

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Subclustering of class I and II type human olfactory receptors (ORs), and detection of conserved motifs using TM-MOTIF tool.

Balasubramanian Nagarathnam*.

Department of Biotechnology, K.S.Rangasamy College of Technology, Tiruchengode-637 215. Tamilnadu.

ABSTRACT

Detection of odors is mediated by olfactory receptors (ORs). Human ORs belong to class A type of GPCRs and are discriminated into class I (sense water-borne odors) and II (sense air-borne odors) type. As odor specificity plays a crucial role in discriminating human ORs as class I and class II type, it is mandatory to identify subclusters of class I and II type human OR clusters. In the current study, sequences of class I OR cluster (54 sequences of one cluster) and class II OR clusters (317 sequences of 9 clusters) from DOR database were used and through NJ method of phylogeny, 45 OR subclusters were obtained, 100 representative ORs have been proposed further for docking studies. Conserved motifs were identified using TM-MOTIF tool to address cluster specificity. To conclude, phylogenetic approach helps to subcluster class I and II type human OR clusters and the findings could be used further to train machine learning approaches to predict unannotated class A type GPCRs. And the proposed representative ORs could be useful for molecular modelling, prediction of binding sites for select ligands to design intra-nasal drugs and to promote aromatherapy.

Keywords: Olfactory receptors, odorant receptors, chemosensory receptors, GPCRs, Class I and II type olfactory receptors, subclusters.

**Corresponding author*

INTRODUCTION

“Sense of smell”- a process of olfaction, is mediated by olfactory receptors (ORs). And ORs belong to class A type of GPCRs, which are G-protein coupled, seven-transmembrane proteins located on the surface of the dendritic cilia of olfactory neurons. Understanding the mechanism of odor recognition, related pathway is not only necessary for biological or chemical perspective, but also for its powerful socio-cultural phenomenon.

Nobel Laureates Buck and Axel have primarily explained the role of olfactory receptors, organization of olfactory system in humans in their landmark paper, published in the year 1991[1]. Thereafter, the olfactory science has brought into the light and tremendous efforts have been made in understanding ORs at various levels such as creating data repositories [2], olfaction in lower chordates, animal models such as *Drosophila* [3-5], *C.elegans* [6], signaling pathway, ion channels and so on. ORs of various taxa show diverse functional mechanisms such as implications in daily life, pharmaceutical Industry (aromatherapy), cosmetic Industry (scent /perfume manufacturing), olfactory-sexual function, olfacto-neural communication, agricultural pest control and olfactory disorders.

Earlier studies showed remarkable findings on identifying and documenting ORs in particular human ORs [7]. Studies on phylogenetic approach in discriminating class I and II type receptors to sense water and air-borne odors in higher eukaryotes, diverse ligands provide essential knowledge on OR class types as well as ligand specificity [8,9], and an extended knowledge on agonist and antagonist of mammalian ORs prevail scope for ORs in pharmacological applications [10]. Prior publications on *in-silico* approaches on ORs such as sequence analysis, inter and intra genome phylogeny, DOR- database [11], identification of conserved motifs and amino acid substitutions (AAS) in predicted helix and loops using TM-MOTIF package [12], cross- genome clustering of GPCRs i.e., human, fly and worm GPCR clusters [13], provide a strong background to employ subclustering approach to OR clusters and to enumerate cluster specificity, OR subclusters.

As odor specificity plays crucial role in discriminating human ORs as class I (to sense water borne odors) and class II type (to sense air borne odors), it is mandatory to subcluster human class I and II type OR clusters and to recommend representative ORs for molecular modelling and virtual screening to design drugs, particularly for intra-nasal drug delivery. Since remote resource/ knowledge is available on OR subfamilies/ OR subclusters so far, the current study on subclustering of class I and II type OR clusters could provide vast scope to identify cluster specificity in terms of subcluster arrangements/distributions, and conserved motifs. Also, the outcome could recommend few representative ORs for molecular modelling and docking studies also to train SVM dataset for the prediction of putative class A GPCRs.

METHODOLOGY

Sequence retrieval from Database of Olfactory receptors – (DOR):

As mentioned in prior literature [11], distinct cluster namely HSC1 related to **class I type receptors** (also known as fish-like ORs to sense water borne odors) and ORs dispersed in other clusters called **class II type receptors** (also known as mammalian-like ORs to sense air-borne odors) in human OR phylogeny were considered for the current study (refer figure 1). Sequences of class I (54 sequences of one cluster) and II (317 sequences of 9 clusters) type ORs were collected from DOR database and saved in separate files such as HSC1.fasta , HSC2.fasta and so on for 10 OR clusters, wherein HS denotes *Homo sapiens* , C1 refers cluster number as one. Naming of clusters, sequence ID were followed as in DOR database and subclusters were denoted as sub_clus followed by the number of subcluster.

Alignment Procedure using MAFFT

MAFFT online alignment server [14] was used to align OR sequences of class I and II type using parameters such as JTT 200 for scoring matrix with gap opening penalty of 1.53. The initial alignment was exported to MEGA 7 software [15] and at this stage, careful editing was done to refine the quality of alignment. The obtained final alignment session was saved in MEGA 7 format (*file.mas*) to construct phylogenetic tree.

Subclustering of class I and II type human OR clusters

The multiple sequence alignment (MSA) of class I and II type human OR clusters were used to construct NJ (neighbor joining) method of phylogeny for 1000 bootstrap replicates. The generated tree session files (file.mts) and radial, circular, dendrogram displays of phylogeny were used to analyze tree topology and to enumerate OR subclusters (refer figure 1). The obtained class I and II type OR subclusters with respective candidate ORs were tabulated (refer table 1) along with proposed representative OR sequences (100 ORs).

Detection of conserved motifs /AAS using TM-MOTIF

Class I (HSC1), II (HSC2-HSC10) type ORs were used to detect conserved motifs / AAS and motifs were identified at 60% level of conservation using TM-MOTIF tool [12].

RESULTS AND DISCUSSION

In the current study, Class I (to sense water-borne odors) and II (to sense air-borne odors) type OR clusters were subjected to subclustering approach to distribute into OR subclusters. The ten OR clusters (namely HSC1 to HSC10) derived from DOR database were subclustered into 45 OR subclusters (refer figures 2.A and 2.B also 3 to 12) and the distributed candidate ORs for respective subclusters have been given in table 1. The resulted OR subclusters not only illustrates the ligand diversity but also cluster/sequence specificity towards odor (s) and property(s). From the obtained subclusters, 100 representative ORs / OR sample sequences have been proposed (marked with * in table 1) and can be further explored for secondary structural details and docking studies with medicinally important phytoconstituents.

Distribution of OR subclusters

The diverse ligands, physiochemical parameters, solubility in particular, odor spectrum, evolutionary trend are the important causes for sub grouping or subclusters of candidate ORs from OR clusters. In the current study, an appreciable no. of OR subclusters have been arrived through subclustering approach and is due to the diverse sequence properties of class I and II type ORs, in otherwords its due to ligand specificity. The distribution of subclusters of HSC1 to HSC10 was as follows: 4, 4, 4, 7, 4, 3, 5, 4, 4, and 6 (refer figure 3-12). Due to subclustering approach, class I OR (HSC1) showed only 4 OR subclusters (refer figure 3), but class II type OR clusters namely HSC4 (refer figure 6), and HSC10 (refer figure 12) showed significant no. of OR distributions and reported seven and six subclusters respectively, which reveals the fact that the odor diversity is higher in class II type ORs than in class I type ORs. It can also be explained in other way that during the process of evolution, the chemical sensing/odor recognition for air soluble odors has increased tremendously than the water soluble odors (fish like ORs) in terrestrial vertebrates i.e., higher eukaryotes like mammals.

Proposed representative OR Sequences

Notably, cluster 6 (HSC6) exhibited three distinct subclusters/sub branches in tree topology (refer figure 8) with limited no. of OR sequences. This particular cluster is a good illustration to explain ligand specificity of OR sequences in turn to refer to "cluster specificity". And the proposed representative ORs / OR sample sequences could be highly useful further to propose for docking studies with diverse odors and to implement for intra-nasal drug delivery method. And the establishment of sequence features such as PSSM profile, conserved motifs of these representative ORs could be highly useful to train the positive dataset for machine learning approaches and further to apply for test datasets with unannotated/putative class A type ORs.

Prediction of conserved motifs using TM-MOTIF tool

- Human class I OR sequences (54 ORs in HSC1) showed average sequence identity as 44% and rest of the nine clusters (HSC2-HSC10) of class II type ORs showed sequence identity ranging from 44 -54 % . As prediction of conserved motifs at 60% of conservation could help to identify the cluster specific motifs/ sequence features, TM-MOTIF tool was utilized to detect cluster specific motifs of class I and II type ORs (refer figure 13). The characteristic motif DRYVAIC was distributed as DRYVAIC, VAICHP, DRYVAICKPLHY,

MAYDRYVAICNPLLY, MAYDRY, MAYDRYAAIC, DRY, YDRVYAIC, DRYVAIC, and YDRYVAICHPL at TM3 helix & ICL2 loop regions of cluster 1 to 10.

- Notably motifs such as PGL (N¹), PMY (ICL1, TM2), IFWF (TM2, ECL1), QMFFIH (TM3), PLRY, ILT (ICL2), LAC (ECL2), YGL (ECL2, TM5), HRFG (ECL3), LNPI (TM7), VKTKQI (C¹) were observed exclusively in cluster I (class I type ORs). KAFSTC (ICL3, TM6) motif was observed in almost all the class II type OR clusters, except clusters 3.
- PKM (TM2, ECL1) motif was observed in most of the class II type OR clusters such as 2, 3, 5, 10. ISF motif (ECL1, TM3) was observed in almost all the clusters of class II type ORs.
- In ECL2 motifs such as LAC, LKL, LAC, LSCSDT, LSC, LSCS, PLL, LAC, KAF, and LAC were observed in class II type receptors. And almost all cluster retained NPXXY motif at TM7 & C¹
- These findings can be used as patterns to train machine learning approaches, which further helps not only to detect class A type of GPCRs, but also helpful to discriminate the unannotated sequence for class I or II type category.

Human ORs with an Out group

An attempt was made to produce a phylogeny for select class I and class II type ORs with an out-group namely OR22a (refer figure 14). In the resulted topology, human ORs were not clustered with insect ORs, since insect OR exhibits “reverse topology”. Also it is an evolutionarily distant candidate to mammalian ORs and it stayed distantly in the tree topology. When an out-group was introduced, class I OR cluster showed “cluster specificity” remarkably and did not co-cluster with available class II type ORs. This illustrates the sequence features of class I type OR and its ligand specificity (sense – water borne odors). Among the HSC2-HSC10 class II type ORs, except few clusters such as cluster 4, 5, 6, rest of all the class II type ORs exhibited its cluster specificity in retaining strong relationship with their respective candidate ORs. Thus an out-group plays a role in understanding the retentionship of cluster specificity, contribution of candidate ORs in organizing OR clusters.

Figure 1: Flow chart for methodology

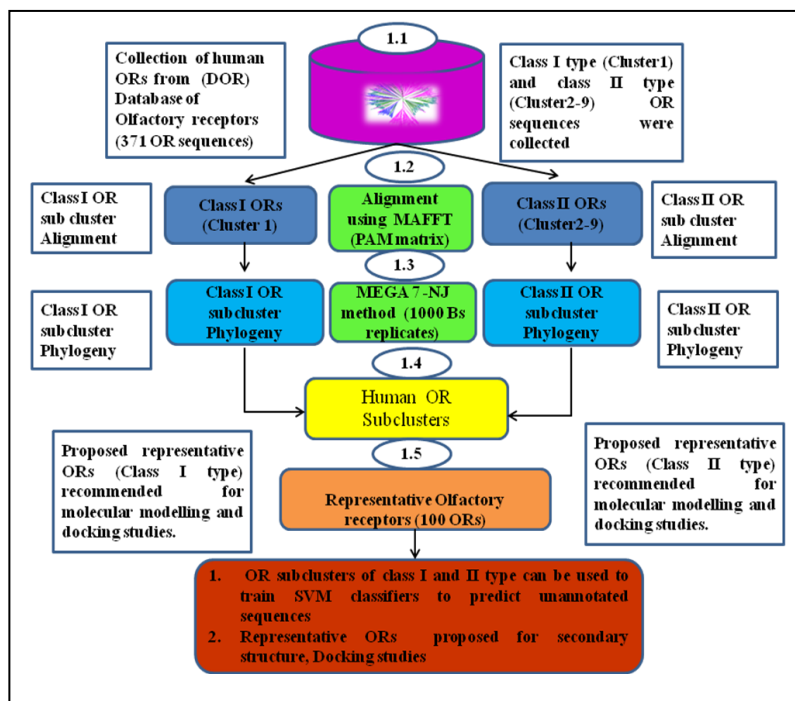


Figure 1 : Flow chart illustrates the steps involved to construct OR subclusters , in which steps such as data collection , alignment procedure, phylogeny, human OR subclusters, representative ORs (100 nos) have been denoted from step 1.1 to 1.5 with respective parameters.

Figure: 2.A: Bar diagram depicts the distribution of OR subclusters and **2.B:** Line diagram shows the no. of proposed representative ORs

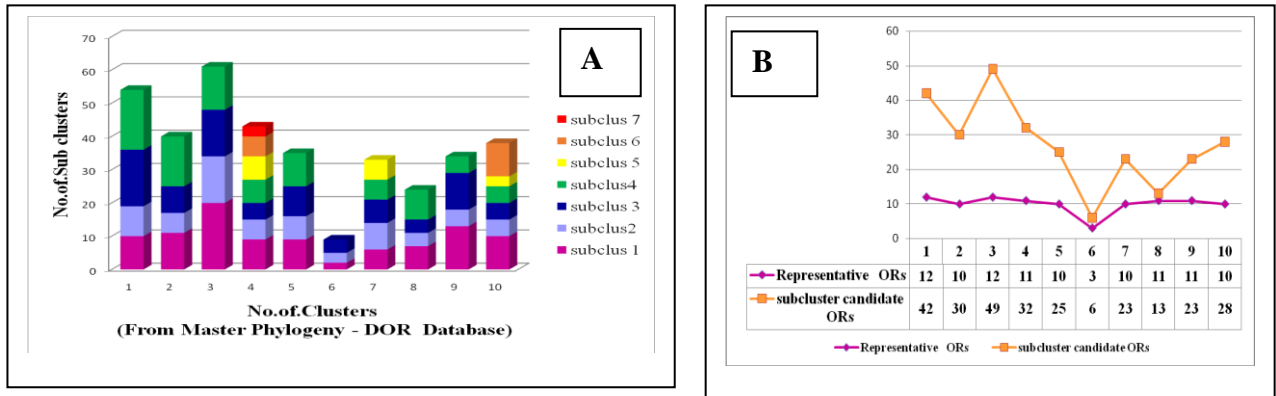


Figure 2.A: Bar diagram illustrates the distribution of OR subclusters, **Figure 2.B:** line diagram shows the distribution of candidate ORs and no. of proposed representative ORs observed in OR subclusters.

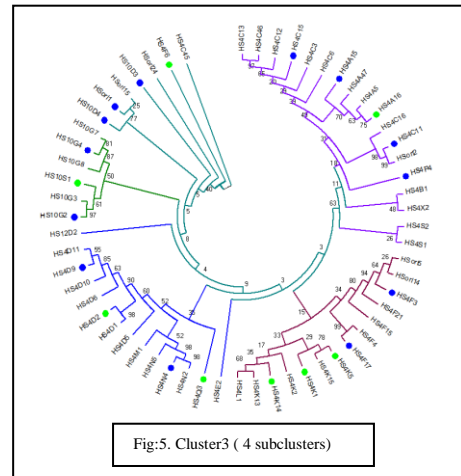
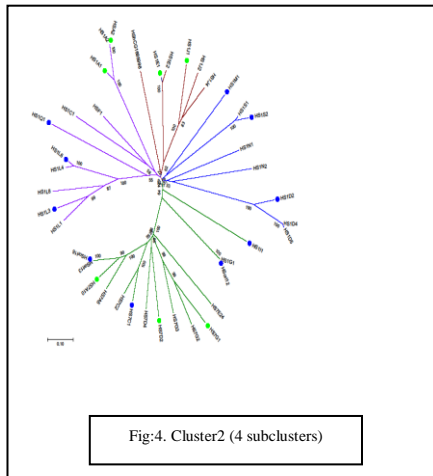
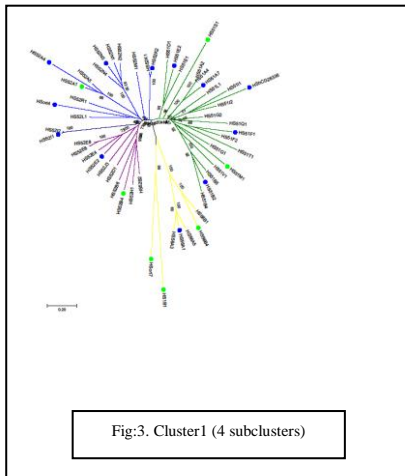


Figure: 3-5 Illustrates the phylogenetic display of subclusters obtained for cluster 1 to 3.

Figure: 6-8 Illustrates the phylogenetic display of subclusters obtained for cluster 4 to 6.

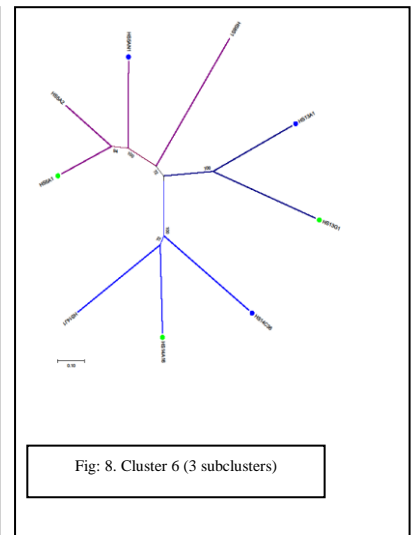
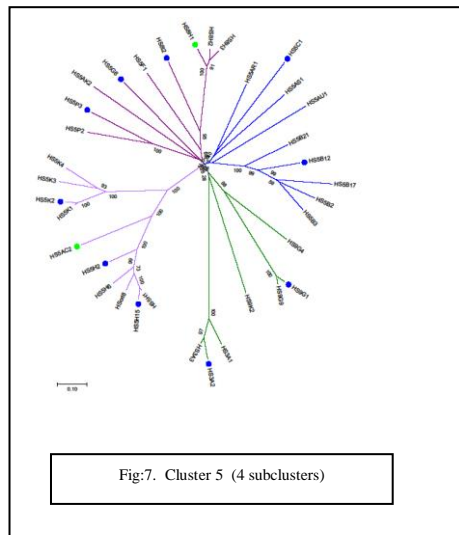
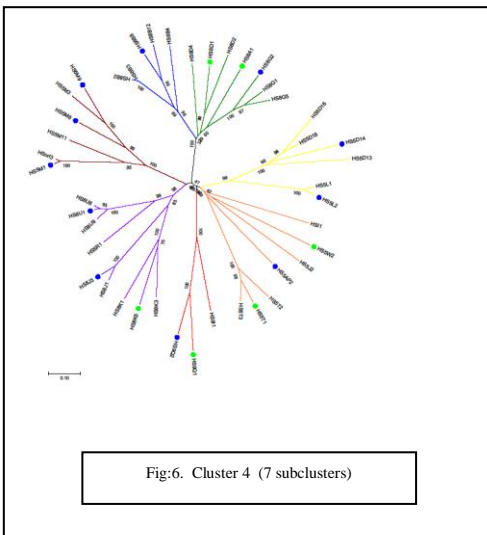


Figure: 9-11 Illustrates the phylogenetic display of subclusters obtained for cluster 7 to 9.

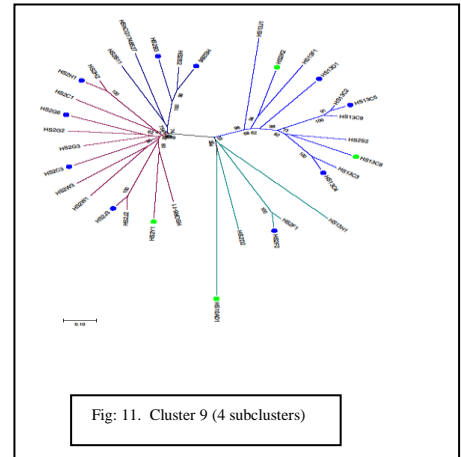
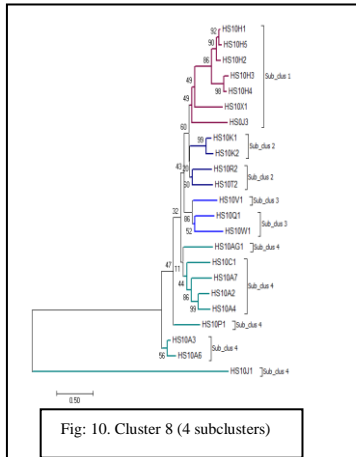
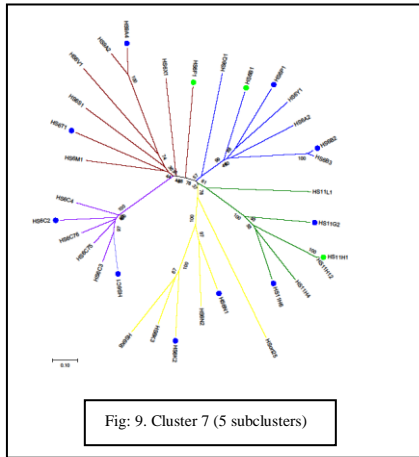


Figure: 12&13 Illustrates the phylogenetic display of subclusters obtained for cluster 10 and detection of conserved motifs using TM-MOTIF tool.

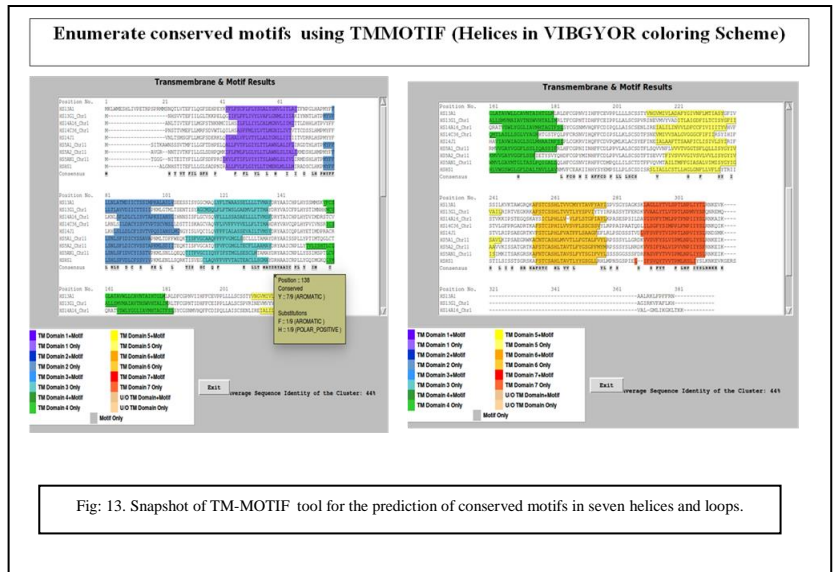
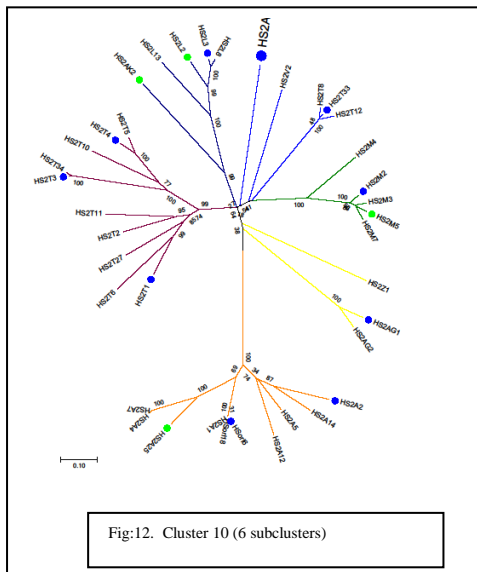
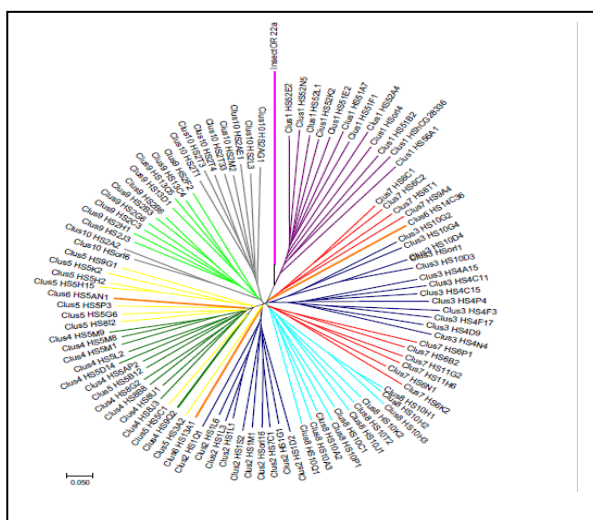


Figure 14 : Phylogeny of human class I and II type ORs with an out-group



Note : The phylogeny with select class I , class II ORs and an out-group clearly illustrates the cluster specificity of Class I and II type OR sequences along with respective cluster no has been given for reference. Distinct cluster, HSC1 - class I type receptors given in purple color and fly OR – an out-group marked in pink color.

S. No	Table 1 : List of OR sequences observed in OR subclusters and * indicates the representative ORs							
Cluster: 01								
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4				
1	HS52B2	*HS52I1	HS51D1	HS56B1				
2	HS52H1	HS52I2	*HS51E2	HS56B4				
3	HS52B4	HS52L1	HS51E1	HS56A5				
4	HS52B6	*HSorI4	HS51S1	*HS56A1				
5	HS52D1	HS52R1	HS51A2	HS56A3				
6	HS52J3	HS52A1	HS51A4	HS1B1				
7	*HS52E2	*HS52A4	*HS51A7	HSorI7				
8	HS52E4	HS52A5	HS51L1					
9	HS52E6	HS52N4	HS51I1					
10	HS52E8	*HS52N5	*HS51CG28336					
11		HS52N1	HS51I2					
12		HS52N2	HS51G2					
13		HS52M1	HS51Q1					
14		HS52K1	*HS51F1					
15		*HS52K2	HS51F2					
16			HS51T1					
17			HS51G1					
18			HS51M1					
19			HS51V1					



20			HS51B5				
21			HS51B4				
22			*HS51B 2				
	Cluster: 02						
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4			
1	HS1L1	HShCG1 685998	*HS1M1	*HS1I1			
2	*HS1L3	HS1E1	HS1S1	HS1G1			
3	HS1L8	HS1E2	*HS1S2	*HSorI 12			
4	HS1L4	HS1J1	HS1N1	HS7E2 4			
5	*HS1L6	HS1J2	HS1N2	HS7G1			
6	*HS1Q1	HS1J4	*HS1D2	HS7G2			
7	HS1C1		HS1D4	HS7G3			
8	HSF1		HS1D5	HS7D2			
9	HS1A1			HS7D4			
10	HS1A2			*HS7C 1			
11	HS1A2			HS7C2			
12				HS7A5			
13				HS7A1 0			
14				HSorI1 3			
15				*HSorI 16			
	Cluster: 03						
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4			
1	HS4C13	*HS4F3	HS4E2	HS12D 2			
2	HS4C46	HSorI5	HS4Q3	HS10G 7			
3	HS4C12	HSorI14	*HS4N4	*HS10 G4			
4	*HS4C15	HS4F21	HS4N2	HS10G 8			
5	HS4C3	HS4F15	HS4N5	HS10S 1			
6	HS4C6	HS4F4	HS4M1	HS10G 3			
7	*HS4A15	*HS4F17	HS4D5	*HS10 G2			
8	HS4A47	HS4K5	HS4D2	*HS10 D3			
9	HS4A5	HS4K15	HS4D1	HSorI2 4			
10	HS4A16	HS4K1	HS4D6	*HS10 D4			
11	HS4C16	HS4K2	HS4D10	HSorI1 5			
12	*HS4C11	HS4K14	*HS4D9	*HSorI 1			
13	HSorI2	HS4K13	HS4D11	HS4C4 5			
14	HS4B1	HS4L1		HS4F6			



15	HS4X2						
16	*HS4P4						
17	HS4S2						
18	HS4S1						
Cluster: 04							
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4	sub cluster 5	sub cluster 6	sub cluster 7
1	HS8K3	*HS5M1	HS8B2	HS8D4	HS5D1 6	HSI1	HS9I1
2	HS8K5	HSorI3	HS8B3	HS8D1	HS5D1 8	HS5W2	HS9Q1
3	HS8K1	HS5M11	*HS8B8	HS8D2	*HS5D 14	HS5J2	*HS9Q 2
4	HS8J1	*HS5M8	HS8B12	HS8A1	HS5D1 3	*HS5A P2	
5	*HS8J3	HS5M3	HS8B4	*HS8G 2	HS5L1	HS5T2	
6	HS5R1	*HS5M9		HS8G1	*HS5L 2	HS5T1	
7	HS8U9			HS8G5		HS5T3	
8	*HS8U1						
9	HS8U8						
Cluster: 05							
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4			
1	HS5H1	HS5P2	HS5AR1	HS9G4			
2	*HS5H15	*HS5P3	*HS5C1	*HS9G 1			
3	HSorI8	HS5AK2	HS5AS1	HS9G9			
4	HS5H6	*HS5G6	HS5AU 1	HS9K2			
5	*HS5H2	HS5F1	HS5B21	HS3A1			
6	HS5AC2	*HS8I2	*HS5B1 2	*HS3A 2			
7	HS5K1	HS8H1	HS5B17	HS3A3			
8	*HS5K2	HS8H2	HS5B2				
9	HS5K3	HS8H3	HS5B3				
10	HS5K4						
Cluster: 06							
	sub cluster 1	sub cluster 2	sub cluster 3				
1	HS5A1	*HS13A 1	*HS14C 36				
2	HS5A2	HS13G1	HS14A1 6				
3	*HS5AN1		HS14J1				
4	HS8S1						
Cluster: 07							



	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4	sub cluster 5		
1	*HS6C1	HS6M1	HS6Q1	HS11L1	HSorl25		
2	HS6C3	*HS6T1	HS6B1	*HS11G2	*HS6N1		
3	HS6C75	HS6S1	*HS6P1	HS11H1	HS6N2		
4	HS6C76	HS6V1	HS6Y1	HS11H12	*HS6K2		
5	*HS6C2	HS9A2	HS6A2	HS11H4	HS6K3		
6	HS6C4	*HS9A4	*HS6B2	*HS11H6	HS6K6		
7		HS6X1	HS6B3				
8		HS6F1					
Cluster: 08							
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4			
1	*HS10H1	HS10K1	HS10V1	HS10AG1			
2	HS10H5	*HS10K2	*HS10Q1	*HS10C1			
3	*HS10H2	HS10R2	HS10W1	HS10A7			
4	*HS10H3	*HS10T2		*HS10A2			
5	HS10H4			HS10A4			
6	HS10X1			*HS10P1			
7	HS0J3			*HS10A3			
8				HS10A6			
9				*HS10J1			
Cluster: 09							
	sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4			
1	HSOR6-11	HS2B11	HS13J1	HS13H1			
2	HS2Y1	HShCG1748527	HS2K2	HS2F1			
3	HS2J2	*HS2B3	HS13F1	*HS2F2			
4	*HS2J3	HS2B2	*HS13D1	HS2D2			
5	HS2W1	*HS2B6	HS13C2	HS10AD1			
6	HS2W3		*HS13C5				
7	*HS2C3		HS13C9				
8	HS2G3		HS2S2				
9	HS2G2		HS13C8				
10	*HS2G6		HS13C3				
11	HS2C1		*HS13C4				
12	*HS2H1						
13	HS2H2						

Cluster: 10						
sub cluster 1	sub cluster 2	sub cluster 3	sub cluster 4	sub cluster 5	sub cluster 6	
1	*HS2T1	HS2AK2	*HS2AE1	HS2M4	HS2Z1	*HS2A2
2	HS2T6	HS2L13	HS2V2	*HS2M2	*HS2AG1	HS2A14
3	HS2T27	HS2L2	HS2T8	HS2M3	HS2AG2	HS2A5
4	HS2T2	*HS2L3	*HS2T33	HS2M5		HS2A12
5	HS2T11	HS2L8	HS2T12	HS2M7		*HSorI6
6	*HS2T3					HSorI18
7	HS2T34					HS2A1
8	HS2T10					HS2A25
9	*HS2T4					HS2A4
10	HS2T5					HS2A7

CONCLUSION

In the current study, NJ method of phylogeny approach, yielded 45 OR subclusters. The observed no. of OR subclusters were high in class II type receptors than in class I type receptors, which refers to the ligand diversity of class II type OR. The recommended OR sequences and identified conserved motifs of class I and II type could be useful to train machine learning approaches to predict putative class A type GPCRs, and mainly useful to discriminate sequences into class I and II type ORs. Separately, a study with an out-group illustrated clearly the cluster specificity of class I and class II type ORs distinctly. The results from sequence analysis like representative ORs could be useful in future for molecular modelling, virtual screening with select ligands to design drugs, in particular for intra-nasal drugs to promote aromatherapy.

ACKNOWLEDGEMENT

I would like to acknowledge DST (WOS-B) scheme for the grant and K.S.Rangasamy College of Technology for infrastructure and other facilities. Thanks to mentors, parents and family.

REFERENCES

- [1] Buck, L. and Axel, R. "A novel multigene family may encode odorant receptors: a molecular basis for odor recognition", Cell, 1991, Apr 5;65(1):175-87. PubMed PMID: 1840504.
- [2] Crasto, C., Marenco, L., Miller, P. and Shepherd, G. "Olfactory Receptor Database: a metadata-driven automated population from sources of gene and protein sequences", Nucleic Acids Res, 2002, Jan 1;30(1):354-60. PubMed PMID: 11752336; PubMed Central PMCID: PMC99065. Vol. 30 pp. 354-60.
- [3] Siddiqi O, Olfaction in *Drosophila*. Chem Senses, 1991, 3:79-96
- [4] Warr, C., Clyne, P., de Bruyne, M., Kim, J. and Carlson, J.R. "Olfaction in *Drosophila*: coding, genetics and e-genetics" ChemSenses, 2001, Feb;26(2):201-6. PubMed PMID:11238252.
- [5] Das A, Chiang A, Davla S, Priya R, Reichert H, Vijayraghavan K, Rodrigues V. "Identification and analysis of a glutamatergic local interneuron lineage in the adult *Drosophila* olfactory system", Neural Syst Circuits, 2011, Jan 26;1(1):4. doi: 10.1186/2042-1001-1-4. PubMed PMID: 22330097; PubMed Central PMCID: PMC3257541.
- [6] Robertson, H.M. and Thomas, J.H. "The putative chemoreceptor families of *C. elegans*", WormBook. 2006 Jan 6:1-12. Review. PubMed PMID: 18050473; PubMed Central PMCID:PMC4781013.

- [7] Glusman G, Yanai I, Rubin I, Lancet D. "The complete human olfactory subgenome", *Genome Res.* 2001 May;11(5):685-702. PubMed PMID: 11337468.
- [8] Zozulya, S., Echeverri, F. and Nguyen, T. "The human olfactory receptor repertoire" *Genome Biol*, 2001, 2001;2(6):RESEARCH0018. Epub 2001 Jun 1. PubMed PMID: 11423007;PubMed Central PMCID: PMC33394.
- [9] Niimura, Y. and Nei, M. "Comparative evolutionary analysis of olfactory receptor gene clusters between humans and mice", *Gene*, 2005, Feb 14;346:13-21. Epub 2005 Jan 7. PubMed PMID: 15716120.
- [10] Kato A, Touhara K. "Mammalian olfactory receptors: pharmacology, G protein coupling and desensitization". *Cell Mol Life Sci*, 2009, Dec; 66(23):3743-53. doi: 10.1007/s00018-009-0111-6. Epub 2009 Aug 4. Review. PubMed PMID: 19652915.
- [11] B. Nagarathnam, Snehal D. Karpe, K. Harini, KannanSankar, Mohammed Iftekar, D. Rajesh, S. Giji, G. Archunan, V. Balakrishnan, M. Gromiha, W. Nemoto, K. Fukui and R. Sowdhamini. "DOR- A Database of Olfactory Receptors- integrated repository for sequence and secondary structural information of olfactory receptors in selected eukaryotic genomes", *Bioinformatics and Biology Insights*, 2014;8 147-158 .
- [12] Balasubramanian Nagarathnam, KannanSankar, Varadhan Dharnidharka, Veluchamy Balakrishnan, Govindaraju Archunan and Ramanathan Sowdhamini, "TM-MOTIF: an alignment viewer to annotate predicted transmembrane helices and conserved motifs in aligned set of sequences", *Bioinformatics*, 2011, 7(5):214-221
- [13] B Nagarathnam, S Kalaimathy, V Balakrishnan and Ramanathan Sowdhamini, "Cross-genome clustering of human and *C. elegans* G-protein coupled receptors", *Evolutionary Bioinformatics*, 2012;8 229-259.
- [14] Katoh K, Kuma K, Miyata T, Toh H. " Improvement in the accuracy of multiple sequence alignment program MAFFT", *Genome Inform*, 2005, 16(1):22-33. PubMed PMID:16362903.
- [15] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods", *Mol Biol Evol*, 2011, Oct;28(10):2731-9. doi: 10.1093/molbev/msr121. Epub 2011 May 4. PubMed PMID: 21546353; PubMed Central PMCID: PMC3203626