Molecular Identification and Characterisation of Olfactory Receptor Genes among Fruit and Insect Eating Bats

Steffi Christiane Ramesh, Sowmya V Bhat and Thangavel Karuppudurai*

Department of Animal Behaviour and Physiology, School of Biological Sciences, Madurai Kamaraj University, Madurai-625 021, India

Abstract

Among vertebrates, olfaction is deemed the oldest and most valuable modes of sensory perception, nevertheless remain the least understood modalities. Most organisms use olfaction at varying grades in all aspects of their life to detect food, avoid predators and for mate choice. The olfactory receptor (OR) gene repertoire is the largest gene family within the mammalian genome. Although the OR gene superfamily constitutes 3 to 6% of mammalian genes and is well annotated and completed in the human and mouse genomes, it is still unclear which odorants bind to which receptors and how this complex process translates into recognising a specific smell. Examining the relatively closely related species that exhibit a high dietary diversity and olfaction can help discern the olfactory receptor genes identification and characterisation. Bats are potentially useful for this purpose. Bats represent one of the most fascinating mammals for studying the OR genes identification and characterisation, since bats exhibit an excellent olfactory performance. The OR genes identification studies among bat species have been rare in India. In fact, until recently, there is no systematic study of OR genes identification and characterisation in bats. In this study, we employ bioinformatics and molecular biology approaches to identify the unique and diverse OR genomic repertoire in bats. Our sequencing results suggest that both fruit and insect eating bat species expressed different OR genes. Over all a total of 37 OR genes (9 gene families) were identified from 10 different bat species. We generated a global multiple alignment of the deduced amino acid sequences of 37 OR genes from 10 bat species. The most variable region was found to be TM segments 3, 4 and 5, within which 17 hypervariable regions were identified. These regions which constitute the odorant complementarity determining regions are the potential sites for ligand binding. Taken together, our results suggest that the total number of OR genes and families vary widely among both fruit and insect eating bats. The wide variety of sensory specializations and modalities in bats could explain the variety and distinctness of the bat OR repertoire if reflected in the OR gene diversity.

Keywords: Chiroptera • Bats • Olfaction • G-protein coupled receptor • OR genes

Introduction

Olfaction is one the most valuable modes of sensory perception and is used in varying stages in all aspects of life [1,2]. Most animals have evolved an acute sense of smell to perceive odors existing in the environment. In vertebrates, the olfactory receptors that are expressed in the olfactory sensory neurons are mainly responsible for detection of these odorant molecules, which are later converted to electrical signals and interpreted by the brain [3]. Olfaction is essential and the most valuable modes of sensory perception within mammals providing the source for extraordinary sensitivity required to distinguish sexual and environmental cues [4,5]. It helps to estimate the favourable and toxic food sources, habitat and mating preferences, predator avoidance, species discrimination, navigation, identifying marked territory etc [4,6-8].

Olfactory Receptor genes (ORs) are seven transmembrane G-protein coupled receptors that detects several odor molecules present in the environment [9-11]. These OR genes encompass the largest gene superfamily with approximately more than 1000 genes, accounting for around 6% of the protein-coding genes in a typical mammalian genome [4,5,12,13]. OR genes in mammals are nearly 1 kb in length, intronless and are found in clusters on most chromosomes [11,14]. The number of mammalian ORs differs greatly between species, with approximately 850 in humans compared to more than 4300 in elephants, typically reflecting the species' dependence on olfaction [1,15]. The number of OR genes differs significantly between species, according to bioinformatics analyses of diverse mammalian genome sequences [8]. It appears that the environment of each species has an impact on the number of OR genes [1].

Based on amino acid sequence, olfactory receptors in mammals can be evidently categorized into two groups, Class I or Class II. Class I ORs tend to bind hydrophilic odorants and Class II ORs hydrophobic odorants, but the functional difference between the two classes is not well understood [16]. These two classes are further divided into four gene families (OR 51, OR 52, OR 55 and OR 56) and nine genes families (OR 1/3/7, OR 2/13, OR 4, OR 5/8/9, OR 6, OR 10, OR 11, OR 12 and OR 14) respectively [17]. Olfactory receptors signal using a "labeled line" system where each olfactory sensory neuron expresses only one OR. Thus, the activated receptor expressed by a particular neuron is indicated by an action potential. Furthermore, ORs in the nose use combinatorial coding system. Instead of each scent being encoded by one neuron, a particular combination of neurons firing at the same time specifies a particular scent. Hence, mammals have the capacity to recognize a massive number of odors [18-20].

Olfactory sensitivity of a species may be influenced by the behavioural importance of an odorant [21,22]. It has also been reported that when fasting, olfactory acuity increases, allowing some mammalian species to identify food and ambient aromas more easily, such as those of predators [23].

Chemicals that bind to olfactory receptors and are transduced into electrical signals are called odorants [24]. In terms of its volatility, each odorant has different attributes. At high concentrations, some odorants smell qualitatively different. Higher concentrations of any odorant are likely to result in more extensive binding to various receptor types [25,26]. Odors stimulate a regimented pattern of activation of the olfactory bulb, but the mechanisms by which this map is transformed into an odor code are still unclear.

The number of OR genes differ significantly between species, according to bioinformatics analyses of diverse mammalian genome sequences [8]. It appears that the environment of each species has an impact on the number of OR genes [1]. Previous studies suggest that mice and rats have 1000 to 1200 functional OR genes in their respective genomes, while cows and opossums have a similar number [4,5]. Higher primates, however, typically have lower proportion of OR genes. The human genome has 400 functional OR genes as well as, intriguingly, 400 OR pseudogenes [27,28]. Orangutans and macaques have even fewer OR genes than chimpanzees, who are almost as likely as humans to have functioning OR genes [28,29].

Bats (Chiroptera) represent one of the most fascinating mammalian groups with distinctive adaptations. Bats use olfaction at a wide range for communication and foraging [30]. Olfaction combined with echolocation and vision, form the set of senses that enable bats to orient in space and find food [31]. Bats possess a distinctive and diverse OR gene repertoire. Till date only a few behavioral and field experimental studies have been conducted to understand the olfactory system in chiropterans. Behavioral trials on Carollia bat species showed that odor plays an important role in the bats' behavioral response to real and artificial fruit and form distinctive patterns in their echolocation behavior during exploration flights, enabling Carollia to find ripe fruit [31]. In a field experiment conducted, it was observed in the frugivorous bats that, they separate foraging and commuting time mediated by an interplay of Chiropteran olfactory ability and the strength of olfactory cues produced by different Chiropterans. Not many studies regarding functional characterization of olfactory genes in bats using molecular approach have been reported. In India, in fact, until recently, there is no systematic study of OR genes identification and characterisation in bats. Therefore, in this paper, we have focused on identifying and expressing the olfactory receptor genes in both fruit and insect-eating bats using bioinformatics and molecular biology approaches to identify the unique and diverse OR genomic repertoire in bats.

Materials and Methods

Sample collection

The present study was carried out from January 2018 to February 2020. The fruit and insect-eating bats (Table 1) were captured from various habitats and foraging areas such as caves, plantations, botanical gardens and orchards in and around Madurai, Tamil Nadu and South India (lat: 9° 58' N; long: 78° 10' E). Nylon mist nets of 9 m x 2.6 m with a mesh size of 38 mm (Avinet-Dryden, New York, USA) have been used. The mist nets were placed away from illuminated areas to avoid the visual detection by bats. Mist nets were set up at 4 m above ground level. They were tied about half an hour before sunset and removed at 0600 h. The bats, which were trapped in the mist net were removed immediately with gloved hands and placed in cloth bags until tissue sample collection. In addition, the bats were captured from the day roosts just before emergence with the help of a hoop net attached to an extendable aluminum pole [32,33]. The entire tree was enveloped with a 6 m x 9 m nylon mist net to prevent bats from escaping. The mouth of the hoop net was placed at the entrance of day roosts. A minor disturbance was caused for trapping the bats [33]. We followed the Institutional Ethical and Bio-safety Committee Guidelines of Madurai Kamaraj University.

Collection of tissue samples from bats

A medical punch was used for the excision of tissue (4 mm^2) and care was taken to place it in an area between the blood vessels to avoid injury (wing membranes healed within 3-4 weeks; [34,35]. After each sampling, the punched hole and the punch was disinfected with 70% ethanol. No negative effects of this treatment on the health of the bats were observed. It should also be noted that the bats frequently have natural injuries of this type in their wing membranes. The collected blood samples were immediately mixed with Anticoagulant Citrate Dextrose (ACD), transferred to microcentrifuge tubes and sealed with parafilm. The blood and tissue samples were stored in ice, transported to the lab and stored at -20°C until DNA extraction [34,35]. No bats were killed or retained as specimens during this project. We followed the institutional ethics committee guidelines (Internal Research Review Board (IRB), Ethical Clearance (EC), Biosafety and Animal Welfare committee approval to Karuppudurai dated 22/06/2015, Madurai Kamaraj University, Madurai.

Genomic DNA isolation and or gene sampling

Genomic DNA was isolated from wing-membrane biopsy samples using standard Proteinase K digestion and phenol: chloroform extraction method. The quality and quantity of the extracted DNA were checked using 0.7% agarose gel electrophoresis and spectrophotometric measurement at A260 and A280 nm (Hitachi U-2000, Tokyo, Japan). Degenerate primers (GPC1& GPC2) were used to amplify 700 bp fragment from genomic DNA of bats species [1,36].

GPC1F, 5'-GCTSCAYSARCCCATGTWYHWYTTBCT-3'

GPC1R, 3'-GGTYYYSAYDCHRTARAYRAYRGGGTT-3'

GPC2F, 5'-GYTNCAYWCHCCHATGTAYTTYTTBCT-3'

GPC2R, 3'-GTTYCTNARGSTRTAGATNANDGGRTT-3'

PCR amplification

PCR reaction was carried out in 25 µl reaction containing template DNA, 12.5 µl of 2X PCR Master Mix, 1 µl of 20 µm of each GPC1/GPC2 primers and 9.5 µl of H$_2$O. All DNA amplifications were performed using an Eppendorf PCR machine, with following cycling conditions including initial denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 40 sec, annealing at 45/48°C for 1 min, extension at 72°C for 1 min and final extension at 72°C for 10 min. The PCR products were visualized on a 1.2% Agarose gel and photographed and analyzed with a Gel Documentation System (Biorad, USA, Model 2000). Only 700 bp products were cut and extracted using gel extraction kit (Qiagen). The pooled OR gene PCR products were cloned into *E. coli* (DH5-Alpha) using a TOPO-TA cloning kit (Invitrogen, USA) to isolate individual OR genes. The colony PCR was carried out to screen the positive clones using M13 vector specific primer of different species. Clones that contain 700 bp inserts were grown in LB media with 100µg/ml Ampicillin. The plasmids were isolated from the positive clones using plasmid extraction kit (FavorPrep). PCR products were purified and Sanger sequenced in forward and reverse directions.

Sequence analysis

After sequencing, the data were analyzed and edited using ABI Analysis Software (version 3.0). Forward and reverse sequences were aligned and checked for ambiguities. Assembly of the consensus sequences was carried out at 99% similarity level to allow for 1% Taq-generated mutations. Each consensus sequence was counted as one gene. The aligned sequenced data was blasted in the NCBI website (https://blast.ncbi.nlm.nih.gov/) and looked for similarity between other organisms. Using the sequenced data as the query, the corresponding Olfactory Receptor (OR) was found for each clone that was sequenced. This analysis was also used to ensure that none of the genes were an artifact of PCR fusion. Finally, each consensus sequence was searched for an uninterrupted open reading frame (ORF) in all six possible frames. Although it is known that there are several highly conserved positions among OR genes, it is not always straightforward to ascertain which, if any, of these positions is necessary to retain function. We therefore chose the most straightforward definition of a pseudogene: a gene without a full ORF [2,36].

Multiple sequence alignment and domain organization

Multiple sequence alignment of amino acid sequences of all bat ORs was performed using Clustal Omega software (1.2.4). The other organism sequences that are required for comparison were retrieved from NCBI database. OR sequences of other organisms taken for sequence alignment are: 17–2 (human, P30953), CFDTMT (dog, P30955), F3 (rat, P23265) and 17–4 (human, P34982). GPCR sequences are dopamine D2DR D (2) receptor, human (P14416), serotonin (5H$_2$C receptor), human (P28335), β2 adrenergic receptor, human (P07550), muscarinic M1 receptor (pig, P04761) and rhodopsin (bovine, P02699). We applied the default pairwise gap opening penalty of 10 and extension penalty of 0.05. We did segmental alignments of ORs with other GPCRs with the Clustal Omega program using a procedure that allows segment-specific gap penalties. For the TM segments, we used a higher gap opening penalty of 20 in order to prevent insertions. The interhelical loops were aligned with the default penalty parameters. In rare cases the alignments were manually edited in TM segments to remove gaps and to ensure that conserved positions were aligned.

Statistical analyses: Principal component analysis (PCA)

To observe variations in our dataset, we performed principal component analysis (PCA). Phylogenetic PCA were performed using normalized frequencies on "all genes," "functional," and "pseudogene" OR data sets. The sequenced 33 functional bat OR genes were aligned using GramAlign 7.1 software. The matrix obtained from the alignment was used for validating principal component analysis. XLSTAT was used to analyse the dataset. Covariance matrix of the data was used in the PCA algorithm.

Heat maps

Heat maps were generated in DISPLAYR tool. The levels of OR genes in each OR gene family across all bat species were visualized. The heat map gradient covered from darker to lighter shades of blue, with darker shade representing high level of OR genes and lighter shade representing lower or no OR genes.

Phylogenetic tree construction

In each data set, deduced amino acid sequences were aligned using BioEdit 7.2.5 software and the alignment was finally inspected and corrected manually. Phylogenetic trees were constructed by the neighbor joining method with matrices of the Poisson corrected amino acid distances, using MEGA 7 software package. The reliability of each nodal relationship was assessed by 1000 bootstrap replications. This analysis involved 33 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option).

Results

Identification of bat or genes

In this study, a total of 37 OR genes from 10 species of bats (3-frugivorous and 7- insectivorous) were successfully amplified and sequenced (Figure 1 and Table 1). A total of 242 clones were sequenced and from which we identified 37 OR genes (belonging to 9 gene families). Most of the genes were expressed in 2 to 8 bat species but none of the gene was expressed in all bat species (Table 2). The fruit bat Rousettus leschenaultii expresses

Table 1. Selected bat species for study the or genes identification and characterisation representing both fruit (s. no: 1-3) and insectivorous (s. no: 4-10) bats.

| S. No.		Scientific Name	Abbreviation	Common Name	Diet	Mode of Detection
1		*Cynopterus sphinx*	CS	Indian short-nosed fruit bat	Fruigivorous	Olfaction
2	Fruit-eating bats	*Pterous giganteus*	PG	Indian flying fox	Fruigivorous	Olfaction
3		*Roussettus leschenaulti*	RL	Fulvous fruit bat	Fruigivorous	Olfaction
4		*Pipistrellus mimus*	PM	Indian pygmy bat	Insectivorous	Echolocation
5		*Megaderma lyra*	ML	Indian false vampire bat	Carnivorous	Echolocation
6		*Hipposideros ater*	HA	Dusky leaf-nosed bat	Insectivorous	Echolocation
7	Insect-eating bats	*Hipposideros fulvus*	HF	Fulvus round leaf bat	Insectivorous	Echolocation
8		*Hipposideros speoris*	HS	Schneider's leaf-nosed bat	Insectivorous	Echolocation
9		*Rhinopoma hardwickei*	RH	Lesser mouse-tailed bat	Insectivorous	Echolocation
10		*Rhinolophus beddomei*	RB	Lesser woolly horseshoe bat	Insectivorous	Echolocation
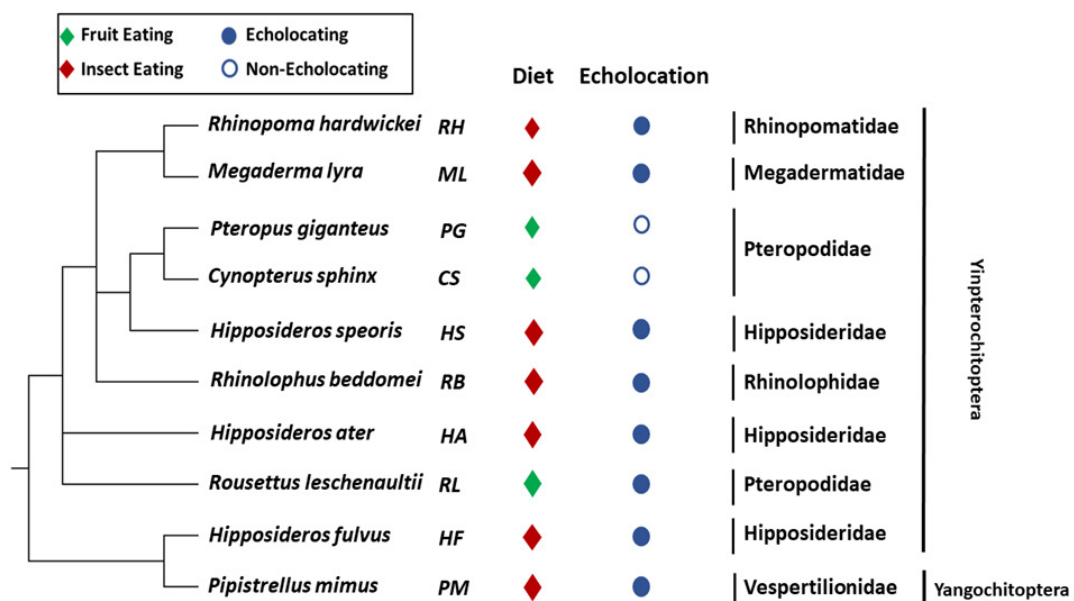
Figure 1. Phylogenetic tree of evolutionary relationship and orderly classification of species with respect to individual dietary preference and echolocation.

minimum of 4 OR genes belonging to 4 different OR gene families while another fruit bat Pteropus giganteus expresses maximum of 6 OR genes belonging to 5 different OR gene families. The insect-eating bats Megaderma lyra and Rhinolophus beddomei express minimum of 2 OR genes from 2 different gene families respectively. The maximum of 5 OR genes from 5 different gene families is expressed by another insect eating bat Pipistrellus mimus (Table 2). The total number of OR genes and families vary widely among both fruit and insect eating bats. In addition, we identified 4 OR pseudogenes from 3 fruit bat species. Fruit eating bats display diversity in OR genes when compared to insect eating bats (Table 2). Sequences for all 37 OR genes from 10 bat species were deposited to NCBI with accession numbers MN635464-MN635493 and MT998249-MT998255.

Domain organization and putative odorant CDRS of bats or gene

Multiple sequence alignment was performed for the deduced amino acid sequences of 37 bat OR genes from 10 different species and a consensus sequence that matches the currently known ORs was identified. The bat OR consensus sequence that was created as well as several other known ORs from other organisms was aligned against other GPCRs (Figure 2). Due to the presence of highly conserved positions common to the entire GPCR family, sequences were aligned most similarly in all the seven TM segments

Table 2. Number and name of the or genes identified from both fruit (s. no: 1-3) and insectivorous (s. no: 4-10) bats.

S. No.	Bat Species (Scientific Name)	No. of Genes Identified	Name of the Genes
1.	*Cynopterus sphinx*	5	OR1(Pseudo), OR2, (Pseudo), OR2, OR11 & OR56
2.	*Pteropus giganteus*	6	OR6, OR7 (2-genes), OR11, OR13 (Pseudo) & OR52
3.	*Rousettus leschenaultii*	4	OR1, OR5, OR10 & OR13 (Pseudo)
4.	*Pipistrellus mimus*	5	OR1, OR10, OR13, OR51 & OR52
5.	*Megaderma lyra*	2	OR11 & OR13
6.	*Hipposideros ater*	3	OR2, OR8 & OR13
7.	*Hipposideros fulvus*	3	OR2, OR51 & OR52
8.	*Hipposideros speoris*	4	OR1, OR6, OR7 & OR51
9.	*Rhinopoma hardwickei*	3	OR8, OR13 & OR52
10.	*Rhinolophus beddomei*	2	OR1 & OR5

except TM1 between ORs and other GPCRs ("G" marked in Figure 2). Also, several highly conserved positions unique to ORs with 90% plurality were observed in the alignment ("O" marked in Figure 2). Since the bats ORs share considerable length of conserved sequences with other GPCRs such as the adrenergic, dopaminergic, serotonergic, histaminergic and rhodopsin receptors were used for the alignment (Figure 2). Further, we observed that out of 36 hypervariable residues in the entire OR protein, 26 are in three variable helices, most of which (20 residues) are on the predicted inner surface of the receptor barrel (Figure 2). Of these 20 amino acid positions, a pronounced majority (17 residues) are located in the extracellular two-thirds of the TM segments, where ligands are known to be bound in other GPCRs. Whereby several variability peaks display an asymmetry within the TM segments. We therefore propose that the subset of 17 positions constitute the olfactory CDRs (Table 3).

Principal component analysis (pca) of or genes in bats

Principal Component Analysis (PCA) was performed in order to visualize and identify significant relationship and differences in all genes, functional and pseudogene dataset between fruit and insect eating bats. There was a considerable level of variability observed between the OR gene repertoires of all the bat species studied (Figure 3). First, the all OR gene dataset was plotted on principal component (PCs) 1 and 2, accounting for 11.62% and 10.38% of the variation respectively and together explaining 22% of total variation (Figure 3A). Second, the functional OR gene dataset was plotted on principal component 1 and 2, accounting for 15.43% and 14.12% of the variation respectively and together explaining 29.35% of total variation (Figure 3B). Finally, the pseudogene dataset was plotted on principal component 1 and 2, accounting for 37.54% and 32.36% of the variation respectively and together explaining 69.9% of total variation (Figure 3C). In the first and second principal component of the all and functional OR genes dataset, 22% and 29% variance was observed in the OR gene repertoires respectively, but, the pseudogene subset showed nearly 70% variance. Analysis across all bat species revealed significant difference only in the PCs of pseudogenes, but all OR genes and the functional OR genes dataset showed no differences.

The PCA analyses distinguished which OR gene families were driving the differences in OR gene repertoire between species. For fruit eating bats, OR gene families 1/3/7 and 2/13 appear to be important while for insect eating
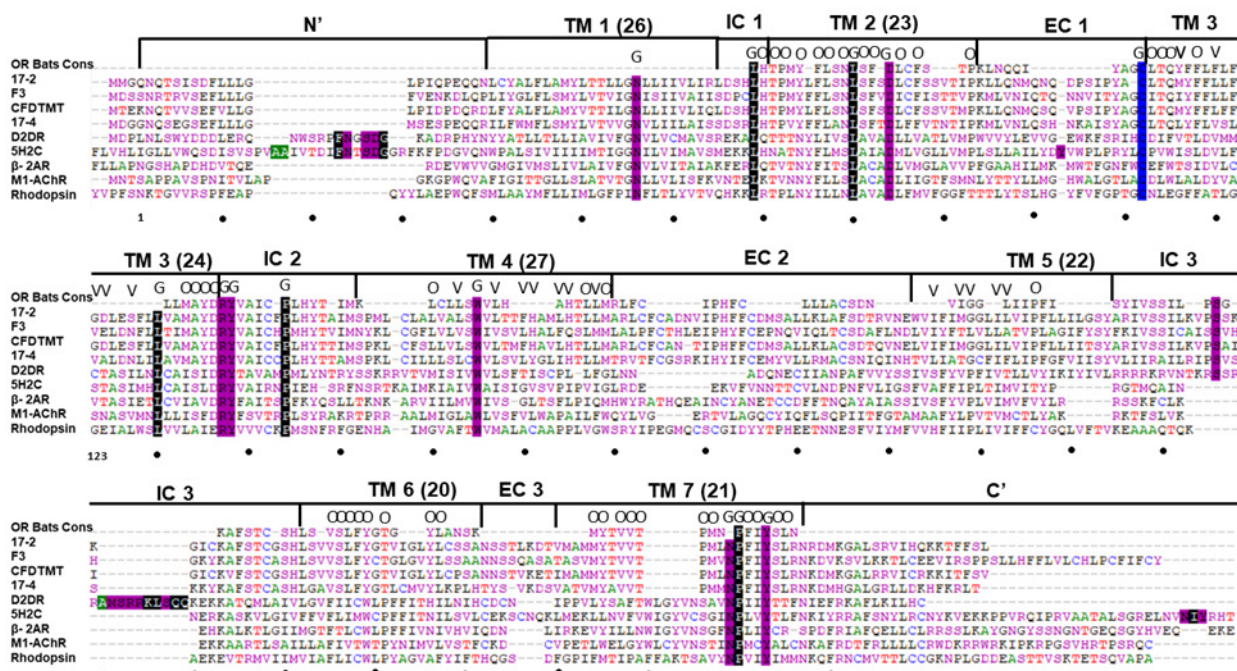

Figure 2. Multiple sequence alignment of five typical OR proteins and non-OR GPCRs in upper and lower rows, respectively. OR Cons represent 37 OR sequences. V: the OR CDR residues;　G: conserved positions among all GPCRs; and O: highly conserved positions unique to ORs.

bats, OR gene families 5/8/9, 51 and 52 appears to be important (Figure 3). For fruit bats, among all genes, the main axis of differentiation showed increase in gene families OR 1/3/7 and 2/13. But this was not observed in case of functional gene dataset. This proposes that the proportion of pseudogenes is more important in discrepancy of OR gene repertoires in

Table 3. The predicted cdr positions. 17 hypervariable cdr positions of bat or proteins in transmembrane 3, 4 and 5 with their corresponding alignment positions.

OR TM Domain	CDS Position	Alignment Position
TM 3	4	116
TM 3	8	120
TM 3	11	123
TM 3	12	124
TM 3	15	127
TM 4	11	163
TM 4	15	167
TM 4	18	170
TM 4	19	171
TM 4	22	174
TM 4	23	175
TM 4	26	178
TM 5	3	215
TM 5	6	218
TM 5	7	219
TM 5	10	222
TM 5	11	223

fruit bats. We observed an increase in the proportion of genes in families OR 1/3/7 and OR 2/13 and a loss of genes in family OR 5/8/9 in both, all genes and functional gene data sets (Figures 3 and 4). Between mega and micro bats, the molecular evolutionary mechanisms through which these patterns have evolved differ to an extent. This pattern is closely consistent when analysing the data as a whole or split into functional and pseudogene and is clearly represented in the heat map (Figure 4).

Characterisation of or genes in bats

To visualize the pattern of OR genes in all genes, functional and pseudogenes OR datasets, heat maps were constructed (Figure 4). Our results showed no differences between the pattern of all OR gene and functional OR gene datasets of insect eating bats (Figures 4A and B, top panel). There were no OR pseudogenes identified in any of the seven insect eating bat species (Figure 4C, top panel). Also, the number of OR genes identified per species is lesser when compared to fruit eating bats. Intestinally, we observed a slight difference between the pattern of all OR and functional OR gene datasets of fruit eating bats (Figures 4A and B, bottom panel). This clarifies that there is no much variation between the OR genes among the fruit and insect eating bats. While, the pseudogenes were identified only in the fruit eating bats, which accounts for its diversity (Figure 4C, bottom panel). Though, the number of bat samples are fewer, the identification of pseudogenes among fruit bats makes this group of bats more diverse when compared to insect eating bats.

Phylogenetic analysis

To investigate and confirm the evolutionary relationship among the OR gene families, a phylogenetic tree was constructed using neighbor-
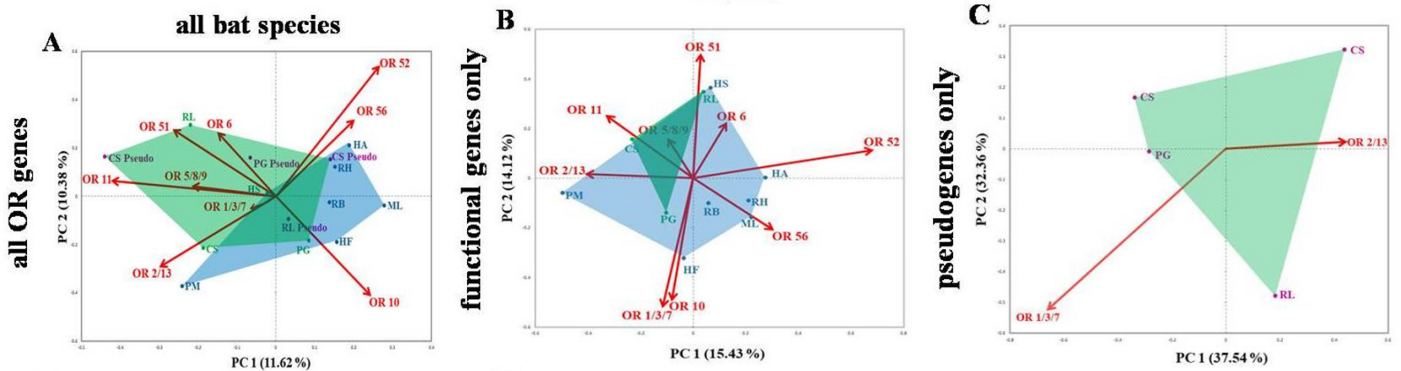

Figure 3. PCA scatter plots. (A) All OR gene subset; (B) Functional OR gene subset and (C) Pseudogene subset. Blue polygons represent insect eating bat species and green polygons represent fruit eating bat species. Red arrows represent the influence of particular OR gene families on the positioning of each species. The variance explained by each PC is indicated in brackets.
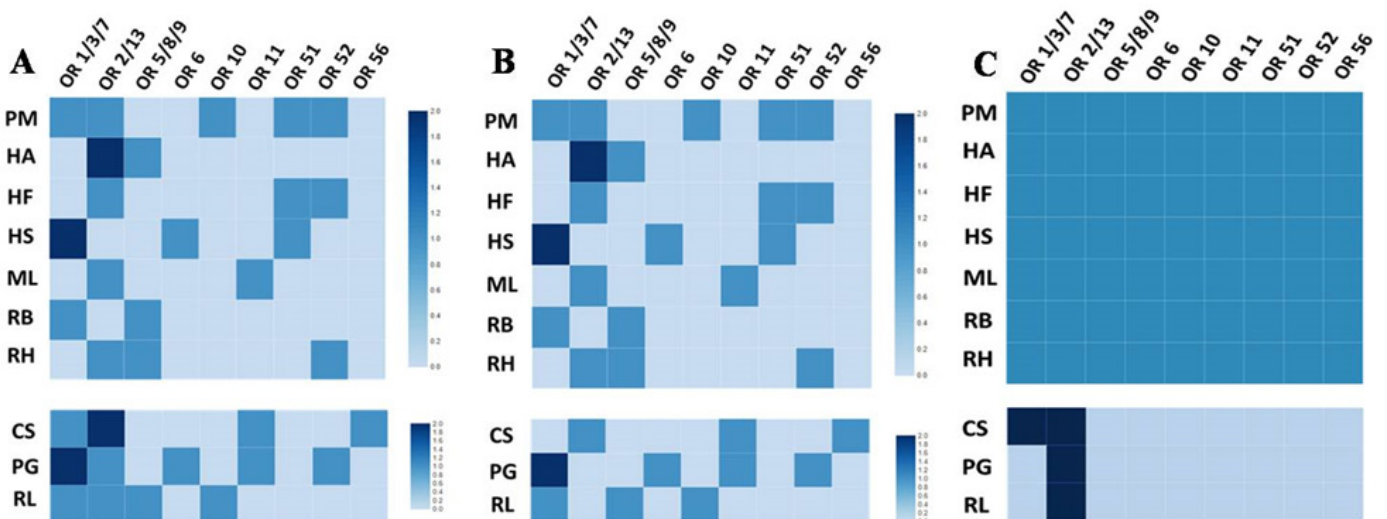

Figure 4. Heat maps of OR gene levels in each OR gene family among the bat species. Darker blue represents a high level of OR genes. Lighter blue represents a low level of OR genes. **(A)** All gene subset, **(B)** Functional subset and **(C)** Pseudogene subset. Species are separated as insect eating (above) and fruit eating (below).
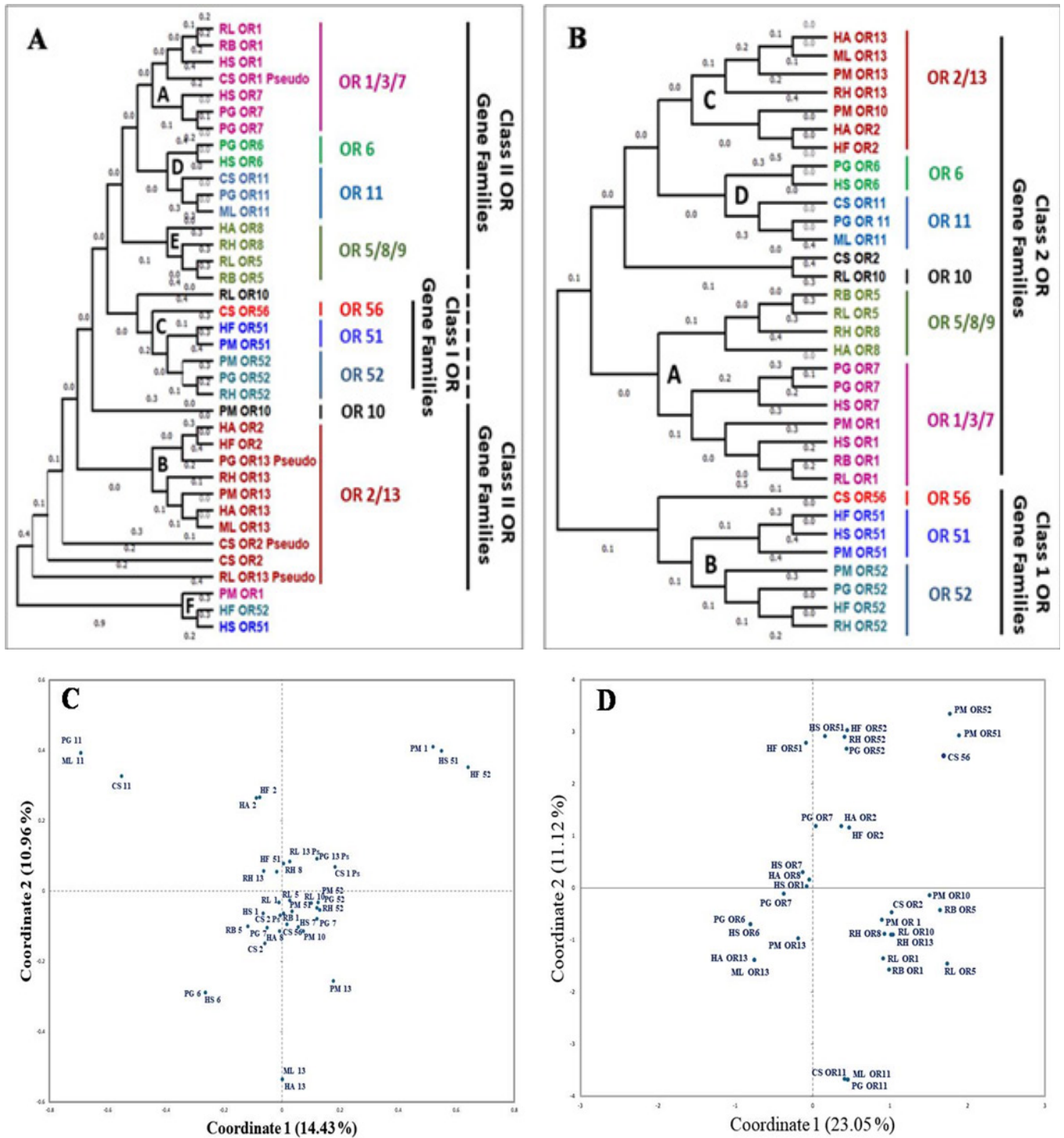
Figure 5. Phylogenetic tree of Neighbor-Joining (NJ) classification of OR gene families based on Maximum Composite Likelihood method. **(A)** All OR genes, **(B)** Functional OR genes, **(C)** Principal coordinates analysis (PCoA) of all OR genes and **(D)** Functional OR genes of fruit and insect eating bats based on Jaccard's similarity matrix.

joining method (Figure 5). We confined our analysis to both functional and pseudogenes. A rectangular cladogram was used to represent the tree (Figures 5A and B). All the 37 OR genes segregated according to their respective gene families. When considering all OR genes dataset, there was no clear division of class I and class II OR genes. Class I OR genes were clutched inside class II OR genes. Further, these classes were subdivided into 6 phylogenetic clades (Clade A-F). The largest clades A and B comprised of OR 1/3/7 and OR 2/13 gene families respectively. The second largest clade C comprised of OR 51, OR 52 and OR 56 gene families. Clade D comprised of OR 6 and OR 11 gene families. Clade E comprised of OR 5/8/9 gene family. Clade F comprised of the repeated OR gene families 51 and 52. The only outnumbered group identified is OR 10 gene family (Figure 5A).

Principal Coordinate Analysis (PCoA) was performed in order to determine the relationships among the OR genes. All OR genes were plotted on principal coordinates 1 and 2, accounting for 14.43% and 10.96% of the variation respectively and together explaining 25.39% of total variation (Figure 5C). The phylogenetic tree and PCoA of all OR gene dataset indicated that 37 all OR genes comprise 6 different clusters.

It is observed that the functional OR gene subset clearly divided into class I and class II OR genes. Class I and class II genes were further subdivided into phylogenetic clades. We considered larger clades when smaller clades were nested inside the larger clades. Hence, we identified 4 large phylogenetic clades and named them from A to D (Figure 5B). These 4 phylogenetic clades comprised 9 OR gene families. The largest clade A

comprised of gene families OR 5/8/9 and OR 1/3/7. The second largest clade B comprised of gene families OR 51 and OR 52. The third largest clade C comprised of gene family OR 2/13 and the smallest clade comprised of OR 6 and OR 11 gene families. The remaining two OR gene families OR 10 and OR 56 outnumbered from the group in different clades (Figure 5B). Principal Coordinate Analysis (PCoA) was performed in order to determine the relationships among the functional OR genes. The functional OR genes were plotted on principal coordinates 1 and 2, accounting for 23.05% and 11.12% of the variation respectively and together explaining 34.1% of total variation (Figure 5D). The phylogenetic tree and PCoA of functional OR gene dataset indicated that 33 functional OR genes comprise four different clusters.

Discussion

This study showed no evidence of sensory trade-off between olfaction and echolocation. There was diversity observed between both fruit and insect eating bats, as fruit bats expressed many OR genes from 8 different OR gene families. For example, the fruit bat *P. giganteus* expressed 6 OR genes belonging to 5 different OR gene families. We identified that, this species is the only fruit bat representing 5 different OR gene families indicating diversity between the 10 different species studied. Apart from the functional OR genes, the proportion of pseudogenes also reflects the level of diversity and evolutionary flexibility of the OR gene repertoire [2,6]. The evolutionary study of OR genes in Chiropterans showed a strong association between OR repertoire diversity and niche specialization even in a small evolutionary timeframe. The percent of pseudogenization of OR genes in frugivorous bats and other mammals is roughly found to be 10-36%, seemingly there is no significant evidence of increase in pseudogenization of OR genes in insectivorous bats that also use echolocation. Rhinolophus hipposideros uses echolocation involving Doppler shift compensation, yet shows only 10% pseudogenization of OR genes [1].

Further examination of different insectivorous and frugivorous bats revealed that echolocating bats did not have an OR gene repertoire that is notably diverse from that of non-echolocating bats and the echolocation abilities could not be attributed to variability in levels of OR pseudogenes [2]. Similarly, in our study, we identified only 4 OR pseudogenes from 3 fruit bat species and no pseudogenes were identified in insectivorous bats. This could possibly add to the fact that insectivorous bats use echolocation and no remarkable diversity in OR gene repertoire was observed between fruit and insect eating bats. Although diversity was observed within the OR gene repertoire of bats, this diversity was not explained either by gene loss or gain in dietary habits or echolocating abilities. In frugivorous bats, vision plays a vital role in olfactory perception, while in insectivorous bats echolocation is crucial. This indicates that more than one sensory signal might be in action for olfactory perception in bats. Previous report suggest that a combination of vision and olfaction is likely being traded for echolocation which resulted in the differences in olfactory bulb and visual brain component sizes in the brains of echolocating and nonecholocating bats [2,37,38]. Similarly, between VNO and OR gene repertoire, sensory trade-off was observed to a little extent as the VNO is known to express ORs [13]. Along with the OR genes, the study of other genes like TRPC2 which is associated vermonasal chemosensation might help elucidating their relationships in olfactory perception in bats [39].

The fruit eating bat *Cynopterus sphinx* was able to discriminate different food odor compounds in a complex olfactory environment. This study reported that seven different odorants were used, to which C. sphinx exhibited quantitative responses with temporal variations. C. sphinx preferred to visit odor containing sample tubes to odorless ones and had a preference for ethyl acetate odors [40]. Similar study was reported in fruit feeding Phyllostomidae bats, including Jamaican fruit eating bat *Artibeus jamaicensis* which visited the new world tree Annona muricata, that contains elevated levels of ethyl acetate in overripe fruits [41,42]. To access the olfactory ability of Asian elephants, similar studies were carried out. Previous report suggests that these elephants successfully differentiated 12 enantiomeric odor pairs and

between 12 other odor pairs of aliphatic alcohols, aldehydes, ketones and carboxylic acids, each of them having only a one-carbon difference between pair members. With these reports, researchers suggest that elephants exhibit olfactory performance similar to mice and better when compared to other mammals including humans, pigtail macaques, or squirrel monkeys [29,43], though no systematic study accessing the olfactory abilities of the African elephants have been published till date [29]. But these elephants were able to distinguish between two ethnic groups of Kenya, the Maasai and the Kamba using their olfactory signals. In addition, these African elephants recognise up to 30 individual family members from olfactory cues in mixtures of urine and earth [44,45]. Similar results have been identified in African elephants which contained approximately 2000 functional genes and more than 2200 genes which is the largest OR gene repertoire among the genome examined. Due to its dependency on olfaction in several circumstances including social communication, reproduction and foraging, the elephant is believed to have the largest OR gene repertoire among mammals [29,46,47].

All olfactory receptors exhibit variability patterns in amino acid sequences to recognise large variety of odorants. Most variable residues were present in three of the seven transmembrane domains, TM 3, 4 and 5. Among these TM, 40 variable residues were identified based on analysis of small sets of OR sequences [4,14]. With much considerable focused picture of the potential odorant binding pocket and utilizing more quantitative variability measures for much larger number of functional OR sequences, 17 hypervariable residues were identified [48]. Based on the previously reported study, we were also able to identify these 17 hypervariable residues in TM3, 4 and 5 of bats. Generally, the conserved amino acids likely mediate interacts with some ligands [49]. TM6 is vital in ligand binding by other GPCRs, since it shows low variability in ORs, enables TM6 to take part in odorant binding as well. Presence of three aromatic residues in TM6 including F/Y motifs has been proposed to interact with aromatic rings in the neurotransmitter ligands and could similarly interact with aromatic moieties of odorant molecules [50]. TM6 in bats also shows the presence of FY motif, with which we can predict that similar kind of interactions also exist in bats.

The presence of consensus motifs LHTPMY in intracellular loop1, MAYDRYVAIC at the end of TM3 and the beginning of intracellular loop 2, SY at the end of TM5, FSTCSSH at the beginning of TM6 and PMLNPF in TM7 are the main features specific to ORs [11,51,52,53]. Despite the fact that these sequences change slightly between species, they have been utilised to identify OR genes in a variety of genomes. More than 80 short motifs have been found through extensive comparative analysis [52,53], some of which are distinctive for different subfamilies or species and have been linked to ligand binding. Seven cysteine residues are highly conserved, with two of them considered to have a role in the protein's structural stability. Two of these (at locations 97 and 179) are found in all GPCRs and are thought to create a disulfide connection between extracellular loops 1 and 2, whereas the other five are found only in ORs. Our result coincides with the previous reported study which suggests several characteristics of olfactory receptors. Our sequence alignment also shows the presence of above-mentioned consensus motifs which confirms that they represent OR gene repertoire of bats. Also, transmembrane 3, 4 and 5 in bat OR gene sequence having hypervariable regions and many cysteine residues indicate that they function as a part the odorant binding pocket.

PCA analysis suggested that for most of the variations observed in the entire dataset of bFats, nine OR gene families were responsible. But, only two OR gene families were responsible for the variations observed among the pseudogene subset. Most fruit bats were distributed in the 1/3/7 OR gene family, indicating that this gene family plays a role in frugivory [2]. Our result slightly deviates from the previous reports which suggested that OR gene families 1/3/7 and 2/13 are linked to frugivory [2]. Though, our result did not contradict the previous reports. Further analysis of OR gene repertoires of bats will help in understanding their mechanisms and the specific odors they bind. This supports in comprehending their direct associations with the habitat. Previous studies reported that OR gene families OR 1/3/7 and OR 2/13 are associated with the OR gene repertoire of frugivorous bats across

two large radiations: the *Yangochiroptera*, including New World fruit bats in the family Phyllostomidae and the *Yinpterochiroptera*, including Old World fruit bats and the family Pteropodidae [2]. Our results also show similar results where OR gene families 1/3/7 and OR 2/13 are associated mostly with fruit eating bats and to a lesser extent in insect eating bats since only a few insect eating bats expressed these two gene families. Since our sample size is low, further screening of bat OR genes becomes a necessary factor in confirming the association of OR gene families 1/3/7 and OR 2/13 with frugivory. Previous reports suggest that fruit bats visit trees whose fruits contain high level of ethyl acetate. C. sphinx preferred to visit odor containing sample tubes to odorless ones and had a preference for ethyl acetate odors [40]. Jamaican fruit eating bat A. *jamaicensis* which visited the new world tree A. *muricata*, that contains elevated levels of ethyl acetate in overripe fruits [41,42]. Behavioural and radio-tracking studies on Australian Flying foxes and Queensland tube-nosed bat, Nyctimene robinsoni bats suggest that these bats have acquired a taste for the fruit of A. *muricata* [54,55]. Concurring to the speculations made by Hayden, et al. 2014, we also suspect that OR genes families OR 1/3/7 and OR 2/13 may be directly involved in the detection of ethyl acetate, though, this will need to be confirmed and explored with future functional assays.

Although it is unclear which aspects of olfactory ability the number of OR genes reflect, it is reasonable to assume that a species with a larger number of OR genes can distinguish between more subtle differences in structurally related odorants and that the resolution of the olfactory world rather than the sensitivity to a particular odour is determined by the number of OR genes [29]. This study clearly illustrates that analyses of OR genes in bats can provide insights into gene function. Nevertheless, more ORs should be deorphanized and more OR gene repertoires should be analysed to further elucidate the differences in the evolutionary fates of genes.

Conclusion

We have identified 37 OR genes belonging to 9 different OR gene families, despite the fact that the sample size is less and the number of samples per species is low. Our result is based on a representative OR gene for a particular species rather than all OR genes in bats. Next-generation sequencing techniques should be employed to get a better sampling of the OR gene repertoire of bat species whose genomes have not yet been sequenced. These methods would allow researchers to examine all OR genes amplified by defective primers. Further sampling of bat OR genes would elucidate the function of these genes and their association to OR gene families. Also, identifying and cloning functional OR repertoire creates a basis for addressing many unresolved issues in bat olfaction. Most importantly, in conjunction with robust heterologous expression and assay systems and high throughput screening of odorant libraries, it will ultimately lead to understanding of structure-function relationships and small molecule recognition by this large group of GPCRs. The impact of genetic polymorphism of ORs on differential olfactory perception in bat population is another exciting topic. Global comparative analysis of functional bat OR candidate gene and pseudogene repertoires will shed light on the evolution of bat olfactory apparatus and its biological consequences. This study represents a further step in revealing the function of bat OR genes and their associations to the gene families

Acknowledgement

Author Contributions

Conceived and designed the experiments: TK RSC. Performed the experiments: RSC SVB. Analyzed the data: TK RSC SVB. Contributed reagents/materials/analysis tools: TK. Wrote the paper: RSC TK.

Conflict of Interest

There are no conflicts of interest by author.

References

1. Hayden, Sara, Michaël Bekaert, Tess A. Crider and Stefano Mariani, et al. "Ecological adaptation determines functional mammalian olfactory subgenomes." *Genome Res* 20 (2010): 1-9.

2. Hayden, Sara, Michaël Bekaert, Alisha Goodbla and William J. Murphy, et al. "A cluster of olfactory receptor genes linked to frugivory in bats." *Mol Biol Evol* 31 (2014): 917-927.

3. Mombaerts, Peter. "Genes and ligands for odorant, vomeronasal and taste receptors." *Nat Rev Neurosci* 5 (2004): 263-278.

4. Buck, Linda and Richard Axel. "A novel multigene family may encode odorant receptors: A molecular basis for odor recognition." *Cell* 65 (1991): 175-187.

5. Niimura, Yoshihito and Masatoshi Nei. "Extensive gains and losses of olfactory receptor genes in mammalian evolution." *PloS One* 2 (2007): e708.

6. Nei, Masatoshi, Yoshihito Niimura and Masafumi Nozawa. "The evolution of animal chemosensory receptor gene repertoires: Roles of chance and necessity." *Nat Rev Gene* 9 (2008): 951-963.

7. Touhara, Kazushige and Leslie B. Vosshall. "Sensing odorants and pheromones with chemosensory receptors." *Annu Rev Psychol* 71 (2009): 307-332.

8. Niimura, Yoshihito. "Olfactory receptor multigene family in vertebrates: From the viewpoint of evolutionary genomics." *Curr Genom* 13 (2012): 103-114.

9. Zozulya, Sergey, Fernando Echeverri and Trieu Nguyen. "The human olfactory receptor repertoire." *Genome Biol* 2 (2001): 1-12.

10. Zhang, Xiaohong, Xinmin Zhang and Stuart Firestein. "Comparative genomics of odorant and pheromone receptor genes in rodents." *Genomics* 89 (2007): 441-450.

11. Gaillard, Isabelle, Sylvie Rouquier and Dominique Giorgi. "Olfactory receptors." *Cell Mol Life Sci* 61 (2004): 456-469.

12. Lindblad-Toh, Kerstin, Claire M. Wade, Tarjei S. Mikkelsen and Elinor K. Karlsson, et al. "Genome sequence, comparative analysis and haplotype structure of the domestic dog." *Nature* 7069 (2005): 803-819.

13. Keller, Andreas and Leslie B. Vosshall. "Better smelling through genetics: Mammalian odor perception." *Curr Opin Neurobiol* 18 (2008): 364-369.

14. Rouquier, Sylvie, Sylvie Taviaux, Barbara J. Trask and Véronique Brand-Arpon,et al. "Distribution of olfactory receptor genes in the human genome." *Nature genetics* 18 (1998): 243-250.

15. Glusman, Gustavo, Itai Yanai, Irit Rubin and Doron Lancet, et al. "The complete human olfactory subgenome." *Genome res* 11 (2001): 685-702.

16. Saito, Harumi, Qiuyi Chi, Hanyi Zhuang and Hiroaki Matsunami, et al. "Odor coding by a Mammalian receptor repertoire." *Sci Signal* 2 (2009): ra9-ra9.

17. Hughes, Graham M., Emma SM Boston, John A. Finarelli and William J. Murphy, et al. "The birth and death of olfactory receptor gene families in mammalian niche adaptation." *Mol Biol Evol* 35 (2018): 1390-1406.

18. Bushdid, Caroline, Marcelo O. Magnasco, Leslie B. Vosshall and Andreas Keller, et al. "Humans can discriminate more than 1 trillion olfactory stimuli." *Science* 343 (2014): 1370-1372.

19. Gerkin, Richard C. and Jason B. Castro. "The number of olfactory stimuli that humans can discriminate is still unknown." *Elife* 4 (2015): e08127.

20. Meister, Markus. "Can humans really discriminate 1 trillion odors?." *ArXivorg* 1411.0165 (2014).

21. Olsson, Peter and Matthias Laska. "Human male superiority in olfactory sensitivity

to the sperm attractant odorant bourgeonal." *Chem Senses* 35 (2010): 427-432.

22. Ferdenzi, Camille, S. Craig Roberts, Annett Schirmer and Sylvain Delplanque, et al. "Variability of affective responses to odors: culture, gender and olfactory knowledge." *Chem Senses* 38 (2013): 175-186.

23. Aimé, P., P. Duchamp-Viret, M.A. Chaput and A. Savigner, et al. "Fasting increases and satiation decreases olfactory detection for a neutral odor in rats." *Behav Brain Res* 179 (2007): 258-264.

24. Chan, Ho Ka, Fabian Hersperger, Emiliano Marachlian and Brian H. Smith, et al. "Odorant mixtures elicit less variable and faster responses than pure odorants." *PLoS Comput Biol* 14 (2018): e1006536.

25. Duchamp-Viret, Patricia, André Duchamp and Gilles Sicard. "Olfactory discrimination over a wide concentration range. Comparison of receptor cell and bulb neuron abilities." *Brain Res* 517 (1990): 256-262.

26. Malnic, Bettina, Junzo Hirono, Takaaki Sato and Linda B. Buck. "Combinatorial receptor codes for odors." *Cell* 96 (1999): 713-723.

27. Niimura, Yoshihito and Masatoshi Nei. "Evolution of olfactory receptor genes in the human genome." Proc Natl Acad Sci 100 (2003): 12235-12240.

28. Matsui, Atsushi, Yasuhiro Go and Yoshihito Niimura. "Degeneration of olfactory receptor gene repertories in primates: No direct link to full trichromatic vision." *Mol Biol Evol* 27 (2010): 1192-1200.

29. Niimura, Yoshihito, Atsushi Matsui and Kazushige Touhara. "Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals." *Genome Res* 24 (2014): 1485-1496.

30. Altringham, John D. and M. Brock Fenton. "Sensory ecology and communication in the Chiroptera." *Bat Ecology* (2003): 90-127.

31. Thies, Wibke, Elisabeth KV Kalko and Hans-Ulrich Schnitzler. "The roles of echolocation and olfaction in two Neotropical fruit-eating bats, Carollia perspicillata and C. castanea, feeding on Piper." *Behav Ecol Sociobiol* 42 (1998): 397-409.

32. Gaisler, J. "Netting as a possible approach to study bat activity." *Period Biol* 75 (1973): 129-134.

33. Gopukumar, N., T. Karuppudurai, P. T. Nathan and K. Sripathi, et al. "Solitary adult males in a polygynous-mating bat (*Cynopterus sphinx*): A forced option or a strategy?." *J Mammal* 86 (2005): 281-286.

34. Karuppudurai, T. and K. Sripathi. "Molecular genetic analysis of male alternative strategy and reproductive success in the polygynous mating bat *Cynopterus sphinx*." *J cell Anim Biol* 4 (2010): 042-050.

35. Karuppudurai, T., K. Sripathi, N. Gopukumar and V. Elangovan, et al. "Genetic diversity within and among populations of the Indian short-nosed fruit bat, *Cynopterus sphinx* assessed through RAPD analysis." *Curr Sci* (2007): 942-950.

36. Zhang, Xiaohong, Omar De la Cruz, Jayant M. Pinto and Dan Nicolae, et al. "Characterizing the expression of the human olfactory receptor gene family using a novel DNA microarray." *Genome Biol* 8 (2007): 1-10.

37. Reep, Roger L. and Kunwar P. Bhatnagar. "Brain ontogeny and ecomorphology in bats." (2000): 93-136.

38. Jones, Kate E. and Ann M. MacLarnon. "Affording larger brains: Testing hypotheses of mammalian brain evolution on bats." *Am Nat* 164 (2004): E20-E31.

39. Zhao, Huabin, Dong Xu, Shuyi Zhang and Jianzhi Zhang, et al. "Widespread losses of vomeronasal signal transduction in bats." *Mol Biol Evol* 28 (2011): 7-12.

40. Elangovan, Vadamalai, Elangovan Yuvana Satya Priya and Ganapathy Marimuthu. "Olfactory discrimination ability of the short-nosed fruit bat *Cynopterus sphinx*." *Acta Chiropt* 8 (2006): 247-253.

41. Goodwin GG and Greenhall AM. "A review of the bats of Trinidad and Tobago." *Bull Am Mus Nat Hist* 122 (1961): 187–301.

42. Márquez C, Carlos J., Aleyda M. Jimenez, Coralia Osorio and José R. Cartagena V, et al. "Volatile compounds during the ripening of colombian soursop (*Annona muricata* L. cv. Elita)." *Vitae* 18 (2011): 245-250.

43. Rizvanovic, Alisa, Mats Amundin and Matthias Laska. "Olfactory discrimination ability of Asian elephants (*Elephas maximus*) for structurally related odorants." *Chem Senses* 38 (2013): 107-118.

44. Bates, Lucy A., Katito N. Sayialel, Norah W. Njiraini and Cynthia J. Moss, et al. "Elephants classify human ethnic groups by odor and garment color." *Curr Biol* 17 (2007): 1938-1942.

45. Bates, Lucy A., Katito N. Sayialel, Norah W. Njiraini and Joyce H. Poole, et al. "African elephants have expectations about the locations of out-of-sight family members." *Biol Letters* 4 (2008): 34-36.

46. Langbauer Jr, W. R. "Elephant communication." Zoo Biology: Published in affiliation with the American Zoo and Aquarium Association 19 (2000): 425-445.

47. Rasmussen, L. E.L., and V. Krishnamurthy. "How chemical signals integrate Asian elephant society: The known and the unknown." *Zoo Biol* 19 (2000): 405-423.

48. Pilpel, Yitzhak and Doron Lancet. "The variable and conserved interfaces of modeled olfactory receptor proteins." *Protein Sci* 8 (1999): 969-977.

49. Singer, Michael S. and Gordon M. Shepherd. "Molecular modeling of ligand-receptor interactions in the OR5 olfactory receptor." *Neuroreport* 5 (1994): 1297-1300.

50. Heitz, Freddy, James A. Holzwarth, Jean-Pierre Gies and Rebecca M. Pruss, et al. "Site-directed mutagenesis of the putative human muscarinic M2 receptor binding site." *Eur J Pharmacol* 380 (1999): 183-195.

51. Probst, William C., Lenore A. Snyder, David I. Schuster and Jürgen Brosius, et al. "Sequence alignment of the G-protein coupled receptor superfamily." *DNA Cell Biol* 11 (1992): 1-20.

52. Liu, Agatha H., Xinmin Zhang, Gustavo A. Stolovitzky and Andrea Califano,et al. "Motif-based construction of a functional map for mammalian olfactory receptors." *Genomics* 81 (2003): 443-456.

53. Zhang, Xinmin and Stuart Firestein. "The olfactory receptor gene superfamily of the mouse." *Nat Neurosci* 5 (2002): 124-133.

54. Spencer, Hugh J. and Theodore H. Fleming. "Roosting and foraging behavior of the queensland tube-nosed bat, nyctimene-robinsoni (Pteropodidae)-preliminary radio-tracking observations." *Wildl Res* 16 (1989): 413-420.

55. Hall, Leslie S. and Gregory Richards. "Flying foxes: fruit and blossom bats of Australia." *UNSW Press*, 2000.