

METHOD OF IDENTIFYING MALODOR MODULATING COMPOUNDS

FIELD

The technical field is directed to odorant and aroma receptors and assays that can
5 be used to identify odorant and/or aroma compounds more specifically inhibitors,
modulators or counteractants of malodor compounds such as dimethyl trisulfide (DMTS).

BACKGROUND

Olfaction is one of the most complex and poorly understood of human sensory
10 systems. From olfactory receptor (OR) activation to perception, there are many steps that
still require further investigation. If we can understand how the OR code for individual
odorants and mixtures translates into perception then we can exploit this knowledge to
bring significant benefit in several areas. These areas include odor modulators like
15 malodor counteractants that block the perception of unpleasant odors, new flavor and
fragrance ingredients that replace non-biodegradable or toxic compounds, and odor
enhancers that would limit our reliance on difficult to source compounds from natural
sources. The 'olfactory code' combinatorial paradigm is centered on the observation that
any single OR may be activated by multiple odorants, and conversely most odorants are
20 capable of activating several ORs. In the mouse genome there are ~1,200 distinct intact
ORs. Humans, by contrast, have ~400. In both cases, the repertoire of ORs is activated
by many thousands of odorants in the world, and it is this combinatorial complexity that
allows for the breadth of olfactory sensations we can perceive. However, odorants or
ligands for only 82 mouse (~8%) and 17 human ORs (~10%) have been identified as of
25 2014 using traditional deorphanization methods. In addition, the physiological relevance
of most ligands for the human ORs, essentially identified *in vitro*, has not been tested.

A method that can rapidly and reliably identify a relatively small subset of ORs,
within the entire repertoire of ORs that exist in an organism that are specifically activated
or inhibited by one or more odorants is described in WO2014/210585. However, using
this method, there remains a need to identify odorant receptors and more particularly
30 malodor receptors that are activated by particular malodor-causing substances.

Malodor-causing compounds such as dimethyl trisulfide (DMTS) and other
closely related polysulfide compounds such as dimethyl disulfide (DMDS) can generate

unpleasant odors that arise, for example, from latrines and other “bathroom” sources that contain fecal matter or from bad breath. Hence, malodor modulators or counteractants that bind, suppress, block, inhibit, and/or modulate the activity of one or more olfactory receptor that is activated by a particular malodor-causing substance such as DMTS or DMDS are desirable. Assays that rely on such new malodor receptors or on malodor receptors newly identified as associated with particular malodors-causing substances to identify new compounds or compounds mixtures that bind to these receptors are further desired.

10 SUMMARY

Provided herein is a non-human host organism or a host cell that has been modified to express a receptor that is activated by DMTS wherein the receptor is selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1.

Further provided is a non-human host organism or a host cell transformed to express a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

Further provided herein is an expression vector comprising a nucleic acid that encodes a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

Also provided herein is an expression vector having a nucleic acid wherein the nucleic acid comprises a nucleotide sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ

ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 39, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, or the reverse complement thereof.

5 Provided herein is a nucleic acid comprising a nucleic acid sequence as recited above, and also provided is a polypeptide comprising an amino acid sequence as recited above.

Further provided herein is a method for identifying a compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of at least one olfactory receptor that is activated by a malodor-causing substance such as dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS), wherein the receptor is a polypeptide comprising
10 an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40, wherein the method comprises:

- a) contacting the receptor, or a chimera or fragment thereof with a compound;
- b) determining whether the compound has an effect on the activity of the receptor.

Also provided is a method for identifying a compound that binds, suppresses, blocks,
20 inhibits, and/or modulates the activity of at least one olfactory receptor that is activated by a malodor-causing substance comprising:

- a. contacting a test substance and a malodor-causing substance to at least one olfactory receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669,
25 Olfr1211, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1;
- b. measuring the response of the olfactory receptor to the malodor-causing substance by measuring the response of the olfactory receptor in the presence and absence of the test substance;
- 30 c. identifying a test substance that modulates the response of the olfactory receptor on the basis of the response that was measured in the presence and absence of the test substance; and

d. selecting the identified test substance as a compound that modulates the response of the olfactory receptor to the malodor-causing substance, wherein the malodor-causing substance is dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS).

5 Further provided is a method for identifying a malodor modulator comprising:

- a. contacting a test substance and a malodor-causing substance to at least one olfactory receptor, wherein the receptor comprises a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40;
- 10 b. measuring the response of the olfactory receptor polypeptide to the malodor-causing substance;
- c. identifying, based on the measured response, a test substance that can suppress the response of the olfactory receptor; and
- d. selecting, as a malodor modulator, the test substance that binds, suppresses, blocks, inhibits, and/or modulates the response of the olfactory receptor,
- 15 wherein the malodor-causing substance is dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS).

25 Additionally provided is a recombinant nucleic acid molecule comprising: a nucleic acid comprising at least one of a Lucy tag, a FLAG[®] tag, and/or a Rho tag and a nucleic acid encoding a receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR4S2, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1 or the reverse complement thereof.

30 Still yet further provided is a non-human host organism or a host cell that is recombinantly modified to express a nucleic acid or a polypeptide as described above.

DETAILED DESCRIPTION OF THE DRAWINGS

Figure 1 shows the pairwise amino acid identity levels between the receptors mentioned herein.

Figure 2 shows DMTS dose response curves of mouse odorant receptors Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738 and Olfr742, and of human odorant receptors OR4S2 and OR52N5.

Figure 3 shows DMTS dose response curves of human odorant receptors OR2L13, OR4C15, OR8H3, OR11G2 and OR52N2.

Figure 4 shows a ranking of volatile compounds with specific inhibitory effects on DMTS receptor activity.

Figure 5 shows DMTS and DMDS dose response curves of mouse odorant receptors Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742 and of human odorant receptor OR4S2.

DETAILED DESCRIPTION

For the descriptions herein and the appended claims, the use of "or" means "and/or" unless stated otherwise. Similarly, "comprise," "comprises," "comprising," "include," "includes," and "including" are interchangeable and not intended to be limiting.

It is to be further understood that where descriptions of various embodiments use the term "comprising," those skilled in the art would understand that in some specific instances, an embodiment can be alternatively described using language "consisting essentially of" or "consisting of."

The following terms have the meanings ascribed to them unless specified otherwise.

"OR" refers to one or more members of a family of G protein-coupled receptors (GPCRs) that are expressed in olfactory cells. Olfactory receptor cells can also be identified on the basis of morphology or by the expression of proteins specifically expressed in olfactory cells. OR family members may have the ability to act as receptors for olfactory signal transduction.

"DMTS OR" or "DMDS OR" refers to a member of the family of G protein-coupled receptors that is expressed in an olfactory cell, which receptors bind and/or are

activated by DMTS or DMDS in a binding or activity assay for identifying ligands that bind and/or activate GPCRs. Such assays are described below. DMTS or DMDS receptors herein will include fragments, variants, including synthetic and naturally occurring, and chimeras or recombinant nucleic acids or proteins that respond to or bind
5 DMTS or DMDS.

“OR” polypeptides are considered as such if they pertain to the 7-transmembrane-domain G protein-coupled receptor superfamily encoded by a single ~1kb long exon and exhibit characteristic olfactory receptor-specific amino acid motifs. The seven domains are called “transmembrane” or “TM” domains TM I to TM VII connected by three
10 “internal cellular loop” or “IC” domains IC I to IC III, and three “external cellular loop” or “EC” domains EC I to EC III. The motifs and the variants thereof are defined as, but not restricted to, the MAYDRYVAIC motif (SEQ ID NO: 53) overlapping TM III and IC II, the FSTCSSH motif (SEQ ID NO: 54) overlapping IC III and TM VI, the PMLNPFIY motif (SEQ ID NO: 55) in TM VII as well as three conserved C residues in EC II, and the
15 presence of highly conserved GN residues in TM I [Zhang, X. & Firestein, S. Nat. Neurosci. 5, 124–133 (2002); Malnic, B., et al. Proc. Natl. Acad. Sci. U. S. A. 101, 2584–2589 (2004)].

“OR” nucleic acids encode a family of GPCRs with seven transmembrane regions that have “G protein-coupled receptor activity,” e.g., they may bind to G proteins in
20 response to extracellular stimuli and promote production of second messengers such as IP₃, cAMP, cGMP, and Ca²⁺ via stimulation of enzymes such as phospholipase C and adenylate cyclase.

“Paralogous” OR genes or “paralogs” are the result of gene duplications and refer to closely related homologous genes within the same species. “Orthologous” OR genes
25 or “orthologs” are defined as phylogenetically linked by a gene present in a common ancestor and refer to closely related homologous genes in other species.

The “N terminal domain” region starts at the N-terminus and extends to a region close to the start of the first transmembrane region. “Transmembrane regions” comprise the seven “transmembrane domains,” which refers to the domain of OR polypeptides that
30 lies within the plasma membrane, and may also include the corresponding cytoplasmic (intracellular) and extracellular loops. The seven transmembrane regions and extracellular and cytoplasmic loops can be identified using standard methods such as hydrophobicity profiles, or as described in Kyte & Doolittle, J. Mol. Biol., 157:105-32

(1982), or in Stryer. The general secondary and tertiary structure of transmembrane domains, in particular the seven transmembrane domains of G protein-coupled receptors such as olfactory receptors, are known in the art. Thus, primary structure sequence can be predicted based on known transmembrane domain sequences, as described in detail below. These transmembrane domains are useful for *in vitro* ligand-binding assays.

The phrase "functional effects" in the context of assays for testing compounds that modulate OR family member mediated olfactory transduction includes the determination of any parameter that is indirectly or directly under the influence of the receptor, e.g., functional, physical and chemical effects. It includes ligand binding, changes in ion flux, membrane potential, current flow, transcription, G protein binding, GPCR phosphorylation or dephosphorylation, signal transduction receptor-ligand interactions, second messenger concentrations (e.g., cAMP, cGMP IP3, or intracellular Ca. ²⁺), *in vitro*, *in vivo*, and *ex vivo* and also includes other physiologic effects such as increases or decreases of neurotransmitter or hormone release.

By "determining the functional effect" or "confirming the activity" in the context of assays is meant assays for a compound that increases or decreases a parameter that is indirectly or directly under the influence of an OR family member, e.g., functional, physical and chemical effects. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties, patch clamping, voltage-sensitive dyes, whole cell currents, radioisotope efflux, inducible markers, oocyte OR gene expression; tissue culture cell OR expression; transcriptional activation of OR genes or activity induced genes such as *egr-1* or *c-fos*; ligand-binding assays; voltage, membrane potential and conductance changes; ion flux assays; changes in intracellular second messengers such as cAMP, cGMP, and inositol triphosphate (IP3); changes in intracellular calcium levels; neurotransmitter release, and the like.

"Binder," "suppressors," "blockers," "inhibitors," "activators," and/or "modulators" of OR genes or proteins are used interchangeably to refer to binding, suppressing, blocking, suppressing, inhibitory, activating, or modulating molecules identified using *in vivo*, *in vitro* and *ex vivo* assays for olfactory transduction, e.g., ligands, agonists, antagonists, enhancers, and their homologs and mimetics. Inhibitors are compounds that, e.g., bind to, partially or totally block stimulation, decrease, suppress,

prevent, delay activation, inactivate, desensitize, or down regulate olfactory transduction, e.g., antagonists. Activators are compounds that, e.g., bind to, stimulate, increase, open activate, facilitate, enhance activation, sensitize, or up regulate olfactory transduction, e.g., agonists. Modulators include compounds that, e.g., alter the interaction of a receptor

5 with: extracellular proteins that bind activators or inhibitor (e.g., odorant-binding proteins, ebnerin and other members of the hydrophobic carrier family, or a member of the lipocalin family); G proteins; kinases (e.g., homologs of rhodopsin kinase and beta adrenergic receptor kinases that are involved in deactivation and desensitization of a receptor); and arrestins, which also deactivate and desensitize receptors. Modulators also

10 include compounds that alter the affinity or the transduction efficacy of an OR altering the effect of an activator on the OR. Modulators can include genetically modified versions of OR family members, e.g., with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing OR family members in cells

15 or cell membranes, applying putative modulator compounds, in the presence or absence of flavor, fragrance or malodour molecules, e.g. a malodour-causing substance such as DMTS, and then determining the functional effects on olfactory transduction, as described above. Samples or assays comprising OR family members that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the

20 inhibitor, activator, or modulator to examine the extent of modulation. Control samples (untreated with modulators) are assigned a relative OR activity value of 100%. Inhibition of an OR is achieved when the OR activity value relative to the control is about 80%, optionally 50% or 25-0%.

The terms "purified," "substantially purified," and "isolated" as used herein refer

25 to the state of being free of other, dissimilar compounds with which the compound of the invention is normally associated in its natural state, so that the "purified," "substantially purified," and "isolated" subject comprises at least 0.5%, 1%, 5%, 10%, or 20%, and most preferably at least 50% or 75% of the mass, by weight, of a given sample. In one preferred embodiment, these terms refer to the compound of the invention comprising at

30 least 95% of the mass, by weight, of a given sample. As used herein, the terms "purified," "substantially purified," and "isolated" "isolated," when referring to a nucleic acid or protein, of nucleic acids or proteins, also refers to a state of purification or concentration different than that which occurs naturally in the mammalian, especially human body. Any

degree of purification or concentration greater than that which occurs naturally in the mammalian, especially human body, including (1) the purification from other associated structures or compounds or (2) the association with structures or compounds to which it is not normally associated in the mammalian, especially human, body, are within the meaning of "isolated." The nucleic acid or protein or classes of nucleic acids or proteins, described herein, may be isolated, or otherwise associated with structures or compounds to which they are not normally associated in nature, according to a variety of methods and processes known to those of skill in the art.

As used herein, the terms "amplifying" and "amplification" refer to the use of any suitable amplification methodology for generating or detecting recombinant of naturally expressed nucleic acid, as described in detail, below. For example, the invention provides methods and reagents (e.g., specific degenerate oligonucleotide primer pairs, oligo dT primer) for amplifying (e.g., by polymerase chain reaction, PCR) naturally expressed (e.g., genomic DNA or mRNA) or recombinant (e.g., cDNA) nucleic acids of the invention *in vivo*, *ex vivo* or *in vitro*.

The term "7-transmembrane receptor" means a polypeptide belonging to a superfamily of transmembrane proteins that have seven domains that span the plasma membrane seven times (thus, the seven domains are called "transmembrane" or "TM" domains TM I to TM VII). The families of olfactory and certain taste receptors each belong to this super-family. 7-transmembrane receptor polypeptides have similar and characteristic primary, secondary and tertiary structures, as discussed in further detail below.

The term "nucleic acid" or "nucleic acid sequence" refers to a deoxy-ribonucleotide or ribonucleotide oligonucleotide in either single- or double-stranded form. The term encompasses nucleic acids, i.e., oligonucleotides, containing known analogs of natural nucleotides. The term also encompasses nucleic-acid-like structures with synthetic backbones. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating, e.g., sequences in which the third position of one or more selected codons is substituted with mixed-base and/or deoxyinosine residues.

In addition to the gene sequences shown in the sequences disclosed herein, it will

be apparent for the person skilled in the art that variants also include DNA sequence polymorphisms that may exist within a given population, which may lead to changes in the amino acid sequence of the polypeptides disclosed herein. Such genetic polymorphisms may exist in cells from different populations or within a population due to
5 natural allelic variation. Allelic variants may also include functional equivalents.

Further embodiments also relate to the molecules derived by such sequence polymorphisms from the concretely disclosed nucleic acids. These natural variations usually bring about a variance of about 1 to 5% in the nucleotide sequence of a gene or in the amino acid sequence of the polypeptides disclosed herein. As mentioned above, the
10 nucleic acid encoding the polypeptide of an embodiment herein is a useful tool to modify non-human host organisms or host cells intended to be used in the methods described herein.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues, whether comprising naturally occurring amino acids or polymers and non-naturally occurring amino acids or polymers. The term
15 "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region
20 from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is defined as an array of nucleic acid sequences that direct
25 transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under
30 most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second

nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

As used herein, "recombinant" refers to a polynucleotide synthesized or otherwise manipulated *in vitro* (e.g., "recombinant polynucleotide"), to methods of using recombinant polynucleotides to produce gene products in cells or other biological systems, or to a polypeptide ("recombinant protein") encoded by a recombinant polynucleotide. "Recombinant" means also the ligation of nucleic acids having various coding regions or domains or promoter sequences from different sources into an expression cassette or vector for expression of, e.g., inducible or constitutive expression of a fusion protein comprising a translocation domain of the invention and a nucleic acid sequence potentially amplified using a primer. "Recombinant" means also modifications obtained by genome editing techniques, such as CRISPR/Cas9, of a cell that leads to stable or transient expression of endogenous genes such as the receptor gene referred to herein.

The term "expression vector" refers to any recombinant expression system for the purpose of expressing a nucleic acid sequence of the invention *in vitro*, *ex vivo*, or *in vivo*, constitutively or inducibly, in any cell, including prokaryotic, yeast, fungal, plant, insect or mammalian cell. The term includes any linear or circular expression systems including but not limited to viral vectors, bacteriophages and plasmids. The skilled person is capable of selecting a suitable vector according to the expression system. The term includes expression systems that remain episomal or integrate into the host cell genome. The expression systems can have the ability to self-replicate or not, i.e., drive transient expression in a cell. The term includes recombinant "expression cassettes" which can contain the minimum elements needed for transcription of the recombinant nucleic acid. The term also covers cassettes or vectors for expression of endogenous genes through, for example, genome editing methods such as CRISPR/Cas9.

By "a non-human organism or a host cell" is meant a non-human organism or a cell that contains a nucleic acid as described herein or an expression vector and supports the replication or expression of the expression vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, HEK-293, and the like, e.g., cultured cells, explants, and cells *in vivo*.

By “tag” or “tag combination” is meant a short polypeptide sequence that can be added to the odorant receptor protein. Typically, the DNA encoding a “tag” or a “tag combination” is added to the DNA encoding the receptor, eventually resulting in a fusion protein where the “tag” or a “tag combination” is fused to the N-terminus or C-terminus of the receptor. Lucy, FLAG[®] and/or Rho tags can enhance the receptor trafficking to the cell membrane, hence they can assist in expression of a functional odorant receptor for *in vitro* cell based assay [Shepard, B. et al. PLoS One 8, e68758–e68758 (2013), and Zhuang, H. & Matsunami, H. J. Biol. Chem. 282, 15284–15293 (2007)].

In one embodiment provided herein is an isolated nucleic acid molecule comprising a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 39, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39 or the reverse complement thereof.

In one embodiment provided herein is an isolated nucleic acid sequence as described above which encodes a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

In a further embodiment provided herein is an isolated polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

In one embodiment, a non-human organism or a host cell is transformed to express a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ

ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

5 In one embodiment, a non-human organism or a host cell is transformed to express a polypeptide comprising an amino acid sequence that is identical to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

10 Further provided herein is an expression vector comprising a nucleic acid that encodes a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.

Also provided herein is an expression vector comprising a nucleic acid that comprises a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39 or the reverse complement thereof.

25 Also provided herein is an expression vector having a nucleic acid wherein the nucleic acid comprises a nucleotide sequence that is identical to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39 or the reverse complement thereof.

30 In one embodiment herein is a method for identifying a compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of an olfactory receptor that is activated by a malodor-causing substance such as dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS), wherein the receptor is a polypeptide comprising an amino acid

sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, 5 SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40, wherein the method comprises:

- a) contacting the receptor, or a chimera or fragment thereof with a compound;
- b) determining whether the compound has an effect on the activity of the 10 receptor.

In one embodiment is a method for identifying a compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of an olfactory receptor that is activated by a malodor-causing substance comprising:

- a. contacting a test substance and a malodor-causing substance to at least one 15 olfactory receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1;
- b. measuring the response of the olfactory receptor to the malodor-causing substance 20 by measuring the response of the olfactory receptor in the presence and absence of the test substance;
- c. identifying a test substance that modulates the response of the olfactory receptor on the basis of the response that was measured in the presence and absence of the test substance; and
- d. selecting the identified test substance as a compound that modulates the response 25 of the olfactory receptor to the malodor-causing substance,

wherein the malodor-causing substance is dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS).

30 In a further embodiment herein is a method for identifying a malodor modulator comprising:

- a. contacting a test substance and a malodor-causing substance to at least one olfactory receptor, wherein the receptor comprises a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%,

- 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40;
- 5 b. measuring the response of the olfactory receptor polypeptide to the malodor-causing substance;
- c. identifying, based on the measured response, a test substance that can modulate or suppress the response of the olfactory receptor; and
- 10 d. selecting, as a malodor modulator, the test substance that binds, suppresses, blocks, inhibits, and/or modulates the response of the olfactory receptor;
- wherein the malodor-causing substance is dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS).

One embodiment is a method for identifying a malodor inhibitor comprising:

- 15 a. contacting a test substance and a malodor-causing substance to at least one olfactory receptor, wherein the receptor comprises a polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40;
- 20 b. measuring the response of the olfactory receptor polypeptide to the malodor-causing substance;
- 25 c. identifying, based on the measured response, a test substance that can suppress the response of the olfactory receptor; and
- d. selecting, as a malodor inhibitor, the test substance that suppresses the response of the olfactory receptor,
- 30 wherein the malodor-causing substance is DMTS and the malodor is a fecal malodor or oral malodor.

One embodiment is a method of identifying a compound that putatively modulates DMTS associated malodor comprising: (i) contacting a cell line that expresses a DMTS

receptor polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, 5 SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40 with at least one compound; (ii) screening for compounds that bind, suppress, block, inhibit, and/or modulate the activity of said olfactory receptor polypeptide; and (iii) identifying a compound that putatively modulates DMTS associated malodor if it binds, suppresses, blocks, inhibits, and/or modulates the activity of said DMTS receptor 10 polypeptide.

In a further embodiment, the malodor-causing substance in the methods described herein is dimethyl trisulfide (DMTS).

Additionally provided is a recombinant nucleic acid molecule comprising

- a. a nucleic acid comprising a tag combination that comprises one or more of a Lucy, 15 FLAG[®], and/or Rho tag; and
- b. a nucleic acid encoding a receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR4S2, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1 or the complement thereof.

20 In a further embodiment the Lucy tag comprises SEQ ID NO: 47, the FLAG[®] tag comprises SEQ ID NO: 43, and the Rho tag comprises SEQ ID NO: 45.

Further provided is any one of a number of malodor modulating compounds that binds, suppresses, blocks, inhibits, and/or modulates the activity of an olfactory receptor that is activated by a malodor-causing substance such as dimethyl trisulfide (DMTS) and 25 that is identified by the methods disclosed herein.

In one embodiment herein is a malodor modulating compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of at least one olfactory receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR52N5, OR2L13, 30 OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1, and that is identified by the methods disclosed herein.

Another embodiment relates to the use of a polypeptide that is or can be activated by DMTS or DMDS comprising an amino acid sequence having at least 70%, 75%, 80%,

85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40 for identifying a malodor modulating compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of an olfactory receptor.

Still yet further provided is a cell that is recombinantly modified to express a polypeptide described above.

10 In one embodiment herein is a non-human host organism or a host cell that has been transformed or modified to express a receptor that is activated by DMTS selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR4S2, OR52N5, OR2L13, OR2AJ1 OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1.

15 In one embodiment herein is a non-human host organism or a host cell that has been transformed or modified to express a receptor that is activated by DMTS, wherein the receptor comprises a polypeptide as described herein or a polypeptide encoded by a nucleic acid described herein.

In a further embodiment herein is a non-human host organism or a host cell that comprises a nucleic acid or expression vector as described herein.

20 In one embodiment provided herein is a cell wherein the cell is a prokaryotic cell. In another embodiment the cell provided herein is a eukaryotic cell. In a particular embodiment, the cell provided herein is selected from a group consisting of a yeast cell and a plant cell. In a more particular embodiment provided herein the cell is selected from the group consisting of HEK293, CHO, Xenopus oocytes, COS, yeast, bacteria and cells derived from the olfactory placode.

25 In order to identify unknown DMTS or DMDS specific receptors, DMTS and DMDS can be used to screen dissociated olfactory sensory neurons (OSNs). DMTS and DMDS can be further used for cell-based dose-response experiments performed on specific DMTS or DMDS receptors to assess both specificity and sensitivity of the receptors.

In one aspect, provided herein are methods to identify mammalian odorant receptors for malodour modulating compounds and the use of the receptors for screening,

particularly for high throughput screening (HTS) of malodor modulators (e.g. that bind, suppress, block, inhibit and/or modulate the activity of an OR).

In particular provided herein are mouse receptors, for example, Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr669, Olfr665, Olfr207, 5 Olfr1211, and their human counterparts OR4S2, OR5T1, OR8H3, OR2L13, OR2AJ1, OR11G2, OR52N5, OR52N2, OR5AC2, and OR4C15, as receptors for the malodour-causing substances DMTS, DMDS or other polysulfide compounds, as shown in Table 1.

Mouse ortholog	Human ortholog	Amino acid identity
Olfr1193	OR4S2	90%
Olfr1093	OR5T1	77%
Olfr1097	OR8H3	70%
Olfr166	OR2L13	85%
Olfr169	OR2AJ1	72%
Olfr738	OR11G2	79%
Olfr742	OR11G2	74%
Olfr669	OR52N5	87%
Olfr665	OR52N2	77%
Olfr207	OR5AC2	64%
Olfr1211	OR4C15	83%

10 Table 1. Orthologous mouse and human odorant receptor pairs and their corresponding identity at the amino acid level.

While not wishing to be bound to any theory, mouse receptor Olfr738 is a paralog of receptor Olfr742 and human OR11G2 is the ortholog of these mouse receptors; and 15 mouse receptor Olfr669 is a paralog of receptor Olfr665 and an ortholog of human receptor OR52N5; and mouse receptor Olfr665 is an ortholog of human receptor OR52N2. Several receptors were previously identified as indole and/or skatole responsive odorant receptors in WO2014/210585. However, these receptors had not been previously associated with the malodour-causing substances DMTS or DMDS or other 20 related polysulfide compounds. Indole and skatole are not chemically related to

polysulfide compounds and therefore approaches based on chemical structure such as structure-activity-relationship (SAR) to identify additional ligands would not define DMTS as a potential agonist of these receptors.

In a further embodiment, indicators for monitoring the activity of olfactory
5 receptors are selected from a fluorescent calcium indicator dye, a calcium indicator protein, a fluorescent cAMP indicator, a cell mobilization assay, a cellular dynamic mass redistribution assay, a label-free cell based assay, a cAMP response element (CRE) mediated reporter protein, a biochemical cAMP HTRF assay, a beta-arrestin assay, or an electrophysiological recording. Particularly, a calcium indicator dye is selected that can
10 be used to monitor the activity of olfactory receptors expressed on the membrane of the olfactory neurons (e.g., Fura-2 AM).

In a particular embodiment, compounds are screened sequentially and the odorant-dependent changes in calcium dye fluorescence are measured using a fluorescent microscope or fluorescent-activated cell sorter (FACS).

15 In a further embodiment, molecular 3D receptor modeling of olfactory receptors is used to assess the binding potential *in silico* and to identify compounds that may activate, mimic, block, inhibit, modulate, and/or enhance the activity of an olfactory receptor.

As an example, olfactory neurons activated by DMTS or DMDS are isolated using either a glass microelectrode attached to a micromanipulator or a FACS machine. Mouse
20 olfactory sensory neurons are screened by Ca^{2+} imaging similar to procedures previously described [Malnic, B., et al. Cell 96, 713–723 (1999); Araneda, R. C. et al. J. Physiol. 555, 743–756 (2004); and WO2014/210585 hereby incorporated by reference in its entirety]. Particularly, a motorized movable microscope stage is used to increase the number of cells that can be screened to at least 1,500 per experiment. Since there are
25 approximately 1,200 different olfactory receptors in the mouse and each olfactory sensory neurons expresses only 1 of 1,200 olfactory receptor genes, this screening capacity will cover virtually the entire mouse odorant receptor repertoire. In other words, the combination of calcium imaging for high-throughput olfactory sensory neuron screening leads to the identification of nearly all of the odorant receptors that respond to a particular
30 profile of odorants. In a particular aspect, odorant receptors that respond to DMTS or DMDS can be isolated. For example, at least one neuron is isolated.

For calcium imaging of olfactory neurons, the main olfactory epithelium may be dissected from a mouse before neuronal dissociation. Dissected olfactory epithelium may

then be transferred to a dissociation buffer for mechanical and enzymatic dissociation. Dissociated neurons may then be seeded onto a coverslip allowing the screening of thousands of cells by fluorescence microscopy and the cells may be loaded with a calcium sensitive dye (Fura-2 AM) for example for about 30 minutes at 31 °C and transferred onto the microscope ready for screening. Cells are stimulated by perfusing diluted solutions of odorants (in physiological saline) over the dissociated olfactory neurons. The rare cells that respond to the malodor compound are identified by for example stimulating the receptors with 50 µm of the malodor compounds and then by monitoring the intracellular Ca²⁺ flux indicated by changes in Fura-2 fluorescence. After analysis, responding cells may be retrieved from a glass coverslip with a suction micropipette. Isolated cells are then pooled into one sample or treated individually for subsequent identification of the odorant receptor genes expressed as mRNA in the responding cells.

In a particular embodiment, the mRNAs of olfactory neurons are purified and amplified according to the method generally described in Marko, N. F., *et al.*, (2005) A robust method for the amplification of RNA in the sense orientation. *BMC genomics*, 6, 27; doi:10.1186/1471-2164-6-27 (Eberwine method). At least a portion of the transcriptome (up to including the entire transcriptome) is sequenced using Next-Generation Sequencing (NGS) technologies or hybridized to known genes using Microarray technologies. NGS is generally discussed and described in Metzker, M. L. Nat. Rev. Genet. 11, 31–46 (2010). In a particular embodiment, a minimum of 5 neurons presenting the same response profile are pooled. The mRNAs are released by cell lysis immediately after picking; no DNase and no purification steps are carried out. The mRNA are amplified by two consecutive rounds of *in vitro* transcription (IVT). The amplification may be done according to MesageAmpII aRNA kit (Ambion, AMA1751) with the following parameters: two rounds of consecutive 14 hour long IVT.

In a further embodiment, the mRNA of a single olfactory neuron is purified and amplified with LD-PCR (Long Distance Polymerase Chain Reaction) based methods such as the one described in NGS-ready kits (e.g., Clontech/Takara, SMARTer® Ultra® Low Input RNA Kit for Sequencing - v3, cat. 634848). Single cell mRNA is first reverse transcribed into the corresponding cDNA, which subsequently is amplified with 18 PCR cycles and serves as NGS sample for transcriptome sequencing.

In yet another embodiment, the identity of a group or gene family of DMTS olfactory receptors is determined (e.g., up to as many as the number of neurons picked)

by comparing the results of the NGS reads obtained from the isolated activated olfactory sensory neurons to a reference genome sequence of the same species. Particularly, the putative DMTS receptors will be the most highly abundant olfactory receptor mRNA in the olfactory neuron-derived NGS sample or present in more than one independent biological replicate. Because of the combinatorial nature of the olfactory code (one compound activates many ORs and one OR can be activated by many compounds), pooling several neurons activated by given compounds allows the retrieval of virtually all of the receptors responsible for the perception of these molecules in a single NGS experiment. Pooling functionally similar neurons thus greatly improves the deorphanization throughput and speed.

Standard bioinformatics tools are then used to identify the most closely related human odorant receptor(s) to other putative mammalian (non-human) DMTS receptor(s) under the assumption that homologous sequence receptors retain similar function. Several methods successfully identify human OR-ligand pairs based on this assumption [Armelin-Correa and Malnic (2017)] and up to 80% of mouse-human orthologs appear to maintain similar functional response profiles [Adipietro, K. A, et al. PLoS Genet. 8, e1002821–e1002821 (2012)]. Default parameters of BLASTP and/or BLASTN algorithm, or other ortholog pair identification algorithms such as InParanoid may be used.

The human or non-human mammalian DMTS receptors may be adapted to a functional assay that can be used to identify compounds that bind, suppress, block, inhibit, and/or modulate the activity of the olfactory receptors. In particular, the assay may be a cell-based assay or a binding assay and the method for identifying compounds may be a high-throughput screening assay. More particularly, provided herein are receptor-based assays adaptable for high-throughput screening of receptors with compound libraries for the discovery of modulating compounds (e.g., binding, blocking, inhibiting, suppressing and masking).

In a particular embodiment, DMTS receptor gene sequences are identified from DMTS-sensitive cells as follows: Pooled neurons are heated to 75°C for 10 minutes to break the cell membrane and render their mRNA available for amplification. This amplification step is important when applying NGS technologies with limited amount of starting material, typically between 1 to 15 cells. A linear amplification according to the Eberwine method (IVT) ensures the maintenance of the relative transcription levels of expressed genes. Two consecutive overnight (14h) rounds of *in vitro* transcription are

used to yield sufficient amounts of cRNA; Amplified cRNA is then used to generate an Illumina HiSeq cDNA library. The resulting short sequences of typically 75 to 150 base pairs (commonly referred to as “reads”) are aligned against the reference genome of the mouse (such as UCSC version mm9 or mm10) in order to build the full transcriptome of these cells. Quantitative analysis of the transcriptome data yields a list of transcribed odorant receptor genes and their respective expression levels. Odorant receptor genes that show the most abundant levels of mRNA (most abundant “reads”) or are present in more than one replicate experiment are considered putative DMTS receptors.

The predicted mouse OR genes are then used to mine the latest versions of both the mouse and human genome databases in order to identify the most closely related receptors (i.e. highest sequence similarity) in mouse (paralogous genes) and in human (orthologous genes). This process may be performed using the BLAST search algorithm (publically available at the NCBI website), a sequence similarity search tool, where every putative gene sequence previously obtained from the initial transcriptome analysis is used as a query sequence. The newly identified genes identified from this data mining process are considered to be potential DMTS receptors under the assumption that paralogous and orthologous genes are highly likely to possess similar activities. In a particular embodiment, pairwise comparison of sequence homology is carried out to identify closely related receptors in mouse and humans and the receptors are identified as described in WO2014/210585. Other approaches may also be used such as RT-PCR and microarray approaches.

In a further embodiment, to complete the deorphanization process, the candidate OR genes are further expressed *in vitro* for confirmation of activity against the compounds used to isolate the olfactory sensory neurons and other structurally-related compounds of interest. The mouse receptors identified from isolated olfactory neurons that respond to DMTS are modified at their N-terminus with short polypeptide sequences (e.g., FLAG[®] (SEQ ID NO: 44), Rho (SEQ ID NO: 46; 20 first amino acids of the bovine rhodopsin receptor), and/or Lucy (SEQ ID NO: 48; cleavable leucine-rich signal peptide sequence) tags), transiently expressed in HEK 293T cells, and stimulated separately with DMTS to confirm their identity as bona fide DMTS receptors. In a further embodiment, an RTP1 gene can also be expressed in the cell lines whether through activation of the endogenous RTP1 gene or through transformation. Co-expression of the human G alpha subunit G α_{olf} in this cell-based assay activates the Gs transduction pathway that leads to

an internal cAMP increase upon binding to the appropriate ligand. Alternatively, co-expression of the human G alpha subunit $G\alpha_{15}$ in the cell based assay activates the Gq transduction pathway that leads to an internal Ca^{2+} increase upon binding to the appropriate ligand. The above process and the results obtained so far serve to validate the process for rapid and reliable identification of mammalian odorant receptors for DMTS or DMDS.

Further provided are assays for identifying compounds that bind to DMTS or DMDS odorant receptors. In a further embodiment provided herein is a malodor modulating compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of at least one olfactory receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, OR52N5, OR2L13, OR4C15, OR5AC2, OR8H3, OR11G2, and OR52N2, and that is identified by the methods described herein, for example, the compounds described in Figure 4.

In one embodiment the activity of the compound is determined by comparing its binding to that of DMTS or DMDS. In another embodiment, the receptor or a chimera or fragment thereof is contacted with a compound in the presence of DMTS or DMDS under conditions that allow for the binding of the compound along with DMTS or DMDS to the receptor.

In a further embodiment, a compound is contacted to a receptor, or a chimera or fragment thereof that is activated by DMTS or DMDS, wherein the receptor, or a chimera or fragment thereof is expressed in a cell that is recombinantly modified to express the receptor polypeptide.

The activity of the compound can be determined using *in vivo*, *ex vivo*, *in vitro* and synthetic screening systems.

In another embodiment, the contacting is performed with liposomes or virus-induced budding membranes containing the polypeptides described herein.

In another embodiment, the methods for identifying compounds that bind, suppress, block, inhibit, and/or modulate the activity of an olfactory receptor that can be activated by DMTS or DMDS, may be performed on intact cells or a membrane fraction from cells expressing the polypeptides described here or on olfactory sensory neurons in culture modified to express endogenous or exogenous odorant receptors.

The 21 ORs described herein are therefore involved in the perception of malodor elicited by DMTS or DMDS and constitute valuable candidate receptors for the

identification of modulators, antagonists and/or blockers that would modulate, reduce, suppress, inhibit and/or block the perception of malodor.

Nucleic acid and amino acid sequences identified and/or used herein are listed below:

5

Olf1193

DNA - SEQ ID NO: 1

at ggcct caaggact t at t ccat ggaagaagt aaat aat gt cact gaat t cat t t t ct t gggt ct
t t ct cagaaccct gaggt t gaaaaagt gt gct t t gt ggt gt t ct cct t ct t t t acat ggt cat t c
10 t gct aggaacct cct cat cat gt t gacagt t t gcagt ggcaat ct t t t caagt t t cccat gt at
t t t t cct caact t t ct gt ct t t t gt ggacat t t gct act cct cagt cacagcaccacaagt gat
t at t gacct gt t agt gaagaaaaagact at at cct at gt ggggt gcat gt t acaact ct t t gt gg
t t cat t t ct t t t ggt t gcaact gagat ct t cat t ct t act gt cat ggcct at gat agat at gt ggcc
at t t gt aaacct ct ccaact at at gact at gat ggaccgggaaagt gcaat aagat gt t gct cgg
15 aacat ggt cgggt ggct t ct t acat t ct at t at ccaagt ggct ct t gt ggt ccagct cccct t t t
gt ggaccgaat gagat t gat cact at t t ct gt gat gt acat cct gt act gaaact t gct gcaact
gacact t acat t gt t ggt at t t t t gt gacagcaaacagt ggcaccat t gcat t gggaagt t t t gt
cat ct t gct gat ct cat acacagt cat t ct cat gt ct ct gagaaagcagt cat ct gaaggcagac
gcaaagct ct ct ccaact t gt ggat cccacat t gct gt t gt cat cat t t t t t t ggccct gt act
20 t t t at gt at at ggggct gacact acct t ct ct gaggacaagt ggt agct at at t t t t t t t t acaccat
t at cact cccat gct gaat cct ct aat t t t acact ct aagaaat gcagaagt aaagaat gcaat ga
gaaact gt gggct agaaagt t t t cct gggaaact act gggaaat ag

Protein - SEQ ID NO: 2

25 MASRTYSMEEVNNVTEFI FLGLSQNPEVEKVFVFSFFYMWI LLGNLLI MLTVCSGNLFKFPMY
FFLNFLSFVDI CYSSVTAPKM I DLLVKKKTI SYVGOMLQLFVWHFFGCTEI FI LTVMAYDRYVA
I CKPLHYMTMMDRERCNKMLLGTW GGFLHSI I QVALVVQLPFCGPNEI DHYFCDVHPVLKLACT
DTYI VGI FVTANS GTI ALGSFVI LLI SYTVI LMSLRKQSSEGRKALSTCGSHI AVVI I FFGPCT
FMYMRPDTTFSEDKMWAI FYTI I TPMLNPLI YTLRNAEVKNAMRKLWARKFSWETTGK
30

Olf1093

DNA - SEQ ID NO: 3

at ggaaaagt cacaat cagct gt ggat gt ccacaat at t ccat t aaagaacat gact gaagccac
cat gt t t at t ct ct t aggat t cacagat gact t t gaact ccaagt ct t cct gt t t t t act gt t t c
35 t t gct at t t at ct ct t cact ct ggt aggaacct t t ggact ggt t gt t t t ggt cat t ggggat t gt
cggct acacaaccccat gt act at t t cct aagt gt t t t gt ct t t cct ggat gct t gct at t ct ac
agt t gt t acacccaaaat gt t ggt caact t t ct aagt gaaaat aagt ccat t t cat t cct t gcat
gt gcaacccaaaat gct t ct ct t t t gt t t t cgt t gggaaccacagaat gct t t ct cct ggcagcaat g
gct t at gaccgat at gt agccat ct acaacccact t ct gt at acagt ggccat gt caccagagt
40 at acct gccact cat cat t gct t cct at gct ggt ggagt t gt gcat ggt gct at ccacacagt gg
cact t t cagt ct gt cct t ct gt ggat ccaat gaaat t aagcat gt ct t ct gt gacat cctt gca
t t gct t gct ct t t ct t gt t ct gat acccacacaaat gagct t ct agt ct t gt act t ggt gggct t
gat t gagat t gt t accat cct gat t gt t ct ggt ct cct at ggat t cat cct ct t t gccat t ct ga
acat gcat t ct gct gaggggt aggaggaaagt gt t ct ct acat gt ggct ct cacct cact ggagt c
45 t ct at t t accat ggt acaat cct t t t cact t at at gaggcct agt t ccagt t at gct t caaat ca
t gacat ggt agt gt caat at t t t t acaccat t gt gat acccat gt t gaat cct at cat ct at agt t
t gaggaacaaagt gt aaaagt agcat t t aat aat t gt ggagaaaat gt gat t cat aa

Protein - SEQ ID NO: 4

50 MEKI TSAVDVHNI PLKNMTEATMFI LLGFTDDFELQVFLFLFLAI YLFTLVGNFGLVVLVI GDC
RLHNPMYFSLSVLSFLDACYSTVWTPKMLVNFLSENKSI SFLACATQMLLFVSLGITTECFLLAAM

AYDRYVAI YNPLLYTVAMSPRVYLPLI I ASYAGGVVHGAI HTVATFSLSFCGSNEI KHVFCDI PA
LLALSCSDTHTNELLVLYLVGLI EI VTI LI VLVSYGFI LFAI LNMHSAEGRRKVFSTCGSHLTGV
SI YHGTI LFTYMRPSSSYASNHDMVSI FYTI VI PMLNPI I YSLRNKDVKVAFNKLWRKDS

5 **Olf1097**

DNA - SEQ ID NO: 5

at gagt gcct gt aat cat acaaat gaacct gagt t cagct t gt gggact gacagact ccaagga
gat t cagct ggt cct ct ct gt t t t gt t t ct cct gat at acat gct cact gt ct t gggaaacat ag
gt at gat act gat cat t cat ct agat gt ccagct ccacact ccaat gt at t t t t t cct caccac
10 t t gt cat t cct t gacct cagt t act caact gt aat cacacct aaaacct t acagaat acgct gac
ct ccat aaaaaat at t t cct t cat gggat gct t caccagct t gt at t t ct t t gt cct ct t ggcag
ct t ct gaat gt t t t at act t t cgt caat ggcct at gaccgct at gt agct at ct gcaacct ct a
cact at ccagt t at t at gt cccct aggcgt cat at act ct cat cact gt gt cct acat gat t gg
agt t t t ggt t ct t ct gt cact gt ct t t t gct t aagcacact ggt t t ct gcaact ccaagt aa
15 t t cat cact t ct t t t gt gacacat t ccaat t t t agct ct gt cct gcagt gat acct at aat gca
gaagccact at at t cgt t t t agct ggt t ccaact ct at t gct gt cgct cat cacgat at cct cat c
ct at gt at ct at t ct ct ct acaat t t t gaagat aaat t ct t ct t caggaaagcacaaagcct t ct
ct acat gt gcct cacat ct t at aggagt cact gt t t t t t at ggt acaat gat ct t t act t at t a
aaaccaagt acgt cct act cctt gggaaaggat caagt agcct ct gt t t t t t at act at agt gat
20 t cccat gct gaaccact t at ct at agt ct caggaaacaagaagt gaaaagt gct gt t gt t agag
t t at gaagaagagagagt gcat ccagaaact agaataa

Protein - SEQ ID NO: 6

MSACNHTNEPEFTLVGLTDSKEI QLVLVFLLI YMLTVLGNIGM LI I HLDVQLHTPMYFFLTH
25 LSFLDLSYSTVI TPKTLQNTLTSI KNI SFMGCTQLYFFVLLAASECFI LSSMAYDRYVAI CNPL
HYPVI MSPRRSYTLI TVSYM GVLDSSVTVFCLSTLDFCNKVI HHFFCDTFPI LALSCSDTYNA
EATI FVLAGSTLLLSLI TI SSSYVSI LSTI LKI NSSSGKHKAFSTCASHLI GVTVFYGTMI FTYL
KPSTSYSLGKDQVASVFYTI VI PMLNPLI YSLRNKEVKSAAVVRVMKKRECI QKLE

30 **Olf166**

DNA - SEQ ID NO: 7

at ggagaaat ggaat cagagct caagt gat t t t act ct gt t aggact gct t ccacaaaaccaaac
aggcct gct act t t t gat gct cat cat ct t t t gt ct t ct ct ggct t t gt gt ggcaact caggaa
t gat ccacct cat t cgt gt ggt ccaaggct ccacaccccat gt act t t ct cct cagt cagct c
35 t ct ct cat ggacct gat gt acat t t ct accact gt t cccaagat ggcata t t aact t cct t t ct gg
ccagaaaagcat ct ct t t t ct gggct gt ggagt gcaat cct t ct t ct t cct gact at ggcata gt t
ct gagggct t gct ct t ggct t ccat ggct t at gat cgt t t t gt ggct at ct gccat cccct t cac
t at cccat t cgcata gagcaaaat aat gt gt ct gaagat gat cat aggat cct ggat at t gggct c
aat caact ct t t agcacat accgt ct at gccct t cat at t cct t act gccat t ct aggt ccat t a
40 accat t t ct t ct gt gat gt t ccagccat gt t gccct ggcct gt at ggacact t ggggt t t at gag
t acat ggt gt t t t gagcacaagcct gt t t ct cct act gcct t t cct t ggt at cacagct t cct a
t ggt cgggt cct t t t t gct gt ct t ccacat gcgct caaaagagggaaagaagaaggcct t cacca
cat gct caact cact t aact gt ggt gacat t t t act at gcacct t t t gt ct at acct at ct t cga
cct aggagt ct t cgt ccccaacagaagat aagat t ct ggct gt t t t ct acact at cct t acccc
45 cat gct caaccccat cat t t at agt ct gaggaat aaggaggt cct gggggccat gacaagagt cc
t t ggt act t t t cct t caact aaaccgt aa

Protein - SEQ ID NO: 8

MEKWNQSSDFTLLGLLPQNTGLLLMLI I FVFSALOGNSGM HLI RVDPRLHTPMYFLLSQL
50 SLMDLMI STTVPKMAFNFLSGQKSI SFLGCGVQSFFFLTMACSEGLLLASIMAYDRFVAI CHPLH
YPI RMSKI MCLKM I GSW LGSI NSLAHTVYALHI PYCHSRSI NHFFCDVPAMLPACNDTWY
YMFVSTSLFLLLPFLGI TASYGRVLFVAFHMRSEKGGKKAFTTCSTHLLTVTFYAPFVYTYLR
PRSLRSPTEDKI LAVFYTI LTPMLNPI I YSLRNKEVLGAMTRVLGTFPSTKP

Olfr169

DNA - SEQ ID NO: 9

at ggaat at gagaact acact t t t aacagcgact t cat cct ct t gggact gt t ct ct t ct t caaa
gacaagct t aact t t t t t ct cat t t at at t t t t cat t t t at t at ggct at aacagaaaat gccc
5 t cat gat cct cct aat ccacagggat t ct cgact ccat accccaat gt at t t cct gct t agt cat
ct ct cct t cat ggat at ct t gcacat t t ccaacat t gt t cct aaaat gat t gct gact t cct ct c
aggcagcagaact at t t cct t t gcaggct gt gcct t ccagat at t t ct ct ct t acct t gct ag
gt ggt gact gcct t ct cct ggcagccat gt cct at gat cgat at gt ggccat ct gccacccact t
cgct accct gt gct gat gagggat aact ccagt aggct cct ggct gcaggct cct ggct ggt ggg
10 gat cct caact ccat agt acacacagt t t t t gcact ccact t t ccct t ct gccact caagagcca
t t gat cact t t t t ct gcgaagt cct gccat gt t gaaat t gt cat gt at agacacaacacact at
gaacgaggcgt t t at gt gact ggcat t at t t t t ct gct gat cccat t t t ccat gat ct ct at at c
t t at gt gcaaat t ct cct cact gt at t ccaaat gcagt cat cagggggcccgcaaaaagt cct t t t
ccacct gt t cct t ccacat ggt t gt t gt cat aat gt act at gggccat t cat t t t t acat at at g
15 agacct cgct cat accacact ccagggcaggat aaat t t t t ggcaat at t ct acaccat cct gac
accacact caaccccat aat ct acagct t t cgt aat aaagat gt cct t at ggct gt gaaaaaca
t cgt ccaaagt aat t t t t t gaat aaaaaat ga

Protein - SEQ ID NO: 10

20 MEYENYTFNSDFI LLGLFSSSKTSLTFFSFI FFI FI MAI TENALM LLI HRDSRLHTPMYFLLSH
LSFMDI LHI SNI VPKM ADFLSGSRTI SFAGCAFQI FLSTLLGGECLLAAMSYDRYVAI CHPL
RYPVLMRDNSSRLLAAGSWLVGI LNSI VHTVFALHFPCHSRAI DHFFCEVPAMLKLSI DTTHY
ERGVYVSGI I FLLI PFSM SI SYVQI LLTVFQMSGARQKSFSTCSFHMMVVI MYGPFI FTYM
RPRSHTPGQDKFLAI FYTI LTPTLNPI I YSFRNKDVLMAVKNI VQSNFLNKK
25

Olfr738

DNA - SEQ ID NO: 11

at gaaagcct t t agcagccccagcaact ccagcat cat cact ggct t cat cct cct gggct t ccc
ct gcccgaaggagggggcaaat cct cct ct t t gt gct ct t ct t cat t at ct acat cct t accct ca
30 t gggcaat gct t ccat cat at gt gct gt gt gct at gat aagaaact t cacagccccat gt acct c
ct gct ggccaact t ct cct t cct agaaat ct ggt at gt cacct ccacagt cccaacat gt t ggc
caact t cct ct ct gacacgaagggt cat ct ct t t ct ct ggat gct t cct gcagt t ct at t t ct t ct
t ct cct t ggggt t ct acagaat gct t t t t cct ggcagt cat ggcat t t gat cgat acct t gccat c
t gcagacct ct acat t at cct t ct ct cat gact gggcgct ct gcaacat cct t gt gat cagt t g
35 ct ggggt gct t ggt t t cct ct ggt t ccct gt t cccat cat cat ct cccaaat gt cct t ct gt g
gat ccagaat t at agaccact t cct gt gt gacccaggccct ct gt t ggccct cacct gt gt gaga
aat t ct t t aat t gagat gact agct ct act t t aagt t ccct gct t t at t t gt t ccat t t t t t t
t at cat ggggt ct t at gct ct agt aat gagggct gt gct cagggt ccct t cagcagct ggacgaa
gaaaggcct t ct ccacct gt ggggt cacact t gact gt ggt t t ct t t t t ct at ggct cagt gat g
40 gt cat gt at gt gagcccaacat ct gaacat gcagct ggagt gcaaaaact t gt gact ct gt t t t a
t t ct gt ggt t act cccct cct t aat cct gt gat at acagt ct gaggaacagagat at gaaacat g
caat gaaaaagt t act gaaaat gt aa

Protein - SEQ ID NO: 12

45 MKAFSSPSNSSI I TGFI LLGFPCPKEGQI LLFVLFFI I YI LTLMGNASI I CAVCYDKKLHSPM/L
LLANFSFLEI WYVTSTVPNMLANFLSDTKVI SFSGCFLQFYFFFSLGSTECFFLAVMAFDRLAI
CRPLHYPSLMTGRLONI LVI SCWLGFLWFPVPI I I SQMSFOGSRI I DHFLCDPGPLLALT/CVR
NSLI EMTSSTLSSLLLFVPFFFI MGSYALVMRAVLRVPSAAGRRAKAFSTCGSHLTVVSLFYGSVM
VMYVSPTSEHAAGVQKLVTLFYSVVTPLLNPI YSLRNRMKHAMKKLLKM
50

Olfr742

DNA - SEQ ID NO: 13

At gaaaaccct cagcagccccagcaact ccagccat cact ggct t cat cct ct t gggct t ccc
ct gcccagggagggggcaaat cct cct ct t t gt gacct t ct t cat t gt t t acat act cat t ct t a

t gggcaat gct t ccat cat ct gt gct gt gt act gt gat cagagcct ccacacccccat gt act t c
 ct gct ggccaact t ct cct t cct ggagat ct ggt at gt cacct ccacagt ccccaacat gt t ggc
 caact t cct t t cagacaccaaggt cat ct ct t t ct ct ggat gct t cct gcagt t ct at t t ct t ct
 t ct cct t t ggt t ct acagaat gct t t t t cct ggcagt cat ggc at t t gat cgat acct t gccat c
 5 t gt aggccact acat t at cct t ct ct cat gact gggcacct ct gcaacat cct t gt gat cagt t g
 ct ggg t gct t ggt t t cct ct ggt t ccct gt acccat cat cat cat ct cccagat gt cct t ct gt g
 ggt ccagaat t at agaccact t cct gt gt gacccaggccct ct t t t ggcct t gcct gt t ccaga
 gccccat t gat ggaggt t t t ct ggacaat t at aat gt ct at gct cct ggt t at t cct t t cct ct t
 cat cat gggaact t acat at t ggt cct aagagct gt gt t t agact t cct t caagagat ggacaaa
 10 aaaaggcct t ct ccact t gcgggt ct cat ct cacagt agt t t cact ct t t t at t gct cagt gat g
 aaaat gt at t t gagcccaacat ct gagcat gaagct ggaat gcagaagct t gt aact ct at t t t a
 t t ct gt ggg t act ccact act t aat cct gt gat at acagt ct gaggaacaaagat at gaaaaat g
 ccct gcagaagat t t t aagaacat aa

15 Protein - SEQ ID NO: 14

MKTLSSPSNSSTI TGFI LLGFPCPREGQI LLFVTFFI VYI LI LMGNASI I CAVYCDQSLHTPMYF
 LLANFSFLEI WYVTSTVPNMLANFLSDTKVI SFSGCFLQFYFFFSFGSTECFFLAVMAFDRLAI
 CRPLHYPSLMTGHLONI LVI SCWLGFLWFPVPI I I SQMSFCGSRI I DHFLCDPGPLLALACSR
 APLMEVFTI I MSMLLVI PFLFI MGTYI LVLRAVFRPLSRDGGQKAFSTCGSHLTVVSLFYCSVM
 20 KMYLSPTSEHAGVQKLVTLFYSVGTPLLNPI YSLRNKDMKNALQKI LRT

Olfr207

DNA - SEQ ID NO: 15

at ggaact gaacaggaccagct gact gaat t t gt t ct cagaggaat aacagat cgt t cagagct
 25 gcaagt cccct gt t cct ggt gt t ct t t ct cat ct at gt t at caccat ggt gggcaacct t ggt
 t aat ct t t gt cat ct ggaaggaccct cat ct t cacacacccat gt acct t t t cct t ggaaat t t g
 gcct t t gct gat gcct gt aat t cat cct ct gt gacaccaaagat gct t at gaaat t t t aaat aa
 gaat gacat gat at ccat ggggt gact gt t t t gct caat t t t at t t ct t t t gt t caagt gt aact g
 cagaagcct t cat t ct ggt agct at ggcct at gaccgt at gt agccat at gcaaacct ct gct c
 30 t at gt agt ggt gat gt ccaacagact ct gt at t cagt t cat aggt gt at cct at ct aat t ggact
 t ct acat ggct t act t cat gt aggat t gt t at t t aggt t aacgt t t t gt agt t ccaat gt aat ag
 at t at t t ct act gt gacat cct gccact t t at aggat t t ct t gcact gacccat cgat caat gt a
 ct ggt agct t t cat t at gggat at t t t at t acaagt gact acct t t at gact at t at agt ct cct a
 t gt ccgt gt cct ct t t gccat cct gagaacaaagt ct gagaggggcagaacaaagcct t ct ct a
 35 ct t gcagt t cccacct gt cat ct gt gt ct t t gt t ct at ggcact ct ct t cat cat at at gt cct c
 t ct ggt ct gacacagat aat t at cagggt aaaat gt at t cact gt t ct at accat t at cat t cc
 t ct gct aaaccct t cat t t acagcct aagaaat aaagaagt cat cgggt gcct t gagaaaagt ca
 gaaaat ga

40 Protein - SEQ ID NO: 16

MELNRTQLTEFVLRGI TDRSELQVPLFLVFFLI YVI TMVGNLGLI FVI WKDPHLHTPMYFLGNL
 AFADACNSSSVTPKMLMKFLNKNDM SMGECFAQFYFFCSSLTAEAFI LVAMAYDRYVAI CKPLL
 YVVMSNRLCI QFI GVSYLEI GL LHGLLHVGLLFRLLFCSSNVI DYFYCDI LPLYRI SCTDPSI NV
 LVAFI MGI LLQVSTFMSI I VSYVRVLF AI LRTKSERGRNKAFTCSSHLSSVSLFYGTLEI I YVL
 45 SGSDTDNYCGKMYSLFYTI I I PLLNPFY YSLRNKEVI GALRKVRK

Olfr665

DNA - SEQ ID NO: 17

at gcct ggggt caat acct ccagcct gacaccaagat act t t at t ct caat gggat t cct ggggt t
 50 ggaagct gcacacat ct ggat ct ct ct gccat t ct t cat t at gt acct cat t gct gt cacaggt a
 act gt ggact t at ct acct cat cagt cat gaggaggct ct gcaccggcccat gt act act t t ct a
 gccat gt t gt ct gct acagat at t t ct ggggt gt aat acaat t gt cccagat at gt t at gcat ct t
 t t ggt t cagt gt caaggagat t gat t t caat gcct gcct t gt acagat gt t t t t cat ccacat gt
 t aacaggcat ggagt ct ggt gt gct cat gct t at ggct ct cgaccgt at gt ggct at at gct at

ccat t acgct at act accat act caccaacact at gat t accaagat t ggat t ggcagcact t gt
t agaagt gt gt t act cat ggt ccct t t t gct t t cct gat caagcgt ct t ccat act gt agaggaa
acct cat ccaacat acct at t gt gat cacat ggct gt ggct aaact at cct gt ggcaat at t aag
at t aat gct at ct at ggt ct t at aat t gct at at t t at t ggggggt t t t gat at at t ct gt at ct c
5 cat gt ct t at gccat gat t at ccat gct gt ggt gaagct at ct t cggcagat gct cgccat aaag
cct t cagt acct gt acat cacacat at gt gct at t gt t at t acct at gt cccagcat t ct t caac
t t ct t t act cat cgct t t gggagaaccact at accccat cat at ccacat t at t at agccaacct
gt at ct at t gct acct cccacct t gaat ccaat t gt at at ggagt aaagaccaagcagat t cgt g
aaggt gt gat caaact gt t t gct agacaaaaagt t gt t t ga

10

Protein - SEQ ID NO: 18

MPGVNTSSLTPRYFI LNGL PGLEAAHI W SLPFFI MYLI AVTGNGLI YLI SHEEALHRPMYFL
AMLSATDI SGNTI VPSMLCI FWFVSKEI DFNAOLVQVFFI HMLTGVESGVLMLMALDRYVAI CY
PLRYTTI LTNTM TKI GLAALVRSVLLMPFAFLI KRLPYCRGNI QHTYCDHIMAVAKLSOGNI K
15 I NAI YGLI I AI FI GCFDI FCI SMSYAM I HAVVKLSSADARHKAFSTCTSHI CAI VI TYVPAFFN
FFTHRFGRITTI PHHI HI I I ANLYLLLPTLNPI VYGVKTKQ REGVI KLFARQKV

Olf669

DNA - SEQ ID NO: 19

20 at gct gat t t ccaacaact cat at gaagccccgcagt ct t t cat t ct t aat ggaat t cct ggt ct
cgaagcagt gcat at at ggat ct ct ct t ccact ct gt acaat gt acat cat ct ccct agt aggca
acct t ggct t gt at at ct cat t t act at gaggaat cct t acat cgcccaat gt at t t ct t t ct g
gccat gct t t ct ct cat agacct gt t t act t gcacaaccact gt ccccaat gccct ct t cat t t t
ct ggt t caaact caaggaaat t aact t cact gct t gcct agt t cagat gt t ct t t gt gcacggat
25 t cacaggt gt ggagt ct ggggt act cat gct cat ggct t ggaccgct at gt ggccat t t gct ac
ccact acgct at gcaacct act t accaaccct gt cat t gccaaagct gggct t gccacct t ct t
gagaggt gt gt t act gat gat t cct t t t ccat t ct t ggt t aaacgt t t gccct t ct gccgaagca
at gt cat ct cccat acat at t gt gaccacat gt ct gt ggt aaagt t at cct gt gccagcat caaa
at caat gt cat ct at ggt ct cat ggt t gcaat t ct gat t ggagt gt t t gacat at gt t gt at at c
30 t gt gt cct acact at gat cct ccgggcagt ggt cagcct gt cct ct gcagat gct cggcagaagg
cct t cagcacct gcacagcccat at ct gccat cat cat t act t at gt t ccagcct t ct t cacc
t t ct t t act cat cgt t t t ggaggt cacaccat ccct cct t ct ct t cat at cat t gt ggct aat ct
t t at ct t ct t ct ccct ccaact ct aaat cccat t gt t t at gggat gaagaccaaacagat cagag
at agt at cat t aaat t ct t t cacggt gaaaaaggt t caaggt ga

35

Protein - SEQ ID NO: 20

MLI SNNSYEAPQSFI LNGL PGLEAVHI W SLPLCTMYI I SLVGNLGLVYLI YVEESLHRPMYFFL
AMLSLI DLFTCTTTVPNALFI FWFKLKEI NFTAOLVQVFFVHGFTGVESGVLMLMALDRYVAI CY
PLRYATI LTNPVI AKAGLATFLRGVLLM PFPFLVKRLPFCRSNI SHTYCDHIMSVVKLSCASI K
40 I NVI YGLMALLI GVFDI CCI SVSYTM LRAVVSLSADARQKAFSTCTAHI SAI I I TYVPAFFT
FFTHRFGGHTI PPSLHI I VANLYLLLPTLNPI VYGMKTKQ RDSI I KFFHGEKGSR

Olf1211

DNA - SEQ ID NO: 21

45 at gcaaaaccagagt t t t gt aacagaat t cat at t cct t ggact t t cacagaaccct aaagt cca
gaaaat agt t t t t at t gt at t t t t at t t gt ct acat t gcaact gt t gggggcaacat gat aat t g
t ggt gacct t gt ct gt agcccagcat t gat agact gccccat gt act t ct t t t t ggcct t ct t g
t ccct at t ggat gcat gct t ct ct t ct gt cat cacaccaaagat ggt t gt ggact ccct gt at ga
gaagaaaact at ct cct t t gaaggat gt at gat gcagt t at t t gct gagcact t cct t gcagcag
50 t agaagt gat t gt ct t gacagccat ggct at gaccgct at gt agcaat t t gcaagccct t gcac
t act ct t ccat cat gaact ggaggct ct gt ggcacact t at ggggat agcat ggacaggggggt t
ct t gcat t ct at cat acaaat t at ct t cacgt t gcaat t gccct t ct gt ggaccaaat gt cat cg
at cat t t cat gt gt gact t gt t cccat t act ggaact t gcct gcact gat act cat at ct t t ggc
ct t t t agt ggt t gccaacagt ggggt ct at ct gcat cat aat ct t ct ct at t t t gct ggt ct cct a

t ggt gt cat cct gt t ct ct ct gaaagct cacagt t ct gaagggcgat ggaaagct ct ct ccacat
 gt ggaat cccacat t gcagt t gt ggt t t t gt t ct t t gt cccgt gt at t t t at t t at gcacgt cct
 ccat ct gct t t ct cct t t gat aaaat ggt ggcgat at t t t at act at cct aact cct t gct caa
 t cct gt gat t t at act t t t cggaat aaggacat gaaaaat gct at gaagaaagt gt ggaagaggt
 5 t ggcagt ggt t t ct gat ggaaagt ga

Protein - SEQ ID NO: 22

MONQSFVTEFI FLGLSQNPVKQKI VFI VFLFVYI ATVGGNIM I VVTI VCSPALI DCPMYFFLAFL
 SLLDACFSSVI TPKM/VDSL YEKKT I SFEGOMQLFAEHFLAAVEVI VLTAMAYDRYVAI CKPLH
 10 YSSI MNWRLCGTLMGI AWTGGFLHSI I QI I FTLQLPFCGPNVI DHFMDLFPILLELACTDTHI FG
 LLVVANSGLI CI I I FSI LLVSYGVI LFSLKAHSSEGRWKALSTCGSHI AVVVLFFVPCI FI YARP
 PSAFSFDKMWAI FYTI LTPLLNPVI YTFRNKDMKNAMKKVWKRLAVVSDGK

OR52N5

15 DNA - SEQ ID NO: 23

at gcct ct at t t aat t cat t at gct ggt t t ccaacaat t cat gt gact cct ccat ct t t t at t ct
 t aat ggaat acct ggt ct ggaaagagt acat gt at ggat ct cct cccact ct gcacaat gt aca
 t cat ct t cct t gt ggggaat ct t ggt ct t gt gt acct cat t t at t at gaggagt cct t acat cat
 ccgat gt at t t t t t t t t t ggccat gct ct ct cct cat t gacct cct t acct gcaccaccact ct
 20 acccaat gcact ct gcat ct t ct ggt t cagt ct caaagaaat t aact t caat gct t gct t ggccc
 agat gt t ct t t gt t cat ggggt t cacaggt gt ggagt ct ggggt gct cat gct cat ggct ct agac
 cgct at gt agccat t t gct accct t t gcgt t at gct accacact caccaaccct at cat t gccaa
 ggct gacgt t gccacct t cct gaggggt gt at t gct gat gat t cct t t cccat t ct t ggt t aagc
 gt t t gcct t t ct gccaaagcaat at t at ct cccat acgt act gcgaccacat gt ct gt agt aaag
 25 ct at ct t gt gccagcat caaggt caat gt aat ct at ggt ct aat ggt t gct ct cct gat t ggagt
 gt t t gacat t t gt t gt at at ct t t gt ct t acact t t gat cct caaggcagcgat cagcct ct ct t
 cat cagat gct cggcagaaggct t t cagcacct gcat gccat at at ct gccat cat cat cacc
 t at gt t ccagcat t ct t cact t t ct t t gccaccgt t t t gggggacacacaat t cccct t ct ct
 t cacat cat t gt ggct aat ct t t at ct t ct t t cccccaact ct aaacct at t gt t t at ggag
 30 t aaagacaaaacagat acgcaagagt gt cat aaagt t ct t ccagggt gat aagggt gcaggt t ga

Protein - SEQ ID NO: 24

MPLFNSLCWFPTI HVTPPSFI LNGI PGLERVHWW SLPLCTMYI I FLVGNLGLVYLI YYEESLHH
 PMYFFFGHALSLI DLLTCTTTLPNALCI FWFSLKEI NFNACLAQMFFVHGFTGVESGVLMLMALD
 35 RYVAI CYPLRYATTLTNPI I AKAELATFLRGVLLM PFPFLVKRLPFCQSN I SHTYCDHMSVVK
 LSCASI KVNVI YGLMALLI GVFDI CCI SLSTYLI LKAAI SLSSSDARQKAFSTCTAHI SAI I I T
 YVPAFFTFFAHRFGGHTI PPSLHI I VANLYLLLPPTLNPI VYGVKTKQI RKSVI KFFQGDKGAG

OR2L13

40 DNA - SEQ ID NO: 25

at ggagaaat ggaat cacact t caaat gat t t cat t t t gt t ggggt ct gct t cccccaat caaac
 t ggaat at t t ct ct t gt gcct t at cat cct cat at t ct t t ct ggcct cgggt ggggt aact cggcca
 t gat t cact cat ccacgt ggat cct cgt ct ccacacaccgat gt act t t ct t ct cagccagct c
 t cct t at ggacct gat gt acat ct ccaccaccgt cccaagat ggcgt acaact t cct gt ccgg
 45 ccagaaaggcat ct cct t cct gggat gt ggt gt gcaaagct t ct t ct t cct gacct ggcgt gt t
 ct gaaggct t act cct gacct ccat ggcct acgaccgt t at t ggccat ct gccact ct ct ct at
 t at cct at ccgcat gagt aaaat gat gt gt gt gaagat gat t ggaggct ct t ggacact ggggt c
 cat caact cct t ggcacacacagt ct t t gccct t cat at t cct act gcaggt ct aggggt at t g
 accat t t ct t ct gcat gt cccagccat gt t gct t ct t gcct gt acagat act t ggggt ct at gaa
 50 t at at ggt t t t t gt aagt acaagcct ct t t ct cct t t t cct t t cat t ggcacact t ct t cct g
 t ggccgagt cct at t t gct gt ct at cat at gcat caaaggaggggagaaaaaaggcct t cacca
 ccat t t caacacat t t aact gt agt gat ct t t t act at gcacct t t t gt ct acacct at ct t cgg
 cccaggaat ct ccgct caccagct gaagacaagat cct ggcagt ct t ct acacct cct t acccc

cat gct caat cccat t at ct acagcct gaggaat aaggaagt cct gggggct at gaggagagt gt
t t gggat at t ct ct t t cct gaaagaat aa

Protein - SEQ ID NO: 26

5 MEKWNHTSNDFI LLGLLPPNQTGI FLLCLI I LI FFLASVGNSAM HLI HVDPRLHTPMFLLSQL
SLMDLMIY STTVPKMAYNFLSGQKGI SFLGCGVQSFFFLTMACSEGLLLTSMAYDRYLAI CHSLY
YPI RIMSKMVCVKM GGSWTLGSI NSLAHTVFALHI PYCRSRAI DHFFCDVPAMLLACTDTWYE
YMFVSTSLFLLFPFI GI TSSOGRVLFAVYHMSKEGRKKAFTTI STHLTVVI FYYAPFVYTYLR
PRNLRSPAEDKI LAVFYTI LTPMLNPI I YSLRNKEVLGAMRRVFGI FSFLKE

10

OR2AJ1

DNA - SEQ ID NO: 27

at gagt gt aacagaaaaat acgct cat gat cct cct cat t cgcagt gact cccgact ccacact cc
aat gt at t t t ct gct cagccat ct ct cct t aat ggat at ct t gcat gt t t ccaacat cgt t ccca
15 aat ggt cact aact t t ct gt caggcagcagaact at t t cat t t gcagg t gt ggg t ccagg t a
t t t ct gt cct caccct cct ggg t ggt ggt gcct t ct cct ggct gcaat gt cct gt gat cgct a
t gt ggct at ct gt caccgct gcgct at ccgat t ct t at gaaggagt at gccagcgct ct cat gg
ct ggaggct cct ggct cat t ggggt t t t caact ccacagt ccacacagct t at gcat gcat t t
cct t ct gt ggct ct agggcaat t gat cact t ct t ct gt gaagt cct gccat gt t gaagt t gt c
20 ct gt gcagacacaacacgct at gaacgaggggt t t gt gt aagt gct gt gat ct t cct gct gat cc
ct t t ct cct t gat ct ct gct t ct t at ggccaaat t at t ct t act gt cct ccagat gaaat cat ca
gaggcaaggaaaaagt cat t t t cact t gt t cct t ccacat gat t gt ggt cagat gt act at gg
gccat t t at t t t t acat at at gagacct aaat cat accacact ccaggcaggat aagt t cct gg
caat at t ct at acgat cct cacacccacact caaccct t t cat ct acagct t t aggaat aaagat
25 gt t ct ggcgggt gat gaaaaat at gct caaaagt aact t t ct gcacaaaaaaat gaat aggaaaat
t cct gaat gt gt t ct gt ct at t t ct at gt t aa

Protein - SEQ ID NO: 28

30 MSVTENTLM LLI RSDSRLHTPMFLLSHLSLMDI LHVSNV VPKMVTNFLSGSRTI SFAGCGFQV
FLSLTLLGGECLLLAAMBCDRYVAI CHPLRYPI LMKEYASALMAGGSWLI GVFNSTVHTAYALQF
PFOGSRAI DHFFCEVPAM.LKLSCADTTRYERGVCSAVI FLLI PFSLI SASYGQI I LTVLQMKSS
EARKKSFSFSTCSFHM VVTMYGPFI FTYMRPKSYHTPGQDKFLAI FYTI LTPTLNPI YSFRNKD
VLAVMKNMLKSNFLHKKMNRKI PECVFCLFLC

35

OR4C15

DNA - SEQ ID NO: 29

at gt t ct caat gacaacagaagcact caat aat t t t gcat t ggat gt accaact t gt t aat gac
t at gat accacaaat t gat ct gaagcaaat t t t cct t t gt cct aat t gcagact at acat gat cc
40 ct gt t ggagct t t cat ct t t t cct t gggaacat gaaaaccaaagct t t gt aact gagt t t gt c
ct cct gggact t t cacagaat ccaaat gt t caggaaat agt at t t gt t gt at t t t t gt t t gt ct a
cat t gcaact gt t gggggcaacat gct aat t gt agt aacct t ct cagcagccct gct ct t ct gg
t gt ct cct at gt act t ct t ct t gggct t cct gt cct t cct ggat gcgt gct t ct cat ct gt cat c
accccaaagat gat t gt agact ccct ct at gt gacaaaaacat ct ct t t t gaaggct gcat gat
45 gcagct ct t t gct gaacact t ct t t gct ggggt ggaggt gat t gt cct cacagccat ggcct at g
at cgt t at gt ggccat t t gcaagccct t gcat t act ct t ct at cat gaacaggaggct ct gt ggc
at t ct gat ggggggt agcct ggacagggggcct ct t gcat t ccat gat acaaat t ct t t t act t t
ccagct t cct t t t gt ggccccaat gt cat caat cact t t at gt gt gact t gt acccgt t act gg
agct t gcct gcat gat act cacat ct t t ggcct cat ggt ggt cat caacagt ggg t t at ct gc
50 at cat aaact t ct cct t gt t gct t gt ct cct at gct gt cat ct t gct ct ct ct gagaacacacag
t t ct gaagggcgct ggaaagct ct ct ccacct gt ggat ct cacat t gct gt t gt gat t t t gt t ct
t t gt ccat gcat at t t gt at at acacgacct ccat ct gct t t t cct t gacaaaaat ggcggca
at at t t t at at cat ct t aaat ccct t gct caat cct t t gat t t acact t t caggaat aaggaagt

aaaacaggccat gaggagaat at ggaacagact gat ggt ggt t t ct gat gagaaagaaaat at t a
aact t t aa

Protein - SEQ ID NO: 30

5 MFSMTTEALNNFALGCTNLLMTM PQI DLKQI FLCPNCRLYM PVGAFI FSLGNMQNSFVTEFV
LLGLSQNPVQEI VFVVFVLFVYI ATVGGNMLI VVTI LSSPALLVSPMYFFLGFLSFLDACFSSVI
TPKM VDSLYVTKTI SFEGQMMQLFAEHFFAGVEVI VLTAMAYDRYVAI CKPLHYSSI MNRRLQG
I LMGVAWVGGLHSM QI LFTFQLPFOGPNVI NHFMDLYPLLELACTDTHI FGLMMVI NSGFI C
10 I I NFSLLLVSAYVI LLSLRTHSSEGRWKALSTOGSHI AVVI LFFVPCI FVYTRPPSAFSLDKMAA
I FYI I LNPILLNPLI YTFRNKEVKQAMRRI WNRLMVSDEKENI KL

OR5AC2

DNA - SEQ ID NO: 31

at ggat at at cagagggaaat aagact ct t gt gacagagt t t gt t ct cacaggact t acagat cg
15 accat ggct gcacgt cct ct t ct t t gt t gt t t t t ggt ggt ct at ct cat caccat ggt gggca
acct t ggact gat agt t ct aat t t ggaacgacccccat ct t cat at gccat gt act t at t cct t
ggt ggt t t agcct t t t cagat gct t gt act t caacct ct at aaccct aggat gct ggt caat t t
ct t agacaagact gcaat gat at ccct agct gagt gcat caccag t t t act t t t t gct t cca
gt gcaact acagaat gct t cct cct ggt gat gat ggcct at gaccgt at gt agccat at gt aat
20 ccct t gct t t at ccagt gat gat gt ccaacaaact cagcgt cagt t gct aagt at t t cat at gt
aat t ggt t t cct gcat cct ct ggt t cat gt gagt t t act at t gcgact aact t t ct gcagg t t a
acat aat acat t at t t ct act gt gaaat t t t acaact gt t caaat t t cat gcaat ggt ccat ct
at t aacgcact aat gat at t t at t t t t ggt gct t t t at acaaat accact t t aat gact at cat
aat ct ct t at act cgt gt gct ct t t gat at t ct gaaaaaaaagt ct gaaaagggcagaagcaaaag
25 cct t ct ccacat gcggcgcccat ct gct t t ct gt ct cat t gt act acggaact ct gat ct t cat g
t at gt gcgt cct gcat ct ggct t agct gaagaccaagacaaagt gt at t ct ct gt t t t acacgat
t at aat t cccct gct aaacccat t t at t t acagct t gagaaat aaaaaagt cat gcat gcat t ga
gaagagt t at aaggaagt aa

30 Protein - SEQ ID NO: 32

MDI SEGKTLVTEFVLTGLTDRPWLHVLFFVWFLVYLI TMVGNLGLI VLI WNDPHLHMPMYLFL
GGLAFSDACTSTSI TPRMLVNFLDKTAM SLAECI TQFYFFASSATTECFLLVMAYDRYVAI CN
PLLYPVMMBNKLSAQLLSI SYVI GFLHPLVHVSLLLRLTFORFNI I HYFCEI LQLFKI SCNGPS
I NALM FI FGAFI QI PTLMTI I I SYTRVLFDI LKKKSEKGRSKAFSTCGAHLLSVSLYYGTLI FM
35 YVRPASGLAEDQDKVYSLFYTI I I PLLNPFI YSLRNKKVMHALRRVI RK

OR8H3

DNA - SEQ ID NO: 33

at gat gggg agaaggaat gacacaaat gt ggct gact t cat cct t acgggact gt cagact ct ga
40 agaggt ccagat ggct ct gt t t at gct at t t ct cct cat at acct aat t act at gct ggggaat g
t ggggat gct at t gat aat ccgcct ggacct ccagct t cacact cccat gt at t t t t cct t act
cacct gt cat t t at t gacct cagt t act caact gt cgt cacacct aaaacct t agcgaact t act
gact t ccaact at at t t cct t cacgggct gct t t gccagat gt t ct gt t t t gt ct t ct t gggg a
ct gct gaat gt t at ct t ct ct cct caat ggcct at gat cgct at gcagcgat ct gcagt cct ct a
45 cact acacagt t at t at gcccaaaaggct ct gcct cgct ct cat cact gggcct t at gt gat t gg
ct t t at ggact cct t t gt caat gt ggt t t ccat gacagat t gcat t t ct gt gact caaacat aa
t t cat cact t t t t ct gt gacact t ccccaat t t t agct ct gt cct gcact gacacagacaacact
gaaat gct gat at t cat t at cgct ggt t ccacct gat ggt gt ccct t at cacaat at ct gcat c
ct at gt gt ccat t ct ct ct accat cct gaaaat t aat t ccact t caggaaagcagaaagct t t ct
50 ct act t gcgt ct ct cat ct ct t gggagt caccat ct t ct at ggaact at gat t t t t act t act t a
aagccaagaaagt ct t at t cct t gggaagagat caagt ggct cct gt gt t t t at act at t gt gat
t cccat gct gaat ccact cat t t at agt ct t agaaacagagaagt gaaaaat gct ct cat t agag
t cat gcagagaagacaggact ccaggt ag

Protein - SEQ ID NO: 34

MMGRNDTNVADFI LTGLSDSEEVQVALFMLLLI YLI TMLGNVGMILLI RLDLQLHTPMYFFLT
HLSFI DLSYSTVVTPTKLANLLTSNYI SFTGCFAMFCFVFLGTAECYLLSSMAYDRYAAI CSPL
HYTVI MPKRLCLALI TGPYVI GFMDSFVNVVSMRLHFCDSNI HHFFCDTSPI LALSCTDTDNT
5 EMLI FI AGSTLMVSLI TI SASVSI LSTI LKI NSTSGKQKAFSTCVSHLLGVTI FYGTM FTYL
KPKRSYSLGRDQVAPVFYTI VI PMLNPLI YSLRNREVKNALI RVMQRQRQDSR

OR11G2

DNA - SEQ ID NO: 35

10 at gaaaat ct t caacagccccagcaact ccagcacct t cact ggct t cat cct cct gggct t ccc
t t gccccagggaggggcagat cct cct ct t t gt gct ct t cact gt t gt t t acct cct gacct ca
t gggcaat ggt t ccat cat ct gt gct gt gcact gggat cagagact ccacgccccat gt acat c
ct gct cgccaact t ct cct t ct t ggagat at gt t at gt cacct ccacagt cccagcat gct ggc
caact t cct ct ct gacaccaagat cat ct cgt t ct ct ggct gct t cct ccagt t ct act t t t t ct
15 t ct cct t gggct ct acagaat gct t t t t cct ggcagt t at ggcatt t gat cgat acct t gccat c
t gt cggcct ct acgct at ccaacct t at gaccagacgt ct ct gt accaat ct t gt ggt caat t g
ct gggct act t ggt t t cat ct ggt t ct t gat t cct at cgt caacat ct cccaaat gt cct t ct gt g
gat ct aggat t at t gaccact t cct at gt gaccagct cct ct t ct aact ct cact t gcaaaaaa
ggccct gt gat agagct t gt ct t t t ct gt ct t aagt cct ct gcct gt ct t t at gct ct t t ct t
20 cat t gt ggggt cct at gct ct ggt cgt gagagct gt t t gagggt ccct t cagcagct gggagaa
gaaaggct t t ct ccacct gt ggggt ct cacct ggct gt ggt t t cact gt t ct acggct cagt act g
gt cat gt at gggagcccaccat ct aagaat gaagct gaaagcagaagact gt gact ct gt t t t a
t t ct gt t gt t accccact gct t aaccct gt gat at at agt ct t aggaacaaagat at gagaaaag
ct ct gaagaaat t t t ggggaacat aa

Protein - SEQ ID NO: 36

MKI FNSPSNSSTFTGFI LLGFPOPREGQI LLFVLFTVYLLTLMGNGSI I CAVHWDQRLHAPMYI
LLANFSFLEI CYVTSTVPSMLANFLSDTKI I SFSGCFLQFYFFFSLGSTECFFLAVMAFDRLAI
CRPLRYPTI MTRRLCTNLVNCWLGFI WFLI PI VNI SQMSFOGSRI I DHFLCDPAPLLTLTCKK
30 GPVI ELVFSVLSPLPVFMLFLFI VGSYALVRAVLRVPSAAGRRAKAFSTCGSHLAVVSLFYGSVL
VMYGSPPSKNEAGKQKTVTLFYSVVTPLLNPI YSLRNKDMRKALKKFWGT

OR52N2

DNA - SEQ ID NO: 37

35 at gt ct ggggacaacagct ccagcct gacccccaggat t ct t t at ct t gaat ggcgt t cct gggct
ggaagccacacacat ct ggat ct ccct gccat t ct gct t t at gt acat cat t gct gt cgt gggga
act gt gggct cat ct gcct cat cagccat gaggaggccct gcaccggcccat gt act act t cct g
gcct gct ct cct t cact gat gt cacct t gt gcaccacat ggt acct aat at gct gt gcat at t
ct ggt t caacct caaggagat t gact t t aacgcct gcct ggcccagat gt t t t t t gt ccat at gc
40 t gacaggat ggagt ct ggggt gct cat gct cat ggccct ggaccgt at gt ggccat ct gct ac
ccct t acgct at gccaccat cct t accaaccct gt cat cgccaaggct ggt ct t gccacct t ct t
gaggaat gt gat gct cat cat cccat t cact ct cct caccaagcgct gccct at t gccggggga
act t cat cccccacacct act gt gaccat at gt ct gt ggccaaggat at cct gt ggcaat t t caag
gt caat gct at t t at ggt ct gat ggt t gct ct cct gat t ggt gt t t gat at ct gct gt at ct c
45 t gt at ct t acact at gat t t t gcaggct gt t at ggcct gt cat cagcagat gct cgt cacaag
cct t cagcacct gcacat ct cacat gt gt t ccat t gt gat cacct at gt t gct gct t t t t cact
t t t t t cact cat cgt t t t gt aggacacaat at cccaaccacat acacat cat cgt ggccaacct
t t at ct gct act gcct cct accat gaacccaat t gt t t at ggagt caagaccaagcagat t cagg
aagggt gt aat t aaat t t t t act t ggagacaagggt t agt t t t acct at gacaaat ga

Protein - SEQ ID NO: 38

MSGDNSSSLTPGFFI LNGVPGLEATHI W SLPFCFMYI I AVVGNGLI CLI SHEEALHRPMYFL
ALLSFTDVTLCCTM/PNMLCI FWFNLKEI DFNACLAQMFFVHMLTGMEGVLMLMALDRYVAI CY
PLRYATI LTNPVI AKAGLATFLRNMLI I PFTLLTKRLPYCRGNFI PHTYCDHMSVAKVSGNFK

VNAI YGLMVALLI GVFDI CCI SVSYTM LQAVMSLSSADARHKAFSTCTSHMCSI VI TYVAAFFT
FFTHRFVGHNI PNHI HI I VANLYLLLPTMNP I VYGVKTKQI QEGVI KFLLGDKVSFTYDK

OR5T1

5 DNA - SEQ ID NO: 39

at gt t t at at t aat aagct t cacagaagaat t t gat gt gcaagt ct t cct at t t t t at t at t t t t
agcaat ct at ct at t cact ct aat aggcaat t t agggct ggt t gt accgat cat t ggggat t t ct
ggct t cacagcccaat gt act at t t t ct t ggt gt t t t at cat t ct t ggat gt ct gct at t ct aca
10 gt t gt cact ccaaaaaat gt t ggt caat t t cct ggcaaaaaat aaat ct at t t cat t t ct t ggat g
t gcaacacagat gt t t ct t gct t gt act t t t ggaaccacagaat gct t t ct ct t ggct gcaat gg
ct t at gat cgct at gt agccat ct acaaccct ct cct gt at t cagt gagcat gt caccagagt c
t at gt gccact cat cact gct t cct at gt t gct agcat t t t acat gct act at acat acagt ggc
t acat t t agcct gt cct t ct gt ggat ccaat gaaat t aggcgt gt ct t t t gt aat at gcci cct c
15 t cct t gct at t t ct t gt t ct gacact cacgt aat ccagct t ct at t ct t ct act t t gt gggct ct
at t gagat agt cact at cct gat t gt cct gat ct cct at ggt t t t at t ct gt t ggccat t ct gaa
gat gcagt ct gct gaagggaggagaaaaagt ct t ct ct acat gt ggagct cacct aact ggagt ga
caat t t at cat gggacaat cct ct t cat gt at gt gagaccaagt t ccagct acact t cggacaat
gacat gat agt gt caat at t t t at accat t gt gat t cccat gct gaat cccat cat ct acagt t t
gcggaacaaaagt gt aaaggaggcaat caaaagat t gct t gt gagaaat t ggt t cat aaat aagt
20 t at ag

Protein - SEQ ID NO: 40

MFI LI SFTTEFDVQVFLFLLFLAI YLFTLI GNLGLVPI I GDFWLHSPMYFLGVL SFLDVCYST
VVT PKMLVNFLAKNKSI SFLGCATQVFLACTFGTTECFLLAAMAYDRYVAI YNPLLYSVSMSPRV
25 YVPLI TASYVASI LHATI HTVATFSLSFOGSNEI RHVFONMPPLLAI SCS DTHVI QLLFFYFVGS
I EI VTI LI VLI SYGFI LLAI LKMQSAEGRRKVFSTOGAHLTGVTI YHGTI LFMVVRPSSSYTSDN
DM VSI FYTI VI PMLNPI I YSLRNKDVKEAI KRLLVRNWF I NKL

OR4S2

30 DNA - SEQ ID NO: 41

at ggaaaaaat aaacaacgt aact gaat t cat t t t ct ggggt ct t t ct cagagcccagagat t ga
gaaagt t t gt t t t gt ggt gt t t t ct t t ct t ct acat aat cat t ct t ct gggaaat ct cct cat ca
t gct gacagt t t gcct gagcaacct gt t t aagt caccat gt at t t ct t t ct cagct t ct t gt ct
t t t gt ggacat t t gt t act ct t cagt cacagct cccaagat gat t gt t gacct gt t agcaaaagga
35 caaaacct ct cct at gt ggggt gcat gt t gcaact gt t t ggagt acat t t ct t t ggt t gcat g
agat ct t cat cct t act gt aat ggcct at gat cgt t at gt ggct at ct gt aaacccct acat t at
at gaccat cat gaaccgggagacat gcaat aaaat gt t at t agggacgt ggggt aggt ggggt t ct t
acact ccat t at ccaagt ggct ct ggt agt ccaact accct t t t gt ggaccaat gagat agat c
act act t t t gt gat gt t caccct gt gt t gaaact t gcct gcacagaaacat acat t gt t ggt gt t
40 gt t gt gacagccaacagt ggt accat t gct ct ggggagt t t t gt t at ct t gct aat ct cct acag
cat cat cct agt t t cct gagaaagcagt cagcagaaggcaggcgcaagccct ct ccacct gt g
gct cccacat t gccat ggt cgt t at ct t t t cggccct gt act t t t at gt acat gcgccct gat
acgacct t t t cagaggat aagat ggt ggct gt at t t t acaccat t at cact cccat gt t aaat cc
t ct gat t t at acact gagaaat gcagaagt aaagaat gcaat gaagaaact gt ggggcagaaat g
45 t t t t ct t ggaggct aaagggaat ag

Protein - SEQ ID NO: 42

MEKI NNTEFI FWGLSQSPEI EKVCFWFSFFYI I I LLGNLLI MLTVCLSNLFKSPMYFFLSFLS
FVDI CYSSVTAPKM VDLLAKDKTI SYVGOMQLFGVHFFGCTEI FI LTVMAYDRYVAI CKPLHY
50 MTI MNRETCKMLLGTWGGFLHSI I QVALVQLPFOGPNEI DHYFCDVHPVLKLACTETIYI VGV
VVTANS GTI ALGSFVI LLI SYSI I LVSLRKQSAEGRRKALSTCGSHI AMVI FFGPCTFMVMPD
TTFSEDKMVAIFYTI I TPMLNPLI YTLRNAEVKNAMKKLVGRNVFLEAKGK

FLAG[®] tag

DNA - SEQ ID NO: 43

gat t acaaggacgacgacgat aag

Protein - SEQ ID NO: 44

5 DYKDDDDK

Rho tag

DNA - SEQ ID NO: 45

10 at gaacgggaccgagggcccaact t ct acgt gcct t t ct ccaacaagacgggcgt ggt g

Protein - SEQ ID NO: 46

MNGTEGPNFYVPFSNKTGVV

Lucy tag

15 DNA - SEQ ID NO: 47

at gagacccagat cct gct gct cct ggccct gct gaccct aggcct ggct

Protein - SEQ ID NO: 48

20 MRPQ LLLLALLTLGLA

Human G-protein alpha subunit G_{olf}

DNA - SEQ ID NO: 49

at gggc ct gt gct acagt ct gcgccgct gct t t t cgggggcccaggggacgaccct gcgcggc
ct cggagccgcccgt ggaggacgcgcagcccgccccggccct ggccccagt ccggggcgg
25 ccgcaaggggacacggcccggaccct gct ccct cgggggcggcgaaggagcccgcat gcgct cgg
cccaaaagcagacaagccgaaggagaagcggcagcgcaccgagcagct t agt gccgaggagcgcga
ggcgcccaaggagcgcgagggcgt caaggaggcagggaaagt gagccggggcat cgaccgcat gc
t gcgcgaccagaagcgcgacct gcagcagacgcaccggct cct gct gct cggggct ggt gact ct
gggaaaagcact at cgt caaacagat gaggat cct gcacgt caat gggc t t aat cccgagggaaa
30 gaaacagaaaat t ct ggacat ccggaaaaat gt t aaagat gct at cgt gacaat t gt t t cagcaa
t gact act at aat acct ccagt t ccgct ggccaaccct gaaaaccaat t t cgat cagact acat c
aagagcat agcccc at cact gact t t gaat at t cccaggaat t ct t t gaccat gt gaaaaact
t t gggacgat gaaggcgt gaaggcat gct t t gagagat ccaacgaat accagct gat t gact gt g
cacaat act t cct ggaaagaat cgacagcgt cagct t ggt t gact acacaccacagaccaggac
35 ct cct cagat gcagagt t ct gacat ct gggat t t t t gagaacgat t ccaagt ggacaaagt aaac
t t ccacat gt t t gat gt t ggt ggccagaggat gagaggagaaaat ggat ccagt gct t t aacga
t gt cacagct at cat t t acgt cgcagcct gcagt agct acaacat ggt gat t cgagaagat aaca
acaccaacaggct gagagagt ccct ggat ct t t t t gaaagcat ct ggaacaacagggt ggt t acgg
accat t t ct at cat ct t gt t ct t gaacaaacaagat at gct ggcagaaaaagt ct t ggcagggaa
40 at caaaaat t gaagact at t t cccagaat at gcaaat t at act gt t cct gaagacgcaacaccag
at gcaggagaagat cccaaagt t acaagagccaagt t ct t t at ccgggacct gt t t t t gaggat c
agcacggccaccggg gacgggcaaacat t act gct accgcact t cact gcgccgt ggacacaga
gaacat ccgcagggt gt t caacgact gccgcgacat cat ccagcggat gcacct caagcagt at g
agct ct t gt ga

45

Protein - SEQ ID NO: 50

50 MGLCYSLRPLLFGPGDDPCAASEPPVEDAQAPAPALAPVRAAARDTARTLLPRGEGSPACAR
PKADKPKEKRQRTQLSAEEREAAKEREAVKEARKVSRGI DRMLRDQKRDLCQTHRLLLL GAGES
GKSTI VKQVRI LHVNGFNPEEKKQKI LDI RKNVKDAI VTI VSAMSTI I PPVPLANPENQFRSDYI
KSI API TDFEYSQEFFDHVKKLWDDEGVKACFERSNEYQLI DCAQYFLERI DSVSLVDYTPTDQD
LLRCRVLTSGL FETRFQVDKVNFMFVDVGGQDERRRWI QCFNDVTAI I YVACSSYNMVI REDN
NTNRLRESLDLFESI WNNRMLRTI SI I LFLNKQDMLAEKVLAKGSKI EDYFPEYANYTVPEDATP
DAGEDPKVTRAKFFI RDLFLRI STATGDGKHYCYPHFTCAVDTENI RRVFNDCRDI I QRMHLKQY
ELL

Human G-protein alpha subunit Gα15

DNA - SEQ ID NO: 51

5 at ggccccggt cgct gacct ggcgct gct gccct ggt gcct gacggaggat gagaaggccgcccgc
 ccgggt ggaccaggagat caacaggat cct ct t ggagcagaagaagcaggaccgccccgagct ga
 agct gct gct t t t gggcccaggcgagagcggaagagcacct t cat caagcagat gcggat cat c
 cacggcgccggct act cggaggaggagcgcaagggt t cgggccct ggt ct accagaacat ct t
 cgt gt ccat gcgggccat gat cgaggccat ggagcggct gcagat t ccat t cagcaggccccgaga
 10 gcaagcaccacgccagcct ggt cat gagccaggaccct at aaagt gaccacgt t t gagaagcgc
 t acgt gcggccat gcagt ggct gt ggagggt gccggcat ccggcct gct at gagcgt cggcg
 ggaat t ccact gct cgat t cagcgt gt act acct gt ccacct ggagcgcac caccaggagg
 gct acgt cccacagct caggacgt gct ccgcagccgcat gccaccact ggcac caacgagt ac
 t gct t ct ccgt gcagaaaaccaacct gcggat cgt ggacgt cgggggccagaagt cagagcgt aa
 gaaat ggat ccat t gt t t cgagaacgt gat cgccct cat ct acct ggcct cact gagt gaat acg
 15 accagt gcct ggaggagaacaaccaggagaaccgcat gaaggagagcct cgcac t gt t t gggact
 at cct ggaact acct ggt t caaaagcacat ccgt cat cct ct t t ct caacaaaaccgacat cct
 ggaggagaaaat cccacact cccacact ggct acct at t t cccagct t t ccagggccct aagcagg
 at gct gaggcagccaagaggt t cat cct ggacat gt acacgaggat gt acaccgggt gcgt ggac
 ggccccgagggcagcaagaaggcgacgat cccgacgct ct t cagccact acacat gt gccac
 20 agacacacagaacat ccgcaaggct ct t caaggacgt gcgggact cgggt gct cgcccgct acct gg
 acgagat caacct gct gt ga

Protein - SEQ ID NO: 52

25 MARSLTWRCPPWLTDEKAAARVDQEI NRI LLEQKKQDRGELKLLLLGPESGKSTFI KQVRI I
 HGAGYSEEERKGFRLVYQNI FVSMRAM EAMERLQI PFSRPESKHASLVMSQDPYKVTTFEKR
 YAAAMQWLWRDAGI RACYERRREFHLLDSAVYYLSHLERI TEEGYVPTAQDVLRSRVPPTTG NEY
 CFSVQKTNLRI VDVGGQKSERKKW HOFENVI ALI YLASLSEYDQCLEENNOENRIVKESLALFGT
 I LELPWFKSTSVI LFLNKTDI LEEKI PTSHLATYFSPFQGPQDAEAAKRFI LDMYTRVMTGCVD
 30 GPEGSKKGARSRLFSHYTCATDTQNI RKFVKDVRDSVLARYLDEI NLL

SEQ ID NO: 53

Motif

MAYDRYVAI C

35 SEQ ID NO: 54

Motif

FSTCSSH

SEQ ID NO: 55

40 Motif

PMLNPF I Y

The following examples are illustrative only and are not meant to limit the scope
 45 of invention as set forth in the Summary, Description or in the Claims.

EXAMPLES**Example 1****Identification of novel mouse and human DMTS activated odorant receptors**

The identification of new odorant receptors was performed according to the method disclosed in WO2014/210585. Briefly, murine olfactory sensory neurons were exposed to DMTS and screened using a Ca^{2+} imaging technique. Neurons that were activated by DMTS were further isolated for full transcriptome analysis to identify the responsive odorant receptor. The cDNA corresponding to the isolated cell mRNA was generated and amplified by PCR based method (Clontech/Takara, SMARTer® Ultra® Low Input RNA Kit for Sequencing - v3, cat. 634848). Amplified cDNA was then used to generate an Illumina cDNA library for Next-Generation-Sequencing and the generation of 100 base pair single read sequences. Sequences were aligned to a mouse reference genome (such as UCSC version mm10) in order to generate the full transcriptome. Because only a single odorant receptor (OR) is strongly transcribed per olfactory sensory neuron, the subsequent identification of the DMTS responsive OR can be achieved. Phylogenetic relationship evaluation using sequence similarity searches were then used to identify the corresponding human OR. Similar functional response profiles between orthologous OR pairs are often observed [e.g. Adipietro, K. A, et al. PLoS Genet. 8, e1002821–e1002821 (2012), Sato-Akuhara, N. et al. J. Neurosci. 36, 4482–4491 (2016) and WO2016/201152] and can be used for human OR identification [e.g. Armelin-Correa L. M. and Malnic B. J Agric Food Chem. doi: 10.1021/acs.jafc.6b04998 (2017)]. Figure 1 shows the pairwise identity levels between the receptors mentioned herein. ORs that do not share an orthologous and/or paralogous relationship only share an average of 39% $\pm$ 13.1%. All orthologous and paralogous relationship share an average of 77% $\pm$ 7.5%. All paralogs and orthologs pairs are indicated by a gray shaded cell. They represent the highest identity levels as supported by the amino acid identity levels indicated.

Example 2

Functional characterization of mouse and human DMTS receptors

Functional dose-response experiments were performed to evaluate the level of DMTS activity of the modified cell line expressing individual putative DMTS receptors. Using a cell-based assay, mouse receptor Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738 and Olfr742, and human receptor OR4S2 and OR52N5 were tested in an HEK293T cell line wherein the endogenous RTP1 gene has been activated and the odorant receptor chaperone protein was expressed (described in WO2016/201153). The mouse receptor genes were tagged with a Lucy-FLAG®-Rho tag combination and the

human receptors were tagged with a Rho-FLAG[®] tag combination resulting in a tag::receptor fusion protein for the cell based assay. Receptor genes were co-transfected with the canonical olfactory human G-protein alpha subunit G_{olf} gene and were exposed to increasing concentrations of the malodor odorant. Co-expression of the human G_{olf} activates the Gs transduction pathway that leads to an internal cAMP increase upon binding to the appropriate ligand. Odorant-induced activity was detected by measuring the cAMP increase in the cytosol using an Homogeneous Time Resolved Fluorescence (HTRF) based kit (CisBio, cAMP dynamic 2 kit, cat. 62AM4PEJ). A dose-dependent increase of the receptor activity is observed specifically for DMTS but not for butyric acid, another known malodor used as a negative control (Figure 1). The activity levels are reported by the potency of the DMTS induced- response for each receptor as a measure of the EC₅₀ value (the effective concentration at which the receptor responds at its half-maximal activation efficacy level for a given compound).

15 **Example 3**

Functional characterization of additional human DMTS receptors

Functional dose-response experiments were performed to evaluate the level of DMTS activity of the modified cell lines expressing individual putative DMTS receptors. A cell line stably expressing a human receptor was generated in a HEK293T cell line for each of the following human receptors: OR2L13, OR4C15, OR5AC2, OR8H3, OR11G2 or OR52N2. The receptors were tagged with a FLAG[®]-Rho tag combination and stably co-expressed with the human G-protein alpha subunit G_{α15} for the cell based assay. Co-expression of the human G_{α15} activates the Gq transduction pathway that leads to an internal Ca²⁺ increase upon binding to the appropriate ligand. Receptors were exposed to increasing concentrations of DMTS. Odorant-induced activity was detected by measuring the Ca²⁺ increase in the cytosol using a calcium sensitive fluorescent dye (Molecular Devices, Calcium 5 dye, cat. R8186) and measuring the change in Relative Fluorescence Ratio (RFU) using a fluorometric imaging plate reader (Molecular Devices, FLIPR) following odorant exposure. A dose-dependent increase of receptor activity was recorded and a corresponding dose-response curve is shown for DMTS (Figure 2). A cell line lacking a receptor was used as a control for non-specific activity at high compound concentrations ('No receptor'). The activity levels are reported by the potency of the DMTS induced-response for each receptor as a measure of the EC₅₀ value.

Example 4**Identification of DMTS receptor inhibitors**

The stable cell lines described in Example 3 were used as an antagonist screening platform to identify compounds that have the property to decrease the DMTS induced receptor activity. Each stable cell line expressing a human odorant receptor was screened with a volatile compound library for their inhibitory properties and potential DMTS smell inhibition. First, individual binary mixtures of DMTS with each one of the test compounds were presented to the cells. Single point monitoring of the DMTS induced cell activity in the presence or absence of a test compound allowed for the identification of compounds with a putative suppression or inhibitory effect. These hits were further confirmed in an inhibitory dose-response curve assay to evaluate the potency of activity inhibition as a measure of the IC₅₀ (the inhibitor concentration at which the receptor activity is inhibited by the half-maximal inhibition efficacy level of a given test compound). A dose-dependent decrease of receptor activity was recorded with increasing concentrations of test compounds in the presence of a single activating concentration of DMTS (EC₈₀) and corresponding dose-response inhibition curves were obtained. The compounds in the following table are examples of compounds that decreased the DMTS induced activity of at least one receptor as depicted in Figure 4.

Cyclemonone A	1,2,3,4,5,6,7,8-octahydro-8,8-dimethyl-2-naphthalenecarbaldehyde (A) + (B,C,D) + octahydro-5,5-dimethyl-2-naphthalenecarbaldehyde
Geonol	(+/-)-perhydro-4alpha,8abeta-dimethyl-4a-naphthalenol
Hivernal Neo®	3-(3,3-dimethyl-2,3-dihydro-1H-inden-5-yl)propanal (A) + 3-(1,1-dimethyl-2,3-dihydro-1H-inden-4-yl)propanal (B) + 3-(1,1-dimethyl-2,3-dihydro-1H-inden-5-yl)propanal (C)
Lilyflore®	(+/-)-2,5-dimethyl-2-indanmethanol
Patchouli alcohol	(-)-(3R,6S,8S)-2,2,6,8-tetramethyltricyclo[5.3.1.0~3,8~]undecan-3-ol
Patchouli Oil	patchouli oil
Rosinol Cryst	(+/-)-2,2,2-trichloro-1-phenylethyl acetate
Spiranol®	5RS,6RS)-2,6,10,10-tetramethyl-1-oxaspiro[4.5]decan-6-ol
Terranol	2,2,7,7-tetramethyltricyclo[6.2.1.0~1,6~]undecan-6-ol
Wolfwood®	(+)-(1S,2S,3S,5R)-2,6,6-trimethylspiro[bicyclo[3.1.1]heptane-3,1'-cyclohexane]-2'-en-4'-one

In Figure 4, the length of the bars indicate the potency of activity inhibition (IC₅₀, expressed in negative log molar concentration) of the select compounds for each receptor. The absence of a bar indicates absence of inhibition for the corresponding receptor compound pair. These compounds can specifically bind and block the activity of DMTS or DMDS receptors, and thus may suppress or inhibit the smell of DMTS or DMDS. Such compounds may therefore be used as malodour counteractants for latrine or oral odor control applications.

Example 5

10 **Response profiles of mouse and human DMTS receptors to DMDS**

Functional dose-response experiments were performed to confirm the DMTS activity on the DMTS receptors identified in Example 2 and to further characterize their response to DMDS. Using the same cell-based assay described in Example 2, mouse receptors Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742 and of human odorant receptor OR4S2 were tested with increasing concentration of DMTS, DMDS or butyric acid (control). A strong dose-dependent increase of the receptor activity is observed for DMTS and DMDS relative to the butyric acid control compound. A weak response for OR4S2 and Olfr169 to butyric acid was visible but was too weak to consider butyric acid as a representative ligand for these receptors. The activity levels are reported by the potency of DMTS or DMDS induced-response for each receptor as a measure of the EC₅₀ value, EC₅₀_{DMTS} and EC₅₀_{DMDS} respectively. A mock transfection control experiment in which the cells do not express the odorant receptors did not show activity upon DMTS or DMDS exposure. Competitive antagonists at the malodor binding site of these receptors have the potential to reduce the unpleasant perception of both DMTS and DMDS.

CLAIMS

1. A method for identifying a compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of an olfactory receptor that is activated by a malodor-causing substance comprising:
- 5
- a. contacting a test substance and a malodor-causing substance to at least one olfactory receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1;
 - 10
 - b. measuring the response of the olfactory receptor to the malodor-causing substance by measuring the response of the olfactory receptor in the presence and absence of the test substance;
 - c. identifying a test substance that modulates the response of the olfactory receptor on the basis of the response that was measured in the presence and absence of the test substance; and
 - 15
 - d. selecting the identified test substance as a compound that modulates the response of the olfactory receptor to the malodor-causing substance,
- wherein the malodor-causing substance is dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS).
- 20
2. A method for identifying a malodor inhibitor comprising:
- a. contacting a test substance and a malodor-causing substance to at least one olfactory receptor, wherein the receptor comprises a polypeptide comprising
 - 25
 - an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40;
 - 30

- b. measuring the response of the olfactory receptor polypeptide to the malodor-causing substance;
- c. identifying, based on the measured response, a test substance that can suppress the response of the olfactory receptor; and
- 5 d. selecting, as a malodor inhibitor, the test substance that suppresses the response of the olfactory receptor,

wherein the malodor-causing substance is DMTS or DMDS.

3. A method for identifying a compound that binds, suppresses, blocks, inhibits, and/or modulates the activity of at least one olfactory receptor that is activated by a malodor-causing substance, the method comprising:

- a. contacting the receptor, or a chimera or fragment thereof with a compound; and
- b. determining whether the compound has an effect on the activity of the receptor;

wherein the receptor is a polypeptide

- i) comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40; or
- 20 ii) encoded by a nucleic acid molecule comprising a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39 or
- 25 30 the reverse complement thereof;

wherein the malodor-causing substance is dimethyl trisulfide (DMTS) or dimethyl disulfide (DMDS).

4. A method of identifying a compound that putatively modulates DMTS or DMDS associated malodor comprising: (i) contacting a cell line that expresses a DMTS or DMDS receptor polypeptide with at least one compound; (ii) screening for compounds that bind, suppress, block, inhibit, and/or modulate the activity of said olfactory receptor polypeptide; and (iii) identifying a compound that putatively modulates DMTS or DMDS associated malodor if it binds, suppresses, blocks, inhibits, and/or modulates the activity of said DMTS or DMDS receptor polypeptide,
- wherein the receptor polypeptide
- 10 a. comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40; or
- 15 b. is encoded by a nucleic acid molecule comprising a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 39, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39 or the reverse complement thereof.
- 20
- 25 5. An isolated polypeptide comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, or SEQ ID NO: 40.
- 30
6. An isolated nucleic acid molecule comprising
- a. a nucleic acid sequence encoding the polypeptide of claim 5; or

- 5 b. a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 39, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39 or the reverse complement thereof.
7. A recombinant nucleic acid molecule comprising
- 10 a. a nucleic acid comprising a tag combination that comprises at least one of a Lucy tag, a FLAG[®] tag, and/or a Rho tag; and
- b. a nucleic acid encoding a receptor selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR4S2, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2,
- 15 OR8H3, OR11G2, OR52N2, and OR5T1 or the complement thereof.
8. The recombinant nucleic acid molecule of claim 7, wherein the Lucy tag comprises SEQ ID NO: 47, the FLAG[®] tag comprises SEQ ID NO: 43, and the Rho tag comprises SEQ ID NO: 45.
- 20 9. An expression vector comprising the nucleic acid of any one of claims 6 to 8.
10. A non-human host organism or a host cell that has been modified to express a receptor that is activated by DMTS or DMDS, wherein the receptor
- 25 a. is selected from the group consisting of Olfr1193, Olfr1093, Olfr1097, Olfr166, Olfr169, Olfr738, Olfr742, Olfr207, Olfr665, Olfr669, Olfr1211, OR52N5, OR2L13, OR2AJ1, OR4C15, OR5AC2, OR8H3, OR11G2, OR52N2, and OR5T1; or
- b. comprises the polypeptide of claim 5 or a polypeptide encoded by the
- 30 nucleic acid of any one of claims 6 to 8.
11. A non-human host organism or a host cell comprising
- a. the nucleic acid of any one of claims 6 to 8; or

b. the expression vector of claim 9.

12. The non-human host organism or host cell of claim 10 or 11, wherein the cell is a eukaryotic cell.

5

13. The non-human host organism or host cell of claim 10 to 11, wherein the cell is a prokaryotic cell.

14. The non-human host organism or host cell of any one of claims 10 to 11, wherein
10 the non-human host organism or host cell is selected from the group consisting of HEK293, CHO, *Xenopus oocytes*, COS, yeast and cells derived from the olfactory placode.

15. Use of a polypeptide that can be activated by DMTS or DMDS comprising an
15 amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID
20 NO: 38, or SEQ ID NO: 40 for identifying a malodor modulating compound.