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>	CID Name
8094	HEPTANOIC ACID
20210	BROMOHEXANOIC ACID
8892 16368	HEXANOIC ACID 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 62378 13187	OCTANOIC ACID
62378	DIHYDROJASMONE
13187	2-NONANONE HEXANEDIOIC ACID
190	NONANAL
13167 196 31289 8878 8158 2969	ALLYL HEPTANOATE
8158	NONANOIC ACID
2969	DECANOIC ACID
12813	GAMMA-DECALACTONE
385	HEPTANEDIOIC ACID
8175	DECANAL
548275	
10457	OCTANEDIOIC ACID
4409936 444294	PHENYLACETATE
	(-)-CAMPHOR
19707	2,3-HEXANEDIONE
22227 443157	(+)-DIHYDROCARVONE
14525	CITRONELLA(-) (+)-FENCHONE
6852393	ANDROSTENONE
7012	2-PHENYLBUTYRIC ACID
91604	LYRAL
10886	AMYL HEXANOATE NONANEDIOIC ACID BUTYL BUTYRYLLACTATE 2-HEXANONE
2266	NONANEDIOIC ACID
24114	BUTYL BUTYRYLLACTATE
11000	2 112/0 110112
11614	BUTYL FORMATE

7000	2 LIEDTANIONE
7802 45747	3-HEPTANONE
15717	ALLYL PHENYLACETATE
8174	1-DECANOL
439570	(-)-CARVONE
16724	(+)-CARVONE
8144	OCTANETHIOL
8914	1-NONANOL GAMMA-CAPROLACTONE
12756 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
	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
	ADE 00 f+ : :

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

```
Human np 001004326 maptdllalgsmgrfqlnhtallllpcgllachnflqnftaavpphhcrgfanhteastnd-sgawlratipldqlgapepcrrftkpcwallspnssi-pgaategckdgwynrsvfpstivme 127
  Chimpanzee XP 001169243 MAFTDLLDALGGMGRFQLIHMALLLLPCGLIACHNFLQNFTAAVPPHHCRGPANNTEASTND-SGAWLRATIPLDQLGAPEPCRHFTKPQWALLSPNSSI-PGAATEGCKDGWVYNRSVFPSTIVME 150
 Pygmy chimp xp 003828667 maftdlldalggmgrfqlihmallllpcgllachnflqnftaavpphhcrgpannteastnd-sgawlratipldqlgapepcrhftkpqmallspnssi-pgaategckdgwynrsvppstivme 150
        Horse XP 003362708 MAFTDLLDALGGVGRFQLIYTAMLLPCSLLACHNFLQNFTAAVPRHHCQGPTNHTVAATDN-SGAWLRATVPLDRLGDPEPCRRFTEPQWALLSPNTSV-QGAATEGCRDGWVYDRSVFPSTIVME 125
    Marmoset XP 002755570 MAFTDLLDALGGVGRFQLLHTALLLLPCSLLACHNFLQNFTAAVPPHHCRGPTNHTEA-TND-SENWLRATVPLDQLGAPEPCRRFTKPQWALPNPNSSL-PGAATEGCKDGWVYNRSVFPSTIVME 124
          Pig XP 003122588 MAFTDLLDALGGVGRFQLVYTALLLLPCCLLACHNFLQNFTAAVPRHHCQSPANHTVATDND-SGAWLRATVPLDRLGAPEPCRRFTKPQWALLSPNASV-HGAATEGCHDGWVYDRSVFPSTIVTE 125
      Elephant XP 003419642 MAFTDLLDTIGGYGRFGLYYTALLLLPCSLLACHNFLGNFTAAMPPHHCRGPANHTTAATND-SGAWLRATVPLDQDGAPVTCGRFTEPQMALLGSNTSI-HGVATEGCKDGWYYDRSVFPSTIVME 125
       Galago xp 003798691 maftdlldalggvgrfqfiytvllllpcgllachnflqnftaavprhhcrgptnhseattnd-sgawlraivpldhlgapepcqrftepqwallnsntsv-hrtatesckdgwvydhsifpstivme 125
                          MAFTDLLDSLGGVGRFQLIYTALLLLPCSLLACHNFLQNFTAAVPHHHCRGPTNHTAAATNT-SGAWLRAIVPLDQLGAPEPCRRFTEPQWALLSPNTSA-HEPATEGCKDGWVYDRSVFPSTIVME 125
         Dog XP 854865
       Gibbon XP 003274235 MAFTDLLDALGGIGRFQLIHTALLLLPCGLIACHNFLQNFTAAVPPHHCQGPANHTEASTND-SGAWLRATIPLDQLGAPEPCRRFTKPQWALLSPNSSV-PGAATEGCKDGWYYNRSVFPSTIVME 129
    Guinea Dig XP 003468271 MAFTULLDALGGVGRFOLVYTALLLLPGGLACHNFLONFTAAIPHHHCOHPANHT--ATND-SGAWLRATVPLDOHGVPEPCORFTEPOWALLSPNTSA-CGAATEPCKDGWVYDRSVFPSTIVME 123
          Rat NP 001099797 MAFTOLLDAIGGVGFFGLVYTALLLLPCGLIACHTFLGNFTAAAPPHHCGHPANYTEATTND-SGAWLRATIPLNQHGVPEPCGRYTEPQMALLEPNTSS-HGVATEGCKDGWVYDRSIFPSTIVME 125
     Hampster xp 003515423 maftdlldalggygrfofyytalllpcgliachtflonftaaapphhcwhpanhteastnd-sgawlratiplnogypepcorytepowdlksnass-hgvatedckdgwyydrsifpstivme 125
         COW XP 604969 MAFTDLLDALGGVGRFQLVSTALLLLPCSLLACHNFLQNFTAAVPPHHCQSPANHSAAAVYG-SGAMLRATVPLHPLGLPEPCRRFTRPQMALLSPNASV-DGAATEGCKDGWYYDRSVFPSTIVTE 125
                          MAFTDLLDALGGVGRFQLVYTALLLLPCGLLACHTFLQNFTAAAPPHHCQHPANYTEPTTNV-SGVWLRAAIPLNQHGDPEPCRRYVEPQWALLKPNASS-HGVATEGCKDGWVYDRSIFPSTIVME 125
       Rabbit XP 002723457 MAFTDLLDALGGVGRFQLVYTALLLPCGLLACHNFLQNFTAAMPPHHCRGPANLT---NT-SGVWLRATVPMDRLGALDSCRRFTEPQWVLLLPNASSGHGAATEGCVDGWEYDRSVFPSTIVTE 123
     Opossum xp 003341962 maftelldtiggvgrpgllytalllfpcsllachnflonftaavpohrcqeranstgsanft-aingwrvsvpedqhqrpepclrfvqpqwallhpnatv-rgaategcedgwtydlsifpstivte 125
    Orangutan XP 002821600 -----MGAARGLSS-PLPSNPSISFPL-TIKLEPLLEYWGNSLIFHHCRGPANXCEASTND-SGAWLRATIPLDQLGXPEPCQRSPSLQWPLLSPNSSV-PGAATEGCKDGWVYNRSVFPSTIVME 117
Tasmanian devil XP 003774222 MAFTELLDTIGGVGRFONLYMALLLFPCSLLACHNFLONFTAAVPLHHCRGHANETGGTNAT-EEGWWRASIPADRHORPEPCLRFVRPOWEILDPNATV-DDVATEGCEDGWTYDRSVFPSTIVTE 125
      Platypus XP 003430231 MGFSDLLDSIGGLGTFQMLQTALLLLPGCLFSCQNFLQNFSAAEPDHRCRLPSPDNRSGAGDlTGPLLRASVPLDAQGRPDRCLRYPRPQMQLLPNASA-EGLVTEGCLQGWVYERTVFLSTIVTE 131
```

PeHHCBSPANLIEASIND SGAWLRAIYPLDOLGSPEPCBREIEPQWALLSPNSS EGSATEGCKDGWYYRRSYFPSTIV

B. Lineage report: Conservation of SLC22A20

```
[vertebrates]
Euteleostomi
 Tetrapoda
                 [vertebrates]
   Mammalia
                   [mammals]
     Theria
                     [mammals]
                      [placentals]
    . . Euarchontoglires [placentals
                          [primates]
          Primates
            Simiiformes
                             [primates]
          . . Catarrhini

    Mominoidea

                               [primates]
                                                           177 7 hits
                 Nomascus leucogenys (White-cheeked Gibbon) -
                                                                     [primates]
                                                           175 8 hits
             . . Pan paniscus (bonobo) .....
                                                                     [primates]
                                                           176 8 hits
             . . Pan troqlodytes .....
        . . . . . Homo sapiens (man) .....
                                                           169
                                                                hits
                                                           149
                                                              9 hits
                 Pongo abelii (Orang-utan) .....
               Macaca mulatta (rhesus macaque) -----
                                                            80 1 hit
              Callithrix jacchus (common marmoset)
                                                           161 5
                                                                hits
                                                                     [primates]
          . Otolemur garnettii -
                                                               4 hits
                                                                     [primates]
    . . . Cavia porcellus (guinea pig) ------
                                                           144
                                                                hit
                                                                     [rodents]
     . . . Rattus norvegicus (brown rat) ......
                                                           140
                                                               3 hits [rodents]
    . . . Mus musculus (mouse)
                                                           139
                                                               1 hit
                                                                     [rodents]
    . . . Cricetulus griseus (Chinese hamsters) .....
                                                              3 hits [rodents]
                                                           124 8 hits [rabbits & hares]
         . Oryctolagus cuniculus (domestic rabbit) .....
    . . Sus scrofa (wild boar)
                                                               3 hits [even-toed unqulates]
        Loxodonta africana (African savannah elephant) ......
                                                           152 6
                                                                hits [placentals]
                                                           152
        Equus caballus (equine) .....
                                                              3 hits
                                                                     [odd-toed unqulates]
        Canis lupus familiaris (dogs) .....
                                                           152
                                                                hit
                                                                     [carnivores]
                                                           147
                                                                hits
         Bos taurus (cow) .....
                                                                     [even-toed ungulates]
                                                           113 2 hits
        Ailuropoda melanoleuca .....
                                                                     [carnivores]
    . Sarcophilus harrisii
                                                                hits
                                                                     [marsupials]
                                                            98 4 hits
      Monodelphis domestica ..
                                                                     [marsupials]
                                                            84 4 hits
     Ornithorhynchus anatinus (duck-billed platypus) -----
   <u>Xenopus laevis</u> (common platanna) -----
                                                                 hit
                                                                     [froqs & toads]
 Danio rerio (sebra fish)
```

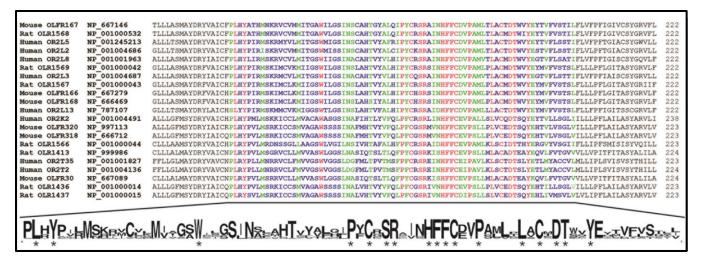
Supplemental Figure 1. Orthologs of Slc22a20 are found in many mammalian genomes and identification of a conserved N-terminus region in SLC22A20/OAT6 orthologs. Full-length amino acid sequences of SLC22A20 from multiple species were aligned. (A) Strong conservation was observed within the N-terminal region of all species examined (A; amino acids are colored based on their physiochemical 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

Α	Description	Max score	Total score	Query	E value	Max	Accession	B Score = 28.9 bits (63), Expec Identities = 22/89 (25%), Posi	tives =	33/89		aps = 22	/89 ((5%)
solute carrie	er family 22 member 20 [Mus musculus]	189	189	100%	4e-58	100%	NP_941052.1	P H+ H	N	++G 1		H		
solute carrie	er family 22, member 27 [Mus musculus]	69.7	69.7	72%	6e-15	55%	NP 599017.1	NP 941052 OAT6 90 WALLKPN	ASSHGVA		TEGCKD	GWVYDRSI	FPSTIV	123
solute carrie	er family 22 member 9 [Mus musculus]	66.6	66.6	72%	7e-14	52%	NP_659034.1	- +AL P	S +			WVY+ ++	F STI+	
solute carrie	er family 22 member 9 [Mus musculus]	66.2	66.2	72%	9e-14	52%	NP 666344.1	C Description	Max	<u>Total</u> score	Query	△ E value	Max	Accession
solute carrie	er family 22 member 6 [Mus musculus]	64.7	64.7	100%	3e-13	46%	NP_032792.2	olfactory receptor 167 [Mus musculus]	166	166	100%	1e-53	100%	NP 667146.1
	/ 3 22 - 1 20 M	000	00.0	700/	0- 40	500/	ND 700204.4	olfactory receptor Olr1568 [Rattus norvegicus]	164	164	100%	2e-52	96%	NP 001000532.1
solute carrie	er family 22. member 29 [Mus musculus]	63.9	63.9	72%	6e-13	52%	NP_766364.1	olfactory receptor 2L5 [Homo sapiens]	135	135	97%	1e-41	79%	NP 001245213.
prostagland	in-specific organic anion transporter [Mus n	62.0	62.0	100%	3e-12	38%	NP_759010.1	olfactory receptor 2L2 [Homo sapiens]	134 133	134	97% 96%	4e-41	80%	NP 001004686.
		00.0	00.0	700/	7. 10	500/	ND 004042042.2	olfactory receptor 2L8 [Homo sapiens] olfactory receptor 2L3 [Homo sapiens]	132	133	95%	9e-41 2e-40	77%	NP 001001963. NP 001004687.
uncharacter	ized protein LOC434674 [Mus musculus]	60.8	60.8	72%	7e-12	50%	NP_001013842.2	olfactory receptor Olr1569 [Rattus norvegicus]	129	129	100%	4e-39	78%	NP 001000042.
ntegral mer	mbrane transport protein UST1R [Mus muse	60.8	60.8	72%	7e-12	50%	NP 795976.1	olfactory receptor Olr1567 [Rattus norvegicus]	128	128	100%	9e-39	77%	NP 001000043.
								olfactory receptor 166 [Mus musculus]	125	125	100%	8e-38	75%	NP 667279.1
solute carrie	er family 22 member 12 [Mus musculus]	59.3	59.3	100%	3e-11	39%	NP_033229.3	olfactory receptor 168 [Mus musculus]	124	124	100%	2e-37	74%	NP 666469.1
solute carrie	er family 22 member 21 [Mus musculus]	49.3	49.3	100%	6e-08	36%	NP 062697.1	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.
solute carrie	er family 22 member 4 [Mus musculus]	47.0	47.0	100%	3e-07	35%	NP_062661.1	olfactory receptor 318 [Mus musculus]	93.2	93.2	100%	6e-26	51%	NP 666712.2
soluto carrio	er family 22 member 8 [Mus musculus]	47.0	47.0	91%	4e-07	39%	NP 112471.3	olfactory receptor 320 [Mus musculus] olfactory receptor Olr1413 [Rattus norvegicus]	93.2 90.5	93.2	100%	6e-26 7e-25	51%	NP 997113.1 NP 999986.1
solute curre	rating 22 member o into moscolosi	41.0	*****	0.70	10 01	0070	1121113	olfactory receptor Olr1413 [Rattus norvegicus]	90.1	90.1	92%	8e-25	51%	NP 001000014.
solute carrie	er family 22 member 5 [Mus musculus]	47.0	47.0	100%	4e-07	35%	NP_035526.1	olfactory receptor Olr1437 [Rattus norvegicus]	90.1	90.1	100%	9e-25	45%	NP 001000015
saluta camir	er family 22 member 7 [Mus musculus]	43.5	43.5	98%	5e-06	35%	NP 659105.2	olfactory receptor 30 [Mus musculus]	90.1	90.1	100%	1e-24	47%	NP 667089.1
Solute Callie	rianny 22 member / [Mus musculus]	40.0	40.0	3070	JE-00	0070	NF 053105.2	olfactory receptor Olr1566 [Rattus norvegicus]	89.4	89.4	100%	1e-24	51%	NP 001000044.
transportin-	3 [Mus musculus]	28.9	28.9	64%	0.33	25%	NP_796270.2	olfactory receptor 316 [Mus musculus]	89,4	89.4	100%	2e-24	48%	NP 001011818.
16 .		00.6	00.0	050	0.00	500/	ND 007440.4	olfactory receptor 2T2 [Homo sapiens]	88.6	88.6	88%	3e-24	51%	NP 001004136.1
olfactory rec	ceptor 167 [Mus musculus]	28.1	28.1	25%	0.60	50%	NP_667146.1	olfactory receptor 2T35 [Homo sapiens]	88.6	88.6	88%	3e-24	51%	NP 001001827.

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 (Olfr166 and Olfr168) for SLC22A20 but these had lower sequence homology scores (not shown). (B) Alignment of N-terminal region of SLC22A20 with OLFR167 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 Olfr167 (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 OLFR167 as a template revealed significant seguence 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



Supplemental Figure 3. Alignment of full-length amino acid sequences of 21 olfactory receptors from human, rat and mouse. Strong conservation was found within the amino acid sequences of the 77 AA fragment of ORs (amino acids are colored based on their conservation—red, identical across all receptors; green, conserved; blue, non-conserved). A Weblogo representation of the conserved 77 amino acid fragment found within the various olfactory receptors is depicted at the bottom; amino acid residues within this fragment which are conserved across all 21 receptors are indicated by an asterisk underneath the residue. Within this 77 amino acid fragment, more than 50% of the amino acid residues were found to be conserved in these ORs.