV that is independent of gene identity. The best-fit regression curve through all points with a pairwise mADV ≥ 250 was calculated using Excel's data analysis feature as $y = 4.6714x^{-0.3544}$ ($R^2 = 0.5838$), where y is the log fold change and x is the mADV. The 95% prediction or confidence interval was then calculated for all values as described by Rosner (2000), which for our HGC data set is approximated by the equation $y = 2.67x^{-0.1984}$.

A total of nine OSNs, seven OPCs, five OBCs, and five AOBCs were used to find differentially enriched transcripts. A transcript of interest was isolated based on three criteria. First, the corresponding probe set had differential expression between OSNs and OPCs that was statistically significant ($p < 0.05$, Student's unpaired t test). Second, for the corresponding probe set, the mADV across all cells analyzed was ≥250. Third, the logFC, based on the ratio of the mADVs of all OSNs and all OPCs for the corresponding probe set, exceeded the fold change threshold defined by the same 95% confidence interval equation described above. For this calculation, the mADV for each probe set across all cells was used to determine the 95% confidence interval that must be exceeded by the observed fold change. A mADV in OSNs or OPCs that was ≤1 was rounded up to 1. All differentially expressed transcripts described in the text were isolated using the criteria defined here unless otherwise noted (see text and Supplemental Table S3 at http://www.neuron.org/cgi/content/full/38/2/161/DC1).

Whole-Tissue Preparation, Hybridization, and Analysis

RNA extraction from whole tissues and microarray hybridization were performed as previously described (Mody et al., 2001). Data analysis was performed as described above.

Probes for Southern Blot and RNA In Situ Hybridization Analysis

All cDNA probes for Southern blot and RNA in situ hybridization analyses were obtained using specific PCR primers spanning most, if not all, of the 3′-most 600 bp of full-length cDNA sequences as annotated in GenBank (http://www.ncbi.nlm.nih.gov). PCR products were obtained from cDNA derived from adult or E15 mouse brain using Superscript II reverse transcriptase (Promega) or from a Lambda-ZAP cDNA library generated from adult MOE (Stratagene). All PCR products were cloned into the pTOPOII vector (Invitrogen). Sequences confirmed by DNA sequencing or restriction enzyme analysis were used as templates for both Southern blot and digoxigenin-labeled RNA probes (Roche Diagnostics).

In Situ Hybridization Analysis

RNA in situ hybridization was performed as described (Schaeren-Weimers and Gerlin-Mose, 1993). MOEs were dissected from young adult (3- to 4-week-old) mice or from the whole heads of E15 mice. OB/AOBs were dissected from e15 mice. Tissues were either fixed in 4% paraformaldehyde or freshly embedded in Tissue-Tek OCT compound (Sakura). Antisense and sense digoxigenin-labeled probes were generated as described above.

Acknowledgments

We wish to acknowledge Lubert Stryer for making this work possible and for inspired support and advice. We thank Renate Hellmiss-

Peralta for artistic work and illustrations and Cecilia Lee for help with the manuscript. We also thank Francois Guillemot for providing Mash1 mutant embryos and helpful discussions, and Erica Pantages, Bryan Dynlacht, Vahan Indjeian, and Kelvin Pham for assistance with some experiments. This work was supported by the Howard Hughes Medical Institute, and by a HFSPO LT fellowship (G.K.) and a Wellcome Trust Functional Genomics Programme grant (066790/E/02/Z to C.D. and G.K.).

Received: November 22, 2002
Revised: March 25, 2003
Accepted: April 9, 2003
Published: April 23, 2003

References

Bae, S., Bessho, Y., Hojo, M., and Kageyama, R. (2000). The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. Development *127*, 2933–2943.

Bertrand, N., Castro, D.S., and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. Nat. Rev. Neurosci. *3*, 517–530.

Beuchle, D., Struhl, G., and Muller, J. (2001). Polycomb group proteins and heritable silencing of *Drosophila* Hox genes. Development *128*, 993–1004.

Buck, L. (2000). Smell and taste: the chemical senses. In Principles of Neural Science, E. Kandel, T. Jessell, and J. Schwartz, eds. (New York: McGraw-Hill Publishing), pp. 625–647.

Cao, Y., and Dulac, C. (2001). Profiling brain transcription: neurons learn a lesson from yeast. Curr. Opin. Neurobiol. *11*, 615–620.

Cau, E., Gradwohl, G., Fode, C., and Guillemot, F. (1997). *Mash1* activates a cascade of bHLH regulators in olfactory neuron progenitors. Development *124*, 1611–1621.

Cau, E., Gradwohl, G., Casarosa, S., Kageyama, R., and Guillemot, F. (2000). Hes genes regulate sequential stages of neurogenesis in the olfactory epithelium. Development *127*, 2323–2332.

Cau, E., Casarosa, S., and Guillemot, F. (2002). Mash1 and Ngn1 control distinct steps of determination and differentiation in the olfactory sensory neuron lineage. Development *129*, 1871–1880.

Dulac, C., and Axel, R. (1995). A novel family of genes encoding putative pheromone receptors in mammals. Cell *83*, 195–206.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531–537.

Guillemot, F., Lo, L.C., Johnson, J.E., Auerbach, A., 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*, 463–476.

Guss, K.A., Nelson, C.E., Hudson, A., Kraus, M.E., and Carroll, S.B. (2001). Control of a genetic regulatory network by a selector gene. Science *292*, 1164–1167.

Heanue, T.A., Reshef, R., Davis, R.J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A.B., and Tabin, C.J. (1999). Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for *Drosophila* eye formation. Genes Dev. *13*, 3231–3243.

Heberlein, U., and Treisman, J.E. (2000). Early retinal development in *Drosophila*. In Invertebrate Eye Development, M.E. Fini, ed. (Berlin: Springer-Verlag), pp. 37–50.

Hinds, J.W. (1968). Autoradiographic study of histogenesis in the mouse olfactory bulb. J. Comp. Neurol. *134*, 305–322.

Hinks, G.L., Shah, B., French, S.J., Campos, L.S., Staley, K., Hughes, J., and Sofroniew, M.V. (1997). Expression of LIM protein genes Lmo1, Lmo2, and Lmo3 in adult mouse hippocampus and other forebrain regions: differential regulation by seizure activity. J. Neurosci. *17*, 5549–5559.

Iacopetti, P., Michelini, M., Stuckmann, I., Oback, B., Aaku-Saraste, E., and Huttner, W.B. (1999). Expression of the antiproliferative gene TIS21 at the onset of neurogenesis identifies single neuroepithelial cells that switch from proliferative to neuron-generating division. Proc. Natl. Acad. Sci. USA *96*, 4639–4644.

Iscove, N.N., Barbara, M., Gu, M., Gibson, M., Modi, C., and Winegarden, N. (2002). Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. Nat. Biotechnol. *20*, 940–943.

Klein, C.A., Seidl, S., Petat-Dutter, K., Offner, S., Geigl, J.B., Schmidt-Kittler, O., Wendler, N., Passlick, B., Huber, R.M., Schlimok, G., et al. (2002). Combined transcriptome and genome analysis of single micrometastatic cells. Nat. Biotechnol. *20*, 387–392.

Koyano-Nakagawa, N., Kim, J., Anderson, D., and Kintner, C. (2000). Hes6 acts in a positive feedback loop with the neurogenins to promote neuronal differentiation. Development *127*, 4203–4216.

Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat. Biotechnol. *14*, 1675–1680.

Malnic, B., Hirono, J., Sato, T., and Buck, L.B. (1999). Combinatorial receptor codes for odors. Cell *96*, 713–723.

Mody, M., Cao, Y., Cui, Z., Tay, K.Y., Shyong, A., Shimizu, E., Pham, K., Schultz, P., Welsh, D., and Tsien, J.Z. (2001). Genome-wide gene expression profiles of the developing mouse hippocampus. Proc. Natl. Acad. Sci. USA *98*, 8862–8867.

Murray, R.C., and Calof, A.L. (1999). Neuronal regeneration: lessons from the olfactory system. Semin. Cell Dev. Biol. *10*, 421–431.

Norlin, E.M., Alenius, M., Gussing, F., Hagglund, M., Vedin, V., and Bohm, S. (2001). Evidence for gradients of gene expression correlating with zonal topography of the olfactory sensory map. Mol. Cell. Neurosci. *18*, 283–295.

Norton, J.D., Deed, R.W., Craggs, G., and Sablitzky, F. (1998). Id helix-loop-helix proteins in cell growth and differentiation. Trends Cell Biol. *8*, 58–65.

Ohto, H., Kamada, S., Tago, K., Tominaga, S.I., Ozaki, H., Sato, S., and Kawakami, K. (1999). Cooperation of six and eya in activation of their target genes through nuclear translocation of Eya. Mol. Cell. Biol. *19*, 6815–6824.

Panda, S., Antoch, M.P., Miller, B.H., Su, A.I., Schook, A.B., Straume, M., Schultz, P.G., Kay, S.A., Takahashi, J.S., and Hogenesch, J.B. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. Cell *109*, 307–320.

Pignoni, F., Hu, B., Zavitz, K.H., Xiao, J., Garrity, P.A., and Zipursky, S.L. (1997). The eye-specification proteins So and Eya form a complex and regulate multiple steps in *Drosophila* eye development. Cell *91*, 881–891.

Porter, F.D., Drago, J., Xu, Y., Cheema, S.S., Wassif, C., Huang, S.P., Lee, E., Grinberg, A., Massalas, J.S., Bodine, D., et al. (1997). Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development. Development *124*, 2935–2944.

Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C., and Melton, D.A. (2002). "Stemness": transcriptional profiling of embryonic and adult stem cells. Science *298*, 597–600.

Riabowol, K., Draetta, G., Brizuela, L., Vandre, D., and Beach, D. (1989). The cdc2 kinase is a nuclear protein that is essential for mitosis in mammalian cells. Cell *57*, 393–401.

Rosner, B. (2000). Fundamentals of Biostatistics, Fifth ed. (Boston: PWS-Kent Publishing Co.).

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 Dev. *13*, 1794–1806.

Schaeren-Weimers, N., and Gerlin-Mose, A. (1993). A single protocol to detect transcripts of various types and expression levels in neural tissue and culture cells: in situ hybridization using digoxigenin-labeled cRNA probes. Histochemistry *100*, 431–440.

Scholzen, T., and Gerdes, J. (2000). The Ki-67 protein: from the known and the unknown. J. Cell. Physiol. *182*, 311–322.

Schumacher, A., and Magnuson, T. (1997). Murine Polycomb- and trithorax-group genes regulate homeotic pathways and beyond. Trends Genet. *13*, 167–170.

Tanabe, Y., and Jessell, T.M. (1996). Diversity and pattern in the developing spinal cord. Science *274*, 1115–1123.

Tanabe, Y., William, C., and Jessell, T.M. (1998). Specification of motor neuron identity by the MNR2 homeodomain protein. Cell *95*, 67–80.

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, W.M. Cowan, T. Jessell, and S.L. Zipursky, eds. (New York: Oxford University Press), pp. 474–508.

Yamagata, M., Weiner, J., and Sanes, J. (2002). Sidekicks: synaptic adhesion molecules that promote lamina-specific connectivity in the retina. Cell *110*, 649–660.

Yasunami, M., Suzuki, K., Houtani, T., Sugimoto, T., and Ohkubo, H. (1995). Molecular characterization of cDNA encoding a novel protein related to transcriptional enhancer factor-1 from neural precursor cells. J. Biol. Chem. *270*, 18649–18654.