

SHARED LIGANDS BETWEEN ORGANIC ANION TRANSPORTERS (OAT1 AND OAT6) AND
ODORANT RECEPTORS

Wei Wu, Kevin T. Bush, Henry C. Liu, Christopher Zhu, Ruben Abagyan and Sanjay K. Nigam

DRUG METABOLISM AND DISPOSITION

Supplemental Table 1

List of 85 Odorant Molecules from OlfactionDB*

| <u>PubChem</u> | <u>CID Name</u> |
|----------------|----------------------|
| 8094 | HEPTANOIC ACID |
| 20210 | BROMOHEXANOIC ACID |
| 8892 | HEXANOIC ACID |
| 16368 | BROMOPENTANOIC ACID |
| 7991 | PENTANOIC ACID |
| 454 | OCTANAL |
| 8908 | HEXYL ACETATE |
| 8130 | HEPTANAL |
| 8093 | 2-OCTANONE |
| 75809 | BROMOBUTANOIC ACID |
| 6184 | HEXANAL |
| 8051 | 2-HEPTANONE |
| 246728 | 3-OCTANONE |
| 8091 | METHYL OCTANOATE |
| 379 | OCTANOIC ACID |
| 62378 | DIHYDROJASMONE |
| 13187 | 2-NONANONE |
| 196 | HEXANEDIOIC ACID |
| 31289 | NONANAL |
| 8878 | ALLYL HEPTANOATE |
| 8158 | NONANOIC ACID |
| 2969 | DECANOIC ACID |
| 12813 | GAMMA-DECALACTONE |
| 385 | HEPTANEDIOIC ACID |
| 8175 | DECANAL |
| 548275 | BROMOOCTANOIC ACID |
| 10457 | OCTANEDIOIC ACID |
| 4409936 | PHENYLACETATE |
| 444294 | (-)-CAMPHOR |
| 19707 | 2,3-HEXANEDIONE |
| 22227 | (+)-DIHYDROCARVONE |
| 443157 | CITRONELLA(-) |
| 14525 | (+)-FENCHONE |
| 6852393 | ANDROSTENONE |
| 7012 | 2-PHENYLBUTYRIC ACID |
| 91604 | LYRAL |
| 10886 | AMYL HEXANOATE |
| 2266 | NONANEDIOIC ACID |
| 24114 | BUTYL BUTYRYLLACTATE |
| 11583 | 2-HEXANONE |
| 11614 | BUTYL FORMATE |

| | |
|---------|--------------------------|
| 7802 | 3-HEPTANONE |
| 15717 | ALLYL PHENYLACETATE |
| 8174 | 1-DECANOL |
| 439570 | (-)-CARVONE |
| 16724 | (+)-CARVONE |
| 8144 | OCTANETHIOL |
| 8914 | 1-NONANOL |
| 12756 | GAMMA-CAPROLACTONE |
| 957 | 1-OCTANOL |
| 7342 | ETHYL ISOBUTYRATE |
| 520196 | NONANETHIOL |
| 62465 | 2-METHOXY-4-ETHYLPHENOL |
| 4763 | HELIONAL |
| 7136 | EUGENYL ACETATE |
| 8129 | 1-HEPTANOL |
| 64832 | BOURGEONAL |
| 8842 | BETA-CITRONELLOL |
| 440917 | LIMONENE |
| 7967 | CYCLOHEXANONE |
| 7895 | 2-PENTANONE |
| 3314 | EUGENOL |
| 1032 | PROPIONIC ACID |
| 8468 | VANILLIC ACID |
| 5485176 | 3,4-HEXANEDIONE |
| 14489 | PRENYL ACETATE |
| 637796 | ISOSAFROLE |
| 8785 | BENZYL ACETATE |
| 8103 | 1-HEXANOL |
| 68382 | 2-COUMARANONE |
| 7144 | 2-METHOXY-4-METHYLPHENOL |
| 637566 | GERANIOL |
| 1549045 | ISOMETHYLEUGENOL |
| 6276 | 1-PENTANOL |
| 3102 | BENZOPHENONE |
| 8467 | ETHYL VANILLIN |
| 240 | BENZALDEHYDE |
| 7410 | ACETOPHENONE |
| 323 | COUMARIN |
| 1183 | VANILLIN |
| 68110 | 4-CHROMANONE |
| 9309 | ALLYL BENZENE |
| 3283 | ETHYL ETHER |
| 241 | BENZENE |

Note: Gray shading indicates odorants with an APF score ≤ -60 ; red font indicates similarity clustering of a subset of 27 out of 35 odorants chosen by pharmacophore virtual screening of OlfactionDB database; green shading indicates six odorants which were obtained for wet-lab uptake/inhibition assays. *, <http://molsim.sci.univr.it/bioinfo/web/>.

SHARED LIGANDS BETWEEN ORGANIC ANION TRANSPORTERS (OAT1 AND OAT6) AND ODORANT RECEPTORS
Wei Wu, Kevin T. Bush, Henry C. Liu, Christopher Zhu, Ruben Abagyan and Sanjay K. Nigam
DRUG METABOLISM AND DISPOSITION

A. Alignment of amino acid sequences of N-terminal regions of SLC22A20 from 20 species

Table with 3 columns: Species, Accession Number, and Amino Acid Sequence. Species include Human, Chimpanzee, Pygmy chimp, Horse, Marmoset, Pig, Elephant, Galago, Dog, Gibbon, Guinea pig, Rat, Hamster, Cow, Mouse, Rabbit, Opossum, Orangutan, Tasmanian devil, and Platypus. Sequences are aligned and color-coded by physicochemical properties.



B. Lineage report: Conservation of SLC22A20

Lineage report table showing taxonomic classification and BLAST hits for SLC22A20. Taxonomy includes Euteleostomi, Tetrapoda, Mammalia, Theria, Eutheria, Euarchontoglires, Primates, Simiiformes, Catarrhini, Hominoidea, and various mammalian orders. Species listed include Nomascus leucogenys, Pan paniscus, Pan troglodytes, Homo sapiens, Pongo abelii, Macaca mulatta, Callithrix jacchus, Otolemur garnettii, Cavia porcellus, Rattus norvegicus, Mus musculus, Cricetulus griseus, Oryctolagus cuniculus, Sus scrofa, Loxodonta africana, Equus caballus, Canis lupus familiaris, Bos taurus, Ailuropoda melanoleuca, Sarcophilus harrisi, Monodelphis domestica, Ornithorhynchus anatinus, Xenopus laevis, and Danio rerio. Hits range from 1 to 75.

Supplemental Figure 1. Orthologs of Slc22a20 are found in many mammalian genomes and identification of a conserved N-terminus region in SLC22A20/OAT6 orthologs. Full-length amino acid sequences of SLC22A20 from multiple species were aligned. (A) Strong conservation was observed within the N-terminal region of all species examined (A; amino acids are colored based on their physicochemical properties—red, small and hydrophobic; blue, acidic; grey, unusual amino acids). (B) A lineage report derived from taxonomic analysis of SLC22A20 in various species. BLAST hits for the N-terminus of human SLC22A20 (indicated by the numbers on the right-hand side of the figure) were found in species ranging from primates to bony fishes and indicates conservation of the N-terminal region of the gene in vertebrate species as far down as frog and fish.

SHARED LIGANDS BETWEEN ORGANIC ANION TRANSPORTERS (OAT1 AND OAT6) AND ODORANT RECEPTORS
Wei Wu, Kevin T. Bush, Henry C. Liu, Christopher Zhu, Ruben Abagyan and Sanjay K. Nigam
DRUG METABOLISM AND DISPOSITION

| A | Description | Max score | Total score | Query cover | E value | Max ident | Accession |
|---|---|-----------|-------------|-------------|---------|-----------|--------------------------------|
| | solute carrier family 22 member 20 [Mus musculus] | 189 | 189 | 100% | 4e-58 | 100% | NP_941052.1 |
| | solute carrier family 22 member 27 [Mus musculus] | 69.7 | 69.7 | 72% | 6e-15 | 55% | NP_599017.1 |
| | solute carrier family 22 member 9 [Mus musculus] | 66.6 | 66.6 | 72% | 7e-14 | 52% | NP_659034.1 |
| | solute carrier family 22 member 9 [Mus musculus] | 66.2 | 66.2 | 72% | 9e-14 | 52% | NP_666344.1 |
| | solute carrier family 22 member 6 [Mus musculus] | 64.7 | 64.7 | 100% | 3e-13 | 46% | NP_032792.2 |
| | solute carrier family 22 member 29 [Mus musculus] | 63.9 | 63.9 | 72% | 6e-13 | 52% | NP_766364.1 |
| | prostaglandin-specific organic anion transporter [Mus musculus] | 62.0 | 62.0 | 100% | 3e-12 | 38% | NP_759010.1 |
| | uncharacterized protein LOC434674 [Mus musculus] | 60.8 | 60.8 | 72% | 7e-12 | 50% | NP_001013842.2 |
| | integral membrane transport protein UST1R [Mus musculus] | 60.8 | 60.8 | 72% | 7e-12 | 50% | NP_795976.1 |
| | solute carrier family 22 member 12 [Mus musculus] | 59.3 | 59.3 | 100% | 3e-11 | 39% | NP_033229.3 |
| | solute carrier family 22 member 21 [Mus musculus] | 49.3 | 49.3 | 100% | 6e-08 | 36% | NP_062697.1 |
| | solute carrier family 22 member 4 [Mus musculus] | 47.0 | 47.0 | 100% | 3e-07 | 35% | NP_062661.1 |
| | solute carrier family 22 member 8 [Mus musculus] | 47.0 | 47.0 | 91% | 4e-07 | 39% | NP_112471.3 |
| | solute carrier family 22 member 5 [Mus musculus] | 47.0 | 47.0 | 100% | 4e-07 | 35% | NP_035526.1 |
| | solute carrier family 22 member 7 [Mus musculus] | 43.5 | 43.5 | 98% | 5e-06 | 35% | NP_659105.2 |
| | transportin-3 [Mus musculus] | 28.9 | 28.9 | 64% | 0.33 | 25% | NP_796270.2 |
| | olfactory receptor 167 [Mus musculus] ← | 28.1 | 28.1 | 25% | 0.60 | 50% | NP_667146.1 |

| B | Score = 28.9 bits (63), Expect = 9.6, Method: Compositional matrix adjust. Identities = 22/89 (25%), Positives = 33/89 (37%), Gaps = 22/89 (25%) |
|----------------------|--|
| NP_941052 OAT6 45 | PPRHQHPANYTEPTINVSQVWLRRAIPLNQHQDPEPCRRYVEPQ 89 |
| NP_667146 OLF167 128 | P H+ H N ++G W+ +I H FLYHATH-MNKRVCVMITGAWLGSINCAHTG----- 160 |
| NP_941052 OAT6 90 | WALLKFNASHGVGVA-----TEGCKDGVVYDRSIFPSTIV 123 |
| NP_667146 OLF167 161 | +AL P S + T C D WYV+ ++F STI+ YALQIPYCRSRAINHFFCDVFMILLACTDITWVYEVFVSTIL 204 |

| C | Description | Max score | Total score | Query coverage | E value | Max ident | Accession |
|---|--|-----------|-------------|----------------|---------|-----------|--------------------------------|
| | olfactory receptor 167 [Mus musculus] | 166 | 166 | 100% | 1e-53 | 100% | NP_667146.1 |
| | olfactory receptor Olr1568 [Rattus norvegicus] | 164 | 164 | 100% | 2e-52 | 96% | NP_001000532.1 |
| | olfactory receptor 2L5 [Homo sapiens] | 135 | 135 | 97% | 1e-41 | 79% | NP_001245213.1 |
| | olfactory receptor 2L2 [Homo sapiens] | 134 | 134 | 97% | 4e-41 | 80% | NP_001004686.1 |
| | olfactory receptor 2L8 [Homo sapiens] | 133 | 133 | 96% | 9e-41 | 80% | NP_001001963.1 |
| | olfactory receptor 2L3 [Homo sapiens] | 132 | 132 | 97% | 2e-40 | 77% | NP_001004687.1 |
| | olfactory receptor Olr1569 [Rattus norvegicus] | 129 | 129 | 100% | 4e-39 | 78% | NP_001000042.1 |
| | olfactory receptor Olr1567 [Rattus norvegicus] | 128 | 128 | 100% | 9e-39 | 77% | NP_001000043.1 |
| | olfactory receptor 166 [Mus musculus] | 125 | 125 | 100% | 8e-38 | 75% | NP_667279.1 |
| | olfactory receptor 168 [Mus musculus] | 124 | 124 | 100% | 2e-37 | 74% | NP_666469.1 |
| | olfactory receptor 2L13 [Homo sapiens] | 122 | 122 | 98% | 2e-36 | 75% | NP_787107.1 |
| | olfactory receptor 2AK2 [Homo sapiens] | 98.6 | 98.6 | 92% | 1e-27 | 59% | NP_001004491.1 |
| | olfactory receptor 318 [Mus musculus] | 93.2 | 93.2 | 100% | 6e-26 | 51% | NP_666712.2 |
| | olfactory receptor 320 [Mus musculus] | 93.2 | 93.2 | 100% | 6e-26 | 51% | NP_997113.1 |
| | olfactory receptor Olr1413 [Rattus norvegicus] | 90.5 | 90.5 | 100% | 7e-25 | 48% | NP_999986.1 |
| | olfactory receptor Olr1436 [Rattus norvegicus] | 90.1 | 90.1 | 92% | 8e-25 | 51% | NP_001000014.1 |
| | olfactory receptor Olr1437 [Rattus norvegicus] | 90.1 | 90.1 | 100% | 9e-25 | 45% | NP_001000015.1 |
| | olfactory receptor 30 [Mus musculus] | 90.1 | 90.1 | 100% | 1e-24 | 47% | NP_667089.1 |
| | olfactory receptor Olr1566 [Rattus norvegicus] | 89.4 | 89.4 | 100% | 1e-24 | 51% | NP_001000044.1 |
| | olfactory receptor 316 [Mus musculus] | 89.4 | 89.4 | 100% | 2e-24 | 48% | NP_00101818.2 |
| | olfactory receptor 2T2 [Homo sapiens] | 88.6 | 88.6 | 88% | 3e-24 | 51% | NP_001004136.1 |
| | olfactory receptor 2T35 [Homo sapiens] | 88.6 | 88.6 | 88% | 3e-24 | 51% | NP_001001827.1 |

Supplemental Figure 2. A 79 amino acid sequence in the N-terminus region of OAT6/SLC22A20 is conserved in other members of the SLC22 drug transporter family as well as in some olfactory receptors. (A) A NCBI blastp query using the amino acid sequence of the N-terminal region of the mouse SLC22A20 gene revealed significant sequence similarities to other mouse SLC22 family members (most of which appear to be organic anion transporters). In addition, the mouse olfactory receptor, Olfr167 (arrow), was also identified as having significant sequence homology. As described, this is particularly interesting since SLC22A20 is not only preferentially expressed in nasal epithelium (along with SLC22A6/OAT1), but has been shown to be capable of interacting with odorant molecules (Kaler et al., 2006; Kaler et al., 2007). Moreover, other odorant receptors also had BLAST hits (Olf166 and Olf168) for SLC22A20 but these had lower sequence homology scores (not shown). (B) Alignment of N-terminal region of SLC22A20 with OLF167 revealed the presence of a 79 amino acid fragment within the N-terminal region of OAT6 (NP_941052—starting at amino acid number 45 and ending at amino acid number 123) which had significant sequence homology to a 77 amino acid fragment in Olf167 (NP_667146—starting at amino acid number 128 and ending at amino acid number 204) (capital letters between the two sequence fragments indicate identical amino acids at that residue, + indicates conserved residues). (C) A NCBI blastp query of human, mouse and rat proteins using the 77 amino acid fragment of mouse OLF167 as a template revealed significant sequence similarities to multiple odorant receptors in these species.

SHARED LIGANDS BETWEEN ORGANIC ANION TRANSPORTERS (OAT1 AND OAT6) AND ODORANT RECEPTORS
 Wei Wu, Kevin T. Bush, Henry C. Liu, Christopher Zhu, Ruben Abagyan and Sanjay K. Nigam
 DRUG METABOLISM AND DISPOSITION

