Published in final edited form as:

Hum Mutat. 2013 January ; 34(1): 32–41. doi:10.1002/humu.22212.

General Olfactory Sensitivity Database (GOSdb): Candidate Genes and their Genomic Variations

Ifat Keydar^{1,*}, Edna Ben-Asher¹, Ester Feldmesser², Noam Nativ¹, Arisa Oshimoto³, Diego Restrepo³, Hiroaki Matsunami⁴, Ming-Shan Chien⁴, Jayant M. Pinto⁵, Yoav Gilad⁶, Tsviya Olender¹, and Doron Lancet¹

¹Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

²Bioinformatics Unit, Department of Biological Services, Weizmann Institute of Science, Rehovot, Israel

³Department of Cell and Developmental Biology, Neuroscience Program, and Rocky Mountain Taste and Smell Center, University of Colorado School of Medicine, Aurora, Colorado

⁴Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina

⁵Section of Otolaryngology-Head and Neck Surgery, University of Chicago, Chicago, Illinois

⁶Department of Human Genetics, University of Chicago, Chicago, Illinois

Abstract

Genetic variations in olfactory receptors likely contribute to the diversity of odorant-specific sensitivity phenotypes. Our working hypothesis is that genetic variations in auxiliary olfactory genes, including those mediating transduction and sensory neuronal development, may constitute the genetic basis for general olfactory sensitivity (GOS) and congenital general anosmia (CGA). We thus performed a systematic exploration for auxiliary olfactory genes and their documented variation. This included a literature survey, seeking relevant functional *in vitro* studies, mouse gene knockouts and human disorders with olfactory phenotypes, as well as data mining in published transcriptome and proteome data for genes expressed in olfactory tissues. In addition, we performed next-generation transcriptome sequencing (RNA-seq) of human olfactory epithelium and mouse olfactory epithelium and bulb, so as to identify sensory-enriched transcripts. Employing a global score system based on attributes of the 11 data sources utilized, we identified a list of 1,680 candidate auxiliary olfactory genes, of which 450 are shortlisted as having higher probability of a functional role. For the top-scoring 136 genes, we identified genomic variants (probably damaging single nucleotide polymorphisms, indels, and copy number deletions) gleaned from public variation repositories. This database of genes and their variants should assist in rationalizing the great interindividual variation in human overall olfactory sensitivity (<http://genome.weizmann.ac.il/GOSdb>).

Keywords

olfactory candidate genes; congenital general anosmia; RNA-seq

© 2012 Wiley Periodicals, Inc.

*Corresponding to: Ifat Keydar, Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel. ifat.keydar@weizmann.ac.il.

The authors have no conflict of interest to declare.

Additional Supporting Information may be found in the online version of this article.

Introduction

The detection of odorant molecules by the human olfactory system is mediated by molecular components localized in specialized cilia of olfactory sensory neurons (OSNs) in the olfactory epithelium [Pifferi et al., 2011; Touhara and Vosshall, 2009]. The first step in this process is stereospecific recognition by a repertoire of ~400 olfactory receptors (ORs) proteins [Buck and Axel, 1991; Nara et al., 2011; Saito et al., 2009]. Subsequently, a signal transduction cascade leads to the generation of action potentials in the OSN axons leading to the brain [Bozza et al., 2009; Kleene, 2008; Mori and Sakano, 2011; Schild and Restrepo, 1998]. Variations in odorant-specific thresholds among healthy individuals typically span several orders of magnitude, showing a monomodal or multimodal distribution [Amoore and Steinle, 1991; Cain and Gent, 1991; Menashe et al., 2007; Stevens et al., 1988; Wysocki and Gilbert, 1989]. The extreme cases pertaining to a given odorant are termed specific hyposmia (diminished sensitivity) and specific hyperosmia (enhanced sensitivity). A complete incapacity to sense a given odorant is termed specific anosmia. It is generally accepted that genetic variations in OR genes contribute significantly to this interindividual phenotypic variation. Thus, a clear genetic factor in odorant-specific thresholds was demonstrated by twin studies [Gross-Isseroff et al., 1992; Whissell-Buechy and Amoore, 1973; Wysocki and Beauchamp, 1984] and in mouse models [Griff and Reed, 1995], and recent studies have shown genetic association of sensitivity to specific odorants with identified OR genes [Keller et al., 2007; Menashe et al., 2007].

The foregoing portrayal suggests that an individual's sensitivity to two unrelated odorants (binding to two different OR proteins) would typically be uncorrelated. However, it has been repeatedly shown that sensitivities to multiple odorants are significantly correlated in a given subject [Brown et al., 1968; Cain and Gent, 1991; Menashe et al., 2007; Punter, 1983; Stevens and Dadarwala, 1993]. This suggests that a cross-odorant sensitivity average is a legitimate quantitative phenotype, with a potential genetic basis. Previously, this phenotype has been termed "general factor of sensitivity" [Cain and Gent, 1991], and it is herein referred to as "General Olfactory Sensitivity" (GOS).

GOS is powerfully affected by nongenetic factors including age, head trauma, viral infections, medications, exposure to chemicals, and menstrual cycle [Doty, 2009; Doty and Cameron, 2009]. Nevertheless, a genetic component for GOS has been demonstrated by twin studies [Finkel et al., 2001; Pause et al., 1998; Segal et al., 1995], as well as by whole genome linkage analysis, even though specific genes were not unequivocally identified yet [Pinto et al., 2008]. A single nucleotide polymorphism (SNP) in the potassium voltage-gated channel *KCNA3* (MIM# 176263) is associated with altered general olfactory function [Guthoff et al., 2009]. A genetic etiology for GOS is also demonstrated by mouse gene knockout models displaying general hyposmia [Layman et al., 2009] or general hyperosmia [Lu et al., 2010; Walz et al., 2006]. Likewise, some human syndromic monogenic disorders are accompanied by general hyposmia [Kulaga et al., 2004], and general hyperosmia [Henkin and Bartter, 1966]. GOS decrements are also associated with complex diseases such as Alzheimer's disease [Wesson et al., 2010], Parkinson's disease [Morley and Duda, 2010], schizophrenia [Atanasova et al., 2008], and autism [Dudova et al., 2011], and in some cases serve as a biomarker. Finally, a clear gender effect for GOS, whereby females show enhanced GOS [Doty and Cameron, 2009] is at least partly ascribable to genetic components.

One well-defined deficit in the human sense of smell is congenital general anosmia (CGA), an inborn and complete incapacity to sense any odorants. Nonsyndromic (isolated) CGA (MIM# 107200), is a rare condition with an estimated frequency of 1:2,000 to 1:10,000

[Feldmesser et al., 2007; Karstensen and Tommerup, 2011]. The genetic etiology of isolated CGA, inferred from its familial occurrences [Gaines, 2010; Ghadami et al., 2004; Glaser, 1918; Lygonis, 1969; Qu et al., 2010; Singh et al., 1970], is yet unidentified [Karstensen and Tommerup, 2011]. In contrast, there are numerous studies showing the involvement of defined genes in syndromic CGA. An early human example for syndromic CGA is *KAL1* gene (MIM# 300836), whose missense mutations are associated with X-linked Kallmann's syndrome, accompanied with anosmia due to the olfactory bulb agenesis [Lewkowicz-Shpuntoff et al., 2012]. Another more recent example is the identification of *SCN9A* (MIM# 603415), the gene that encodes voltage-gated sodium channel Nav1.7, whose loss-of-function mutations cause congenital insensitivity to pain accompanied by anosmia [Weiss et al., 2011]. Underlying genes and mutations have been reported for additional instances of monogenic syndromic CGA [McEwen et al., 2007; Pluznick et al., 2011], including cases of structural genome variations and linkage peaks [Gerber et al., 2011; Ghadami et al., 2004; Pinto et al., 2008; Sobin et al., 2006] where the specific gene is not yet known. Moreover, multiple mouse gene knockout models of core olfactory transduction genes lead to syndromic CGA, among them *ADCY3* (MIM# 600291) [Wong et al., 2000] and *GNAL* (MIM# 139312) [Belluscio et al., 1998].

Previously, it has been suggested that genetic variations in downstream transduction genes or in genes involved in higher neuronal processing may form the explanatory basis for across-odorant sensitivity phenotypes such as GOS and CGA [Feldmesser et al., 2007; Menashe et al., 2007]. Here, we expand this line of thought and employ a systematic search strategy to define candidate genes for these phenotypes. We report the establishment of a database of GOS/CGA candidate genes based on both data mining and experimental research. As the candidate gene approach has proven extremely powerful for studying the genetic architecture of monogenic and complex traits [Dattani, 2009; Ho-Pun-Cheung et al., 2010; Lagendijk et al., 2011; Mirabello et al., 2011], we believe that the resultant database of prioritized gene list should serve as a powerful tool for the study of olfactory genetics. GOSdb (<http://genome.weizmann.ac.il/GOSdb>) integrates GOS candidate genes, their annotation sources and their variations. The mutation nomenclature format follows the general recommendations for LSDB databases suggested by the Human Genome Variation Society (HGVS) [Vihinen et al., 2011].

Materials and Methods

Data Mining (DS1-5, 10, 11)

DS1: Olfactory functional genes identified in vitro were traced by extensive screen of Pubmed articles. An initial gene search was performed for: olfact* AND vertebrate AND (transduction OR development OR cilia) AND (gene OR review). *DS2*: Genes whose mouse ortholog knockout have an olfactory phenotype were traced systematically using the Mouse Genome Informatics (MGI) database via the mammalian phenotype browser, under the sections abnormal olfactory system morphology (MP:0005499) and abnormal olfactory system physiology (MP:0001983). *DS3*: Genes were attained from the OMIM database following discovery of relevant diseases by screen of Pubmed articles for the keywords anosmia or hyposmia. *DS4*: Genes whose encoded proteins are localized in olfactory cilia were obtained from a report on rat proteome analysis [Mayer et al., 2009]. HUGO symbols were allocated for 329 of them using GeneCards. *DS5*: Genes differentially expressed in mouse mature OSN were obtained from published microarray data [Sammata et al., 2007]. We used as reference the data for immature OSN and brain tissues expression, included in the same microarray report. For each gene, the differential expression score was computed as the rank in mature OSN divided by the average rank for the two control tissues. A smaller number indicated a higher degree of differential expression. *DS10*: We performed search in the reported genomic regions [Gerber et al., 2011; Ghadami et al., 2004; Pinto et al., 2008;

Sobin et al., 2006] and excluded genes that were not supported by an additional source. *DS11*: We performed GeneCards search (www.genecards.org) with the keywords: olfact* NOT “olfactory receptor” NOT cortex, and excluded genes already present in DS1. For all genes, aliases were substituted by approved HUGO gene symbols using GeneCards.

RNA-seq Samples

Human olfactory epithelium RNA samples were collected from four autopsies and three consented live biopsies (Supp. Table S1). Mouse olfactory epithelium and olfactory bulb RNA was collected in pool from four adult female *Trpm5*-GFP mice crossed with C57BL/6 for over 10 generations [Lin et al., 2007]. In humans, olfactory epithelium was dissected from the superior ethmoturbinal, the cribriform plate and the superior septum [Nakashima et al., 1984]. Total RNA olfactory epithelium samples were shipped to the Weizmann Institute high-throughput sequencing unit on dry ice for transcriptome sequencing. mRNA was prepared by oligo-dT affinity separation, chemically sheared, and cDNA synthesized by random priming, blunt-ended and ligated to specific adaptors, as prescribed by the Illumina method. RNA samples were evaluated by ultraviolet spectroscopy for purity and concentration (NanoDrop, Wilmington, DE) and were assessed further for RNA integrity on the Agilent Bioanalyzer (Santa Clara, CA). cDNA of each individual was sequenced by NGS in separate lane. The sequencing was done using the Illumina Genome Analyzer platform (Illumina GA IIX) as prescribed by the Illumina method. Libraries were prepared using the Illumina mRNA-seq Sample Preparation Kit (San Diego, CA). Briefly, poly(A)+ RNA was recovered from 1 μ g of total RNA using two rounds of isolation with oligo-dT-coated Sera-Mag magnetic beads. The recovered poly(A)+ RNA was then chemically fragmented. RNA fragments were converted to cDNA using SuperScript II and random primers. The second strand was synthesized using RNaseH and DNA Pol I. The ends of the cDNA were repaired using T4 DNA polymerase, T4 polynucleotide kinase, and Klenow DNA polymerase. A single adenosine was added to the 3' end using Klenow fragment (3' to 5' exo minus). Adaptors were attached to the ends of the cDNA using T4 DNA ligase. About 300 bp fragments were extracted from a 2% low-range ultra agarose sizing gel. The 300 bp fragment was then amplified by 15 cycles of polymerase chain reaction (PCR) using Phusion DNA polymerase. Libraries were validated with an Agilent Bioanalyzer (Santa Clara, CA). Libraries were diluted to 10 pM and applied to an Illumina flow cell using the Illumina Cluster Station. Sequencing was performed on an Illumina GAIIx. Sequences were 76 cycle single read except for the third flowcell, which was 70 cycles. Sequencing resulted in an average of 22 million reads per lane (8–32 M) (Supp. Table S1). We used as controls data of 8 human tissues mRNA, sequenced by the Illumina human body map 2.0 project, in FASTQ format (Illumina BodyMap 2.0 transcriptome (SRA accession ERP000546)). Tissues included in our analysis were adrenal, brain, colon, kidney, liver, lung, testes, and white blood cells.

Human Olfactory Epithelium RNA-seq (DS8)

We originally carried out RNA-seq experiments of seven human olfactory epithelium samples (four autopsies and three biopsies). Preliminary comparison to mouse olfactory epithelium data suggested that six of these contain a low amount of olfactory sensory cells, based on an unsatisfactory olfactory epithelial transcriptome signature (Supp. Fig. S1). This likely stems from the fact that the structure of human olfactory epithelium is often patchy and contains a significant component of respiratory epithelium [Morrison and Moran, 1995; Witt et al., 2009]. We selected OE7 (a biopsy sample) as the only bone-fide human olfactory epithelial sample. Samples OE1-3 were sequenced with low coverage (Supp. Table S1) and were, therefore, excluded from further analyses. We selected sample OE6 (an autopsy sample), which was the most remote from acceptable olfactory signature, as a putative

respiratory epithelium and included it as a ninth control in the differential expression analysis.

Human RNA-seq analyses were performed at the transcript level. Reads were aligned to the hg19 version of the human genome assembly using Tophat [Trapnell et al., 2009]. We allowed reads to be mapped up to 40 times to the Genome with three mismatches. Identification and differential expression analysis of splice isoforms and novel genes was performed using Cufflinks [Langmead et al., 2009], in comparison with UCSC known genes transcripts. Cufflinks analysis generated 145,000 gene transcripts based on 34,607 UCSC transcripts. Normalization was performed by reads per kilobase of exon model per million mapped reads (RPKM). Transcripts below 1 RPKM were excluded. Transcripts that were mapped to unannotated regions, many of them seem as short and unspliced fragments, were eliminated from our current analysis. For the remaining 78,900 transcripts we performed nine pairwise comparisons using Cuffdiff of OE7 versus 9 controls, and filtered out transcripts that not fulfilling the differential expression criterion of being expressed $>X4$ in OE7 compared to and having an absolute OE7 expression level in OE7 of >2 RPKM. This resulted in defining 352 genes differentially expressed in the chemosensory epithelium sample OE7, of them 290 genes being expressed >10 -fold change in OE7 compared to highest value in controls. For DS5–DS9, 250 genes with the highest relative ranks were selected and scored from 1 to 0 based on their order. The somewhat arbitrary number 250 is aimed to generate a sufficiently strong, unbiased and equal representation for each source type as a basis for the summed score (see the section “Results”). The alternative method of using a specific score cutoff was deemed less favorable because different data sources use different scales.

Mouse RNA-seq (DS6, 9)

Mouse RNA-seq data were analyzed at the gene rather than transcript level. Human and mouse RNA-seq reads were aligned to the hg19 version of the human genome assembly or NCBI m37 version of the mouse genome assembly, using Bowtie short read alignment program allowing up to three mismatches. Only uniquely mapping reads were used in this analysis. PCR duplicates were removed using Picard’s command line tool MarkDuplicates (<http://picard.sourceforge.net>). Reads were assigned to Refseq transcripts using Partek. Human orthologs for such mouse genes were determined using GeneALaCart symbol converter in GeneCards. This resulted in the definition of 17,147 mouse genes that have human orthologs. For differential expression analyses ORs and miR-NAs were excluded, the later because our poly(A)+ RNA method did not retain them [Morin et al., 2008]. DS6, 9: For each gene, the differential expression score was computed as the rank in mouse olfactory tissue divided by the average rank for the nine control tissues. DS7: For mouse single OSN analysis, we selected 1000 genes with highest expression levels. Ranks were normalized by average rank of nine controls.

Single Mouse OSN RNA-seq (DS7)

Single cell-derived cDNA was conducted as described [Saito et al., 2004]. Briefly, olfactory epithelium from 3-week-old mice were dissected and enzymatically dissociated. Single OSNs were picked with micropipettes and transferred to reaction tubes. Oligo d(T)-primed cDNAs were synthesized and amplified by PCR. Amplified samples with olfactory marker protein (*Omp*), olfactory type guanine nucleotide binding protein (*Gnal*), cyclic nucleotide gated channel alpha 2 (*Cnga2*), and adenylate cyclase 3 (*Adcy3*) were selected. Sequence libraries for Illumina short read sequencing were constructed from the selected amplified cDNAs. The sequenced results from individual OSNs were mapped to mouse RefSeq database with a threshold <10 repetitive-mapped locations by using Bowtie alignment

program. The reads of the transcripts from the same genes were merged for the further analysis at the gene level.

Genomic Variations

Genomic variations for the auxiliary olfactory genes were obtained from the NCBI dbSNP database for validated SNPs and indels, using dbSNP build 134, Genome build 37.2. We considered variations with no reported frequency as $MAF < 0.01$. Missense variations were filtered by PolyPhen-2 score (genetics.bwh.harvard.edu/pph2), which represents the probability that a substitution is damaging. We used a stringent inclusion criterion selecting only missense SNPs with a PolyPhen-2 score > 0.95 (“probably damaging”), that may be considered as likely deleterious with a higher confidence. Exonic copy number deletion variants were mined using DGV (projects.tcag.ca/variation).

Results

Gene Data Sources

Our strategy for seeking genes with a possible role in the GOS and CGA phenotypes involved the integration of eleven assorted data sources (Table 1). Four data sources (DS1–DS3, DS11) were based on literature data mining of gene-focused studies. Two additional data sources (DS4–DS5) were obtained via the analysis of high-throughput data from previously reported studies of OSN gene expression. Four data sources (DS6–9) were obtained experimentally via our own RNA-seq of human olfactory epithelium as well as of mouse olfactory epithelium and bulb. Lastly, published genetic linkage analyses performed in hyposmic and anosmic cohorts constituted an additional gene source (DS10). This diverse ensemble of data sources ensured a balanced view that includes both peripheral and central mechanisms, and addresses a wide range of biological processes.

Functional studies (DS1)—We performed an extensive Pubmed search, which identified 181 genes via functional studies (Supp. Table S2). This identified genes in three broad categories: 108 mediating odorant signal transduction, modulation and termination, 66 involved in sensory neuronal development and connectivity, of which 24 code for transcription factors, and 8 underlying intracellular trafficking in these neurons and their ciliary organelles.

Mouse knockouts (DS2)—We extracted from MGI 88 human genes whose mouse orthologous gene knockouts manifest an olfactory phenotype (Supp. Table S3). The olfactory phenotypes include impaired olfactory system physiology (e.g., reduced sensitivity and general anosmia) and morphology (e.g., absent nasal placodes and fewer sensory neurons). Of these, 77 genes present multiple severe phenotypes coupled to the olfactory dysfunction. For example, *GNAL* knockout mice also have neurological syndromes [Belluscio et al., 1998]. Multiple genes are embryonically lethal (e.g., *SOX2*, *TCOF1*), and their role in olfaction is deduced from grossly abnormal olfactory organ structure. For *CNGA2* (MIM# 300338) knockout, mice suffer from postnatal lethality and growth retardation but this phenotype has been ascribed to their inability to locate the nipple, which requires olfaction [Brunet et al., 1996].

Human diseases (DS3)—Here, we include genes whose mutations cause a human genetic disorder/syndrome accompanied by olfactory impairment. Our data mining revealed 16 diseases (Table 2) associated with syndromic olfactory impairments for which the gene variation has been identified, encompassing a total of 60 genes (Supp. Table S4). Six of these syndromes are identified as ciliopathies, a broad class of human diseases that involve cilia-specific genes and pathways [Jenkins et al., 2009]. Most of the diseases are

heterogenic, that is, more than one gene may cause the disease phenotype, and olfactory dysfunction was only tested for a subset of the disease subtypes (Table 2, marked *). For 22 of the genes (marked in †), parallel evidence exists from mouse knockout. The remaining genes were not yet directly proven to be related to olfactory phenotype, and are included in our list as merely candidates with potential to affect olfaction functionally.

Ciliary proteome (DS4)—This data source includes 377 genes whose encoded proteins were identified by mass-spectrometry-based proteomics in rat olfactory cilia [Mayer et al., 2009]. For 329 genes, an orthologous human gene symbol was retrieved (Supp. Table S5). The genes were scored by differential expression in comparison to RNA-seq values from the nine control tissues, so as to minimize the inclusion of general housekeeping genes.

Microarray expression (DS5)—Here we included ~10,000 genes whose transcripts were identified in mature OSNs by Affymetrix mouse microarrays [Sammata et al., 2007]. The genes were scored by differential expression in comparison to two internal controls, immature OSNs and brain, and the top scoring 250 genes were selected.

RNA-seq of olfactory tissues (DS6, 8, 9)—These data sources constitute results of our RNA-seq analyses of olfactory tissues, in comparison to the 9 control tissues (Supp. Table S6). We sought genes differentially expressed in mouse olfactory epithelium (DS6), human olfactory epithelium (DS8), and mouse olfactory bulb (DS9). These as well as DS7 data were scored by differential expression as in DS4 and the top 250 genes were selected. Figure 1A shows the expression patterns for the 30 olfactory-implicated genes (DS1-3) with the highest score in DS6 (see below), whereas Figure 1B shows the 30 top scoring genes not in DS1-DS3 (novel olfactory-implicated genes). All 250 genes contributed by DS6 are shown in Supp. Fig. S2.

RNA-seq of isolated mouse OSNs (DS7)—We extracted from these data (from the laboratory of HM, Supp. Table S7) the 1,000 genes with highest expression in these neurons. The genes were prioritized as for DS6, 8, and 9.

Linkage peak genes (D10)—We identified the Refseq genes located in regions reported as linkage peaks for CGA on Chr18p11-q12 [Ghadami et al., 2004] and Chr6q27 [Gerber et al., 2011]; and for general hyposmia on Chr4q22-q31 [Pinto et al., 2008] and on Chr22q11.2 [Sobin et al., 2006] (Supp. Table S8). Only genes already identified in additional data sources were included.

Genecards search (D11)—Data source 11 contributed genes identified via computer search. Genes already identified in DS1 were excluded here (Supp. Table S9).

Gene Prioritization

Scoring the genes—Each of the 11 data sources resulted in a roster of genes. For each data source gene a score was ascribed as follows: For data sources 1–3, all genes received score = 1.0; for data sources 4–9, genes were assigned a score between 0.0 and 1.0, based on their differential expression ranks (top rank assigned 1.0); For data sources 10 and 11, all genes received a score of 0.5 (Table 1 and see section “Materials and Methods”). Subsequently, we integrated the 11 data sources to generate a full “long list” of 1,680 candidate auxiliary genes. These were ranked by a summed score of all 11 data source scores (Supp. Fig. S3, Supp. Table S10). The summed score provides an integrated tool for assessing the significance of genes across all data sources. The overlap relationships of the annotation data sources DS1-3 and the differential expression data sources are shown in Figure 2. The mutual overlap relationships among some of the data sources and the

distribution of the number of sources contributing to a gene are depicted in Supp. Figure S4. Subsequently, two sublists were generated to afford certain detailed analyses as shown herein on a more limited number of genes: The “medium list” (450 genes) consist of genes with score 1 or genes reported by two or more data sources (Supp. Fig. S5, Supp. Table S11); and the “short list” (136 genes) consisting of genes with score 1.7 (Fig. 3, Supp. Table S12). The short list does not contain olfactory annotated genes that are supported by merely a single source. This selection possibly eliminates important genes yet focus on the most potent candidates for olfactory specific phenotype.

Notably, among the top 19 genes with scores higher than 3.0 (indicating evidence from 4 sources, including at least one gene expression source), a great majority are widely documented to have a central role in olfaction, supporting our analytic strategy. Such genes include *CNGA2* (#1 in Supp. Table S12) and *CNGA4* (#2 in Supp. Table S12; MIM# 609472), the cyclic nucleotide gated channel $\alpha 2$ and $\alpha 4$ subunits; *OMP* (#5 in Supp. Table S12; MIM# 164340), olfactory marker protein; the receptor transporter proteins *RTP1* (#8 in Supp. Table S12; MIM# 609137) and *RTP2* (#3 in Supp. Table S12; MIM# 609138), as well as the olfactory cAMP phosphodiesterase *PDE1C* (#4 in Supp. Table S12; MIM# 602987), G-protein α subunit *GNAL* (#10 in Supp. Table S12) and adenylyl cyclase *ADCY3* (#14 in Supp. Table S12).

Syndromic auxiliary genes—The short list is also a rich source of information regarding olfactory-related genes for which human disease and mouse knockout phenotypes are the only indicator for functional relevance. These are genes that are not in DS1 and DS11 but seen in DS3 (“syndromic genes,” Supp. Table S4). Such genes include *RD3* (#15 in Supp. Table S12; MIM# 180040) which underlies Leber congenital amaurosis (LCA12), not hitherto annotated in the olfactory context; *PKD1* (#100 in Supp. Table S12; MIM# 601313) and *PKD2* (#42 in Supp. Table S12; MIM# 173910) which underlies polycystic kidney disease, *MKKS* (#91 in Supp. Table S12; MIM# 604896) which underlies McKusick–Kaufman syndrome, *MKS1* (#92 in Supp. Table S12; MIM# 609883) which underlies Meckel–Gruber syndrome and Bardet–Biedl syndrome (BBS 13), and *PEX7* (#98 in Supp. Table S12; MIM# 601757) and *PHYH* (#99 in Supp. Table S12; MIM# 602026) which underlies Refsum disease. Notably, all these diseases are ciliopathies.

Transcriptome-implicated auxiliary genes—Of importance are genes for which information is derived solely from transcriptome analyses (DS5–9), including our RNA-seq experiments, which may be considered newly discovered auxiliary olfactory genes. These include *NEU2* (#32 in Supp. Table S12; MIM# 605528) that functions in removing sialic acid residues from glycoproteins; *SLC17A6* (#37 in Supp. Table S12; MIM# 607563), a sodium-phosphate cotransporter; *SP8* (#43 in Supp. Table S12; MIM# 608306), a developmental transcription factor; *CTXN3* (#47), a poorly annotated neuronal development gene; and *SEC14L3* (#51 in Supp. Table S12; MIM# 612824) a probable hydrophobic ligand-binding protein and involved in golgi vesicle biogenesis. Other interesting (mostly not on the short list) are genes, which have a very strong differential expression in the two murine samples (olfactory epithelium and bulb, DS6 and DS9). Such genes include *BOLA2* (#106 in Supp. Table S12; MIM# 613182) probably involved in cell cycle and cell proliferation; *RPS17* (#110 in Supp. Table S12; MIM# 180472), with proposed role in retinoic acid signaling; and *GTPBP6* (#111 in Supp. Table S12; MIM# 300124), a putative GTP-binding protein.

Catalog of Genomic Variations in the Short-List Genes

The gene catalogue was a first step en route to finding functionally relevant variations in the genome affecting olfaction. We generated a catalog of 1,818 likely deleterious genomic

variations for the 136 genes in the short list, using data from public databases (dbSNP and DGV). This included 1253 likely deleterious missense SNPs, 143 nonsense SNPs, 178 frame-disrupting indels, and 244 deletion CNVs (Supp Table S13). Among the 1,574 non-CNV variations, 65 are reported in Human Gene Mutation Database (HGMD, www.hgmd.org) as causing some of the diseases in DS3 genes, as well as implicated to some nonolfactory diseases (Supp. Table S13). We further identified 1,357 rare variants (minor allele frequency < 0.01), including 1,055 likely deleterious missense, 128 stop-gain and 174 frameshifts (Fig. 4) (CNVs are excluded because allele frequencies are not readily available for them). These rare variations are presumed candidates for underlying the monogenic CGA, both non-syndromic and syndromic. We view the remaining 217 non-CNV variations with minor allele frequency > 0.01 as better candidates for the quantitative trait of GOS.

Some genes at the top-scored short list show an unusually high prevalence of catalogued likely deleterious variations (Supp. Table S12). Prominent examples include the cAMP-gated channel subunit *CNGA2*, with 30 variations; *SCN9A*, the olfactory axonal Na channel with 31 variations; and *ACSM4* (MIM# 614360), an olfactory coenzyme A ligase with 9 variations (Supp. Fig. S6). Other examples: *CEP290* (MIM# 610142), involved in ciliary transport processes with 30 variations; the cAMP-gated channel subunit *CNGA4* with 11 variations, and *RPGRIP1L* (MIM# 610937), a ciliopathy gene with 32 variations. Five of these mutations (in *CEP290* and *RPGRIP1*) are documented in the context of human diseases by HGMD (Supp. Table S12, DC column).

Stop-gain variations with high minor allele frequency (>0.01) have been widely documented [MacArthur et al., 2012], with one of the earliest and prominent examples being segregating pseudo-genes in OR genes [Menashe et al., 2003]. We have identified such variations in 15 of the shortlist genes (Supp. Table S13), constituting prominent candidates for CGA as well as contributors to GOS.

Discussion

Study Design Consideration

We integrated and scored 1,680 candidate auxiliary olfactory genes, and extracted two progressively prioritized sublists: a medium list with 450 genes, including all those implicated by human disease and mouse knockouts, and a short list with 136 genes with highest probability of olfactory function, and for which we present full documentation of likely deleterious variations. We note that some genes previously annotated as related to olfaction are not included in the short list. This is because this list includes only genes supported by a summed score >1.7, hence genes only reported by one of the sources DS1–3 are excluded (summed score = 1.0). An example is *KAL1*, known to underlie human syndromic CGA, but otherwise not appearing in the literature in an olfactory context, perhaps due to the absence of a rodent ortholog.

An important aspect of our endeavor is the utilization of RNA-seq differential expression analyses, which in the example of the short list contributes 40 genes not previously studied, and implicated in olfaction solely via differential expression in olfactory tissues. The auxiliary gene list is unique in that it contributes extra dimensions to the investigation of the genetics of human olfaction, as it combines assorted functional annotation sources, proteomic and genetic linkage data together with differential expression data. Yet, multiple factors may result in inclusion as false positives or false negatives. For that reason, we generated a procedure that combines permissive inquiry followed by a scoring system and a demand for consensus of two sources or more to support inclusion of a gene. This minimizes, but, of course, does not eliminate erroneous gene placement. We are aware of

several sources of false positive inclusion, such as spurious cellular colocalization with functional gene products; inclusion of several paralogs due to the lack of specific olfactory-implicating information; using genes of heterogenic human disease with olfactory phenotype for which some genes were not directly studied in the olfactory context; using mouse/rat models which are not fully identical to human; differential expression signals in heterogeneous tissues such as olfactory epithelium; insufficient control tissues for differential expression determination; the obvious inclusion of multiple false positives when using linkage peaks containing numerous genes. Potential sources for false negatives include: RNA-seq bias toward highly expressed genes and long transcripts leading to the omission of potentially important genes that happen to have low expression levels; using mouse/rat models which are not fully identical to human; and using information with a limited gene spectrum such as olfactory cilia.

Predicted Effect of the Variations

We present here lists of variations for the 136 genes in our short list. These variation rosters are obtained from dbSNP and DGV, and are not aimed to be comprehensive. For example, for the gene *CEP290* a specialized mutation database exists under the auspices of locus-specific mutation Database (<http://medgen.ugent.be/cep290base>) that presents a more comprehensive set of variations. The GOSdb database includes deep links to all relevant LSDB variation databases.

For the great majority of our reported variations, there is no direct evidence for an olfactory phenotype. For some of the genes, however, we can attempt to make an educated guess, based on the information in our data sources, for how a single deleterious genetic variation might affect olfactory function. One obvious example is variations in *DS3* genes, shown to underlie human monogenic olfactory phenotypes. Some of the mutations documented here are the actual syndrome-causing variations as previously reported. In addition, there are 610 other variations in such genes, and all are rather likely to result in an olfactory phenotype. As most database variations are purportedly derived from healthy individuals, such variations are more likely to contribute to the GOS human quantitative trait.

We highlight 74 genes based on their involvement in mouse knockouts with olfactory phenotype. Variation in these genes will also have a high probability of underlying GOS or CGA. This is based on the evidence for human mouse similarity, whereby an olfactory phenotype is recapitulated in a mouse knockout [Kulaga et al., 2004; Weiss et al., 2011]. Although the effect of full gene loss is not always similar to that of a hypomorphic mutation in an otherwise intact gene, some examples show that such inference is legitimate, for example, for the *CEP290* gene [McEwen et al., 2007]. There are 10 cases of *DS2* genes that report pure (nonsyndromic) mouse olfactory phenotype (Supp. Table S3), as exemplified by signaling genes such as *ADRBK2*, *CNGA2*, *CNGA4*, *OMP*, and *PDE1C*. Notably, only one of these genes causes complete anosmia (*CNGA2*), whereas others mainly cause hyposmia. Many of these genes are thus most likely to cause nonsyndromic GOS.

We find 16 genes supported by in vitro studies (*DS1*) in conjunction with differential expression (*DS5*–*9*), but not by mouse and human phenotypes (*DS2*–*3*). These are high-quality novel candidates, and therefore the 96 mutations identified within such genes deserve attention. This is despite some counter examples, where *DS1* genes do not show a mouse olfactory phenotype, for example, *PDE4A* [Cygnar and Zhao, 2009], sometimes because of redundant paralogs. Another 40 genes are solely supported by transcriptome analyses (*DS5*–*9*), 10 of which also count as ciliary genes (*DS4*). These are newly identified candidates for generating human olfactory phenotypes. We note that expression evidence alone may be relatively weak, as among 126 genes with genetic evidence (*DS2*, 3) only 19

showed differential expression (15%), a rather low overlap. Yet, high differential expression is generally a predictor of higher probability for nonsyndromic olfactory phenotype.

A small, highly interesting class of genes appears to increase olfactory sensitivity (general hyperosmia) in knockout mice (*NCAM2*, *GABRR1*, and *KCNA3*). For another gene, *ACSM4*, mice have enhanced sensitivity to aversive odors [Kobayakawa et al., 2007]. It is not unlikely that the frequent (MAF = 0.05) nonsense variation in *ACSM4* (rs7485773: NM_001080454.1:c.1069C>T) will cause a similar phenotype in humans. The *KCNA3* gene (sodium-activated potassium channel) is an example of correspondence between mouse knockout [Lu et al., 2010] and a human cohort study [Guthoff et al., 2009]. A nonreference variant in human *KCNA3* (rs2821557, NC_000001.10:g.111218920C>T) was associated with decreased GOS. However, since the nonreference variant is ancestral, it is legitimate to view the derived reference variant as causing hyperosmia. Possibly, the frequent nonsense variation in *KCNA3* (rs35154586, NM_002232.3:c.1516C>T) appearing in our list will result in a similar hyperosmia phenotype.

Olfactory Genetics at a Systems Level

Olfactory perception relies on an elaborate interplay of numerous genes. The most obvious candidates, begun to be deciphered nearly 30 years ago, are members of the immediate postreceptor transduction, adaptation and termination chain in the sensory neurons: olfactory G-protein, adenylyl cyclase, and cAMP-gated cation channel. However, recent years have seen an explosion of information regarding genes and proteins necessary for the development, maintenance, and mutual interactions of the different cellular and molecular components. The formation and maintenance and transport mechanisms within olfactory cilia [McEwen et al., 2008] and the pathfinding of sensory axons in the glomeruli of the olfactory bulb [Imai and Sakano, 2011] are examples, which have attracted considerable attention. Interestingly, 64% of all the olfactory gene knockouts summarized here result in an abnormal developmental sequence, adding an entire layer of complexity to the understanding of olfactory phenotypes. The gene list provided here encompasses sufficient organellar, cellular, and anatomical sophistication to serve as a useful reference for numerous future studies.

The category of olfactory auxiliary genes is currently not well annotated in existing gene databases. For example, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway of Olfactory Transduction (hsa04740) contains only 30 olfactory transduction genes, whereas we report 108 genes. The Gene Ontology (GO) term Sensory Perception of Smell (GO: 0007608) contains chiefly ORs. Other GO terms such as Olfactory Nerve Formation (GO: 0021628) and Olfactory Bulb Axon Guidance (GO:0071678) are empty, although here we report of multiple genes under these categories.

A major aim of the presently portrayed list of human auxiliary olfactory genes and their variations is to help decipher the genetic basis of two nonsyndromic olfactory phenotypes, CGA and GOS. These two afflictions likely represent two different genetic models: monogenic inheritance for CGA, where a single, rare, and deleterious mutation (not necessarily the same in all affected families) causes the extreme sensory phenotype; and polygenic inheritance for GOS, with etiology stemming from several frequent genetic variants with additive genetic influences. In parallel, these results will assist in providing novel candidate genes for syndromic anosmia, either in monogenic or polygenic form, thus allowing the study of a plethora of undeciphered Mendelian and complex diseases associated with olfactory disorders. Finally, tools might become available to decipher the genetics of a specific GOS manifestation, the enhanced overall olfactory performance of females relative to males [Doty and Cameron, 2009].

Thus, the emerging knowledge about the ensemble of auxiliary olfactory genes and genetic variation within them may assist in rationalizing the interindividual phenotype variation in human GOS, and contribute to the development of personalized genomics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Contract grant sponsor: NIH (grants DC04657 and DC006070 to D.R. and A.O.); NIA (grant K23 AG036762 to J.P.); Crown Human Genome Center; ISF Legacy grant.

Thanks to Marilyn Safran and Tsippi Iny Stein from GeneCards for their help. Mouse single olfactory sensory neuron RNA-seq analysis was contributed by Hiroaki Matsunami. Mouse olfactory epithelium and bulb samples were contributed by Diego Restrepo. Human olfactory epithelium samples were contributed by Yoav Gilad and Jayant Pinto.

References

- Ahdab-Barmada M, Claassen D. A distinctive triad of malformations of the central nervous system in the Meckel–Gruber syndrome. *J Neuropathol Exp Neurol*. 1990; 49:610–620. [PubMed: 2230839]
- Amoore JE, Steinle S. A graphic history of specific anosmia. *Chem Senses*. 1991; 3:331–351.
- Atanasova B, Graux J, El Hage W, Hommet C, Camus V, Belzung C. Olfaction: a potential cognitive marker of psychiatric disorders. *Neurosci Biobehav Rev*. 2008; 32:1315–1325. [PubMed: 18555528]
- Belluscio L, Gold GH, Nemes A, Axel R. Mice deficient in G(olf) are anosmic. *Neuron*. 1998; 20:69–81. [PubMed: 9459443]
- Bergman JE, Bocca G, Hoefsloot LH, Meiners LC, van Ravenswaaij-Arts CM. Anosmia predicts hypogonadotropic hypogonadism in CHARGE syndrome. *J Pediatr*. 2010; 158:474–479. [PubMed: 20884005]
- Bozza T, Vassalli A, Fuss S, Zhang JJ, Weiland B, Pacifico R, Feinstein P, Mombaerts P. Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. *Neuron*. 2009; 61:220–233. [PubMed: 19186165]
- Brown KS, Maclean CM, Robinette RR. The distribution of the sensitivity to chemical odors in man. *Hum Biol*. 1968; 40:456–472. [PubMed: 5715213]
- Brunet LJ, Gold GH, Ngai J. General anosmia caused by a targeted disruption of the mouse olfactory cyclic nucleotide-gated cation channel. *Neuron*. 1996; 17:681–693. [PubMed: 8893025]
- Buck L, Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*. 1991; 65:175–187. [PubMed: 1840504]
- Cain WS, Gent JF. Olfactory sensitivity: reliability, generality, and association with aging. *J Exp Psychol Hum Percept Perform*. 1991; 17:382–391. [PubMed: 1830082]
- Cygnar KD, Zhao H. Phosphodiesterase 1C is dispensable for rapid response termination of olfactory sensory neurons. *Nat Neurosci*. 2009; 12:454–462. [PubMed: 19305400]
- Dattani MT. The candidate gene approach to the diagnosis of monogenic disorders. *Horm Res*. 2009; 71(Suppl 2):14–21. [PubMed: 19407492]
- Doty RL. The olfactory system and its disorders. *Semin Neurol*. 2009; 29:74–81. [PubMed: 19214935]
- Doty RL, Cameron EL. Sex differences and reproductive hormone influences on human odor perception. *Physiol Behav*. 2009; 97:213–228. [PubMed: 19272398]
- Doty RL, Fernandez AD, Levine MA, Moses A, McKeown DA. Olfactory dysfunction in type I pseudohypoparathyroidism: dissociation from Gs alpha protein deficiency. *J Clin Endocrinol Metab*. 1997; 82:247–250. [PubMed: 8989268]
- Dudova I, Vodicka J, Havlovicova M, Sedlacek Z, Urbanek T, Hrdlicka M. Odor detection threshold, but not odor identification, is impaired in children with autism. *Eur Child Adolesc Psychiatr*. 2011; 20:333–340.

- Fath MA, Mullins RF, Searby C, Nishimura DY, Wei J, Rahmouni K, Davis RE, Tayeh MK, Andrews M, Yang B, Sigmund CD, Stone EM, et al. Mks-null mice have a phenotype resembling Bardet-Biedl syndrome. *Hum Mol Genet.* 2005; 14:1109–1118. [PubMed: 15772095]
- Feldmesser E, Bercovich D, Avidan N, Halbertal S, Haim L, Gross-Isseroff R, Goshen S, Lancet D. Mutations in olfactory signal transduction genes are not a major cause of human congenital general anosmia. *Chem Senses.* 2007; 32:21–30. [PubMed: 17008420]
- Finkel D, Pedersen NL, Larsson M. Olfactory functioning and cognitive abilities: a twin study. *J Gerontol B Psychol Sci Soc Sci.* 2001; 56:226–233.
- Gaines AD. Anosmia and hyposmia. *Allergy Asthma Proc.* 2010; 31:185–189. [PubMed: 20615320]
- Gerber JC, Neuhaus TM, Tyshchenko N, Smitka M, Hackmann K. Expanding the clinical and neuroradiological phenotype of 6q27 microdeletion: olfactory bulb aplasia and anosmia. *Am J Med Genet A.* 2011; 155A:1981–1986. [PubMed: 21744487]
- Ghadami M, Majidzadeh AK, Morovvati S, Damavandi E, Nishimura G, Komatsu K, Kinoshita A, Najafi MT, Niikawa N, Yoshiura K. Isolated congenital anosmia with morphologically normal olfactory bulb in two Iranian families: a new clinical entity? *Am J Med Genet A.* 2004; 127A:307–309. [PubMed: 15150785]
- Gibberd FB, Feher MD, Sidey MC, Wierzbicki AS. Smell testing: an additional tool for identification of adult Refsum's disease. *J Neurol Neurosurg Psychiatry.* 2004; 75:1334–1336. [PubMed: 15314127]
- Glaser O. Hereditary deficiencies in the sense of smell. *Science.* 1918; 48:647–648. [PubMed: 17781959]
- Griff IC, Reed RR. The genetic basis for specific anosmia to isovaleric acid in the mouse. *Cell.* 1995; 83:407–414. [PubMed: 8521470]
- Gross-Isseroff R, Ophir D, Bartana A, Voet H, Lancet D. Evidence for genetic determination in human twins of olfactory thresholds for a standard odorant. *Neurosci Lett.* 1992; 141:115–118. [PubMed: 1508392]
- Guthoff M, Tschritter O, Berg D, Liepelt I, Schulte C, Machicao F, Haering HU, Fritsche A. Effect of genetic variation in Kv1.3 on olfactory function. *Diabetes Metab Res Rev.* 2009; 25:523–527. [PubMed: 19489042]
- Henkin RI, Bartter FC. Studies on olfactory thresholds in normal man and in patients with adrenal cortical insufficiency: the role of adrenal cortical steroids and of serum sodium concentration. *J Clin Invest.* 1966; 45:1631–1639. [PubMed: 5925519]
- Ho-Pun-Cheung A, Assenat E, Bascoul-Mollevi C, Bibeau F, Boissiere-Michot F, Thezenas S, Cellier D, Azria D, Rouanet P, Senesse P, et al. A large-scale candidate gene approach identifies SNPs in SOD2 and IL13 as predictive markers of response to preoperative chemoradiation in rectal cancer. *Pharmacogenomics J.* 2010; 11:437–443. [PubMed: 20644561]
- Imai T, Sakano H. Axon-axon interactions in neuronal circuit assembly: lessons from olfactory map formation. *Eur J Neurosci.* 2011; 34:1647–1654. [PubMed: 22103421]
- Jenkins PM, McEwen DP, Martens JR. Olfactory cilia: linking sensory cilia function and human disease. *Chem Senses.* 2009; 34:451–464. [PubMed: 19406873]
- Karstensen H, Tommerup N. Isolated and syndromic forms of congenital anosmia. *Clin Genet.* 2011
- Keller A, Zhuang H, Chi Q, Vosshall LB, Matsunami H. Genetic variation in a human odorant receptor alters odour perception. *Nature.* 2007; 449:468–472. [PubMed: 17873857]
- Kleene SJ. The electrochemical basis of odor transduction in vertebrate olfactory cilia. *Chem Senses.* 2008; 33:839–859. [PubMed: 18703537]
- Kobayakawa K, Kobayakawa R, Matsumoto H, Oka Y, Imai T, Ikawa M, Okabe M, Ikeda T, Itoharu S, Kikusui T, et al. Innate versus learned odour processing in the mouse olfactory bulb. *Nature.* 2007; 450:503–508. [PubMed: 17989651]
- Kulaga HM, Leitch CC, Eichers ER, Badano JL, Lesemann A, Hoskins BE, Lupski JR, Beales PL, Reed RR, Katsanis N. Loss of BBS proteins causes anosmia in humans and defects in olfactory cilia structure and function in the mouse. *Nat Genet.* 2004; 36:994–998. [PubMed: 15322545]
- Lagendijk AK, Smith KA, Bakkers J. Genetics of congenital heart defects: a candidate gene approach. *Trends Cardiovasc Med.* 2011; 20:124–128. [PubMed: 21335282]

- Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009; 10:R25. [PubMed: 19261174]
- Layman WS, McEwen DP, Beyer LA, Lalani SR, Fernbach SD, Oh E, Swaroop A, Hegg CC, Raphael Y, Martens JR, et al. Defects in neural stem cell proliferation and olfaction in *Chd7* deficient mice indicate a mechanism for hyposmia in human CHARGE syndrome. *Hum Mol Genet.* 2009; 18:1909–1923. [PubMed: 19279158]
- Lazic SE, Goodman AO, Grote HE, Blakemore C, Morton AJ, Hannan AJ, van Dellen A, Barker RA. Olfactory abnormalities in Huntington's disease: decreased plasticity in the primary olfactory cortex of R6/1 transgenic mice and reduced olfactory discrimination in patients. *Brain Res.* 2007; 1151:219–226. [PubMed: 17400200]
- Lewkowicz-Shpuntoff HM, Hughes VA, Plummer L, Au MG, Doty RL, Seminara SB, Chan YM, Pitteloud N, Crowley WF Jr, Balasubramanian R. Olfactory phenotypic spectrum in idiopathic hypogonadotropic hypogonadism: pathophysiological and genetic implications. *J Clin Endocrinol Metab.* 2012; 97:E136–E144. [PubMed: 22072740]
- Lin W, Margolskee R, Donnert G, Hell SW, Restrepo D. Olfactory neurons expressing transient receptor potential channel M5 (TRPM5) are involved in sensing semiochemicals. *Proc Natl Acad Sci USA.* 2007; 104:2471–2476. [PubMed: 17267604]
- Lu S, Das P, Fadool DA, Kaczmarek LK. The slack sodium-activated potassium channel provides a major outward current in olfactory neurons of *Kv1.3*–/– super-smeller mice. *J Neurophysiol.* 2010; 103:3311–3319. [PubMed: 20393063]
- Lygonis CS. Familiar absence of olfaction. *Hereditas.* 1969; 61:413–416. [PubMed: 5399208]
- MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science.* 2012; 335:823–828. [PubMed: 22344438]
- Mayer U, Kuller A, Daiber PC, Neudorf I, Warnken U, Schnolzer M, Frings S, Mohrlen F. The proteome of rat olfactory sensory cilia. *Proteomics.* 2009; 9:322–334. [PubMed: 19086097]
- McEwen DP, Jenkins PM, Martens JR. Olfactory cilia: our direct neuronal connection to the external world. *Curr Top Dev Biol.* 2008; 85:333–370. [PubMed: 19147011]
- McEwen DP, Koenekoop RK, Khanna H, Jenkins PM, Lopez I, Swaroop A, Martens JR. Hypomorphic CEP290/NPHP6 mutations result in anosmia caused by the selective loss of G proteins in cilia of olfactory sensory neurons. *Proc Natl Acad Sci USA.* 2007; 104:15917–15922. [PubMed: 17898177]
- Menashe I, Abaffy T, Hasin Y, Goshen S, Yahalom V, Luetje CW, Lancet D. Genetic elucidation of human hyperosmia to isovaleric acid. *PLoS Biol.* 2007; 5:e284. [PubMed: 17973576]
- Menashe I, Man O, Lancet D, Gilad Y. Different noses for different people. *Nat Genet.* 2003; 34:143–144. [PubMed: 12730696]
- Mirabello L, Yu K, Berndt SI, Burdett L, Wang Z, Chowdhury S, Teshome K, Uzoka A, Hutchinson A, Grotmol T, et al. A comprehensive candidate gene approach identifies genetic variation associated with osteosarcoma. *BMC Cancer.* 2011; 11:209. [PubMed: 21619704]
- Mori K, Sakano H. How is the olfactory map formed and interpreted in the mammalian brain? *Annu Rev Neurosci.* 2011; 34:467–499. [PubMed: 21469960]
- Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S, Marra M. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques.* 2008; 45:81–94. [PubMed: 18611170]
- Morley JF, Duda JE. Olfaction as a biomarker in Parkinson's disease. *Biomark Med.* 2010; 4:661–670. [PubMed: 20945979]
- Morrison, EE.; Moran, DT. Anatomy and ultrastructure of the human olfactory neuroepithelium. In: Doty, RL., editor. *Handbook of olfaction and gustation*. New York: Marcel Dekker; 1995. p. 75–101.
- Nakashima T, Kimmelman CP, Snow JB Jr. Structure of human fetal and adult olfactory neuroepithelium. *Arch Otolaryngol.* 1984; 110:641–646. [PubMed: 6477257]
- Nara K, Saraiva LR, Ye X, Buck LB. A large-scale analysis of odor coding in the olfactory epithelium. *J Neurosci.* 2011; 31:9179–9191. [PubMed: 21697369]

- Pause BM, Ferstl R, Fehm-Wolfsdorf G. Personality and olfactory sensitivity. *J Res Personality*. 1998; 32:510–518.
- Pifferi, S.; Menini, A.; Kurahashi, T. Signal transduction in vertebrate olfactory cilia. In: Menini, A., editor. *The neurobiology of olfaction*. Boca Raton, FL: CRC Press; 2011.
- Pinto JM, Thanaviratananich S, Hayes MG, Naclerio RM, Ober C. A genome-wide screen for hyposmia susceptibility Loci. *Chem Senses*. 2008; 33:319–329. [PubMed: 18245794]
- Pluznick JL, Rodriguez-Gil DJ, Hull M, Mistry K, Gattone V, Johnson CA, Weatherbee S, Greer CA, Caplan MJ. Renal cystic disease proteins play critical roles in the organization of the olfactory epithelium. *PLoS One*. 2011; 6:e19694. [PubMed: 21614130]
- Punter PH. Measurement of human olfactory thresholds for several groups of structurally related compounds. *Chem Senses*. 1983; 7:215–235.
- Qu Q, Liu J, Ni D, Zhang Q, Yang D, Wang N, Wu X, Han H. Diagnosis and clinical characteristics of congenital anosmia: case series report. *J Otolaryngol Head Neck Surg*. 2010; 39:723–731. [PubMed: 21144370]
- Ribeiro M, Machado A, Soares-Fernandes J. Septo-optic dysplasia with olfactory tract hypoplasia. *J Pediatr Neurosci*. 2009; 4:49. [PubMed: 21887179]
- Saito H, Chi Q, Zhuang H, Matsunami H, Mainland JD. Odor coding by a mammalian receptor repertoire. *Sci Signal*. 2009; 2:ra9. [PubMed: 19261596]
- Saito H, Kubota M, Roberts RW, Chi Q, Matsunami H. RTP family members induce functional expression of mammalian odorant receptors. *Cell*. 2004; 119:679–691. [PubMed: 15550249]
- Sammata N, Yu TT, Bose SC, McClintock TS. Mouse olfactory sensory neurons express 10,000 genes. *J Comp Neurol*. 2007; 502:1138–1156. [PubMed: 17444493]
- Sanka M, Tangsinmankong N, Loscalzo M, Sleasman JW, Dorsey MJ. Complete DiGeorge syndrome associated with CHD7 mutation. *J Allergy Clin Immunol*. 2007; 120:952–954. [PubMed: 17931565]
- Schild D, Restrepo D. Transduction mechanisms in vertebrate olfactory receptor cells. *Physiol Rev*. 1998; 78:429–466. [PubMed: 9562035]
- Segal NL, Topolski TD, Wilson SM, Brown KW, Araki L. Twin analysis of odor identification and perception. *Physiol Behav*. 1995; 57:605–609. [PubMed: 7538679]
- Singh N, Grewal MS, Austin JH. Familial anosmia. *Arch Neurol*. 1970; 22:40–44. [PubMed: 5409599]
- Sisodiya SM, Free SL, Williamson KA, Mitchell TN, Willis C, Stevens JM, Kendall BE, Shorvon SD, Hanson IM, Moore AT, et al. PAX6 haploinsufficiency causes cerebral malformation and olfactory dysfunction in humans. *Nat Genet*. 2001; 28:214–216. [PubMed: 11431688]
- Sobin C, Kiley-Brabeck K, Dale K, Monk SH, Khuri J, Karayiorgou M. Olfactory disorder in children with 22q11 deletion syndrome. *Pediatrics*. 2006; 118:e697–e703. [PubMed: 16908619]
- Stevens JC, Cain WS, Burke R. Variability of olfactory thresholds. *Chem Senses*. 1988; 13:643–653.
- Stevens JC, Dadarwala AD. Variability of olfactory threshold and its role in assessment of aging. *Percept Psychophys*. 1993; 54:296–302. [PubMed: 8414888]
- Touhara K, Vosshall LB. Sensing odorants and pheromones with chemosensory receptors. *Annu Rev Physiol*. 2009; 71:307–332. [PubMed: 19575682]
- Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009; 25:1105–1111. [PubMed: 19289445]
- Valkov IM, Dokumov SI, Genkova PI, Dimov DS. Olfactory, auditory, and gustatory function in patients with bonadal dysgenesis. *Obstet Gynecol*. 1975; 46:417–418. [PubMed: 1165876]
- Vihinen M, den Dunnen JT, Dagleish R, Cotton RG. Guidelines for establishing locus specific databases. *Hum Mutat*. 2011; 33:298–305. [PubMed: 22052659]
- Walz A, Mombaerts P, Greer CA, Treloar HB. Disrupted compartmental organization of axons and dendrites within olfactory glomeruli of mice deficient in the olfactory cell adhesion molecule, OCAM. *Mol Cell Neurosci*. 2006; 32:1–14. [PubMed: 16531066]
- Weiss J, Pyrski M, Jacobi E, Bufo B, Willnecker V, Schick B, Zizzari P, Gossage SJ, Greer CA, Leinders-Zufall T, et al. Loss-of-function mutations in sodium channel Nav1.7 cause anosmia. *Nature*. 2011; 472:186–190. [PubMed: 21441906]

- Wesson DW, Wilson DA, Nixon RA. Should olfactory dysfunction be used as a biomarker of Alzheimer's disease? *Expert Rev Neurother*. 2010; 10:633–635. [PubMed: 20420483]
- Whissell-Buechy D, Amoore JE. Odour-blindness to musk: simple recessive inheritance. *Nature*. 1973; 242:271–273. [PubMed: 4696259]
- Witt M, Bormann K, Gudziol V, Pehlke K, Barth K, Minovi A, Hahner A, Reichmann H, Hummel T. Biopsies of olfactory epithelium in patients with Parkinson's disease. *Mov Disord*. 2009; 24:906–914. [PubMed: 19205070]
- Wong ST, Trinh K, Hacker B, Chan GC, Lowe G, Gaggar A, Xia Z, Gold GH, Storm DR. Disruption of the type III adenylyl cyclase gene leads to peripheral and behavioral anosmia in transgenic mice. *Neuron*. 2000; 27:487–497. [PubMed: 11055432]
- Wysocki CJ, Beauchamp GK. Ability to smell androstenone is genetically determined. *Proc Natl Acad Sci USA*. 1984; 81:4899–4902. [PubMed: 6589634]
- Wysocki CJ, Gilbert AN. National Geographic Smell Survey. Effects of age are heterogenous. *Ann N Y Acad Sci*. 1989; 561:12–28. [PubMed: 2735671]

Figure 1.

A: The 30 top DS6 annotated genes. These genes are differentially expressed in mouse olfactory epithelium, and also annotated in the olfactory context by other sources (DS1, 2, 3, 11). **B:** The 30 top DS6 novel genes. These genes are differentially expressed in mouse olfactory epithelium, but are not annotated in the olfactory context by DS1, 2, 3, 11. Genes implicated in genetic diseases include ASAH2 (MIM# 611202), associated with Alzheimer's disease and Niemann–Pick disease, SULT6B1 associated with schizophrenia, AMBN (MIM# 601259), associated with amelogenesis imperfecta, and AMLEX (MIM# 300391) associated with amelogenesis imperfecta as well.

Figure 2.

The relative contribution of the 3 data sources DS1-3 to the long list of 1,680 auxiliary olfactory genes. In parentheses: the number of genes that are also found differentially expressed in olfactory tissues via DS5-9.

Figure 3.
Scores for short list genes.

Figure 4. Histogram of minor allele frequency for different variation types. Deleterious missense include SNPs probably damaging according to PolyPhen-2 (Score > 0.95).

Table 1

Data Sources

ID	Data source	Experimental method	Source	Total Genes	Long list	Medium list	Short list	Score
DS1	In vitro	functional studies, heterologous expression, inhibitors, in situ hybridization	Pubmed	181	181	181	85	1.0
DS2	Knockout	gene knockout in mouse models	MGI	88	88	88	74	1.0
DS3	Disease	identification of genes involved in human disease with olfactory phenotype	Pubmed OMIM	60	60	60	28	1.0
DS4	Ciliome	proteome analysis of rat OSN cilia	[Mayer et al., 2009]	377	329	58	24	0–1.0
DS5	OSN microarray	genes identified in mouse OSN	[Sammata et al., 2007]	11,596	250	66	17	0–1.0
DS6	OE NGS mouse	RNA-seq of mouse OE	our experimental data	16,647	250	141	57	0–1.0
DS7	OSN NGS	RNA-seq of mouse single OSNs	our experimental data	1,000	250	85	32	0–1.0
DS8	OE NGS human	RNA-seq of human OE	our experimental data	16,647	250	49	16	0–1.0
DS9	Bulb NGS	RNA-seq of mouse olfactory bulb	our experimental data	16,647	250	86	29	0–1.0
DS10	Linkage	Whole genome linkage analysis of anosmic and hyposmic individuals	[Gerber et al., 2011; Ghadami et al., 2004; Pinto et al., 2008; Sobin et al., 2006]	596	56	56	8	0.5
DS11	GeneCards	Computerized curation of genes, excluding genes present in DS1	GeneCards	260	175	36	8	0.5
Total					1,680	450	136	

MGI, Mouse Genome Informatics; NGS, next generation sequencing; OE, olfactory epithelium; OMIM, Online Mendelian Inheritance in Man; OSN, olfactory sensory neuron; RNA-seq, next-generation transcriptome sequencing.

Table 2**Genetic Diseases with Olfactory Phenotype**

#	Disease	Olfactory impairment	Associated genes	Reference
1	Kallmann syndrome	Anosmia, hyposmia	<i>KAL1</i> [*] , <i>AKAP2</i> , <i>FGF8</i> ^{*†} , <i>FGFR1</i> ^{*†} , <i>KISS1R</i> , <i>PROK2</i> ^{*†} , <i>PROKR2</i> ^{*†} , <i>NELF</i> [*] , <i>CHD7</i> [*] , <i>TAC3</i> , <i>TACR3</i>	[Lewkowitz-Shpuntoff et al., 2012]
2	Bardet–Biedl syndrome (BBS)	Anosmia, hyposmia	<i>BBS1</i> ^{*†} , <i>BBS2</i> [†] , <i>ARL6</i> , <i>BBS4</i> ^{*†} , <i>BBS5</i> , <i>MKKS</i> [†] , <i>BBS7</i> , <i>TTC8</i> [†] , <i>BBS9</i> , <i>BBS10</i> , <i>TRIM32</i> , <i>BBS12</i>	[Kulaga et al., 2004]
3	Meckel–Gruber syndrome (MKS)	Anosmia	<i>MKS1</i> [†] , <i>MKS2</i> , <i>TMEM67</i> [†] , <i>CEP290</i> ^{*†} , <i>BBS1</i> ^{*†} , <i>BBS4</i> ^{*†} , <i>NPHP3</i>	[Ahdab-Barmada and Claassen, 1990]
4	Polycystic kidney disease (PKD)	Anosmia	<i>PKD1</i> [†] , <i>PKD2</i> [†] , <i>NPHP3</i>	[Pluznick et al., 2011]
5	Leber congenital amaurosis (LCA)	Anosmia	<i>CEP290</i> ^{*†} , <i>GUCY2D</i> , <i>RPE65</i> , <i>SPATA7</i> , <i>AIPL1</i> , <i>LCA5</i> , <i>RPGRIP1</i> [†] , <i>CRX</i> , <i>CRB1</i> , <i>LCA9</i> , <i>IMPDH1</i> , <i>RD3</i> , <i>RDH12</i> , <i>LRAT</i> , <i>TULP1</i> , <i>LCA3</i>	[McEwen et al., 2007]
6	Refsum disease (RD)	Anosmia	<i>PEX7</i> [*] , <i>PHYH</i> [*]	[Gibberd et al., 2004]
7	McKusick-Kaufman syndrome (MKS)	Hyposmia	<i>MKKS</i> [†]	[Fath et al., 2005]
8	Aniridia	Anosmia, hyposmia	<i>PAX6</i> ^{*†}	[Sisodiya et al., 2001]
9	Congenital analgesia	Anosmia	<i>SCN9A</i> ^{*†}	[Weiss et al., 2011]
10	CHARGE syndrome	Anosmia, hyposmia	<i>CHD7</i> [*] , <i>SEMA3E</i>	[Bergman et al., 2010]
11	Septo-optic dysplasia (SOD)	Anosmia	<i>HESX1</i> ^{*†}	[Ribeiro et al., 2009]
12	Pseudohypoparathyroidism (PHP) 1a	Impaired discrimination	<i>GNAS</i> [†]	[Doty et al., 1997]
13	Huntington's disease	Impaired discrimination	<i>HTT</i> ^{*†}	[Lazic et al., 2007]
14	DiGeorge syndrome (Chromosome 22q11 deletion)	Hyposmia	Chromosome 22q11 deletion [*] , <i>CHD7</i> [*]	[Sanka et al., 2007; Sobin et al., 2006]
15	6q27 microdeletion	Anosmia	Chromosome 6q27 deletion [*]	[Gerber et al., 2011]
16	Turner syndrome (45XO karyotype)	Hyposmia	Chromosome X deletion [*]	[Valkov et al., 1975]

Some genes are associated with more than one disease;

* indicates that olfactory performance was tested in humans with known genotype;

† knockout.

Diseases #2–7 are ciliopathies. MKS is the acronym for two diseases (#3,7).