

# Microarray-based discovery of highly expressed olfactory mucosal genes: potential roles in the various functions of the olfactory system

Mary Beth Genter,<sup>1</sup> Paul P. Van Veldhoven,<sup>2</sup> Anil G. Jegga,<sup>3</sup> Bhuvana Sakthivel,<sup>3</sup> Sue Kong,<sup>3</sup> Kristin Stanley,<sup>3</sup> David P. Witte,<sup>3</sup> Catherine L. Ebert,<sup>3</sup> and Bruce J. Aronow<sup>3</sup>

<sup>1</sup>Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio 45267-0056;

<sup>2</sup>Katholieke Universiteit Leuven, Department of Molecular and Cellular Biology, Pharmacology

Division, B-3000 Leuven, Belgium; and <sup>3</sup>Children's Hospital Medical Center, Cincinnati, Ohio 45229

Submitted 17 June 2003; accepted in final form 8 October 2003

**Genter, Mary Beth, Paul P. Van Veldhoven, Anil G. Jegga, Bhuvana Sakthivel, Sue Kong, Kristin Stanley, David P. Witte, Catherine L. Ebert, and Bruce J. Aronow.** Microarray-based discovery of highly expressed olfactory mucosal genes: potential roles in the various functions of the olfactory system. *Physiol Genomics* 16: 67–81, 2003. First published October 21, 2003; 10.1152/physiolgenomics.00117.2003.—We sought to gain a global view of tissue-specific gene expression in the olfactory mucosa (OM), the major site of neurogenesis and neuroregeneration in adult vertebrates, by examination of its overexpressed genes relative to that in 81 other developing and adult mouse tissues. We used a combination of statistical and fold-difference criteria to identify the top 269 cloned cDNAs from an array of 8,734 mouse cDNA elements on the Incyte Mouse GEM1 array. These clones, representing known and poorly characterized gene transcripts, were grouped according to their relative expression patterns across the other tissues and then further examined with respect to gene ontology categories. Approximately one-third of the 269 genes were also highly expressed in developing and/or adult central nervous system tissues. Several of these have been suggested or demonstrated to play roles in neurogenesis, neuronal differentiation, and/or neuronal migration, further suggesting that many of the unknown genes that share this expression pattern may play similar roles. Highly OM-specific genes included a palate, lung, and nasal epithelium carcinoma-associated gene (*Plunc*); sphingosine phosphate lyase (*Sgpl1*), and paraoxonase 1 (*Pon1*). Cell-type-specific expression within OM was established using in situ hybridization for several representative expression pattern clusters. Using the ENSEMBL-assembled mouse genome and comparative genomics analyses to the human genome, we assigned many of the unknown expressed sequence tags (ESTs) and poorly characterized genes to either novel or known gene products and provided predictive classification. Further exploration of this database will provide additional insights into genes and pathways critical for olfactory neurogenesis, neuronal differentiation, olfaction, and mucosal defense.

cDNA microarray; gene expression profiles; gene discovery; bioinformatics; neurodevelopment; *Plunc*; sphingosine phosphate lyase; paraoxonase

THE OLFACTORY MUCOSA (OM) is positioned at the junction of the respiratory tract and nervous system and performs functions integral to both organ systems. Olfactory sensory neurons are responsible for odorant transduction. In addition, inhaled air can be detoxified of toxic agents by mucociliary mechanisms and metabolic enzymes in the nasal respiratory and olfactory tissues. The OM is unique with respect to the rest of the

nervous system, as it has a basal cell population that provides a continuous supply of new olfactory neurons. Thus neurons lost to injury or toxic insult can be replaced (13, 14, 27, 33). At the present time, we lack an understanding of the overall genomic expression patterns that allow for this to be accomplished.

We have used the power of microarray-based genomics to examine gene expression in the mouse OM relative to that of many other tissues. For example, other studies utilizing the University of Cincinnati-Children's Hospital Medical Center (UC-CHMCC) Mouse Tissue Specific Gene Expression Database described gene expression in liver development and regeneration (42) and in various anatomical segments of the gastrointestinal tract (6). In the present study we are specifically interested in understanding gene expression patterns that distinguish the OM, with its capacity to support continuous neurogenesis, from more quiescent regions of the brain, as well as from adjacent tissue in the nasal cavity, namely, the nasal respiratory mucosa. We hypothesized that the molecular definition of these different, but closely related, tissues could be performed by comparing the expression and function of large numbers of genes in OM, nasal respiratory mucosa, and brain. This approach revealed a group of genes with expression highly restricted to the OM. Other clusters of genes were highly expressed in OM and other tissues, including in 1) OM and olfactory bulb; 2) OM and lung; 3) OM and liver; 4) OM and respiratory mucosa; and 5) an OM and skin gene clusters. We also demonstrated the utility of this approach for gene discovery with the detection of novel genes likely to be relevant in the life cycle of olfactory neurons.

## MATERIALS AND METHODS

**Animals, tissue samples, and RNA preparation.** Adult OM, olfactory bulb, and nasal respiratory mucosa were obtained from 6- to 8-wk-old male C57BL/6 mice (Jackson Laboratories) as part of an institutional consortium effort to develop a generalized mouse gene expression database. This database includes a wide range of developing, normal adult, and diseased adult mouse tissues, including multiple brain subregions, dorsal root ganglion, heart, lung, kidney, liver, reproductive and endocrine organs, immune organs, muscle, and skin. The ethmoid turbinates and caudal one-third of the nasal septum served as the source of OM RNA, and the nasoturbinates, maxillo-turbinates, and anterior two-thirds of the nasal septum were the source of respiratory mucosal RNA (for anatomy, see Ref. 87). Mice were killed by carbon dioxide inhalation according to institutional and national animal care guidelines, and tissues were rapidly dissected and snap frozen in liquid nitrogen. Frozen tissues, pooled from eight mice, were homogenized on ice with a stainless steel tissue grinder, and total RNA was prepared using TRI Reagent (Molecular Research Center,

Article published online before print. See web site for date of publication (<http://physiolgenomics.physiology.org>).

Address for reprint requests and other correspondence: M. B. Genter, Dept. of Environmental Health, Univ. of Cincinnati, Cincinnati, OH 45267-0056 (E-mail: MaryBeth.Genter@UC.edu).

Cincinnati, OH), following the manufacturer's protocol. Total RNA was precipitated with ethanol/sodium acetate and resuspended in DEPC-treated water. Polyadenylated [poly(A)<sup>+</sup>] RNA was prepared from total RNA using Oligotex (Qiagen, Valencia, CA). Poly(A)<sup>+</sup> RNA was quantitated using RiboGreen dye (Molecular Probes, Eugene, OR) and checked for RNase degradation using an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). A total of 600 ng of poly(A)<sup>+</sup> RNA (50 ng/μl) for each tissue was submitted for cDNA labeling and microarray hybridization.

**Fluorescent probe preparation and microarray hybridization.** Probe preparation and microarray hybridization were performed by Incyte Genomics (Palo Alto, CA) using the GEMBright random primer reverse-transcription labeling kit (Incyte Genomics). Labeled cDNA was prepared from each poly(A)<sup>+</sup> RNA sample using nucleotides labeled with the fluorescent dye Cy5. Labeled cDNAs were then subjected to competitive hybridization to mouse GEM1 microarrays, composed of spotted cDNAs corresponding to 8,638 sequence-verified IMAGE expressed sequence tag (EST) clones (500–5,000 nucleotides in length) with a low redundancy rate with respect to their corresponding gene products (a complete list of the genes represented on the GEM1 microarray can be found at [http://microarray.uc.edu/DataBases/Incyte\\_Mouse\\_GEM1.xls](http://microarray.uc.edu/DataBases/Incyte_Mouse_GEM1.xls)). For all samples in the database, hybridizations were performed in competition with Cy3-labeled mRNA from whole postnatal *day 1* (*P1*) mouse. This reference sample was chosen based on independent tests of its ability to generate reproducible competitive hybridizations from three different *P1* mouse isolates (data not shown). For each mRNA sample, duplicate arrays were hybridized. Additional dye-flip control experiments indicated a low dye effect, particularly compared with other Cy3/Cy5 labeling methods (data not shown). Dye effect was, however, subtracted from relative sample-specific gene expression based on per gene median normalization across a large number of samples (see below). Fluorescence intensity analyses and background subtraction were performed using an Axon Instruments scanner and GenePix software.

**Data analyses.** Primary quantitative data, spot geometry, and background fluorescence were examined using Incyte GEMTools software. Defective cDNA spots (irregular geometry, scratched, or less than 40% area compared with average) were eliminated from the data set. All data sets were normalized first using a balancing coefficient of the median of all Cy5 channel measurements divided by the median of all Cy3 channel measurements. Each microarray contained 192 control genes present as either nonmammalian single gene “spikes” or complex targets consisting of pools of genes expressed in most cell types, and each experimental mRNA sample was augmented with incremental amounts of nonmammalian gene RNA (2-fold, 4-fold, 16-fold, etc.) to permit assessment of the dynamic range attained within each microarray. Less than twofold variation was observed across the microarray series with respect to the 192 control genes (data not shown), providing additional support for the feasibility of interarray comparisons to detect genes regulated in a tissue-specific manner. Second-stage data analyses were carried out using GeneSpring software (Silicon Genetics, Redwood City, CA). A total of 269 cDNAs corresponding to genes highly expressed in the adult mouse OM were identified based on expression that was  $\geq 1.7$ -fold their median expression in all other tissues in the database. We found that this cutoff was capable of identifying many known important genes and was more efficient than ANOVA-based approaches (59), principally because of the number of genes expressed in OM that are also expressed, albeit relatively selectively, in other compartments such as liver and central nervous system.

The relative expression patterning of the 269 OM-overexpressed cDNAs was determined using the hierarchical tree clustering algorithm as implemented in the GeneSpring program using Pearson correlation applied to the log ratio of gene expression values. Relative gene expression ratio values were obtained using a three-step normalization strategy consisting of a per-spot normalization in which

sample channel was divided by the whole mouse reference channel followed by a LOWESS (“locally weighted scatter plot smoothing”) intensity-dependent normalization, and then per gene across the series of probe arrays using the median expression ratio observed for the gene across the entire University of Cincinnati-Children's Hospital Medical Center (UC-CHMCC) Mouse Tissue Specific Gene Expression Database.

The measured intensity of each gene was divided by its control channel value in each sample. When the control channel value was below 10.0, the data point was considered invalid. Intensity-dependent normalization was also applied, where the ratio was reduced to the residual of the LOWESS fit of the intensity vs. ratio curve as implemented in GeneSpring 4.2.1. The 50th percentile of all measurements was used as a positive control for each sample; each measurement for each gene was divided by this synthetic positive control, assuming that this was at least 0.01. Each gene was normalized to itself by making a synthetic positive control for that gene and dividing all measurements for that gene by this positive control, assuming it was at least 0.01. This synthetic control was the median of the gene's expression values over all the samples.

Hierarchical tree clustering allowed grouping of genes based on relative expression across the sample series; we used this approach to partition genes highly expressed in the OM into those that were also highly expressed in other tissues. For example, this allowed for clear identification of 105 of the 269 genes with strong expression in both OM and nervous system regions. A similar result could be obtained by using a Venn diagram function within GeneSpring to obtain the conjunction (Venn intersection) of the 269 OM genes and those identified by Student's *t*-test ANOVA as being nervous system-expressed vs. non-nervous system-expressed.

**Sequence analysis, identification, and functional classification of “olfactorome” genes.** The genes represented among the 269 olfactorome gene cDNAs were identified and categorized using a variety of resources. The Incyte GEM1 microarray contains a set of sequenced cDNA clones that have been mapped to representative UniGene sequences (<http://www.ncbi.nlm.nih.gov/UniGene/Mm.Home.html>) and BLASTN searches (<http://www.ncbi.nlm.nih.gov/BLAST/>) vs. the nonredundant nucleotide database over a period from July to November, 2002. This procedure allowed 1) matching of the ESTs to full-length gene products (principally using UniGene), 2) their assignment to corresponding representative mRNA from RefSeq (<http://www.ncbi.nlm.nih.gov/RefSeq/>), and then 3) classification of the encoded protein function using SwissProt (<http://us.expasy.org/sprot/>), Gene Ontology (<http://www.geneontology.org>), BLINK (<http://www.ncbi.nlm.nih.gov/sutils/static/blinkhelp.html>), HomoloGene (<http://www.ncbi.nlm.nih.gov/HomoloGene>), and additional Medline-based literature analyses (<http://www.ncbi.nlm.nih.gov/PubMed>). For unannotated EST sequences, additional BLASTN and BLAT (<http://genome.ucsc.edu>) searches were performed to identify genomic region hits, and potential gene hits were based on the use of University of California, Santa Cruz (UCSC) Genome Browser (<http://genome.ucsc.edu>) and ENSEMBL (<http://www.ensembl.org/>) resources (December, 2002).

Gene annotations (functional classification and groupings) were based on high scoring homologies of the corresponding encoded proteins to known proteins using biochemical function and molecular process concepts represented in SwissProt and Gene Ontology into the following categories: cytoskeletal, metabolic enzymes, xenobiotic metabolizing enzymes, intracellular protein trafficking, apoptosis related, cell cycle control related, immune system/inflammatory response related, signal transduction, extracellular matrix/cell adhesion, proteinase inhibitors, DNA/RNA binding/transcription factors, calcium homeostasis, membrane proteins, and others (Table 1). From the 269 gene cDNAs that we designated as top members of the “olfactorome,” we identified 41 cDNAs/ESTs that lacked useful genomic annotations or designation as a gene. We subcategorized these 41

Table 1. Characterization of highly expressed olfactory mucosal genes (with GenBank/RefSeq accession numbers)

Cytoskeletal	Fibroblast growth factor receptor 1 [ <i>Fgfr1</i> ; BC010200 (GenBank); NM_010206 (RefSeq)]
Kinesin heavy chain 1A ( <i>Kif1a</i> ; NM_008440)	Guanine nucleotide binding protein alpha o [ <i>Gnao</i> ; AK047197 (GenBank); NM_010308 (RefSeq)]
Adducin 2 $\beta$ ( <i>Add2</i> ; NM_013458)	Extracellular matrix/cell adhesion
Myosin light chain ( <i>Mylpf</i> ; NM_016754)	Procollagen type IX $\alpha$ -1 ( <i>Col9a1</i> ; NM_007740)
Tau ( <i>Mapt</i> ; NM_010838)	Lumican ( <i>Lum</i> ; NM_008524)
Dynein heavy chain 11, axonal ( <i>Dnahc11</i> ; NM_01006)	Procollagen type I $\alpha$ -1 ( <i>Colla1</i> ; NM_009926)
Kinesin light chain 2 ( <i>Klc2</i> ; NM_008451)	Secreted phosphoprotein 1 ( <i>Spp1</i> ; NM_009263)
Myozenin ( <i>Myoz1</i> ; NM_021508)	A disintegrin and metalloprotease domain 22 ( <i>Adam22</i> ; AB009674)
Septin 3 ( <i>Sept3</i> ; NM_011889)	Neural cell adhesion molecule 1 [ <i>Ncam1</i> ; X15052 (GenBank); NM_010875 (RefSeq)]
Profilin 2 ( <i>Pfn2</i> ; NM_019410)	Parvin [ <i>Parva</i> ; AF237774 (GenBank); NM_020606 (RefSeq)]
Metabolic enzymes	Surfactant associated protein C ( <i>Sftpc</i> ; NM_011359)
Sphingosine phosphate lyase [ <i>Sgpl1</i> ; BC026135 (GenBank); NM_009163 (RefSeq)]	Fibromodulin ( <i>Fmod</i> ; NM_021355)
Prostaglandin D2 synthase ( <i>Ptgsd</i> ; AB006361)	Proteinase inhibitors
Betaine-homocysteine methyltransferase ( <i>Bhmt</i> ; NM_016668)	Nexin (Clade E) ( <i>Serpine2</i> ; NM_009255)
Carboxylesterase 3 ( <i>Ces3</i> ; NM_053200)	Extracellular matrix proteinase inhibitor ( <i>Expi</i> ; NM_007969)
Glycine methyltransferase ( <i>Gnmt</i> ; NM_010321)	$\alpha$ -2-Macroglobulin ( <i>A2m</i> ; NM_007376)
Plasminogen ( <i>Pgl</i> ; NM_008877)	Clade F ( <i>Serpinf2</i> ; NM_008878)
Fatty acid amide hydrolase [ <i>Faah</i> ; AK079598 (GenBank); NM_010173 (RefSeq)]	Kininogen ( <i>Kng</i> ; NM_023125)
$\beta$ -Glucuronidase ( <i>GusB</i> ; NM_010368)	DNA/RNA binding/transcription factor related
Lipase ( <i>Lipc</i> ; NM_008280)	Inhibitor of DNA binding 4 ( <i>Idb4</i> ; NM_031166)
Calpain 6 ( <i>Capn6</i> ; NM_007603)	Cold inducible RNA binding protein ( <i>Cirbp</i> ; NM_007705)
Aminolevulinic acid synthase 1 ( <i>Alas1</i> ; NM_020559)	Paternally expressed 3 ( <i>Peg3</i> ; AB003040)
3- $\beta$ -Hydroxysteroid dehydrogenase ( <i>Hsd3b4</i> ; NM_008294)	High mobility group nucleosomal binding domain 3 ( <i>Hmgn3</i> ; NM_026122)
Xenobiotic metabolizing enzymes	Zinc finger protein 238 ( <i>Zfp238</i> ; NM_013915)
Alcohol dehydrogenase 1 ( <i>Adh1</i> ; NM_007409)	Eukaryotic translation elongation factor 1 $\alpha$ -2 ( <i>Eef1a2</i> ; NM_007906)
Alkaline phosphatase 2 ( <i>Akp2</i> ; NM_007431)	Myeloid/lymphoid or mixed lineage leukemia ( <i>Mll-1</i> ; L17069)
Esterase 22 ( <i>Es22</i> ; NM_133660)	P68 D-E-A-D box polypeptide 5 ( <i>Ddx5</i> ; NM_007840)
3'-Phosphoadenosine 5'-phosphosulfate synthase 2 ( <i>Paps2</i> ; NM_011864)	Max dimerization protein 4 ( <i>Mad4</i> ; NM_010753)
Cyp2f2 ( <i>Cyp2f2</i> ; NM_007817)	Amyloid $\beta$ (A4) precursor ( <i>Aplp2</i> ; NM_009691)
Paraoxonase 1 ( <i>Pon1</i> ; NM_011134)	Calcium homeostasis
Amine-N-sulfotransferase ( <i>Sultn</i> ; NM_016771)	Ca <sup>2+</sup> channel $\beta$ 3 subunit ( <i>Cacnb3</i> ; NM_007581)
Esterase 1 ( <i>Es1</i> ; NM_007954)	Voltage-dependent Ca <sup>2+</sup> channel $\alpha$ 2 $\delta$ subunit 1 ( <i>Cacna2d1</i> ; NM_009784)
Lung $\alpha/\beta$ hydrolase 3 ( <i>Abhy3</i> ; NM_134130)	Calsequestrin 2 ( <i>Casq2</i> ; NM_009814)
Mitochondrial aldehyde dehydrogenase 2 ( <i>Aldh2</i> ; NM_009656)	Troponin C2 ( <i>Tnnc2</i> ; NM_009394)
Superoxide dismutase ( <i>Sod1</i> ; CA478361)	Calcium channel, voltage-dependent $\gamma$ subunit 1 ( <i>Cacng1</i> ; NM_007582)
Intracellular protein trafficking	Membrane proteins
Rab3A ( <i>Rab3A</i> ; NM_009001)	ABC family G member 1 ( <i>Abcg1</i> ; NM_009593)
Amyloid (A4) precursor protein A2 ( <i>Apba2</i> ; L34676)	Solute carrier family 4 (anion exchanger) ( <i>Slc4a1</i> ; NM_011403)
Vesicle-associated membrane protein 2 [ <i>Vamp2</i> ; AK090178 (GenBank); NM_009497 (RefSeq)]	ATP binding cassette subfamily A (ABC1), member 2 ( <i>Abca2</i> ; NM_007379)
Nuclear transport factor 2 (placental protein 15) [ <i>Nutf2</i> ; BC003955 (GenBank); NM_026532 (RefSeq)]	Solute carrier family 5 (sodium, glucose cotransport) ( <i>Slc5a1</i> ; NM_019810)
Blocked early in transport 1 [ <i>Bet1</i> ; BC005572 (GenBank); NM_009748 (RefSeq)]	Solute carrier family 27 (fatty acid transporter) ( <i>Slc27a2</i> ; NM_011978)
Apoptosis related	Erythrocyte protein band 7.2 ( <i>Epb7.2</i> ; NM_013515)
Cell death including DNA fragmentation factor ( <i>Cidea</i> ; NM_007702)	Amyloid $\beta$ (A4) precursor like protein 1 ( <i>Aplp1</i> ; NM_007467)
Deoxyribonuclease II ( <i>Dnase2a</i> ; NM_010062)	Pantophysin ( <i>Pphn</i> ; NM_013635)
Fas apoptotic inhibitory molecule ( <i>Faim</i> ; NM_011810)	Interferon induced transmembrane protein 3-like ( <i>Ifitm3l</i> ; NM_030694)
Cell cycle control	Other
Pleiotrophin ( <i>Ptn</i> ; NM_008973)	Olfactory marker protein ( <i>Omp</i> ; NM_011010)
Immune system/inflammatory response related	Fetuin beta ( <i>Fetub</i> ; NM_021564)
Rat generating islet derived homolog ( <i>Reg3g</i> ; NM_011260)	Lipin 1 ( <i>Lpin1</i> ; NM_015763)
Polymeric immunoglobulin receptor ( <i>Pigr</i> ; NM_011082)	Uteroglobin (Clara cell secretory protein) ( <i>Scgb1a1</i> ; NM_011681)
Immunoglobulin heavy chain 1 (serum IgG2a) ( <i>Igh-1</i> ; BC018365)	Major urinary protein 2 ( <i>Mup2</i> ; M16356)
Lactotransferrin ( <i>Ltf</i> ; NM_008522)	Crystallin lambda ( <i>Cryl1</i> ; NM_030004)
Hemochromatosis ( <i>Hfe</i> ; NM_010424)	Crystallin $\mu$ ( <i>Crym</i> ; NM_016669)
Immunoglobulin superfamily containing leucine-rich repeat ( <i>Islr</i> ; NM_012043)	Crystallin gamma S ( <i>Crygs</i> ; NM_009967)
Palate, lung, and nasal epithelium carcinoma associated ( <i>Plunc</i> ; NM_011126)	Fatty acid binding protein 7, brain 0 ( <i>Fabp7</i> ; NM_021272)
Cathelin-like protein ( <i>Camp</i> ; NM_009921)	Synuclein ( <i>Suca</i> ; NM_009221)
Signal transduction	ATP binding cassette subfamily C (CFTR/MRP), member 1b ( <i>Abcc1</i> ; NM_008576)
Homer 2 (Cupidin) ( <i>Homer2-pending</i> ; AB017136)	Hemoglobin alpha chain complex ( <i>Hba</i> ; AA109900)
Regulator of G protein signaling 5 ( <i>Rgs5</i> ; NM_009063)	Plexin A3 ( <i>Plxna3</i> ; NM_008883)
	Osteoblast-specific factor (fasciclin I-like) ( <i>Osf2-pending</i> ; NM_015784)
	Dihydropyrimidinase-like 3 ( <i>Dpysl3</i> ; NM_009468)

cDNAs/ESTs into 6 major groups, principally based on their expression pattern: highly expressed in 1) OM alone; 2) olfactory bulb and OM; 3) OM and respiratory mucosa; 4) whole brain and olfactory bulb and OM; 5) whole brain and olfactory bulb; and 6) a mixed patterns. Each of these cDNAs/ESTs was then subjected to a BLAT search against the UCSC Genome Browser (on Mouse Feb. 2003 Freeze) and BLASTN searches against Celera (<http://www.celera.com/>) and ENSEMBL databases for full-length known and putative mouse mRNAs.

**Identification of cis-elements in coordinately expressed genes.** To identify putative cis-acting regulatory regions and elements in genes coordinately expressed in the olfactory system, we selected three genes from each of the first three expression subcategories above and subjected each gene to mouse-human ortholog pair analysis for conserved cis-regulatory elements (40). To do this, the complete genomic sequences of the selected mouse genes and the corresponding human orthologs were extracted from the Celera/ENSEMBL/UCSC mouse and human databases, respectively. Advanced Pip-Maker (<http://bio.cse.psu.edu>) was used to compute alignments of similar regions in the mouse-human orthologous DNA sequences using the chaining option (67). The TRANSFAC database, a database of eukaryotic transcription factors and DNA-binding site profiles (<http://transfac.gbf.de>) (88), was utilized in the program MatInspector (Professional ver. 4.3, 2000; <http://www.genomatix.de/>) to locate putative transcription factor binding sites in each of these genomic sequences. We then compared genomic regions within and across the three groups to identify transcription factor binding sites in common.

**In situ hybridization analysis of highly expressed OM genes.** In situ hybridization was performed for localization of selected highly expressed genes using <sup>35</sup>S-labeled probes which were generated from the Incyte clones (Incyte Pharmaceuticals; St. Louis, MO) available from the University of Cincinnati Microarray Core (<http://microarray.uc.edu>). The probes were as follows: W08172 for sphingosine phosphate lyase (*Sgpl1*); AA028678 for *Plunc*; W17967 for *Pon1*; AA220582 for cytochrome P-450 2f2 (*Cyp2f2*); AA498773 for crp ductin (*Dmbt1*); AA474964 for lactotransferrin (*Ltf*); and W11846, which was originally an EST, but which has been associated with the recently characterized *Alms1* locus (16, 17). We sequenced each clone from the T3 and T7 primer sites and found two chimeric clones: W08172 contained an unidentified sequence at the T3 (antisense) end, and AA220582 had high identity with *CYP2f2* from the T3 (antisense) end and with *Dmbt1* from the T7 (sense) end. To generate a vector containing only the *Sgpl1* sequence for use as a probe, a restriction map of the *Sgpl1* sequence was generated (based on Ref. 65), and based on known restriction sites, the *NotI* linearized plasmid was cut with *BbsI*, blunt-end subcloned, and validated for contaminating fragment removal using *EcoRI/BamHI* and *HindIII/DraI*. This modified clone was used as a template for <sup>35</sup>S-labeled sense and antisense probe synthesis.

**Sphingosine phosphate lyase activity measurements.** Sphingosine phosphate lyase (*Sgpl1*) activity was measured in tissues isolated from male Long-Evans rats because of the relative abundance of several tissues, including OM, in rat vs. mouse, and for ease of comparison with previously published results (79). Tissues were homogenized in 0.25 M sucrose, 5 mM MOPS-NaOH buffer, pH 7.2, 1 mM DTT, 1 mM EDTA, either with a Dounce homogenizer or, for more refractive tissues, a Polytron PT7 device (Kinematica Benelux). *Sgpl1* activity was measured as described before using [4,5-<sup>3</sup>H]sphinganine-1-phosphate (80).

**Sgpl1 antiserum preparation.** Antiserum against (human) *Sgpl1* was obtained as follows. A His-tagged human *Sgpl1* (amino acids 59–568) fusion protein, expressed in DH5 $\alpha$  *Escherichia coli* cells transformed with plasmid pVB001 (81) and induced with 0.2% (wt/vol) D-arabinose, was purified by means of Ni-NTA agarose (Qiagen), followed by SDS-PAGE and blotting to nitrocellulose. Portions of the blot containing ~100  $\mu$ g poly-His-fusion protein were

dissolved in DMSO (1), mixed with an equal volume of Freund's complete adjuvant and injected subcutaneously in a New Zealand White rabbit after procurement of an aliquot of preimmune serum. Booster injections were given every 4 wk with 50  $\mu$ g of protein mixed with incomplete adjuvant. Ten days after the third booster, the animal was bled. The final preparation recognized a single band of 65 kDa in blots prepared from human, rat, and mouse liver (P. P. Van Veldhoven, unpublished data).

**Immunohistochemistry.** Sections (5  $\mu$ m in thickness) were cut from archived paraffin-embedded nasal cavity sections from untreated male Long-Evans rats and male C57BL/6J mice. Tissues had been prepared for embedding by fixation in 10% neutral buffered formalin and decalcified in either 10% formic acid or 0.3 M EDTA (pH 7.4). Anti-Cyp2f2 antiserum (previously described, see Ref. 63) was provided by Dr. Gerold Yost (University of Utah). Immunohistochemistry was performed with 1:100 dilution of preimmune serum or 1:100 dilution of *Sgpl1* antiserum or 1:2,500 dilution of anti-Cyp2f2 antiserum. Immunoreactivity was detected using anti-rabbit HRP-conjugated secondary antibody (Dako, Carpinteria, CA) and aminoethyl-carbazole as chromogen as previously described (29).

**Data archive.** Gene identities, expression data, cluster groups, and the functional categories of the dynamically regulated genes are available on our microarray database web server [<http://genet.cchmc.org>]; log in (as guest or with username: olfepi; password: olfepi), select "IncyteMouseGEM1 (127\_CloneID)," then select the "Olfactorome" subdirectory from the "Gene Lists" folder; select the list of genes of interest, and click on the paperclip icon to the right of the list.

## RESULTS

**Identification of genes highly expressed in adult OM.** To define genes important for neurogenesis and OM homeostasis, we identified arrayed cDNAs that hybridized to adult mouse OM cDNA at a level at least 1.7-fold higher than their median expression in all other tissues. We found 269 cDNAs and ESTs in OM with increased expression in replicate hybridizations compared with reference. Of the known genes, the highest expression in OM was noted for *Plunc* (86). *Sgpl1*, *Pon1*, *Cyp2f2*, and *Ltf* were among other highly expressed OM genes. We subjected the log<sub>2</sub>-transformed ratios of the overexpressed genes to hierarchical tree cluster analysis and clustered the genes into those that were highly expressed only in OM vs. those also highly expressed in other tissues (Fig. 1).

The 269 highly expressed OM genes showed both temporally and spatially restricted expression in other regions of the central nervous system. Seventy-seven genes were highly expressed in OM and in at least one brain region. Some of these genes exhibited very different behavior from one brain region to another. For example, lumican (*Lum*) and procollagen, type I,  $\alpha$ 1 (*Colla1*) were highly expressed only in OM, postnatal day 2 brain and dorsal root ganglion, and significantly underexpressed in many other brain regions represented in the analysis. Genes that were highly expressed in both olfactory bulb and OM encoded many cytoskeletal proteins and proteins associated with vesicular trafficking [e.g., RAB3B (*Rab3b*); vesicle-associated membrane protein 2 (*Vamp2*); pantophysin (*Pphn*); and amyloid- $\beta$  (A4) precursor protein binding (*Aplp2*) (35, 47, 53, 75)], axonal guidance [e.g., plexin A3 (*Plxna3*) and dihydropyrimidinase-like 3 (*Dpysl3*); NM\_009468; orthologous to unc-33, which is associated with axonal guidance in *Caenorhabditis elegans* (12, 52, 68)], and signal transduction [e.g., the neuronal early immediate gene homer (*Homer2-pending*)

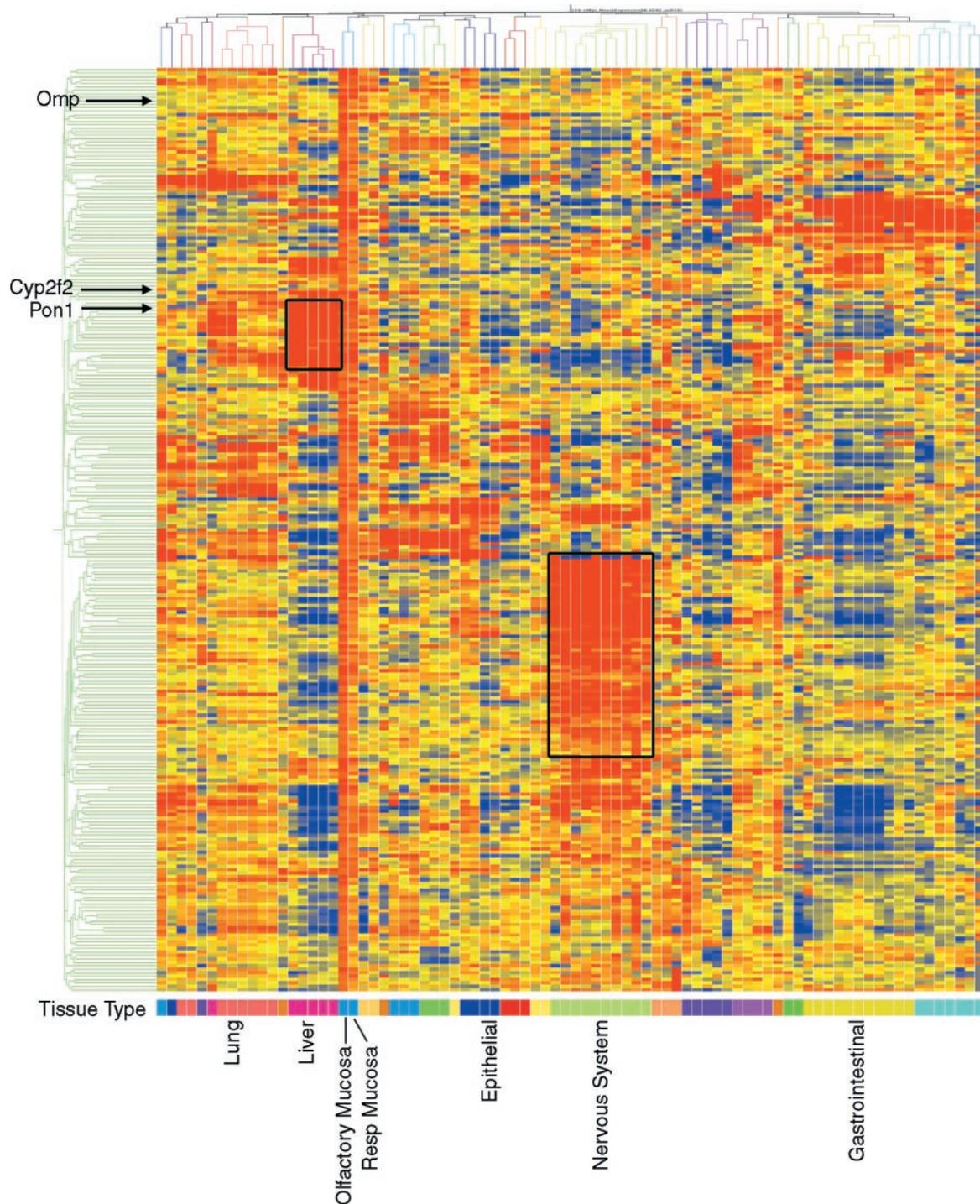


Fig. 1. Hierarchical tree clustering of the 269 highly expressed olfactory genes and expressed sequence tags (ESTs). Intensity in the red and blue color range indicates upregulated and downregulated RNAs, respectively. Each row represents a gene. Each column represents an individual tissue. Each tissue sample RNA was hybridized in duplicate with the mean value shown. The position of the olfactory marker protein gene (*Omp*), a marker of mature olfactory neurons, is indicated. Note that there is a large cluster of olfactory mucosal (OM) genes that are also highly expressed in the brain (larger box). Brain regions included in this cluster are embryonic *day 18* brain, postnatal *day 2* brain, and the following regions of the adult mouse brain: hippocampus, hypothalamus, nucleus accumbens, olfactory bulb, striatum, cerebellum, spinal cord, and dorsal root ganglion. Genes included in this cluster include synuclein- $\alpha$  (*Snca*; NM\_009221); calcium channel  $\beta 3$ -subunit (*Cacnb3*; NM\_007581); leucine-rich repeat protein 1, neuronal (*Lrrm1*; NM\_008516); *Tau4* (NM\_010838); and adducin 2 $\beta$  (*Add2*; NM\_013458). There are also separate clusters of genes with high coexpression in OM and liver [smaller box; e.g., paraoxonase 1 (*Pon1*; NM\_011134, shown); aldehyde dehydrogenase 1 (*Adh1*; NM\_007407); and glycine *N*-methyltransferase (*Gnmt*; NM\_010321)]; gastrointestinal tract, skin, and lung (e.g., cytochrome *P*-450 2f2; *Cyp2f2*; NM\_007817, shown). Both genes and samples are clustered using Pearson correlation as implemented in GeneSpring 6.

Table 2. *Sphingosine phosphate lyase (Sgpl1) gene expression and activity in selected tissues*

Tissue	Normalized Gene Expression Level	P Value (t-test)	Activity, (pmol·min <sup>-1</sup> ·mg <sup>-1</sup> )
Thymus	5.56	0.004	
Jejunum	4.92	0.007	
Ileum	4.90	0.022	
Olfactory mucosa	4.88	0.008	216.1 ± 1.2
Skin (postnatal day 2)	2.10	0.020	
Nasal respiratory mucosa	1.69	0.044	49.0 ± 15.5
Liver	1.38	0.105	35.7 ± 1.7
Lung	0.82	0.225	23.9 ± 9.2
Cerebellum	0.70	0.036	45.3 ± 16.2
Olfactory bulb	0.62	0.147	74.2 ± 4.6
Hippocampus	0.40	0.002	
Nucleus accumbens	0.40	0.002	

Normalized expression is derived from the raw expression of a gene of interest (*Sgpl1* in this case) compared with the average expression of all genes in the UC-CHMCC gene expression database. *P* values refer to the difference in expression between the respective tissues and the reference control (*P* < 0.05 was considered significant). *Sgpl1* activity measurements (means ± SD; *n* = 2) were performed on the respective tissues from male Long-Evans rats.

and calcium channel β3-subunit (*Cacnb3*)]. Fifteen genes and ESTs were highly expressed in both OM and in *E18.5* mouse brain. Of the known genes highly coexpressed in OM and in developing brain, eight are clearly implicated in nervous system function, neuronal development, and/or epithelial differentiation. These are tau (*Mapt*), stathmin (*Stmn4*; 50, 51, 57, 60), fatty acid binding protein 7 (*Fabp7*; 23), *Meg3* (66), pleiotrophin (*Ptn*; 74), *Mad4* (39), *Dpysl3*, and *Plxa3*.

*Sgpl1* activity measurements. Several tissues in the mouse gene expression database displayed relatively high *Sgpl1* gene expression, including OM, thymus, jejunum, ileum, and skin (Table 2). In a screening of the main organs of the rat, highest activities were previously detected in intestinal mucosa, Harderian gland, and liver (79). Measurement of the specific activity

of *Sgpl1* in rat OM in the present study confirmed that it was indeed high compared with other rat tissues (Table 2).

*In situ hybridization and immunohistochemistry studies.* We performed *in situ* hybridization to independently confirm results of the gene expression analysis and to localize several previously undescribed OM genes, *Sgpl1*; W11846, originally an EST that has recently been associated with the *Alms1* locus (16, 17); *Dmbt1*, *Cyp2f2*, and *Pon1*. The expression of *Pon1* and *Cyp2f2* was highly localized to the OM airway surface and in acinar cells and ducts of Bowman's glands (Fig. 2). This distribution was further confirmed by immunohistochemistry for *Cyp2f2*. This is a classic distribution for many xenobiotic metabolizing enzymes in OM (10). In contrast, *Sgpl1* and *Alms1* were distributed in the olfactory epithelium (i.e., above the epithelial basement membrane) but not in the subepithelial compartment (Fig. 2). Immunohistochemistry suggested *Sgpl1* protein localization to the perinuclear region of mature olfactory neurons (Fig. 2); colocalization studies with olfactory marker protein (OMP) to confirm localization to mature neurons were deemed unnecessary based on the localization of the histochemical reaction products in tissue sections. *Dmbt1* expression was localized in the lateral nasal gland (Fig. 2). *Plunc* expression was included as a positive control for the *in situ* methodology, as its high and punctate expression has been previously described in both developing and adult OM (86) (Fig. 2).

*Functional classification of genes highly expressed in the adult mouse olfactory system.* We classified each of the highly expressed genes based on biochemical function of the encoded protein (Table 1). Overall, the function of ~60% of the genes/ESTs could be identified based on public databases. The largest fraction of identified genes is related to metabolism, with the encoded proteins participating in metabolism of both endogenous and exogenous substrates (although we separated metabolic enzymes and xenobiotic metabolizing enzymes into separate categories, we realize that these distinctions are not

Table 3. *Shared clusters of cis-elements in the conserved (human-mouse) noncoding regions of coexpressed genes in different segments of the olfactory system*

Segment	Genes	Putative Cis-Regulatory Modules in the Upstream 3-kb Region and Occurring Within a 200-bp Window
OM	<i>Tpd52</i> <i>1700017B05Rik</i> <i>4921524P20Rik</i>	1) CLOX, OCT1, BRNF, FKHD 2) CREB, BRNF, SORY, OCT 3) MEIS, ETSF, OCT1, ETSF
OM and respiratory mucosa	<i>1110017O10Rik</i> <i>Sec14l2</i> <i>BF232834</i>	1) GATA, FKHD, CREB, ETSF 2) GATA, GATA, GFII, ETSF 3) GATA, SMAD, CREB 4) GATA, MEF2, CREB
OM and brain	<i>Nup88 (Prei2)</i> <i>Gpm6a</i> <i>Rsec15l1</i>	1) CREB, NKXH, CLOX, MEF2, BRNF, HNF1, FKHD 2) MEF2, GATA, SRFF, NKXH, NKXH, BRNF 3) OCT1, CREB, NKXH, CLOX, LEFF, FKHD, MYT1 4) SORY, GATA, SRFF, GATA, NKXH, NKXH, MEIS

Genes highly expressed in olfactory mucosa (OM): *Tpd52*, tumor protein D52; *1700017B05Rik*, XM\_134943; *4921524P20Rik*, XM\_143313, WD40 repeats domain containing gene. Genes highly expressed in OM and respiratory mucosa: *1110017O10Rik*, XM\_130346, domain found in Dishevelled, Egl-10, and Pleckstrin; *Sec14l2*, NM\_144520, SEC14-like 2 (*S. cerevisiae*); *BF232834*, aldolase 1 epimerase like. Genes with high expression in OM and brain: *Nup88 (Prei2)*, nucleoporin 88-kDa (preimplantation protein 2); *Gpm6a*, glycoprotein m6a; *Rsec15l1*, BF581861. Abbreviations for the modules follow here. CLOX, CLOX and CLOX homology (CDP) factors; OCT1, octamer factor 1; BRNF, Brn POU domain factors; FKHD, fork head domain factor; CRBP (CREB), cAMP-responsive element binding proteins; SORY, *sox/sry-sex/testis* determining and related HMG Box factors; MEIS, homeodomain factor aberrantly expressed in myeloid leukemia; ETSF, human and murine ETS1 factors; GATA, GATA-binding factor; GFII, growth factor independence-transcriptional repressor; SMAD, vertebrate SMAD family of transcription factors; MEF2, myocyte-specific enhancer-binding factor; NKXH, NKX/DLX homeodomain sites; HNF1, hepatic nuclear factor; SRFF, serum response element binding factor; LEFF, TCF/LEF-1, involved in the Wnt signal transduction pathway;

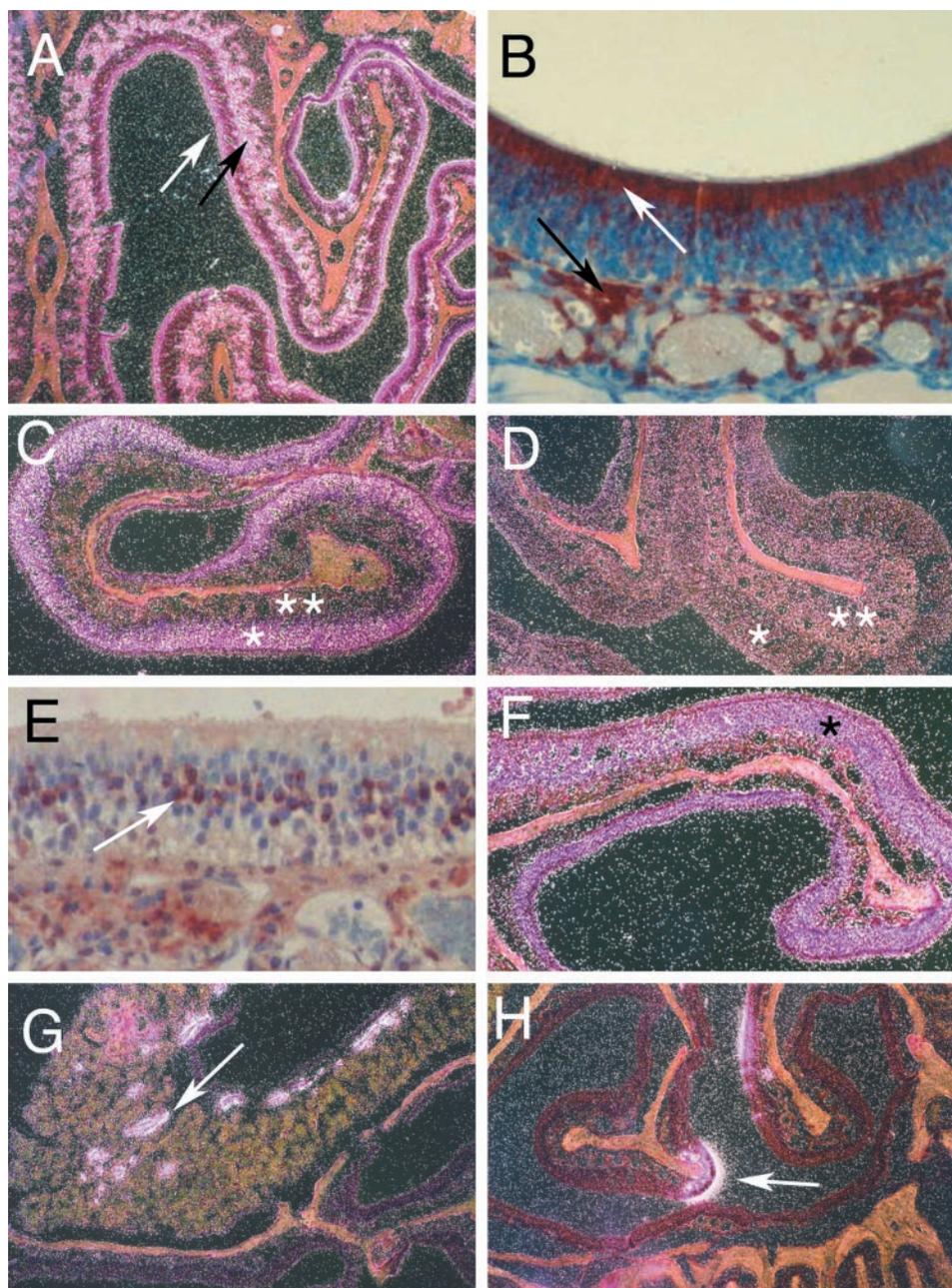


Fig. 2. Localization of novel genes and ESTs in the OM using in situ hybridization and immunohistochemistry. *A*: cytochrome *P*-450 2f2 (*Cyp2f2*) is shown by in situ hybridization in this bright-field photomicrograph to localize in the Bowman's glands (black arrow) and along the epithelial surface (white arrow) of the dorsal septum and the turbinates lining the dorsal medial airways of the mouse. *B*: immunohistochemistry using an anti-*Cyp2f2* antibody confirms localization in apical cytoplasm of sustentacular cells (white arrow) and duct and acinar cells of Bowman's glands (black arrow). *C*: in situ hybridization with antisense *Sgpl1* reveals expression in the olfactory epithelial layer (\*), but not in the subepithelial compartment (\*\*), suggesting that *Sgpl1* is present in either olfactory sensory neurons or in sustentacular (supporting) cells. *D*: in situ hybridization with *Sgpl1* sense probe does not result in hybridization in the olfactory epithelial layer (\*) or the subepithelial compartment (\*\*). *E*: immunohistochemistry with anti-*Sgpl1* antibody reveals localization in the perinuclear layer of mature olfactory sensory neurons (example indicated with arrow). *F*: *Alms1* antisense probe similarly localizes to the olfactory epithelial compartment (\*), but not subepithelial structures (\*\*). *G*: *Dmbt1* localizes to the lateral nasal gland, but not in OM (white arrow). *H*: *Plunc* expression (arrow) was the highest of any gene in the OM compared with other genes in the mouse gene expression database. Its expression is punctate and extremely high in olfactory epithelium (often displaying high expression at the junctions of OM and respiratory mucosa), consistent with previous localization studies (86).

absolute, i.e., a given enzyme might well metabolize endogenous and exogenous substrates). The predicted protein products of these genes bear homology to members of protein families that participate in DNA-dependent regulation of transcription, cell differentiation, positive regulation of cell proliferation, carbohydrate transport, cell adhesion, metabolism, or signal transduction.

**Gene discovery in the olfactory system.** The use of EST and cDNA microarrays corresponding to unknown genes provides the opportunity for gene discovery relevant to the multiple cellular components of OM. As shown in Fig. 3, a total of 137 cDNAs and ESTs out of the 269 could be associated with a well-annotated mouse gene (having a gene symbol and an mRNA). Of these 137 genes, 103 had a corresponding NCBI

RefSeq protein entry for which an ortholog could be identified. Of the 34 mouse genes that did not show a human entry in the initial search, 27 of them could be mapped to the human genome to a likely human ortholog. However, there were 7 mouse genes for which no human ortholog could be found after BLASTN and BLAT searches against human genome. Of the remaining 120 cDNA/ESTs, of 269 that were poorly annotated, 47 could be assigned to a gene with a valid symbol and an mRNA entry in the RefSeq database of GenBank (largely representing Riken long clones). To find the human orthologs for these 47 mRNAs, BLAT and BLASTN searches were performed against the human genome using corresponding full-length Riken clone sequence or mouse genomic sequence. Taken together, of the total 269 olfactorome cDNA/ESTs,

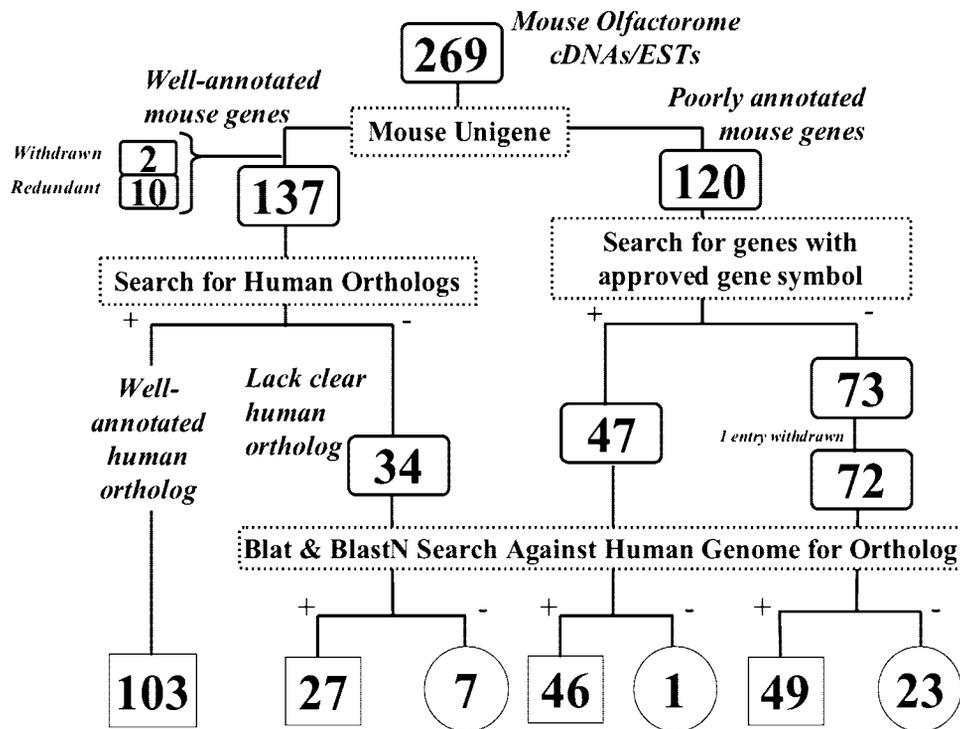


Fig. 3. Flow chart for the annotation of 269 gene elements of the olfactorome. A total of 137 cDNA/ESTs out of the 269 could be associated with a well-annotated mouse gene, and of these 137 genes, for 103 the human ortholog could be identified. Of the 34 mouse genes which did not show a human entry in the initial search, 27 of them could be mapped to the human genome. However, there were 7 mouse genes for which no human ortholog was found even after a BLASTN and BLAT search against human genome. Of the remaining 120 cDNA/ESTs out of 269, and that were poorly annotated, there were 47 that had an approved gene symbol and an mRNA entry in the RefSeq database of GenBank. To find the human orthologs for these 47 mRNAs, a BLAT and BLASTN search was performed against the human genome. There were 46 human hits, and for one there was no human ortholog identified. The 73 cDNA/ESTs out of 120 with no approved gene symbol association were also subjected to a BLAT and BLASTN search against human genome. For 49, we could identify an associated human counterpart. However, 23 were unmatched. One (CA490663) was withdrawn from the GenBank database. Thus, of the total 269 olfactorome cDNA/ESTs, there were 31 for which no human ortholog is available in the present databases. The complete table for these annotations can be obtained at the <http://genet.cchmc.org> data archive ("269.xls") as an attachment to the gene list file titled "269\_treeOrder".

there were 31 for which no human ortholog is available in the present databases.

To further investigate the 120 poorly annotated genes, we selected a representative group of 41 clones likely to represent real genes but could not be otherwise assigned using established annotations (summarized in Fig. 4; complete table is posted on the microarray database web server at <http://genet.cchmc.org>; for login details, see *Data archive*, in MATERIALS AND METHODS, above). Many of these could be credibly annotated using mouse-human homology and EST-to-genome mapping. For example, sequence BF232834 could first be mapped onto aldose-1-epimerase-like gene in mouse, and the novel human ortholog for this gene was then identified. Sequence BC006717 could be mapped to *CPR8* (cell cycle progression 8 protein) in human, and we could then identify the novel mouse ortholog for this gene by human-mouse homology. A search for functional domains in the proteins encoded by these genes

revealed that 23 of the 41 did not have any known functional protein domain. Of the 18 that mapped onto a known, functional domain, there were 3 (2 with high expression in OM) WD-repeat protein-encoding genes. Overall, based on public databases [LocusLink (<http://www.ncbi.nlm.nih.gov/LocusLink/>) and SwissProt], functional annotations (Gene Ontology Attributes) could be found for only 9 of 41 clones. Interestingly, many of the unknown ESTs represented on the Mouse GEM1 cDNA microarray appear to encode entirely novel proteins of potential importance for OM. In addition, there were nine mouse genes for which no human ortholog was found. A BLASTN and BLAT search of these nine mouse mRNAs against human genome also did not yield any significant hits. For instance, there was no human ortholog reported for *Expi* (extracellular proteinase inhibitor, with high expression in OM and respiratory mucosa). Likewise, the EST sequence AA060184 mapped onto the intergenic region of a putative gene (NM\_025582)

Fig. 4. Gene discovery from olfactorome-implicated ESTs. Forty-one genes expressed across brain olfactory bulb, lung, skin, OM, and respiratory mucosa. Genes within the table are organized by clone ID and expression pattern and are tabulated for BLAST-implicated hits to the mouse genome. Sequence-predicted features of the implicated proteins are also listed. A more complete version of this annotation table is downloadable from the Expression Data Server located at <http://genet.cchmc.org> (see Excel worksheet attached to the gene list file titled "41\_olfactorome\_gene-discovery"; username: olfepi; password: olfepi).

Clone Identification and Expression Pattern										Blast-implicated hits to the mouse genome.								
Incyte Accession	Genbank Accession	BrE18	BrD2	Off Bulb	Lung Ad	Lung E16.5	Skin P2	Skin E16.5	Off Epi	Resp Epi	Extra Sites Expression	Description	Mouse Chr	human ortholog	Refseq	Encoded protein	Protein Domains	
443638	AK006055											None	Unknown gene with Pro rich & DQNB depend Rect domains	9	yes	XM_134943	XP_134943	WD domain, ortholog to abnormal chemotaxis protein 2 (C. elegans)
481773	AA060184											None	3'UTR of transcript for OSC255, Neprilysin-like metalloproteinase 1	4	yes	NM_025582	NP_079858	No known domains
573324	AK011391											Thymus	WD, g-beta repeat containing protein	2	yes	XM_130321	XP_130321	Ubox - Modified RING finger domain, SAM - Sterile alpha motif & WD40 domains
737605	BL693493											None	Highly similar to KIAA0800	9	yes	None	None	No known domains
597415	AK014241											None	3' intron within ninein centrosomal protein homolog (human)	12	yes	None	None	No known domains
<b>Offactory Mucosa &amp; Respiratory Mucosa</b>																		
404169	AK017304											Islets	homolog to Potential Ligand Binding protein, similar to Rat Fy5	2	yes	XM_130612	XP_130612	Herpes_glycop_D - Herpes virus glycoprotein D
679341	BF232834											Liver, GI and GL_canc	UDP-glucose-4-epimerase (LOC209922)	17	yes	XM_140179	XP_140179	Aldose 1-epimerase
622257	BF140337											None	Similar to leashirt 2, MGC:41494	2	yes	XM_130644	XP_130644	No known domains
747543	BC005759											Liver and proximal jejunum	SECT14-like 2	11	yes	XM_126025	NP_653103	CRAL_TRIO - CRAL/TRIO domain, SECT14 & CRAL_TRIO_N - CRAL/TRIO N-terminus
832158	NM_007969											Mammary and GI_canc	extracellular proteinase inhibitor; no human ortholog	11	no	NM_007969	NP_031995	WAP domain & WAP-type (Whey Acidic Protein) 'four-disulfide core
662909	BG801546											None	transcript for unknown protein with Herpesvirus glycoprotein D domain	11	no	XM_125992	XP_125992	No known domains
532961	BC066717											None	CCPB-like (Coiled coil protein like, Human CPB (Cell cycle progression 8))	9	yes	BC066717.1	AAH06717	LEA - Late embryogenesis abundant protein. Different types of LEA proteins are expressed at different stages of late embryogenesis in higher plant seed. Domain found in Drosophila, Egr-10, and Pleckstrin. Domain of unknown function present in signalling proteins that contain PH, rasGEF, rhoGEF, rhoGAP, RGS, PDZ domains,
598411	BC026382											GI_Duo	unknown protein; DEP dishevelled domain signaling associated	2	yes	XM_130346	XP_130346	
<b>OffMucosa &amp; Brain</b>																		
735527	BC024759											None	6 exon gene similar to human Neuronal membrane GPM6A	8	yes	XM_134147	XP_134147	Myelin_PLP & PLP (proteolipid protein) domain
374564	AK002643											None	Preimplantation protein 2	11	yes	XM_109854	XP_109854	ERM - Ezrin/radixin/moesin family
333307	W15872											None	synaptotagmin I	10	yes	None	None	No known domains
335943	W18385											None	Similar to RAS related protein RAB (ENSMUSG00000032549)	9	no	XM_150236	XP_150236	No known domains
475403	NM_013865											None	Nmyc downstream regulated 3	2	yes	NM_013865	NP_038893	Ndr & ahydrolase - alpha/beta hydrolase fold
959804	NM_023328											Testes	ATP/GTP binding protein 1	13	yes	NM_023328	NP_075817	Zn_pept domain
669869	AV221896											None	Intronic region of Nptr gene	15	yes	None	None	No known domains
337572	W29432											None	WD domain containing gene, homologous to human hypothetical protein	10	yes			Wd domain
<b>Brain &amp; OlfactoryBulb</b>																		
464134	BG261601											None	3' end of ANK2 gene, sodium, calcium exchanger, IP3 receptor	3	yes	XM_149388	XP_149388	No known domains
571503	AK005148											None	Similar to P53 inducible protein in human	11	yes	NM_133769	NP_598530	No known domains
<b>OlfactoryBulb &amp; Of Mucosa</b>																		
456344	BB653922											None	unknown protein, similar to human hematopoietic zinc finger (LOC151129)	2	yes	mCP48893		Znf U1 & Zinc finger C2H2 1 domains
463342	BC017636											Islets and Lung E13.5	serologically defined colon cancer antigen 33 complete CDS	18	yes	XM_129060	XP_129060	HomeoDomain
<b>Others</b>																		
761264	AV254040											Testes	no similarity found	13	no	None	None	No known domains
634733	BI078179											None	ZYG homolog	2	yes	None	None	None
642382	BI151173											None	No significant similarity found	6	no	None	None	No known domains
777490	BC017523											None	large unknown TPR-repeat containing protein	12	yes	XM_127058	XP_127058	TPR-Repeat
735673	BM945494											GL_ColorCanc	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme	10	yes	None	None	No known domains
732290	AK009964											KidneyAdult and P14	Similar to Cgl78 (ENSMUSG00000032371)	9	yes	NM_027169	NP_081445	No known domains
644840	BE534526											None	Similar to haman KIAA0471 protein	1	no	None	None	No known domains
774890	BF581861											None	Rsect15-like; similar to yeast vesicular traffic control protein SEC15	6	yes			No known domains
464442	AK018080											MammaryLac	Similar to human SOX11 gene	12	no	None	None	No known domains
472562	BM114590											None	Similar to AK056003	15	yes	None	None	No known domains
657528	NM_025613											Heart E14.5	Similar to human CREBBP/EP300 inhibitory protein 1	2	yes	NM_025613	NP_079889	No known domains
748241	NM_025670											None	Homolog to human 6 exon gene Chr.16 ORF5 (c16orf5)	16	yes	NM_025670	NP_079946	PRICHTEXTENS, Pro-rich region & LTAF domains
808083	NM_008680											Heart, D1Bladder and epididymis	ENAH, enabled homolog (drosophila)	1	no	NM_010135.1	NP_034265.1	No known domains
620240	BC013814											Thymus and testes	Hypothetical protein with WD 40 domain	3	yes	XM_143313	XP_143313	WD40 repeats, WD domain and G-beta repeat
638244	NM_009412											multi-system, proliferation	Tumor protein D52, G-protein beta family	3	yes	NM_009412	NP_033438	DNA_topoisomIV & ERM - Ezrin/radixin/moesin family
779165	NM_133762											Spleen, LymphNode Spleen and Heart E14.5	Homologous to human hypothetical protein FLJ20311	12	yes	NM_133762	NP_598523	No known domains

and *Mell1* in mouse. However, the intergenic space in human is smaller, and no human ortholog was found.

*Common cis-elements in coordinately expressed genes.* Based on in situ hybridization analysis, several of the genes that are coordinately expressed across multiple mouse tissues appeared to be localized within the same OM cell types. We therefore sought to test the hypothesis that cell-type-specific gene expression within the OM could be due to shared *cis*-acting regulatory elements. To do this, we selected three different gene group clusters reflecting expression in 1) OM only; 2) OM plus respiratory mucosa; and 3) whole brain, olfactory bulb, and OM. The complete genomic sequences of the selected mouse genes and their corresponding human orthologs were extracted from the Celera/ENSEMBL/UCSC mouse and human databases, respectively. We next identified clusters of regulatory elements conserved between each pair of mouse and human orthologs within their conserved noncoding region using the TraFaC server (<http://trafac.chmcc.org>; Ref. 40). Using an extension of this system, we next identified *cis*-element modules consisting of compositionally similar *cis*-element clusters that occurred between three different coordinately regulated ortholog pairs.

Analysis of the upstream 3-kb region for the gene group showing high levels of expression in OM revealed *cis*-elements CLOX, OCT1, BRNF, FKHD, CREB, SORY, MEIS, and ETSF, occurring in clusters of four elements each with in a maximum span of 200 bp. Likewise, common *cis*-elements were found among each group of genes examined: OM-specific and respiratory mucosa-specific genes showed *cis*-elements GATA, FKHD, CREB, ETSF, GFI1, SMAD, and MEF2. Genes of the group OM and brain shared *cis*-elements CREB, NKXH, CLOX, MEF2, BRNF, HNF1, FKHD, LEFF, MYT1, SORY, and MEIS (Fig. 5).

Since the potential regulatory switches that direct tissue-specific expression are not always limited to promoter region alone, but are also known to occur in the intronic, downstream, and upstream regions, we extended the query to identify compositionally similar *cis*-regulatory element clusters across these regions. The window size ranged from 200 to 300 bp. Two of the genes (*Tpd52* and *1700017B05Rik*) with high expression in OM shared *cis*-elements NOLF, ETSF, and HNF1, whereas the genes with high expression in both OM and respiratory mucosa shared *cis*-elements HNF4, NOLF, and MYT1. MYT1, a *Xenopus* C2HC-type zinc finger protein, has been associated with a regulatory function in neuronal differentiation (7). NOLF ("neuron-specific olfactory factor") has been previously reported in olfactory-neuron-specific genes (85).

## DISCUSSION

The results presented in this study have identified a relatively large number of genes strongly expressed in the OM, and we have classified them by function and by their relative expression patterning across a large number of other tissues. However, it should be noted that there are several limitations at the present stage of analyses. First, as for any microarray study, the findings are necessarily a reflection of the content of the array, as well as the precise samples that have been profiled.

The Incyte Mouse GEM1 array is relatively old and incomplete. For example, candidate odorant receptor genes, which are undoubtedly among the highly expressed OM genes and comprise ~1% of the mammalian genome (46), are not present on the GEM1 microarray. This result illustrates the limitation of using less-than-whole genome representation on a particular microarray platform. A recent publication details changes in gene expression in the senescence-accelerated mouse, using Affymetrix U74Av2 GeneChip oligonucleotide arrays (30). These investigators noted significant expression and aging-related changes in a number of chemosensory receptor genes. However, many genes on the GEM1 microarray are not present on the Affymetrix U74Av2 array. Of more interest, many genes and gene groups that are differentially regulated in senescence-accelerated mice correspond well to the genes that we found to be highly expressed in the mucosa; e.g., a large number of immune system-related genes and metabolic enzyme genes. Thus the results of our study serve more as model of the power that can be obtained in the definition of tissue-specific genomic anatomy if a very large number of different biological samples of diverse differentiation could be carefully intercompared. In addition, our assay of mucosal gene expression is relatively coarse-grained because of the use of tissue samples with mixed cell types. Very interesting results could also be obtained from isolated individual cell types of the mucosa as a number of technical hurdles are progressively overcome.

Gene expression patterns in the OM strongly reflect its unique anatomy and physiology. The cellular anatomy of the OM includes basal cells, supporting (sustentacular) cells, both immature and mature olfactory neurons, plus subepithelial Bowman's glands and nerve bundles. Olfactory neurons interact directly with the external environment by means of their surface dendrites. Therefore, olfactory neurons are in contact with, and potentially susceptible to damage by, environmental agents. This provides a strong rationale for the high-level coexpression of a series of xenobiotic metabolizing enzyme genes in OM that are also strongly expressed in liver and/or lung. Since the OM continuously generates new olfactory neurons from neuronal progenitor cells, accounting for its regenerative capacity following traumatic or toxicant-induced damage, we hypothesized that a significant component of OM gene expression might overlap with genes that participate in neurodevelopment. Consistent with this hypothesis, we observed a significant overlap between the OM and the embryonic and early postnatal brains with respect to genes involved in neurogenesis, neuronal differentiation, and migration. Finally, the OM contains a family of odorant receptors, as well as the signal transduction machinery necessary to conduct odorant-related information to the rest of the nervous system.

A major goal of the present study was to identify genes involved in neurogenesis and neuronal development, using the OM as a model system, given that the OM supports sustained generation of mature neurons from neuronal progenitor cells. A number of genes that had not previously been localized to the OM were identified as a result of these efforts, although their roles, if any, in neurogenesis are still to be determined. Among these are *Sep3* (NM\_011889; Refs. 49 and 83), *Mad4*, and *Pdgfrec* ("platelet-derived endothelial growth factor/thymidine phosphorylase/gliostatin"; NM\_138302). *Pdgfrec* had previ-

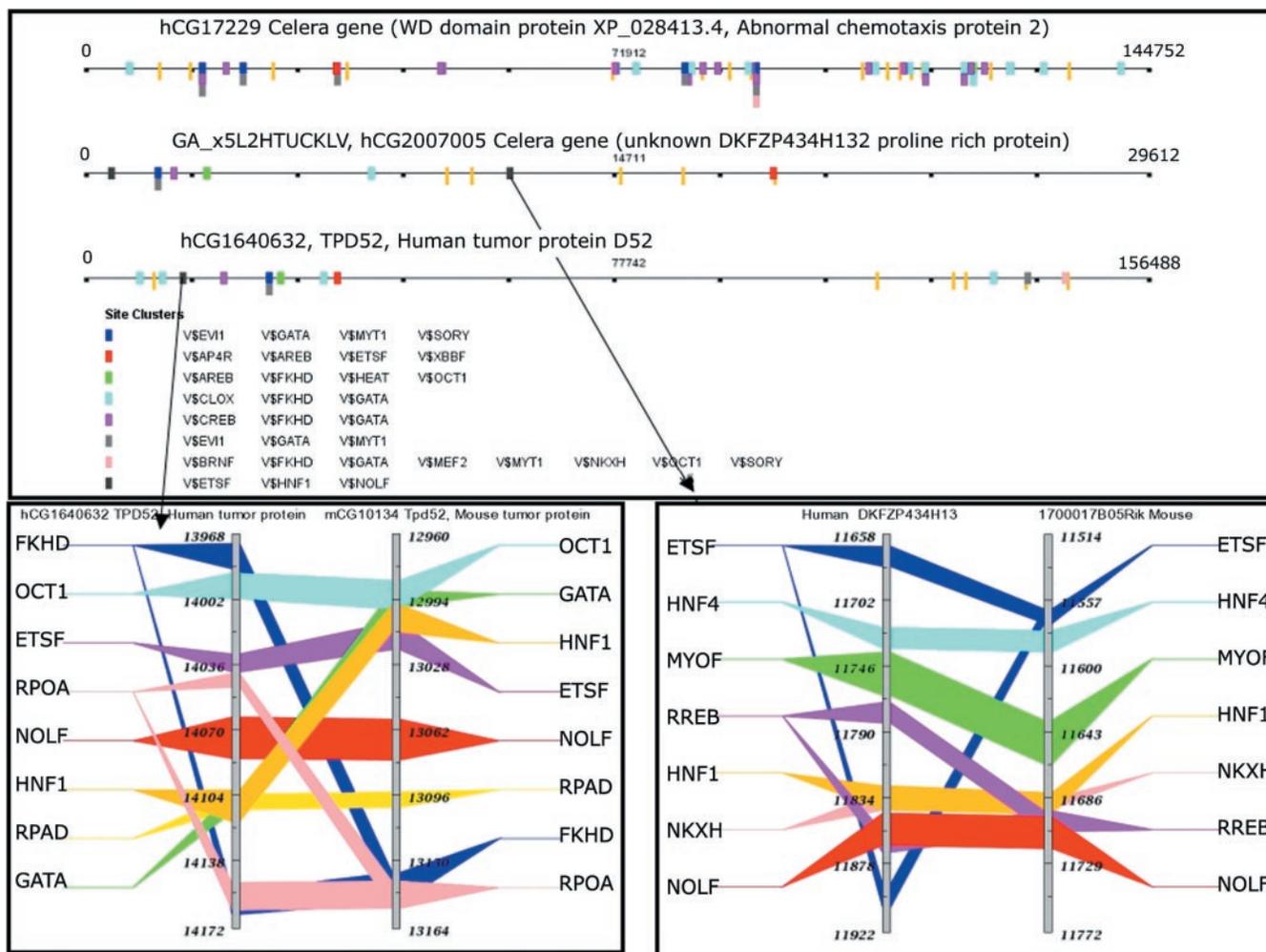


Fig. 5. Shared *cis*-clusters in OM-specific genes. Compositionally similar *cis*-regulatory element clusters that occur in groups of coregulated genes within each of their ortholog-pair evolutionarily conserved *cis*-regulatory regions were identified by comparative genomics-based *cis*-element analyses (<http://trafac.cchmc.org>). The composition of each cluster (a set of two or more *cis*-elements occurring in conserved sequence region between human and mouse) is displayed at *bottom*. The *cis*-clusters are represented as different color squares along the genomic sequence. The vertical yellow bars are the exons of the human gene. The numbers indicate nucleotide coordinates of the human genomic sequence. Two of the genes (XM\_134943 and XM\_143313) encode hypothetical proteins, while the third is *TPD52* (“tumor protein D52”). The mouse clone ID 443638 (XM\_134943) is orthologous to human hCG17229 Celera gene (unknown WD domain protein XP\_028413.4), a conserved ortholog to “Abnormal chemotaxis protein 2” [*Caenorhabditis elegans*] AAK68383.1. The Riken mouse gene 1700017B05Rik is the ortholog of the hCG2007005 Celera gene, which corresponds to unknown human proline-rich protein *DKFZP434H132*. TraFaC images display the composition of the *cis*-clusters conserved in human and mouse. The two gray vertical bars are the two genes that are compared. The TF-binding sites occurring in both the genes are highlighted as various colored bars drawn across the two genes. The numbers indicate the nucleotide coordinates of the genomic sequences. Two of the genes (*Tpd52*, first intron; and 1700017B05Rik, second intron) with high expression in OM shared *cis*-elements NOLF, ETSF, and HNF1.

ously been localized to the peripheral nervous system (20), identified as a protein that has a potential role in development and regeneration of the central nervous system (4), and determined to be highly expressed in invasive colorectal tumors (44); therefore, this gene/protein may regulate the switch between proliferation and differentiation in the process of neurogenesis in the OM. *Mad4* is highly induced in differentiated cells (39, 43), so we would expect to find this protein localized to one or more nonproliferating cell populations in the OM. In contrast to a previous report, in which brain fatty acid binding protein 7 (*Fabp7*; NM\_021272) was found to be absent from OM (but present in olfactory bulb; Ref. 45), we

found expression in OM to be well above our cutoff criterion, with an expression level of approximately one-half that of olfactory bulb and one-tenth that of whole embryonic and postnatal *day 2* brain. This protein has been characterized as a potential brain morphogen during development (45).

The olfactory system requires a high degree of synaptic plasticity, as even in the absence of toxicant exposure mature neurons die and new olfactory neurons are continuously generated. As a result, synapses between olfactory sensory cell axons and mitral cells in the olfactory bulb are continuously lost and reformed. Consequently, there is the constant need for axonal guidance to direct the axons of newly generated neurons

to the olfactory bulb. Cell adhesion molecules are important in this process, and we observed high expression of NCAM in OM, olfactory bulb, cerebellum, and postnatal *day 2* brain. *Dpysl3* (NM\_009468), another highly expressed OM gene, is orthologous to the *C. elegans* axonal guidance phosphoprotein *unc-33* (12). Similarly, *Plxna3*, together with neuropilins, acts as a receptor for class 1 semaphorins; these aid in axonal guidance, likely through growth avoidance cues (12, 58, 71).

Three different amyloid precursor protein (APP)-related transcripts were highly expressed in OM. While multiple functions for amyloid-related proteins have been proposed in nervous system homeostasis and pathology, of particular interest, APP is highly expressed in the developing brain, and specifically in the olfactory system (15, 38, 48, 78). *Aplp2* has previously been reported to be abundant in olfactory sensory axons and in postsynaptic connections in the olfactory bulb and is proposed to play a role in axonal pathfinding, axogenesis, and/or synaptogenesis (15, 76). APP-like proteins were also highly expressed in the regenerating OM following administration of the reversible OM toxicant diethylthiocarbamate (70). However, amyloid  $\beta$ -peptide was found to inhibit neurogenesis in the subventricular zone of mice and in human cortical neuronal precursor cells *in vitro* (36). Furthermore, OM and olfactory bulb contain high levels of  $\beta$ -amyloid and amyloid precursor proteins in patients with Alzheimer's disease, Parkinson's disease, and Down syndrome (18, 84). Therefore,  $\beta$ -amyloid-related proteins are clearly pleiotropic, given their expression during brain development, in multiple regions of the adult brain, and in degenerative nervous system conditions.

Many of the highly expressed OM genes are associated with signal transduction. Examples of these genes include those related to  $\text{Ca}^{2+}$  ion channels and  $\text{Ca}^{2+}$  homeostasis; G-protein-coupled receptor-related genes; and homer, the neuronal early immediate gene associated with glutamate signaling (77). Following interaction of odorants with cell surface receptors, two signal transduction pathways are activated. One class of odorants activate G-protein-mediated formation of adenosine 3',5'-cyclic monophosphate (cAMP), causing opening of cAMP-gated channels and influx of calcium. The increase in intracellular  $\text{Ca}^{2+}$ , in turn, causes opening of  $\text{Ca}^{2+}$ -activated chloride channels, further increasing membrane depolarization (41). Other odorants stimulate G-protein-mediated formation of inositol-1,4,5-triphosphate ( $\text{IP}_3$ ). An increase in intracellular  $\text{IP}_3$  causes the opening of a  $\text{Ca}^{2+}$  conductance and a nonspecific cation conductance and may also activate  $\text{Ca}^{2+}$ -activated  $\text{K}^+$  channels (41). The existence of an  $\text{IP}_3$ -mediated pathway in mammals has been questioned, whereas it has been well accepted in nonmammalian species (e.g., Refs. 11, 22, 32). The rich representation of genes within these categories provides strong support for the role of their corresponding pathways in OM physiology.

The neurotransmitter(s) used by olfactory sensory neurons has also been a source of controversy over the years. Initially, the tripeptide carnosine was regarded as the primary neurotransmitter released following activation of odorant receptors (64). It now appears to be more widely accepted that glutamate is the excitatory olfactory neurotransmitter (8, 24, 34). L-Glutamate acts at both ligand-gated (ionotropic) and G-protein-coupled (metabotropic) glutamate receptors (5). *Homer2* en-

codes a protein that functions directly at the synapse, linking group 1 metabotropic glutamate receptors with  $\text{IP}_3$  receptors (77).

Among the genes most highly coexpressed in OM, nasal respiratory mucosa, and lung are those encoding proteins with epithelial defense and immune function, including multiple immunoglobulin-related genes, *Ltf*, and a cathelin-like protein (a member of a family of antimicrobial peptides; Ref. 25). *Ltf* is an iron-binding multifunctional protein, produced primarily by neutrophils, and is believed to be important in defense against bacteria, fungi, and viruses (3). Plunc proteins have recently been found to be associated with respiratory tract irritation and may also function in mucosal defenses (9, 31).

Genes encoding xenobiotic-metabolizing enzymes were highly expressed in OM, including multiple cytochromes *P-450* (CYP), sulfotransferases, and paraoxonase 1. These observations corroborate previous reports documenting sulfotransferases in OM, with some specifically expressed in OM (72, 73). Although the presence of *Cyp2f2* in OM had been inferred based on enzymatic activity (69), we demonstrate herein that *Cyp2f2* is localized by *in situ* hybridization and immunohistochemistry to sustentacular cells and Bowman's glands. *Cyp2f2* has been demonstrated to bioactivate naphthalene in the lung; its presence in OM is not surprising, given that naphthalene induces lung and OM tumors in rodents (54, 55, 69). *Pon1*, which detoxifies organophosphorous pesticides and nerve gases (2), was similarly highly expressed in the Bowman's glands, a distribution shared among many metabolic enzymes (10, 28). The  $\alpha/\beta$ -hydrolase catalytic domain is found in a wide range of enzymes, including esterases and epoxide hydrolases (37). The esterase 1 and 22 transcripts were highly expressed in OM in the present study, and both esterases and microsomal epoxide hydrolase have previously been localized to OM (10, 28).

Our results strongly demonstrate the power of large-scale microarray profiling approaches for compartment-specific gene discovery. One gene exhibiting a particularly interesting expression profile encodes sphingosine phosphate lyase (*Sgpl1*). *Sgpl1* expression was high in thymus, multiple regions of the gastrointestinal tract, skin, and OM. Direct *Sgpl1* activity measurements confirmed the high expression in OM, relative to other tissues (Table 2). Although the aforementioned tissues are sites of relatively high cell proliferation, *Sgpl1* expression was not exclusively associated with proliferative states, as *Sgpl1* expression was higher in unmanipulated mouse liver than in regenerating liver in our database (confirmed by unpublished enzymatic activity measurements from the laboratory of P. P. Van Veldhoven). The role of *Sgpl1* in OM is currently unclear, but it may play a role in the kinetics of neuronal proliferation and apoptosis; a similar role for *Sgpl1* has been proposed in skin (26). A function of *Sgpl1* is the degradation of sphingosine-1-phosphate (S-1-P), which is not only a sphingolipid catabolite, but also a potent second messenger with roles in proliferation and differentiation in various model systems (56, 62, 82). In many cell types, S-1-P is associated with inhibition of apoptosis (21); the high *Sgpl1* activity we have observed in OM may be important in modulating cell kinetics in OM, perhaps regulating apoptosis in the oldest population of olfactory neurons that are targeted for destruction. Our studies show that Fas apoptotic inhibitory

molecule (*Faim*; NM\_011810) and *Sgpl1* share a common pattern of high expression in OM, suggesting that mitochondria-dependent Fas-induced apoptosis may be a major signaling pathway in olfactory neuronal homeostasis (19).

The recent completion and large-scale annotation of the mouse genome has greatly improved our ability to characterize EST gene sequences. We originally selected an EST (W11846) for localization within OM because of its relatively high expression and because it appeared to segregate with *Sgpl1* by hierarchical tree analysis. This gene has recently been associated with the *Alms1* locus (2p13); mutations of this locus cause Alstrom syndrome, with a wide range of symptoms, including neurosensory degeneration (16, 17). The sodium bicarbonate transporter NBC4 also maps to chromosome 2p13 in humans and is regarded as a new candidate gene for Alstrom syndrome (61). The function of this protein in OM is currently unknown.

The importance of independent verification of genes that appear to be highly expressed in microarray experiments cannot be overemphasized. We chose to do so by *in situ* hybridization, enzymatic activity measurements, and immunohistochemistry. In our experiments, crp ductin (*Dmbt1*) appeared to be among the highest expressed genes in OM. However, *in situ* hybridization studies revealed negligible *Dmbt1* expression in OM but very high expression in the nearby lateral nasal gland (Fig. 2). Furthermore, the rat homolog of crp ductin, ebnerin, has recently been immunolocalized to nasal respiratory mucosa and alachlor-induced OM tumors but was not detected in control OM (29). Therefore, we suspect that the speed required for isolation of tissue for preparation of high-quality RNA resulted in accidental collection of the lateral nasal gland from one or more of the mice. This tissue localization issue would not have been resolved by use of reverse transcription, followed by polymerase chain reaction, as the same RNA would have been used for the confirmatory studies. Although the function of crp ductin in the lateral nasal gland remains unknown, we do now know that we should eliminate this gene/protein from those candidates for serving a role in neurogenesis.

Finally, we have also sought to exploit the identification of OM-specific genes to aid in the identification of potential regulatory regions responsible for cell- and region-specific gene expression. Switches that direct tissue-specific expression are not always limited to promoter region alone, but are also known to occur in the intronic, downstream, and upstream regions also. Hence, we extended the query to identify compositionally similar *cis*-regulatory element clusters within each of their ortholog-pair evolutionarily conserved *cis*-regulatory regions. The window size ranged from 200 to 300 bp. Genes with high expression in OM shared *cis*-elements NOLF, ETSF, and HNF1, while the genes with high expression in both olfactory and respiratory mucosa shared *cis*-elements HNF4, NOLF, and MYT1. MYT1, a *Xenopus* C2HC-type zinc finger protein, has been associated with a regulatory function in neuronal differentiation (7). NOLF ("neuron-specific olfactory factor") has been reported in olfactory-neuron-specific genes (85). These results support the hypothesis that the approach of tissue-specific expression mining coupled to detailed genomic analyses will be very helpful in the identification of compartment-specific gene regulatory regions.

## ACKNOWLEDGMENTS

We thank the many contributors to the UC-CHMCC Mouse Tissue Specific Gene Expression Database.

## GRANTS

This work was supported by the Howard Hughes Medical Institute; the Children's Hospital Research Foundation; the University of Cincinnati Center for Environmental Genetics (ES-06096); and by National Institutes of Health Grants U01-ES-11038 (to B. J. Aronow) and R01-ES-08799 (to M. B. Genter) and by Grant G.0405.02 from the Fonds Wetenschappelijk Onderzoek-Vlaanderen (to P. P. Van Veldhoven).

## REFERENCES

1. **Abou-Zeid C, Filley E, Steele J, and Rook GA.** A simple new method for using antigens separated by polyacrylamide gel electrophoresis to stimulate lymphocytes *in vitro* after converting bands cut from Western blots into antigen-bearing particles. *J Immunol Methods* 98: 5–10, 1987.
2. **Akgur SA, Ozturk P, Sozmen EY, Delen Y, Tanyalcin T, and Ege B.** Paraoxonase and acetylcholinesterase activities in humans exposed to organophosphorous compounds. *J Toxicol Environ Health A* 58: 469–474, 1999.
3. **Andersen JH, Osbakk SA, Vorland LH, Traavik T, and Gutteberg TJ.** Lactoferrin and cyclic lactoferricin inhibit the entry of human cytomegalovirus into human fibroblasts. *Antiviral Res* 51: 141–149, 2001.
4. **Asai K, Nakanishi K, Isobe I, Eksioglu YZ, Hirano A, Hama K, Miyamoto T, and Kato T.** Neurotrophic action of gliostatin on cortical neurons. Identity of gliostatin and platelet-derived endothelial cell growth factor. *J Biol Chem* 267: 20311–20316, 1992.
5. **Bailey A, Kelland EE, Thomas A, Biggs J, Crawford D, Kitchen I, and Toms NJ.** Regional mapping of low-affinity kainate receptors in mouse brain using [<sup>3</sup>H](2S,4R)-4-methylglutamate autoradiography. *Eur J Pharmacol* 431: 305–310, 2001.
6. **Bates MD, Erwin CR, Sanford LP, Wiginton D, Bezerra JA, Schatzman LC, Jegga AG, Ley-Ebert C, Williams SS, Steinbrecher KA, Warner BW, Cohen MB, and Aronow BJ.** Novel genes and functional relationships in the adult mouse gastrointestinal tract identified by microarray analysis. *Gastroenterology* 122: 1467–1482, 2002.
7. **Bellefroid EJ, Bourguignon C, Hollemann T, Ma Q, Anderson DJ, Kintner C, and Pieler T.** X-MyT1, a *Xenopus* C2HC-type zinc finger protein with a regulatory function in neuronal differentiation. *Cell* 87: 1191–1202, 1996.
8. **Berkowitz DA, Trombley PQ, and Shepherd GM.** Evidence for glutamate as the olfactory receptor cell neurotransmitter. *J Neurophysiol* 71: 2557–2561, 1994.
9. **Bingle CD and Craven CJ.** PLUNC: a novel family of candidate host defence proteins expressed in the upper airways and nasopharynx. *Hum Mol Genet* 11: 937–943, 2002.
10. **Bogdanffy MS.** Biotransformation enzymes in the rodent nasal mucosa: the value of a histochemical approach. *Environ Health Perspect* 85: 177–186, 1990.
11. **Brunet LJ, Gold GH, and Ngai J.** General anosmia caused by a targeted disruption of the mouse olfactory cyclic nucleotide-gated cation channel. *Neuron* 17: 681–693, 1996.
12. **Byk T, Dobransky T, Cifuentes-Diaz C, and Sobel A.** Identification and molecular characterization of Unc-33-like phosphoprotein (Ulip), a putative mammalian homolog of the axonal guidance-associated unc-33 gene product. *J Neurosci* 16: 688–701, 1996.
13. **Camara CG and Harding JW.** Thymidine incorporation in the olfactory epithelium of mice: early exponential response induced by olfactory neuroectomy. *Brain Res* 308: 63–68, 1984.
14. **Cancelon P.** Degeneration and regeneration of olfactory cells induced by ZnSO<sub>4</sub> and other chemicals. *Tissue Cell* 14: 717–733, 1982.
15. **Clarris HJ, Key B, Beyreuther K, Masters CL, and Small DH.** Expression of the amyloid protein precursor of Alzheimer's disease in the developing rat olfactory system. *Brain Res Dev Brain Res* 88: 87–95, 1995.
16. **Collin GB, Marshall JD, Ikeda A, So WV, Russell-Eggitt I, Maffei P, Beck S, Boerkoel CF, Siculo N, Martin M, Nishina PM, and Naggert JK.** Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alstrom syndrome. *Nat Genet* 31: 74–78, 2002.
17. **Collin GB, Marshall JD, Boerkoel CF, Levin AV, Weksberg R, Greenberg J, Michaud JL, Naggert JK, and Nishina PM.** Alstrom

- syndrome: further evidence for linkage to human chromosome 2p13. *Hum Genet* 105: 474–479, 1999.
18. **Crino PB, Martin JA, Hill WD, Greenberg B, Lee VM, and Trojanowski JQ.** Beta-amyloid peptide and amyloid precursor proteins in olfactory mucosa of patients with Alzheimer's disease, Parkinson's disease, and Down syndrome. *Ann Otol Rhinol Laryngol* 104: 655–661, 1995.
  19. **Cuvillier O, Edsall L, and Spiegel S.** Involvement of sphingosine in mitochondria-dependent Fas-induced apoptosis of type II Jurkat T cells. *J Biol Chem* 275: 15691–15700, 2000.
  20. **Eccleston PA, Funa K, Heldin CH.** Neurons of the peripheral nervous system express thymidine phosphorylase. *Neurosci Lett* 192: 137–141, 1995.
  21. **Edsall LC, Cuvillier O, Twitty S, Spiegel S, and Milstien S.** Sphingosine kinase expression regulates apoptosis and caspase activation in PC12 cells. *J Neurochem* 76: 1573–1584, 2001.
  22. **Fadool DA and Ache BW.** Plasma membrane inositol 1,4,5-phosphate-activated channels mediate signal transduction in lobster olfactory receptor neurons. *Neuron* 9: 907–918, 1992.
  23. **Feng L, Hatten ME, and Heintz N.** Brain lipid-binding protein (BLBP): a novel signaling system in the developing mammalian CNS. *Neuron* 12: 895–908, 1994.
  24. **Ferraris N, Perroteau I, De Marchis S, Fasolo A, and Bovolín P.** Glutamatergic deafferentation of olfactory bulb modulates the expression of mGluR1a mRNA. *Neuroreport* 8: 1949–1953, 1997.
  25. **Gallo RL, Kim KJ, Bernfield M, Kozak CA, Zanetti M, Merluzzi L, and Gennaro R.** Identification of CRAMP, a cathelin-related antimicrobial peptide expressed in the embryonic and adult mouse. *J Biol Chem* 16: 13088–13093, 1997.
  26. **Geilen CC, Barz S, and Bektas M.** Sphingolipid signaling in epidermal homeostasis. Current knowledge and new therapeutic approaches in dermatology. *Skin Pharmacol Appl Skin Physiol* 14: 261–71, 2001.
  27. **Genter MB, Llorens J, O'Callaghan JP, Peele DB, Morgan KT, and Crofton K.** Olfactory toxicity of  $\beta,\beta'$ -iminodipropionitrile (IDPN) in the rat. *J Pharmacol Exp Ther* 263: 1432–1439, 1992.
  28. **Genter MB, Owens DM, and Deamer NJ.** Distribution of microsomal epoxide hydrolase and glutathione S-transferase in the rat olfactory mucosa: relevance to distribution of lesions caused by systemically-administered olfactory toxicants. *Chem Senses* 20: 385–392, 1995.
  29. **Genter MB, Burman DM, Vijayakumar S, Ebert CL, and Aronow BJ.** Genomic analysis of alachlor-induced oncogenesis in rat olfactory mucosa. *Physiol Genomics* 12: 35–45, 2002. First published November 5, 2002; 10.1152/physiolgenomics.00120.2002.
  30. **Getchell TV, Peng X, Stromberg AJ, Chen KC, Paul Green C, Subhedar NK, Shah DS, Mattson MP, and Getchell ML.** Age-related trends in gene expression in the chemosensory-nasal mucosae of senescence-accelerated mice. *Ageing Res Rev* 2: 211–243, 2003.
  31. **Ghafouri B, Stahlbom B, Tagesson C, and Lindahl M.** Newly identified proteins in human nasal lavage fluid from non-smokers and smokers using two-dimensional gel electrophoresis and peptide mass fingerprinting. *Proteomics* 2: 112–120, 2002.
  32. **Gold G.** Controversial issues in vertebrate olfactory transduction. *Annu Rev Physiol* 61: 857–871, 1999.
  33. **Graziadei PP and Graziadei GA.** Neurogenesis and neuron regeneration in the olfactory system of mammals. I. Morphological aspects of differentiation and structural organization of the olfactory sensory neurons. *J Neurocytol* 8: 1–18, 1979.
  34. **Griff ER, Greer CA, Margolis F, Ennis M, and Shipley MT.** Ultrastructural characteristics and conduction velocity of olfactory receptor neuron axons in the olfactory marker protein-null mouse. *Brain Res* 866: 227–236, 2000.
  35. **Haass NK, Kartenbeck AM, and Leube RE.** Pantophysin is a ubiquitously expressed synaptophysin homologue and defines constitutive transport vesicles. *J Cell Biol* 134: 731–746, 1996.
  36. **Haughey NJ, Liu D, Nath A, Borchard AC, and Mattson MP.** Disruption of neurogenesis in the subventricular zone of adult mice, and in human cortical neuronal precursor cells in culture, by amyloid beta-peptide: implications for the pathogenesis of Alzheimer's disease. *Neuro-molecular Med* 1: 125–135, 2002.
  37. **Heikinheimo P, Goldman A, Jeffries C, and Ollis DL.** Of barn owls and bankers: a lush variety of  $\alpha/\beta$  hydrolases. *Structure* 7:R141–146, 1999.
  38. **Hsu LJ, Mallory M, Xia Y, Veinbergsm I, Hashimoto M, Yoshimoto M, Thal LJ, Saitoh T, and Masliah E.** Expression pattern of synucleins (non-A $\beta$  component of Alzheimer's disease amyloid precursor protein/ $\alpha$ -synuclein) during murine brain development. *J Neurochem* 71: 338–344, 1998.
  39. **Hurlin PJ, Queva C, Koskinen PJ, Steingrimsson E, Ayer DE, Copeland NG, Jenkins NA, and Eisenman RN.** Mad3 and Mad4: novel Max-interacting transcriptional repressors that suppress c-myc dependent transformation and are expressed during neural and epidermal differentiation. *EMBO J* 15: 2030, 1996.
  40. **Jegga AG, Sherwood SP, Carman JW, Pinski AT, Phillips JL, Pestian JP, and Aronow BJ.** Detection and visualization of compositionally similar cis-regulatory element clusters in orthologous and coordinately controlled genes. *Genome Res* 12: 1408–1417, 2002.
  41. **Kaur R, Zhu XO, Moorhouse AJ, and Barry PH.** IP3-gated channels and their occurrence relative to CNG channels in the soma and dendritic knob of rat olfactory receptor neurons. *J Membr Biol* 181: 91–105, 2001.
  42. **Kelley-Loughnane N, Sabla GE, Ley-Ebert C, Aronow BJ, and Bezerra JA.** Independent and overlapping transcriptional activation during liver development and regeneration in mice. *Hepatology* 35: 525–534, 2002.
  43. **Kime L and Wright SC.** Mad4 is regulated by a transcriptional repressor complex that contains Miz-1 and c-Myc. *Biochem J* 370:291–298, 2003.
  44. **Kitabatake T, Kojima K, Fukasawa M, Beppu T, and Futagawa S.** Correlation of thymidine phosphorylase staining and the Ki-67 labeling index to clinicopathologic factors and hepatic metastasis in patients with colorectal cancer. *Surg Today* 32: 322–328, 2002.
  45. **Kurtz A, Zimmer A, Schnutgen F, Bruning G, Spener F, and Muller T.** The expression pattern of a novel gene encoding brain-fatty acid binding protein correlates with neuronal and glial cell development. *Development* 120: 2637–2649, 1994.
  46. **Lane RP, Cutforth T, Young J, Athanasiou M, Friedman C, Rowen L, Evans G, Axel R, Hood L, and Trask BJ.** Genomic analysis of orthologous mouse and human olfactory receptor loci. *Proc Natl Acad Sci USA* 98: 7390–7395, 2001.
  47. **Leenders AGM, Lopes da Silva FH, Ghijsen WWEJM, and Vehrage M.** Rab3A is involved in transport of synaptic vesicles to the active zone in mouse brain nerve terminals. *Mol Biol Cell* 12: 3095–3102, 2001.
  48. **Loffler J and Huber G.** Beta-amyloid precursor protein isoforms in various rat brain regions and during brain development. *J Neurochem* 59: 1316–1324, 1992.
  49. **Methner A, Leyboldt F, Joost P, and Lewerenz J.** Human septin 3 on chromosome 22q13.2 is upregulated by neuronal differentiation. *Biochem Biophys Res Commun* 283: 48–56, 2001.
  50. **Moreno FJ and Avila J.** Phosphorylation of stathmin modulates its function as a microtubule depolymerizing factor. *Mol Cell Biochem* 183: 201–209, 1998.
  51. **Mori M and Morii H.** SCG10-related neuronal growth-associated proteins in neural development, plasticity, degeneration, and aging. *J Neurosci Res* 70: 264–273, 2002.
  52. **Murakami Y, Suto F, Shimizu M, Shinoda T, Kameyama T, and Fujisawa H.** Differential expression of plexin-A subfamily members in the mouse nervous system. *Dev Dyn* 220: 246–258, 2001.
  53. **Naruse S, Thinakaran G, Luo JJ, Kusiak JW, Tomita T, Iwatsubo T, Qian X, Ginty DD, Price DL, Borchelt DR, Wong PC, and Sisodia SS.** Effects of PS1 deficiency on membrane protein trafficking in neurons. *Neuron* 21: 1213–1221, 1998.
  54. **National Toxicology Program.** Toxicology and carcinogenesis studies of naphthalene (CAS No. 91-20-3) in B6C3F1 mice (inhalation studies). TR-410 <http://ntp-server.niehs.nih.gov/hdocs/LT-studies/tr410.html>, 1992.
  55. **National Toxicology Program.** Toxicology and carcinogenesis studies of naphthalene (CAS No. 91-20-3) in F344/N rats (inhalation studies). TR-500 <http://ntp-server.niehs.nih.gov/hdocs/LT-studies/tr500.html>, 2000.
  56. **Olivera A and Spiegel S.** Sphingosine-1-phosphate as second messenger in cell proliferation induced by PDGF and FCS mitogens. *Nature* 365: 557–559, 1993.
  57. **Ozon S, El Mestikaway S, and Sobel A.** Differential, regional, and cellular expression of the stathmin family transcripts in the adult rat brain. *J Neurosci Res* 56: 553–564, 1999.
  58. **Pasterkamp RJ, Ruitenber MJ, and Verhaagen J.** Semaphorins and their receptors in olfactory axon guidance. *Cell Mol Biol (Noisy-le-grand)* 45: 763–779, 1999.
  59. **Pavlidis P and Noble WS.** Analysis of strain and regional variation in gene expression in mouse brain. *Genome Biol* 2: research0042, 2001. (<http://genomebiology.com/2001/2/10/research/0042>)

60. **Pellier-Monnin V, Astic L, Bichet S, Riederer BM, and Grenningloh G.** Expression of SCG10 and stathmin proteins in the rat olfactory system during development and axonal regeneration. *J Comp Neurol* 433: 239–254, 2001.
61. **Pushkin A, Abuladze N, Newman D, Lee I, Xu G, and Kurtz I.** Cloning, characterization and chromosomal assignment of NBC4, a new member of the sodium bicarbonate cotransporter family. *Biochim Biophys Acta* 1493: 215–218, 2000.
62. **Pyne S and Pyne NJ.** Sphingosine 1-phosphate signaling in mammalian cells. *Biochem J* 349: 385–402, 2000.
63. **Ritter JK, Owens IS, Negishi M, Nagata K, Sheen YY, Gillette JR, and Sasame HA.** Mouse pulmonary cytochrome P-450 naphthalene hydroxylase: cDNA cloning, sequence, and expression in *Saccharomyces cerevisiae*. *Biochemistry* 30: 11430–11437, 1991.
64. **Rochel S and Margolis FL.** Carnosine release from olfactory bulb synaptosomes is calcium-dependent and depolarization-stimulated. *J Neurochem* 38: 1505–1514, 1982.
65. **Saba JD, Nara F, Bielawska A, Garrett S, and Hannun YA.** The BST1 gene of *Saccharomyces cerevisiae* is the sphingosine-1-phosphate lyase. *J Biol Chem* 272: 26087–26090, 1997.
66. **Schuster-Gossler K, Bilinski P, Sado T, Ferguson-Smith A, and Gossler A.** The mouse Gtl2 gene is differentially expressed during embryonic development, encodes multiple alternatively spliced transcripts, and may act as an RNA. *Dev Dyn* 212: 214–228, 1998.
67. **Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, and Miller W.** PipMaker: a web server for aligning two genomic DNA sequences. *Genome Res* 10: 577–586, 2000.
68. **Shimizu M, Murakami Y, Suto F, and Fujisawa H.** Determination of cell adhesion sites of neurophilin-1. *J Cell Biol* 148: 1283–1293, 2001.
69. **Shultz MA, Choudary PV, and Buckpitt AR.** Role of murine cytochrome P-450 2F2 in metabolic activation of naphthalene and metabolism of other xenobiotics. *J Pharmacol Exp Ther* 290: 281–288, 1999.
70. **Struble RG, Dhanraj DN, Mei Y, Wilson M, Wang R, and Ramkumar V.** Beta-amyloid precursor protein-like immunoreactivity is upregulated during olfactory nerve regeneration in adult rats. *Brain Res* 780: 129–137, 1998.
71. **Tamagnone L, Artigiani S, Chen H, He Z, Ming GI, Song H, Chedotal A, Winberg ML, Goodman CS, Poo M, Tessier-Lavigne M, and Comoglio PM.** Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates. *Cell* 99: 71–80, 1999. (Erratum in *Cell* 104: following 320, 2001.)
72. **Tamura HO, Harada Y, Miyawaki A, Mikoshiba K, and Matsui M.** Molecular cloning and expression of a cDNA encoding an olfactory-specific mouse phenol sulphotransferase. *Biochem J* 331: 953–958, 1998.
73. **Tamura H, Miyawaki A, Inoh N, Harada Y, Mikoshiba K, and Matsui M.** High sulfotransferase activity for phenolic aromatic odorants present in the mouse olfactory organ. *Chem Biol Interact* 104: 1–9, 1997.
74. **Tanaka M, Maeda N, Noda M, and Marunouchi T.** A chondroitin sulfate proteoglycan PTP $\zeta$ /RPTP $\beta$  regulates the morphogenesis of Purkinje cell dendrites in the developing cerebellum. *J Neurosci* 23: 2804–2814, 2003.
75. **Thakker-Varia S, Alder J, Crozier RA, Plummer MR, and Black IB.** Rab3A is required for brain-derived neurotrophic factor-induced synaptic plasticity: transcriptional analysis at the population and single-cell levels. *J Neurosci* 27: 6782–6790, 2001.
76. **Thinakaran G, Kitt CA, Roskams AJ, Slunt HH, Maslah E, von Koch C, Ginsberg SD, Ronnett GV, Reed RR, and Price DL.** Distribution of an APP homolog, APLP2, in the mouse olfactory system: a potential role for APLP2 in axogenesis. *J Neurosci* 15: 6314–6326, 1995.
77. **Tu JC, Yuan JP, Lanahan AA, Leoffert K, Li M, Linden DJ, and Worley PF.** Homer binds a novel proline-rich motif and links group I metabotropic glutamate receptors with IP3 receptors. *Neuron* 21: 717–726, 1998.
78. **Utsumi M, Sato K, Tanimukai H, Kudo T, Nishimura M, Takeda M, and Tohyama M.** Presenilin-1 mRNA and beta-amyloid precursor protein mRNA are expressed in the developing rat olfactory and vestibulocochlear system. *Acta Otolaryngol* 118: 549–553, 1998.
79. **Van Veldhoven PP and Mannaerts GP.** Sphingosine phosphate lyase. *Adv Lipid Res* 26: 69–98, 1993.
80. **Van Veldhoven PP.** Sphingosine-1-phosphate lyase. *Methods Enzymol* 311: 244–254, 2000.
81. **Van Veldhoven PP, Gijbbers S, Mannaerts GP, Vermeesch JR, and Brys V.** Human sphingosine-1-phosphate lyase: cDNA cloning, functional expression studies and mapping to chromosome 10q22(1). *Biochim Biophys Acta* 1487: 128–134, 2000.
82. **Vogler R, Sauer B, Kim DS, Schafer-Korting M, and Kleuser B.** Sphingosine-1-phosphate and its potentially paradoxical effects on critical parameters of cutaneous wound healing. *J Invest Dermatol* 120: 693–700, 2003.
83. **Xiong JW, Leahy A, and Stuhlmann H.** Retroviral promoter-trap insertion into a novel mammalian septin gene expressed during mouse neuronal development. *Mech Dev* 86: 183–191, 1999.
84. **Yamagishi M, Ishizuka Y, and Seki K.** Pathology of olfactory mucosa in patients with Alzheimer's disease. *Ann Otol Rhinol Laryngol* 103: 421–427, 1994.
85. **Wang MM, Tsai RY, Schrader KA, and Reed RR.** Genes encoding components of the olfactory signal transduction cascade contain a DNA binding site that may direct neuronal expression. *Mol Cell Biol* 13: 5805–5813, 1993.
86. **Weston WM, Trzyna W, McHugh KM, Lafferty CM, Ma L, and Greene RM.** Differential display identification of plunc, a novel gene expressed in embryonic palate, nasal epithelium, and adult lung. *J Biol Chem* 274: 13698–13703, 1999.
87. **Wetmore BA, Mitchell AD, Meyer SA, and Genter MB.** Evidence for site-specific bioactivation of alachlor in the olfactory mucosa of the Long-Evans rat. *Toxicol Sci* 49: 202–212, 1999.
88. **Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Pruss M, Reuter I, and Schacherer F.** TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res* 28: 316–319, 2000.