

## Genomics of selected human odorant receptors

Kirsten Knape<sup>1</sup>, Anton Beyer<sup>2</sup>, Anna Stary<sup>2,3</sup>, Gerhard Buchbauer<sup>1</sup>, Peter Wolschann<sup>2</sup>

<sup>1</sup> Department of Clinical Pharmacy and Diagnostics, University of Vienna, Vienna, Austria

<sup>2</sup> Institute of Theoretical Chemistry, University of Vienna, Vienna, Austria

<sup>3</sup> Present address: Computational Biomolecular Dynamics Group,  
Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany

Received 17 April 2008; Accepted 26 April 2008; Published online 9 July 2008

© Springer-Verlag 2008

**Abstract** Single nucleotide polymorphisms (SNPs) in odorant receptor genes may influence the protein sequence and consequently also the function of the receptors. An analysis of the HapMap data for human OR3A1 was performed and provided evidence that genetic differences subject to ancestry and gender can be recognized. A genomic comparison of individuals shows the diversity of odorant receptor genes and therefore potentially the variety of the sense of smell. At this time, two complete human genomes are available in public domain, which we used for this purpose.

**Keywords** Odorant receptor; OR3A1; Genomics; Single nucleotide polymorphism.

### Introduction

Olfaction is one of the most ancient senses [1] for humans as well as for animals. During evolution this sense was important for the finding and the identification of nutrition, for the warnings of danger (rotten food or enemies) as well as for reproduction. Moreover, human sensual pleasure (*e.g.*, the odor of perfume) is one important aspect of life quality. So this primal sense can be designated as a major agent for communication between the individual and its environment [2].

Odorants are volatile chemical compounds that have to possess some special molecular properties to reach the *regio olfactoria* – a small area of about 2.5 cm<sup>2</sup> located at the roof of the two nasal cavities of the human nose. Odorants are usually compounds with low molecular weight (up to ~300), with a relatively high vapor pressure, with some lipophilicity, but they have to be water soluble to some extent.

At the *regio olfactoria*, the odorant molecules have to pass a layer of mucous. The layer is enriched with odorant binding proteins (OBPs), which belong to the family of lipocalins that help to transport the odor molecules to the olfactory receptors (ORs) [3]. The reception occurs on the olfactory cilia – every olfactory neuron possesses 8–20 [1] of them – and so sensory transmissions can be started. Signals that our brain interprets as odor are produced by the ORs and transmitted to the olfactory bulb through neurons.

Odorant receptors are G-protein coupled receptors (GPCRs) and belong to class A sub-family (rhodopsin-like receptors, *e.g.* rhodopsin, olfactory, hormone, and cannabinoid receptors). GPCRs consist of seven membrane-spanning helices, which are connected by three extracellular and three intracellular loops. The G-protein binds on the intracellular side of the receptor. The seven transmembrane helices, which are arranged anticlockwise as viewed from the extracellular receptor side, are responsible for the anchoring of the receptor in the cell membrane. Especially helices 3–6 contain binding sites for ligands [4]. In

Correspondence: Anton Beyer, Institute of Theoretical Chemistry, University of Vienna, A-1090 Vienna, Austria. E-mail: Anton.Beyer@univie.ac.at

contrast to the intracellular and extracellular loops, the transmembrane domains of GPCRs are highly conserved. Some motifs in the amino acid sequence of the helices can be found in nearly all rhodopsin-like receptors. These motifs are important for the activation of the receptor.

In 2000, the first crystal structure of a mammalian GPCR – that of bovine rhodopsin – has been resolved (PDB ID: 1U19) [5]. Seven years later, the first structure of a human GPCR – the beta-2-receptor – has been solved (PDB ID: 2R4R, 2R4S) [6]. A higher resolution (2.40 Å) structure of the same receptor has been published just recently (PDB ID: 2RH1) [7]. This structure is highly similar to bovine rhodopsin in terms of the relative orientation of the seven helices.

No experimental structures of odorant receptors are known up to now. Nevertheless, for some receptors homology models based on rhodopsin as template have been generated and used for docking studies of related odor molecules [8–10]. As the sequence similarity between rhodopsin and the investigated receptors is rather poor (around 23%), evolutionary criteria have been additionally applied in some cases to support the alignment of the sequences and the structures. Such homology models have also been used to explain the association of helional and lilyal to odor receptors OR1A2 and Olfr49 [10, 11].

Generally, the odor impression results from the interaction of the odor molecules with various odorant receptors. The combination of the activation of several ORs leads to a distinct fragrance recognition in the brain [12, 13]. In special cases the fragrance molecules form association complexes with the odorant receptors without further activation, and this is the reason for a competitive hindrance of the association of activating odorants [14]. Such odor molecules act as antagonists.

Olfactory receptor genes constitute the basis for the sense of smell and are encoded by the largest mammalian gene superfamily. The human genome contains around 850 odorant receptor genes, from which approximately 400 are functional (The Human Olfactory Data Exploratorium (HORDE) [15]). The rest are pseudogenes, which are non-functional resulting from nonsense mutations, deletions, and frameshift [16]. For comparison, macaques possess approximately 700 functional olfactory receptor genes [17] and mice, rats, and dogs even about 1000 [18, 19].

The OR genes are present on all human chromosomes, except on chromosomes 20 and Y. Chromosome 11 contains by far the most OR genes.

Generally, ORs are distributed in genomic clusters with a typical intergenic distance of ~20 kb, only 48 ORs are singletons.

Recently, our group has performed studies on the odorant receptor cluster on human chromosome 17, which contains 18 odorant receptor genes [20]. After the extended search for orthologues of those odorant receptor genes in chimpanzee, rat, and mouse genome, comparisons subject to genomic location, similarities, and function were made. These analyses provide some interesting insights into the evolutionary relation between the mentioned species. To explore further details about the evolutionary development of odorant receptors we studied in detail single nucleotide mutations of the odorant receptor OR3A1 recognizing helional. Additionally, the comparison between two individual genomes with respect to various odorant receptor families is described.

## Results and discussion

### *Single nucleotide polymorphisms (SNPs) in odorant receptor genes*

A key aspect of research in genetics is associating sequence variations with heritable phenotypes. The most common variations are single nucleotide polymorphisms (SNPs), mutations of one single base, which occurs every 100–300 bases in the human genome and makes up about 90% of all variations [21]. For a variation to be considered as SNP, it must appear in at least 1% of the population.

SNPs can be located in both coding and non-coding regions of the gene. About 60% of all SNPs involve the replacement of cytosine (C) by thymine (T). The type of SNP can be – depending on the location – synonymous or non-synonymous. Synonymous variations do not have any consequences for the protein sequence (so called silent mutations) whereas non-synonymous variations cause changes in the protein sequence and therefore may alter the function of the protein.

Many SNPs have no detectable effect on cell function, but some others can help to determine the likelihood of developing a particular disease (*e.g.*, *Alzheimer's* disease [22]). Moreover, they can have a major impact on the human response to environmental insults, drugs, and other therapies. For that reason, SNPs are of great interest for biomedical research and for developing pharmaceutical products and medical diagnostics.

SNPs do not change much from generation to generation [21], but they are supposed to differ in various ethno geographic groups. Therefore, it can be hypothesized that there is a diversity of olfactory sensitivity amongst people, suggesting that also genetic differences can be found subject to different ancestry.

All known SNPs for human OR3A1 have been examined and will be discussed in detail in the following part. Human OR3A1 (synonym hOR17-40) genes are not only expressed in the nose but also in testes, and so they are supposed to play a role in sperm chemotaxis [23]. For that reason, also differences between men and women were regarded.

All SNPs annotated for OR3A1 were looked up using the SNP database (<http://www.ncbi.nlm.nih.gov/SNP/index.html>) and basic facts, like ID number, location, and concerned nucleic acids and protein residues, were collected for each SNP. Then the diversity in five different ancestries (European, Han Chinese, Japanese, Sub-Saharan African, and African American) as well as in both genders was analyzed and percentages of occurring SNPs were calculated. That was possible by the results of the International HapMap Project, an organization whose goal is to develop a haplotype map of the human genome (the HapMap) and which examined about 90 people (state: January 2008) of each descent and annotated the SNPs for each tested individual [24].

In the following Table 1, collected data for all 9 SNPs are presented, including SNP ID and location. Above mentioned percentages of occurrence of variation are given in the tables for every ancestry, distinguishing men and women. First numbers stand for total occurrence, first values in parenthesis for occurrence in all tested women, and second numbers in parenthesis for occurrence in all tested men. The first column shows the original genetic situation, the second column a mutation in one gene, and the third column a mutation in both alleles.

There are six SNPs in OR3A1 that are located in the non-coding region that means they are silent mutations. The other three SNPs (Rs16952828, Rs7218125, and Rs703903) are found in the coding region and, therefore, cause a mutation in the protein sequence. The degree of influence on olfactory perception depends on the location of the altered residue. In general, it can be assumed that mutations which take place in the seven transmembrane helices might influence the sense of smell because parts of these regions are important for ligand binding. Figure 1 shows an overview of the location of the

three SNPs on the coding region in OR3A1 on chromosome 17 in a receptor model developed recently by our group.

Rs16952828 – situated in the last third of helix 2 – might affect ligand binding. Nevertheless, extraordinary changes caused by a mutation in this place cannot be expected, because according to *Leffingwell et al.* [1] the largest binding site cavity falls in the transmembrane domains TM 3–7. Although it is not absolutely clear, it can be supposed that Rs703903 has an impact on olfaction because it is located at a highly conserved position at the end of TM 3.

One of those three mutations, Rs7218125, causes a loss of protein function that means it creates a pseudogene. Figure 1 shows that this SNP is located in the extracellular loop between helix 2 and 3. This region might be important for ligand binding. It is obvious, that none of the tested individuals possesses this variation on both genes (*A/A*) regardless of which ancestry. Another interesting fact is, that this SNP does not concern tested Europeans as well as no Han Chinese and Japanese women. Regarding tested Sub-Saharan Africans and African Americans it is apparent that men are stronger affected by the mutation than women (only 8.3% of all tested Sub-Saharan African women show this variation whereas 18.2% of all tested men of same descent).

Another SNP situated on the coding region is Rs703903. It leads to a change from glutamine to arginine in the protein sequence. Surprisingly, all analyzed Han Chinese and Japanese people possess a mutation on both genes (*T/T*). Also Sub-Saharan African and European individuals have a variation on one gene (*C/T*) to a higher percentage than primary genetic composition (*C/C*). Europeans also show a quite high occurrence (24.4%) of a mutation on both genes, Sub-Saharan Africans only 3.3%.

The same phenomenon can be observed in Rs16952819 and Rs9893392, both located in the non-coding region. Especially Rs9893392 seems to be under a high pressure of genetic drift. Thus, there are no European, Han Chinese, and Japanese people who feature the original code (*C/C*), and only 17.4% of African Americans. Instead, the majority exhibits a mutation on both genes. 64% of Europeans, about 93% of Han Chinese people, and even all tested Japanese people show this double mutation.

It is noticeable, that some populations are more burdened with mutations than other ones. This fact was described in a work by *Lohmueller et al.* [25]. Having analyzed the *DNA* of 20 European Americans

**Table 1** SNPs in OR3A1 with their location, concerned amino acids, and percental occurrence in different populations. Total percentage (occurrence in all tested women; occurrence in all tested men)

Rs16952828

Genomic location: Chr. 17, 3142369 to 3142420 (–)

Alternate residues: S, G

	<i>T/T</i>	<i>C/T</i>	<i>C/C</i>
European	100%	0%	0%
Han Chinese	100%	0%	0%
Japanese	100%	0%	0%
Sub-Saharan African	93.8% (93.3%; 82.0%)	6.2% (2.7%; 18.0%)	0%
African American	91.3% (83.4%; 100%)	8.7% (16.6%; 0%)	0%

Rs7218125

Genomic location: Chr. 17, 3142327 to 3142378 (–)

Alternate residues: K, \*

	<i>T/T</i>	<i>A/T</i>	<i>A/A</i>
European	100%	0%	0%
Han Chinese	95.6% (100%; 90.9%)	4.4% (0%; 9.1%)	0%
Japanese	95.5% (100%; 90.9%)	4.5% (0%; 9.1%)	0%
Sub-Saharan African	90% (94.6%; 86.8%)	10% (5.4%; 13.2%)	0%
African American	87% (91.7%; 81.8%)	13% (8.3%; 18.2%)	0%

Rs703903

Genomic location: Chr. 17, 3142227 to 3142278 (–)

Alternate residues: Q, R

	<i>C/C</i>	<i>C/T</i>	<i>T/T</i>
European	27.9% (17.4%; 38.6%)	47.7% (58.7%; 36.4%)	24.4% (23.9%; 25.0%)
Han Chinese	0%	0%	100%
Japanese	0%	0%	100%
Sub-Saharan African	47.8% (56.8%; 41.5%)	48.9% (37.8%; 56.6%)	3.3% (5.4%; 1.9%)

Rs16952819

Genomic location: Chr. 17, 3139769 to 3139820 (–) (non-coding region)

	<i>C/C</i>	<i>C/T</i>	<i>T/T</i>
European	0%	50.0% (66.6%; 52.2%)	50.0% (33.4%; 47.8%)
African American	52.3% (8.4%; 18.3%)	13.0% (66.6%; 36.3%)	34.7% (25.0%; 45.4%)

Rs9893392

Genomic location: Chr. 17, 3140447 to 3140498 (–) (non-coding region)

	<i>C/C</i>	<i>C/T</i>	<i>T/T</i>
European	0%	36.0% (29.8%; 41.1%)	64.0% (70.2%; 58.0%)
Han Chinese	0%	6.7% (8.7%; 4.5%)	93.3% (91.3%; 95.5%)
Japanese	0%	0%	100%
African American	17.4% (16.6%; 18.1%)	43.5% (50%; 36.4%)	39.1% (33.4%; 45.5%)

Rs7214042

Genomic location: Chr. 17, 3142676 to 3142727 (–) (non-coding region)

	<i>C/C</i>	<i>A/C</i>	<i>A/A</i>
European	100%	0%	0%
Han Chinese	100%	0%	0%
Japanese	100%	0%	0%
Sub-Saharan African	100%	0%	0%
American African	100%	0%	0%

(continued)

**Table 1** (continued)

Rs4790459

Genomic location: Chr. 17, 3139926 to 3139977 (–) (non-coding region)

	G/G	A/G	A/A
European	100%	0%	0%
Han Chinese	86.7% (78.3%; 95.5%)	13.3% (21.7%; 4.5%)	0%
Japanese	97.8% (100%; 95.7%)	2.2% (0%; 4.3%)	0%
Sub-Saharan African	100%	0%	0%

Rs2676604

Genomic location: Chr. 17, 3142696 to 3142747 (–) (non-coding region)

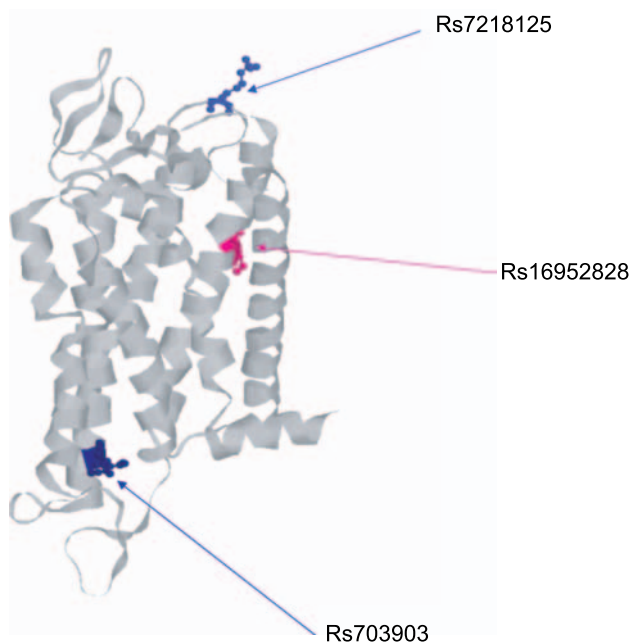
	C/C	C/T	T/T
European	100%	0%	0%
Han Chinese	100%	0%	0%
Japanese	100%	0%	0%
Sub-Saharan African	90% (89.2%; 90.6%)	10% (10.8%; 9.4%)	0%

Rs2317892

Genomic location: Chr. 17, 3140682 to 3140733 (–) (non-coding region)

	G/G	A/G	A/A
European	85.1% (87.7%; 82.5%)	14.9% (12.3%; 17.5%)	0%
Han Chinese	26.8% (39.2%; 14.2%)	46.6% (39.1%; 54.5%)	26.6% (21.7%; 31.8%)
Japanese	26.2% (22.8%; 26.1%)	51.1% (54.5%; 47.8%)	24.4% (22.7%; 26.1%)
Sub-Saharan African	100%	0%	0%
African American	91.3% (83.4%; 100%)	8.7% (16.6%; 0%)	0%

and 15 African Americans, these researchers found that Europeans contain a lower genetic diversity compared to African Americans. These results are

**Fig. 1** Location of SNPs in receptor model of human OR3A1

in agreement with the hypothesis that after migrating out of Africa, human groups experienced a population “bottleneck” in which their genetic diversity was reduced. Expanding quickly, the European population accumulated new mutations before the old ones that cause negative implication could be weeded out [26]. This might be an explanation for the above mentioned fact that Africans are supposed to have a larger functional olfactory receptor repertoire than other populations.

But the question remains, why just the strongest genetic drift can be observed in non-coding regions and whether mutations in this area are important for odorant receptors.

Some groups showed that in general changes in non-coding DNA can have a tremendous impact on phenotype including drug response [27], disease susceptibility [28–30], and the evolution of differences between species [31]. In terms of odorant receptors in special, the dimension of influence by mutations in non-coding regions is not proved yet, and much more experiments and analysis will be required to draw conclusions how SNPs and odorant perception are related to each other.

### Comparison of two human genomes

Just seven years ago, the first sequence and analysis of the human genome has been published simultaneously in two draft versions, one in *Nature* [32] and the other one in *Science* [33]. Both of these versions are composite sequences derived from haploid genomes of mostly anonymous donors [34].

But such studies cannot detect rare alleles related to diseases or reflecting individual peculiarities. Only sequencing the genome of individuals who are willing to reveal relevant details about their personal life will offer the opportunity to explore the complex interactions between different alleles, their regulatory elements and the individual's environment, which determines typical physical characteristics, and therefore make it possible to correlate genomic variations with a specific phenotype [35].

Now, two entire genome sequences of two individuals, *Craig Venter* and *James Watson*, are accessible for public and will help to encourage the development of an area of personalized genomic medicine with the intent to identify and prevent diseases and to create personalized medical therapies.

In May 2007, the sequencing of the six billion base pairs of *James Watson's DNA* has been finished and released to the public through the web site [www.jimwatsonsequence.cshl.edu](http://www.jimwatsonsequence.cshl.edu). Only a few months later, in September 2007, researchers of the *Craig Venter* Institute in Rockville, USA, have published the diploid genome sequence of *Craig Venter*, in which his two sets of chromosomes (one inherited from his mother and one from his father) are represented [35]. It is available in GenBank under the code ABBA00000000, though not yet assembled. This genome project shows that an individual human has about 4.1 million *DNA* variants (including 1.2 million variants that have never been seen before) and 3.2 million of these are SNPs. It conveys that human to human variation is 5 to 7 times greater compared to previous human genome analyses [35]. Over the next 3 years, an international team plans to sequence the genomes of 1000 individuals. The so called "1000 Genome Project" will serve as gold-standard reference for analysis of human variations [36].

For odorant receptor research it would be interesting to have the opportunity to explore why individual people have a different sense of smell. So, maybe in future it will be possible to create the most pleasant fragrance for an individual or a special popu-

lation after having analyzed their odorant receptor genes and built the ligand that stimulates the desired receptor(s) most efficiently.

To show that there exist vast variations from human to human, a choice of olfactory receptors genes of *Venter's* and *Watson's* genome was compared with each other. Additionally, they were compared to the sequences from HORDE as reference. For this comparison, the six human odorant receptors with known ligands plus four additional receptors including one pseudogene (OR10B1P) were selected.

First of all, *Venter's* and *Watson's* nucleic acid sequences were searched with the help of NCBI Blast respectively *Jim Watson* Genome Browser, using HORDE sequences as query. After having aligned them with ClustalW, they were transformed into protein sequences with Transeq in case that the alignments showed some differences in the nucleotide sequence. The protein sequences were aligned to identify some silent mutations and to examine the variations in detail.

When comparing all those odorant receptors, it is conspicuous that *James Watson's* sequences are always completely concordant with the HORDE references. Unlike that, *Craig Venter's* olfactory genes show many variations in comparison to the reference and therefore also to *Watson's* OR genes. All found differences concerning *Venter* are summed up in Table 2.

As apparent, three of the examined receptors (OR1D2, OR1G1, and OR7D4) are completely alike both in *Venter's* and *Watson's* genome. Some others show many differences, like OR51M1 with seven

**Table 2** Variations in *Craig Venter's* genome (compared to HORDE sequences)

Odorant receptor	Number of variations in nucleotide sequence	Number of variations in protein sequence	Number of silent mutations
OR1A1	3	1	2
OR1A2	4	2	2
OR3A1	1	1	0
OR1D2	0	0	0
OR1G1	0	0	0
OR7D4	0	0	0
OR52D1	5	1	4
OR52J3	4	2	1
OR51M1	7	6	1
OR10B1P	7	?	?

OR10B1P = pseudogene

variations in nucleotide sequence and still six differences in protein sequence, or OR1A2, that features four mutations in nucleic acid sequence and two in protein sequence. This receptor shall serve as example to examine the variations in detail.

There are four mutations in *Venter's* olfactory gene. The first mutation, on position 15, is annotated in SNP Database as Rs2241093. This synonymous variation feature 35% of Caucasian people like *Craig Venter* [37]; 65% possess the ancestral base thymine. The next mutation, which is named Rs2241092 and located at position 264, is also a synonymous one. Only 7.5% of people of *Venter's* ancestry hold this genomic variant. Mutations in positions 766 and 879, Rs2241091 and Rs12150427, are both non-synonymous and can be found in around 30% of Caucasian people. These two variations determine differences in the protein sequences.

*Craig Venter* possesses a cysteine residue in position 256 instead of a glycine, and another cysteine in position 293 in place of a tryptophan residue. Regarding chemical properties, the change of those amino acids causes substantial differences. In both cases, a non-polar residue was replaced by a polar one, which contains a sulfide and therefore has a higher chemical reactivity. Moreover, varieties can be suggested considering the size of residues. Tryptophan is a large amino acid containing an aromatic indole group while cysteine is somewhat smaller.

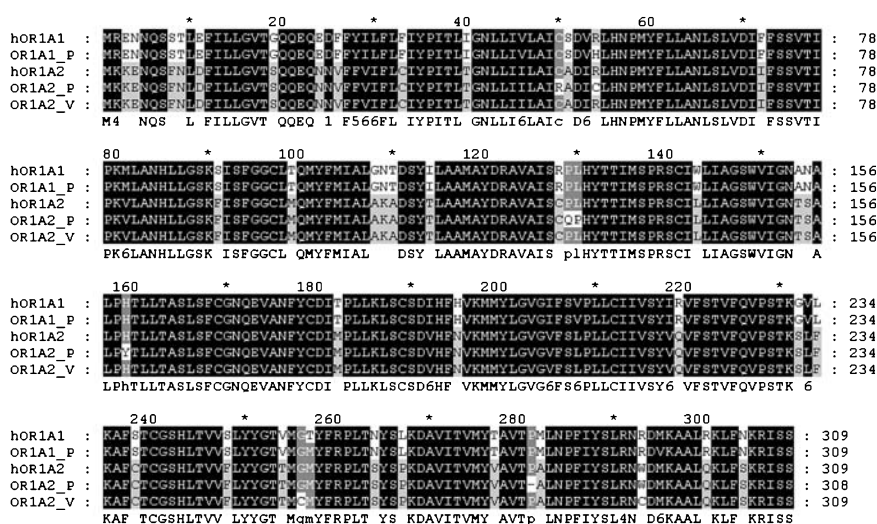
Probably even more important than the kind of substitution is the location. If the mutation takes

place in a highly variable region it might not cause any shifts in receptor function whereas changes concerning binding pockets will influence the ligand binding and therefore the receptor function and sense of smell.

Position 293 is situated at the intracellular C-terminus, which follows TM7. This region is not very important for ligand binding but for binding of the G-protein and other signalling molecules. However, it can be assumed that this variation might not influence the function of OR1A2. In contrast, position 256 is located near a putative binding site in helix 6. Therefore, it is possible, that mutations in this position might influence the ligand binding indirectly.

To get an idea about the conservation of those positions, an alignment of protein sequences of OR1A2 and the closely related OR1A1 were made, containing human sequences as well as chimpanzee ones because of the evolutionary relation. Obviously, all sequences feature a glycine in position 256 except *Craig Venter*. Residues in position 293 vary: human and chimpanzee OR1A1 show an arginine residue, but OR1A2 sequences all have a tryptophan with the exception of *Craig Venter*. This might confirm the hypothesis that position 256 is of higher importance than position 293.

This example shows the possible differences in odorant receptor function. Regarding the fact that there are around 800 human odorant receptors, of those many contain mutations from human to human, it is virtually certain that two individuals like *Venter*



**Fig. 2** Alignment of protein sequences of OR1A1 and OR1A2: HORDE sequence (hOR), chimpanzee sequence (P), and *Venter's* sequence (V)

and *Watson* have a different sense of smell with unequal preferences and disfavours, which also cause different patterns in behavior. To entirely discover the correlation between the complex sense of smell and inter-individual variations it will be a long way but the “1000 Genome Project” will be one step forward.

## Acknowledgements

We are grateful to Ms. *Martina Ziehegraser* for technical assistance.

## References

- Leffingwell JC *et al.* (2002) Leffingwell Reports, vol. 2 (No. 1), [www.leffingwell.com](http://www.leffingwell.com)
- Zarzo M (2007) *Biol Rev Camb Philos Soc* 82:455
- Pelosi P, Baldaccini NE, Pisanelli AM (1982) *Biochem J* 201:245
- Filmore D (2004) *Modern Drug Discovery* 7(11):24
- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) *Science* 289:739
- Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK (2007) *Nature* 450:383
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) *Science* 318:1258
- Singer MS (2000) *Chem Senses* 25:155
- Katada S, Hirokawa T, Oka Y, Suwa M, Touhara K (2005) *J Neurosci* 25:1806
- Sary A, Suwattanasophon C, Wolschann P, Buchbauer G (2007) *Biochem Biophys Res Commun* 361:941
- Schmiedeberg K, Shirokova E, Weber HP, Schilling B, Meyerhof W, Krautwurst D (2007) *J Struct Biol* 159:400
- Malnic B, Hirono J, Sato T, Buck LB (1999) *Cell* 96:713
- Buck LB (2004) *Nutr Rev* 62:S184
- Oka Y, Nakamura A, Watanabe H, Touhara K (2004) *Chem Senses* 29:815
- Safran M, Chalifa-Caspi V, Shmueli O, Olender T, Lapidot M, Rosen N, Shmoish M, Peter Y, Glusman G, Feldmesser E, Adato A, Peter I, Khen M, Atarot T, Groner Y, Lancet D (2003) *Nucleic Acids Res* 31:142
- Vanin EF (1985) *Annu Rev Genet* 19:253
- Gilad Y, Przeworski M, Lancet D (2004) *PLoS Biol* 2:E5
- Rouquier S, Blancher A, Giorgi D (2000) *Proc Natl Acad Sci USA* 97:2870
- Rat Genome Sequencing Project Consortium (2004) *Nature* 428:493
- Knappe K, Beyer A, Sary A, Buchbauer G, Wolschann P (2008) submitted
- SNP Fact Sheet: [http://www.ornl.gov/sci/techresources/Human\\_Genome/faq/snps.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/faq/snps.shtml)
- Genome News Network, [www.genomenewsnetwork.org/resources/whats\\_a\\_genome/Chp1\\_1\\_1.shtml](http://www.genomenewsnetwork.org/resources/whats_a_genome/Chp1_1_1.shtml)
- Vosshall LB (2004) *Curr Biol* 14:918
- Thorisson GA, Smith AV, Krishnan L, Stein LD (2005) *Genome Res* 15:1591
- Lohmueller KE, Indap AR, Schmidt S, Boyko AR, Hernandez RD, Hubisz MJ, Sninsky JJ, White TJ, Sunyaev SR, Nielsen R, Clark AG, Bustamante CD (2008) *Nature* 451:994
- Check-Hayden E (2008) *Nature* 451:876
- Marcuello E, Altes A, del Rio E, Cesar A, Menoyo A, Baiget M (2004) *Int J Cancer* 112:733
- Ueda H, Howson JMM, Esposito L, Heward J, Snook H, Chamberlain G, Rainbow DB, Hunter KMD, Smith AN, Di Genova G, Herr MH, Dahlman I, Payne F, Smyth D, Lowe C, Twells RCJ, Howlett S, Healy B, Nutland S, Rance HE, Everett V, Slink L, Lam AC, Cordell HJ, Walker NM, Bordin C, Hulme J, Motzo C, Cucca F, Hess JF, Metzker ML, Rogers J, Gregory S, Allahabadi A, Nithiyanathan R, Tuomilehto-Wolf E, Tuomilehto J, Bingley P, Gillespie KM, Undlien DE, Ronningen KS, Guja C, Ionescu-Tirgoviste C, Savage DA, Maxwell AP, Carson DJ, Patterson CC, Franklyn JA, Clayton DG, Peterson LB, Wicker LS, Todd JA, Gough SCL (2003) *Nature* 423:506
- Lazzaro BP, Scurman BK, Clark AG (2004) *Science* 303:1873
- Tournamille C, Colin Y, Cartron JP, Levankim C (1995) *Nat Genet* 10:224
- Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yoh C, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Rust AG, Pan ZJ, Schilstra MJ, Clarke PJC, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H (2002) *Science* 295:1669
- Human genomes, public and private (2001) *Nature* 409(6822):745
- The human genome (2001) *Science genome map*. *Science* 291(5507):1218
- Gross L (2007) *PLoS Biol* 5:e266. doi: 10.1371/journal.pbio.0050266
- “Fact Sheet – First Publication of an Individual Diploid Human Genome Sequence”, <http://www.jcvi.org/>
- Kaiser J (2008) *Science* 319:395
- Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AWC, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC (2007) *PLoS Biol* 5:2113