Taste and odorant receptors of the coelacanth - a gene repertoire in transition

Barbara Picone[1+], Uljana Hesse[1+], Sumir Panji[1+*], Peter Van Heusden[1], Mario Jonas[1], Alan Christoffels[1*]

[+] These authors contributed equally to this work

[1]South African National Bioinformatics Institute, University of the Western Cape, 7535 Bellville, South Africa

*Present address: Computational Biology Unit, Institute of Infectious Disease and Molecular Medicine, University of Cape Town Health Sciences Campus, Anzio Road, Observatory 7925, South Africa

6 figures in text
1 table in text
3 supplementary figures
3 supplementary tables

Abbreviated Title
Chemosensory receptors in coelacanths

*Correspondence to:
Alan Christoffels
South African National Bioinformatics Institute
University of the Western Cape
Modderdam Road
7535 Bellville
South Africa
Email: alan@sanbi.ac.za
Telephone: +27 21 959 3645
Telefax: +27 21 959 2592

Supporting Grant:
The South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation: Grant 64751

Abstract

G-protein coupled chemosensory receptors (GPCR-CRs) aid in the perception of odors and tastes in vertebrates. So far, six GPCR-CR families have been identified that are conserved in most vertebrate species. Phylogenetic analyses indicate differing evolutionary dynamics between teleost fish and tetrapods. The coelacanth *Latimeria chalumnae* belongs to the lobe-finned fishes, which represent a phylogenetic link between these two groups. We searched the genome of *L. chalumnae* for GPCR-CRs and found that coelacanth taste receptors are more similar to those in tetrapods than in teleost fish: two coelacanth T1R2s co-segregate with the tetrapod T1R2s that recognize sweet substances, and our phylogenetic analyses indicate that the teleost T1R2s are closer related to T1R1s (umami taste receptors) than to tetrapod T1R2s. Furthermore, coelacanths are the first fish with a large repertoire of bitter taste receptors (58 T2Rs). Considering current knowledge on feeding habits of coelacanths the question arises if perception of bitter taste is the only function of these receptors. Similar to teleost fish, coelacanths have a variety of olfactory receptors (ORs) necessary for perception of water-soluble substances. However, they also have seven genes in the two tetrapod OR subfamilies predicted to recognize airborne molecules. The two coelacanth vomeronasal receptor families are larger than those in teleost fish, and similar to tetrapods, form V1R and V2R monophyletic clades. This may point to an advanced development of the vomeronasal organ as reported for lungfish. Our results show that the intermediate position of *Latimeria* in the phylogeny is reflected in its GPCR-CR repertoire.

Introduction

Chemosensory receptors (CRs) enable animals to detect chemical signals from their environment. In vertebrates, perception of tastes, odorants, and pheromones is conducted predominantly through G-protein coupled receptors (GPCRs), cell membrane proteins with seven transmembrane domains that, upon ligand binding, activate signal transduction pathways. To date, the GPCR-CRs have been classified into six multigene families. These include two types of taste receptors (T1Rs and T2Rs) that detect sweet, umami, and bitter tasting substances; as well as olfactory receptors (ORs), trace amine associated receptors (TAARs), and two types of vomeronasal receptors (V1Rs and V2Rs) that recognize odorants and pheromones.

Phylogenetic analyses indicate that these families appeared independently at different time points during evolution (Grus and Zhang, 2009, Hussain et al., 2009, Libants et al., 2009; Niimura, 2009). According to these studies, the ORs and V1Rs emerged in the common ancestors of chordates and vertebrates, respectively. TAARs, V2Rs, and T1Rs appeared first in the common ancestor of jawed vertebrates, and T2Rs evolved in the ancestor of bony vertebrates. Although representatives of all families are present in teleost fish and in tetrapods, the evolutionary dynamics within the receptor families differ between these two groups of bony vertebrates. The discrepancies are attributed to the different habitats and diets, and to differences in the development of the olfactory organs.

Latimeria inhabits a unique phylogenetic position; it belongs to the lobe-finned fishes, in-between the teleost fishes and tetrapods. Despite its fish-like appearance, the living coelacanth shares several characteristics with land vertebrates (most notably the bony limbs and the tetrapod-like movement of paired fins, but also enameled teeth and use of urea for disposal of nitrogenous waste). The genome and gene repertoire of *L.*

chalumnae have recently been published (Amemiya et al., 2013), which provided the opportunity to conduct a genome-wide analysis on the GPCR-CRs of this organism. We found that the intermediate position of *Latimeria* in the phylogenetic tree is reflected in its GPCR-CR repertoire. Coelacanths have teleost-like as well as tetrapod-like receptors, and display evolutionary patterns characteristic of both vertebrate groups.

Materials and Methods

Identification of *L. chalumnae* chemosensory receptors

For this study we used the *L. chalumnae* ENSEMBL gene predictions version 66 published recently by Amemiya et al. (2013). Our previous analyses on gene duplication events in this dataset resulted in the construction of 18,307 OrthoMCL clusters (Amemiya et al., 2013). To identify all chemosensory receptors in the coelacanth genome we first selected all coelacanth genes annotated by InterProScan (Zdobnov and Apweiler, 2001) as G-protein coupled receptors (e.g. IPR000276, IPR017452, IPR000337, IPR017978, IPR011500), and searched this subset for InterPro annotations associated with the respective CR domains (ORs: IPR000725, IPR019424, IPR019430; TAARs: IPR009132, IPR009133; V1Rs: IPR004072; V2Rs: IPR004073, IPR001828; T1Rs: IPR001828; T2Rs: IPR007960). We then identified the OrthoMCL clusters of genes thus annotated which included all un-annotated genes from these clusters into the coelacanth GPCR-CR subset. For validation, we compared all coelacanth GPCRs against the sequences in GPCRDB (Vroling et al., 2010) and verified the number of transmembrane domains using TMHMM 2.0 (Krogh et al., 2001). Coelacanth proteins with less than six or more than eight transmembrane domains were considered pseudogenes.

Chemosensory receptor sequences from other organisms

CR-protein sequences from other species were derived from literature or downloaded from CRDB (Dong et al., 2012), UNIPROT (www.uniprot.org), and NCBI (www.ncbi.nlm.nih.gov). For some species, several datasets per GPCR-CR family were available. To account for differences in GPCR-CR identification methods between the various sources, we aligned the sequences from all sources of a receptor family by species. One representative sequence per gene was chosen, aiming at maximum sequence diversity without mixing datasets. In the process, we found that a specific gene identifier did not always correspond to the same gene in different sources. Since we had to rename some of the proteins to avoid name duplication we provide a mapping of our protein identifiers to the original gene/protein identifiers and list the source for each protein sequence in Table S2.

Phylogenetic Analysis

Multiple sequence alignments of the proteins of each receptor family were generated using M-Coffee (Notredame et al., 2000). ProtTest 3 (Abascal et al., 2005) was used to select the amino acid substitution model that would fit the protein alignment best. Trees were reconstructed using maximum likelihood methods in PhyML 3.0 (Guindon et al., 2010) with rate variation among sites and proportion of invariant sites optimized. Statistical confidence values for branches were calculated using the aLRT model (Anisimova and Gascuel, 2006). The resulting ML trees were visualized using FigTree v1.3.1.

Results

Taste receptor family 1 (T1Rs)

All studied land vertebrates have three T1R genes each belonging to a different class (T1R1, T1R2, and T1R3, respectively). Teleost fish also have one T1R1 and one T1R3, but multiple T1R2s. In contrast to previous studies (Ishimaru et al., 2005; Shi and Zhang, 2006), our phylogenetic analyses indicate that the teleost T1R2s are more closely related to the T1R1s than to the tetrapod T1R2s (Fig. 1). As confirmed by the approximate likelihood-ratio test (Anisimova and Gascuel, 2006), this tree topology provides a significantly better phylogenetic model for the given T1R dataset than the alternate tree topology described in previous analyses (Table S3). *L. chalumnae* has five T1Rs (Table S1): one T1R1, two T1R2s that group with the tetrapod genes, and two T1R3s. We did not find teleost-like T1R2s in the *Latimeria* genome.

Taste receptor family 2 (T2Rs)

L. chalumnae is the first fish found to have an extensive repertoire of bitter taste receptors. The genome harbors 58 potentially functional T2Rs, all of which contain the T2R signature motif (Table S1). Our previous analyses grouped these genes into one OrthoMCL cluster with 18 confirmed coelacanth-specific gene duplication events (Amemiya et al., 2013). As shown in Figure 2, coelacanths form two T2R clades, both of which group with T2R clades from teleost fish. Similar to land vertebrates, most coelacanth T2Rs are clustered in a large monophyletic clade. Within the genome, most of the coelacanth T2Rs occurs in small clusters of two to five genes that are distributed over several genome contigs.

Olfactory receptors (ORs)

We identified 56 putatively functional and nine pseudogenized ORs in the *L. chalumnae* genome (Table S1). Nine of the potentially functional ORs are predicted to have more than one exon in the coding region which is not typical for members of this gene family. While intron introduction is possible, empirical sequence verification should be conducted to rule out prediction errors resulting from genome assembly and annotation mistakes. Nevertheless the underlying conserved protein aligment that underpins the phylogenetic tree remains unchanged. Niimura and Nei (2005) classified the ORs of jawed vertebrates into Type 1 with six subgroups (α-ζ) and Type 2 with five subgroups (η-λ). Genes of the groups θ, κ, and λ are phylogenetically nested within vertebrate OR genes but apparently do not encode ORs (Niimura, 2009). Since this paper focuses on *Latimeria* GPCRs potentially involved in olfaction, pheromone perception and/or taste we excluded these groups from our analysis. Figure 3 shows the phylogenetic tree with all currently known potentially functional coelacanth, zebrafish, and human ORs (see also Figure S1). Our analyses confirmed all subgroups but ζ. Most coelacanth ORs form clades in one of the previously described subgroups, including α, γ, δ, ζ, and η, and only two coelacanth ORs clustered separately. The genes of the individual coelacanth clades belong to distinct OrthoMCL clusters (Table S1), indicating that our previously described method for identifying paralogs in *L. chalumnae* was successful (Amemiya et al., 2013). Those analyses provided evidence for coelacanth-specific gene duplication events in the subgroups α, γ, δ, ζ2, and η. Within the genome, most of the ORs cluster loosely by subgroup; the largest five clusters each consist of seven to nine genes that are separated on average by 50 kbp.

A previous study describes eight potentially functional *Latimeria* ORs (Freitag et al, '98). We identified four of the genes in the current genome assembly (LCor21, LCor23, LCor25, LCor30) and classified them into the groups γ and ζ (Table S1). The remaining four genes (LCor16, LCor17, LCor27, LCor99) are missing from the current genome assembly, which may be due to incomplete genome coverage. However, it appears peculiar that the missing genes are those four assigned by Freitag et al, ('98) into the tetrapod class (here subgroup α), which has many pseudogenes (Niimura and Nei, 2005). Therefore, polymorphisms among coelacanths can also not be ruled out.

Trace amine associated receptors (TAARs)

The current *L. chalumnae* gene set contains only four potentially functional TAARs (Table S1). All four genes consist of a single exon, and contain the characteristic TAAR fingerprint (NSXXNPXX[YH]XXX[YF]XWF) as well as the aminergic ligand motif necessary for amine ligand binding (Huang, 2003). We analyzed their location in the phylogenetic tree relative to seven vertebrate species (Fig. 4). To date, TAARs have been categorized into three classes with 28 subfamilies (Hussain et al., 2009). Our study confirmed the evolutionary patterns for classes I and III. The subfamilies of class II segregated into three subclasses, as exemplified by the nodes A, B, and C. Nevertheless, we were able to determine that three of the four coelacanth TAARs belong to the subfamily TAAR1 (class I). These three genes belong to one OrthoMCL cluster for which previous analyses indicated a coelacanth specific gene duplication event. They are located on one genome contig, separated by 35 and 120 kbp, respectively. The fourth coelacanth TAAR (LCH_01339) segregates with the shark gene of class II and the zebrafish genes of TAAR13.

Vomeronasal receptor family 1 (V1Rs/ORAs)

Our analyses indicate that *L. chalumnae* has 15 potentially functional V1Rs (Table S1). These V1Rs contain eight of the 13 teleost-specific amino acids, four amino acids conserved in mouse, and six amino acids found in both teleost fish and mouse (Fig. S2). Previous studies have shown that the tetrapod V1Rs form a subclade within the paraphyletic Ora gene family (Saraiva and Korsching, 2007; Gruz and Zhang, 2009). Our phylogenetic analyses on vertebrate V1Rs/ORAs (Fig. 5) support the described tree topologies: the teleost *ora* genes segregate into three main clades (*ora*1-*ora*2, *ora*3-*ora*4, and *ora*5-*ora*6), and most of the tetrapod V1Rs group within the *ora*1-*ora*2 genes forming several species-specific subclades. The coelacanth V1Rs appear in both teleost and tetrapod clades. Five coelacanth V1Rs are located in the teleost-dominated Ora clades (*ora*1, *ora*2, and *ora*3-4) and another eight coelacanth V1Rs form a monophyletic group within the *ora*1-*ora*2 clade.

We observed that two frog (XTR+V1R1, XTR+V1R4) and one lizard V1Rs (ACA+V1R1) identified by Dong et al. (2012) grouped separately from all other tetrapods. Whether these genes represent annotation mistakes or functional V1Rs requires further verification. Moreover, recent investigations indicated the presence of at least five V1Rs in sea lamprey, which are expressed in the olfactory organ of this organism (Gruz and Zhang, 2009; Libants et al., 2009). Although these lamprey proteins contain some of the amino acids conserved in vertebrate V1Rs, the sequences are so divergent that their phylogenetic location remains unresolved as indicated by low bootstrap values (Fig. S3).

Vomeronasal receptor family 2 (V2Rs) and other OlfCs

In this study, 61 potentially functional and five pseudogenized coelacanth V2Rs were identified (Table S1). Our phylogenetic analyses with five other vertebrate species

indicate the existence of three well-supported V2R clades (Fig. 6)[1]. The basal V2R clade has representatives from all fish and tetrapod species. The fish-specific V2R clade contains representatives from elephant shark, coelacanth and teleost fish, with few sequences forming species-specific monophyletic subclades. The third clade comprises coelacanth and tetrapod V2Rs, which predominantly segregate by species. The coelacanth V2Rs of the basal and the fish-specific V2R clades form two clusters in the genome. For these genes, the number of exons (6) is highly conserved. Most of the 56 coelacanth genes that belong to the coelacanth/tetrapod V2R clade are dispersed throughout the genome and exon numbers vary considerably (1 – 11 exons). Since this clade is experiencing accelerated evolution as indicated by numerous recent gene duplication events (Amemiya et al., 2013), intron introduction and gene truncations are possible. However, to rule out genome assembly and annotation mistakes these gene structures should be verified by targeted genome and cDNA sequencing in further studies.

In addition, we identified three potentially functional coelacanth GPCRs that show high similarity to an experimentally characterized odorant receptor in goldfish 5.24, and that cluster with the zebrafish, fugu and mouse homologs in OlfC Group III (Fig.6). The functions of these coelacanth genes remain to be investigated, since the human and mouse homologs are not bona fide olfactory receptors (Wellendorph and Bräuner-Osborne, 2004; Kuang et al., 2005).

[1] So far, the phylogenetic classification of the V2Rs and other olfactory receptors from the GPCR Family C has not been resolved satisfactorily. We refrain from introducing yet another classification system, but include and discuss below the OlfC classification proposed by Alioto and Ngai (2006).

Discussion

Taste receptors

Our analyses indicate that *L. chalumnae* may have a taste perception that is more similar to tetrapods than to fish. Taste recognition occurs via specialized epithelial sensory cells, termed taste receptor cells (TRCs) that are located in taste buds. In fish, these taste buds are found not only in the mouth, they may also occur on the barbels, in the gills, on the head, and even all over the body. Whether *L. chalumnae* or lungfish have external taste buds has not been reported. Two families of taste receptors are currently known: T1Rs and T2Rs. Recent investigations have shown that in fish not all TRCs express T1Rs or T2Rs, indicating the existence of other taste receptors (Ohmoto et al., 2011). This may explain why so far the total number of known taste receptors in teleost fish is relatively low in comparison to land vertebrates (Table 1).

Three types of T1Rs are known: T1R1, T1R2, and T1R3. These proteins form heterodimers. In mammals T1R1/T1R3 receptors perceive umami taste, while T1R2/T1R3 receptors recognize sweet tasting molecules. In teleost fish, both heterodimers are formed, but recent investigations on T1R ligands indicate that the T1R2/T1R3 dimers recognize amino acids, similar to T1R1/T1R3 receptors (Oike et al., 2007). This supports our result that the teleost fish T1R2s are closer related to the T1R1s than to the tetrapod T1R2s (Fig. 1). It also means that *L. chalumnae* may be the first fish to recognize sweet substances: two coelacanth genes grouped together with the tetrapod T1R2s, but none grouped with the teleost T1R2s.

T2Rs are bitter taste receptors, and because many toxic plant compounds taste bitter to mammals, perception and aversion to bitter substances are thought to represent a defense mechanism against ingestion of toxins. The coelacanth is the first fish species

found to have a large repertoire of T2Rs: fish genomes investigated to this date contain few if any T2Rs, whereas most studied land vertebrates harbor ~20-60 T2Rs (Table 1). Can coelacanths detect or discriminate between a wide range of bitter tasting substances and choose food accordingly? Recognition of bitter tasting substances via T2Rs and aversion to them has been reported for zebrafish (Oike et al., 2007). However, studies on stomach contents and feeding habits (swift feeding strikes without evident chewing) show that coelacanths are not prey specific (Uyeno and Tsutsumi, '91; Fricke and Hissmann, 2000). Furthermore, prey items of coelacanths include octopus, squid, and cuttlefish, and recent investigations imply that most cephalopods may produce poison to aid predation (Fry et al., 2009). Interestingly, an increasing body of evidence for mammals suggests that T2Rs have multiple additional functions (Finger and Kinnamon, 2011). So, they were found to participate in signal transduction during digestive and metabolic processes affecting absorption of bitter molecules in the gut (Behrens and Meyerhof, 2011; Jeon et al., 2011). In humans T2Rs are expressed in the nasal cavity where they help to identify bacteria and activate immune responses (Shah et al., 2009; Tizzano et al., 2010). Therefore, coelacanth T2Rs may also have very different functions besides or instead of taste. One bitter tasting substance vital for coelacanth life is urea. In general, teleost fish excrete nitrogenous waste (ammonia) directly into the water. Tetrapods dispose of toxic nitrogenous waste by converting ammonia to urea or uric acid, a process considered to represent a key adaptation to terrestrial life (Wright, '95). Coelacanths depend on urea not only for detoxification of ammonia, but also for osmoconformation, and an involvement of T2Rs in the regulation of these processes seems an intriguing hypothesis.

Chemosensory receptors of the main olfactory system (ORs and TAARs)

Table 1 Number of potentially functional GPCR-CRs in lamprey, fish and tetrapod species

species	T1Rs	T2Rs	ORs	TAARs	V1Rs/ORAs	V2R/V2R-like (OlfC I, II, IV, V)	OlfC III
lamprey	0[1,2,4]	0[1,2]	27[1]-32[3,4]	0[5]*	5[1,2]	0[1,2]	-
shark	2>[2]	0[2]	1[3]	2[5]	2[2]	32[2]	-
zebrafish	1[10]- 4[4]	4[4,10]	143[11]-154[3,4]	107[4]-112[5]	2[4,9]- 6[6]	40[4]- 44[9]-53[12]	1[12]
fugu	4[4,10]	4[4,10]	44[11]-47[3,4]	2[4]-18[5]	1[4,9]- 5[6]	18[4,9,12]	1[12]
coelacanth	**5**	**58**	**56**	**4**	**15**	**61**	**3**
frog	0[10]	49[4,10]	824[3,4]	3[5]-5[4]	21[9]- 23[7]	248[4]- 249[9]	-
lizard	3[4]	37[4]	112[3,4]	3[4]	1[4]	16[4]	-
mouse	3[4,10]	33[4]-36[8]	1035[3]-1037[4]	10[4]-15[5]	187[9]-211[4]	70[9]-121[4]	1[13]
human	3[4,10]	24[4]-25[10]	396[3,4]	5[4]-6[5]	4[4]- 5[9]	0[4,9]	1[14]

* Hussain et al. (2009) showed that the TAAR-like sequences from sea lamprey are indeed aminergic receptors,. therefore the putative lamprey TAARs published by Libants et al. (2009) and Dong et al. (2012) are omitted in this table 1=Libants et al., 2009; 2=Grus and Zhang, 2009; 3=Niimura, 2009; 4=Dong et al., 2012; 5=Hussain et al., 2009; 6=Saraiva and Korsching, 2007; 7=Date-Ito et al., 2003; 8=Wu et al., 2005; 9=Shi and Zhang, 2007; 10=Shi and Zhang, 2006; 11=Alioto and Ngai, 2005; 12=Alioto and Ngai, 2006; 13=Kuang et al., 2005; 14=Wellendorph and Bräuner-Osborne, 2004

In vertebrates, the main olfactory system (MOS) consists of olfactory sensing neurons located in the nasal olfactory epithelium with axons reaching through the olfactory nerve into the olfactory bulb of the brain. In general, each neuron expresses only a single odorant receptor gene (either OR or TAAR). However, each receptor may recognize several types of odor molecules and a single odorant molecule can activate various receptors. This motif-like recognition method allows for identification of a vast number of odorants (reviewed in Mombaerts, 2004; Tirindelli et al., 2009). It explains why the ORs are not only one of the largest, but also most diverse gene families in vertebrate genomes. Their numbers vary from two in elephant shark (Niimura, 2009) to ~1300 in pig (Groenen et al., 2012). In general, teleost fish have

fewer numbers, but a higher diversity of ORs than tetrapods. To date, the ORs of jawed vertebrates have been classified into Type 1 with six OR subgroups (α-ζ) and Type 2 with one OR subgroup (η) (Niimura and Nei, 2005; Niimura, 2009). ORs from the groups α and γ are present in great numbers in tetrapods but seem to be absent in teleost fish (except for one zebrafish gene). These are predicted to detect airborne odorants. In contrast, ORs from the groups δ, ε, ζ, and η are found in teleost fish and amphibians, but not in reptiles, birds or mammals. These are predicted to detect water-soluble odorants. Our genome-wide analysis revealed that *L. chalumnae* has a variable repertoire of ORs. With 64 genes in six subgroups, the number of putative ORs in *L. chalumnae* is comparable to those in teleost fish, and the variability of these proteins is similar to that in amphibians. Consistent with its habitat, most of the coelacanth genes grouped with ORs predicted to detect water-soluble odorants. Seven coelacanth ORs formed basal clades in the groups α and γ. Together with the single zebrafish gene in group γ, these genes potentially represent the ancestral clades of ORs hypothesized to detect airborne odorants.

The trace amine associated receptors, which are also expressed in the MOS, serve as specialized receptors for trace amines and related compounds (Liberles and Buck, 2006). Hussain et al. (2009) categorized the TAARs into three classes with 28 subfamilies. Their studies indicate that class III seems to be a new receptor family that is rapidly evolving in teleost fish and may not recognize amines due to loss of the aminergic ligand-binding motif. Not counting the members of class III, the numbers of TAARs are similar in teleost fish (4-25) and land vertebrates (3-17). Our phylogenetic analyses indicate that *L. chalumnae* may have only one TAAR functioning as an olfactory receptor. Three of the four coelacanth TAARs grouped with proteins from TAAR1, a highly conserved subfamily with a single ortholog in

cartilaginous fish, teleost fish, and tetrapod species, respectively. Expression analyses indicate that this particular gene subfamily is generally not involved in olfaction, and as in most tested species TAAR1 is not expressed in the olfactory epithelium, but in the brain (Liberles and Buck, 2006; Hashiguchi and Nishida, 2007; but see Gliem et al., 2009). The sole remaining coelacanth TAAR (LCH_01339) grouped with the zebrafish proteins of subfamily TAAR13. Interestingly, one of the *D. rerio* receptors from this subfamily (Q5QNP2) was shown to respond to cadaverine and putrescine, two compounds that are generated during putrefication of dead fish (Hussain, 2010). These two compounds increased feeding activity in goldfish (Rolen et al., 2003) and induced avoidance behavior in zebrafish (Hussain, 2010) and therefore represent important feeding and behavior cues. The *L. chalumnae* TAAR repertoire is therefore most similar to the one from shark that has one gene in the subfamily TAAR1 and one gene with basal location relative to TAAR13.

Vomeronasal receptor gene families

Most land vertebrates possess an accessory olfactory system, the vomeronasal system (VNS). It consists of a vomeronasal organ and an accessory olfactory bulb that are anatomically distinct from the organs of the MOS. Similar to the olfactory neurons, each vomeronasal neuron expresses only one type of receptor (V1R, V2R, or OR). Current investigations indicate that V1Rs predominantly recognize volatile compounds, while V2Rs mainly respond to water-soluble molecules (reviewed in Mombaerts, 2004; Tirindelli et al., 2009). VNS specific genes, including V1Rs, V2Rs, and the Trpc2 channel protein were also found in teleost fish, cartilaginous fish, and some even in sea lamprey (a jawless vertebrate), and expression patterns imply their

involvement in olfaction (Pfister and Rodriguez, 2005; Gruz and Zhang, 2006; Gruz and Zhang, 2009). However, a morphologically distinct VNS is missing in these species. Our results indicate that *L. chalumnae* not only retained some of the ancestral vomeronasal genes present in teleost fish, it also experienced a species-specific expansion of vomeronasal receptors as observed in land vertebrates. So, the number of potentially functional V1Rs in *L. chalumnae* is more similar to the frog, than to the elephant shark or teleost fish (Table 1, Figure 5). While five coelacanth V1Rs appear in the conserved teleost clades *ora*1, *ora*2, and *ora*3-*ora*4, another eight coelacanth V1Rs form a monophyletic group. Similarly, our V2R analyses show that *L. chalumnae* genes do not only appear in the fish-specific clade and in the tetrapod clade; they also mimic the respective evolutionary patterns (high diversity in the fish-specific V2R clade and formation of a large monophyletic subclade in the tetrapod V2R clade). A possible explanation for the observed increase in numbers and diversity of vomeronasal receptors in *L. chalumnae* in comparison to other fish could be an advance development of the VNS in lobe-finned fish. Recently, a primordial vomeronasal system was discovered in lungfish, the only other living group of lobe-finned fish (Gonzalez et al., 2010; Nakamuta et al., 2012). However, further analyses are required to verify that these genes are actually involved in olfaction, and whether *L. chalumnae* also has a primordial VNS. If confirmed, the expansion of V1R genes in lobe-finned fish may question the hypothesis that these receptors exclusively recognize airborne molecules.

As mentioned above, the classification of the V2Rs and other olfactory receptors of the C family GPCRs is not yet resolved satisfactorily. Yang et al. (2005) classified only three mammalian V2R subclades into the families A, B, and C, ignoring other vertebrate V2R clades. Alioto and Ngai (2006) proposed a more extensive

classification system (V2R-like/OlfC). However, they include a clade, OlfC group III, which is more closely related to T1Rs (ergo not V2R-like), and that contains mammalian GPCRs not involved in olfaction (ergo not OlfCs). Here, we refrain from introducing a new classification system, but suggest that this matter should be addressed by a consortium of experts, which should take into account the multitude of functions recently discovered for various GPCR-CRs.

Conclusions

Our analyses on the chemosensory receptors of *L. chalumnae* revealed a gene repertoire in transition, with teleost-like and tetrapod-like genes. Recent analyses on these receptors in mammals indicate that many have additional functions apart from smell or taste perception. Whether this multi-functionality is associated with the species-specific expansion of gene families in tetrapods or is already present in fish remains to be seen. Studies on expression and functions of CRs in teleost and lobe-finned fish may shed light on this matter.

Acknowledgments

We would like to thank J-B Domelevo Entfellner for advise on statistical verification of the T1R tree. This work was supported by The South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation: Grant 64751.

References

Alioto TS and Ngai J. 2005. The odorant receptor repertoire of teleost fish. BMC Genomics 6:173-187.

Alioto TS and Ngai J. 2006. The repertoire of olfactory C family G protein-coupled receptors in zebrafish: candidate chemosensory receptors for amino acids. BMC Genomics 7:309-327.

Amemiya CT, Alföldi J, Lee AP, Fan S, Philippe H, MacCallum I, Braasch I, Manousaki T, Schneider I, Rohner N, Organ C, Chalopin D, Smith JJ, Robinson M, Dorrington RA, Gerdo M, Aken B, Biscotti MA, Barucca M, Baurain D, Berlin AM, Blatch GL, Buonocore F, Burmester T, Campbell MS, Canapa A, Cannon JP, Christoffels A, De Moro G, Edkins AL, Fan L, Fausto AM, Feiner N, Forconi M, Gamieldien J, Gnerre S, Gnirke A, Goldstone JV, Haerty W, Hahn ME, Hesse U, Hoffmann S, Johnson J, Karchner SI, Kuraku S, Lara M, Levin JZ, Litman GW, Mauceli E, Miyake T, Mueller MG, Nelson DR, Nitsche A, Olmo E, Ota T, Pallavicini A, Panji S, Picone B, Ponting CP, Prohaska SJ, Przybylski D,

Saha RN, Ravi V, Ribeiro FJ, Sauka-Spengler T, Scapigliati G, Searle SMJ, Sharpe T, Simakov O, Stadler PF, Stegeman JJ, Sumiyama K, Tabbaa D, Tafer H, Turner-Maier J, van Heusden P, White S, Williams L, Yandell M, Brinkmann H, Volff1JN, Tabin CJ, Shubin N, Schartl M, Jaffe DB, Postlethwait JH, Venkatesh JH, Di Palma F, Lander ES, Meyer A and Lindblad-To K. 2013. Analysis of the African coelacanth genome sheds light on the tetrapod evolution. Nature 196:311-316.

Anisimova M and Gascuel O. 2006. Approximate likelihood ratio test for branches: A fast, accurate and powerful alternative. Syst Biol 55:539-552.

Abascal F, Zardoya R, Posada D. 2005. ProtTest: Selection of best-fit models of protein evolution. Bioinformatics 21:2104-2105.

Behrens M, Meyerhof W. 2011. Gustatory and extragustatory functions of mammalian taste receptors. Physiol Behav 105:4-13.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: A sequence logo generator. Genome Res 14:1188-1190.

Dong D, Jin K, Wu X, Zhong Y. 2012. CRDB: Database of Chemosensory Receptor Gene Families in Vertebrate. PLoS One 7(2): e31540.

Finger TE, Kinnamon SC. 2011. Taste isn't just for taste buds anymore. Biol Rep 3:20-27.

Freitag J, Ludwig G, Andreini I, Rossler P, Breer H. 1998. Olfactory receptors in aquatic and terrestrial vertebrates. J Comp Physiol [A] 183:635-650.

Fricke H, Hissmann K. 2000. Feeding ecology and evolutionary survival of the living coelacanth Latimeria chalumnae. Marine Biol 136:379-386.

Gliem S, Schild D, Manzini I. 2009. Highly specific responses to amine odorants of individual olfactory receptor neurons *in situ*. Eur J Neurosci 29:2315-2326.

Gonzalez A, Morona R, Lopez JM, Moreno N, Northcutt RG. 2010. Lungfishes, like tetrapods, possess a vomeronasal system. Front Neuroanat 4 (130): 1-11.

Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, Li S, Larkin DM, Kim H, Frantz LA, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi SH, Chow W, Clark RC, Clee C, Crooijmans RP, Dawson HD, Dehais P, De Sapio F, Dibbits B, Drou N, Du ZQ, Eversole K, Fadista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JG, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu ZL, Humphray SJ, Hunt T, Hornshøj H, Jeon JT, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim JH, Kim KW, Kim TH, Larson G, Lee K, Lee KT, Leggett R, Lewin HA, Li Y, Liu W, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L, McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni P, Oh SJ, Onteru S, Panitz F, Park EW, Park HS, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang Y, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang G, Zhang J, Zhang J, Zhao S, Rogers J, Churcher C, Schook LB. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491:393-398.

Grus WE, Zhang J. 2006. Origin and evolution of the vertebrate vomeronasal system viewed through system-specific genes. Bioessays 28:709-718.

Grus WE, Zhang J. 2009. Origin of the genetic components of the vomeronasal system in the common ancestor of all extant vertebrates. Mol Biol Evol 26:407-419.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and Methods to estimate Maximum-Likelihood Phylogenies. Syst Biol 59:307-321.

Hashiguchi Y, Nishida M. 2007. Evolution of Trace Amine-Associated Receptor. TAAR. Gene Family in Vertebrates: Lineage-Specific Expansions and Degradations of a Second Class of Vertebrate Chemosensory Receptors Expressed in the Olfactory Epithelium. Mol Biol Evol 24:2099-2107.

Huang ES. 2009. Construction of a sequence motif characteristic of aminergic G protein-coupled receptors. Protein Sci 12:1360-1367.

Hussain A. 2010. The teleost taar family of olfactory receptors: From rapidly evolving receptor genes to ligand-induced behavior. PhD Thesis. Universität zu Köln.

Hussain A, Saraiva LR, Korsching SI. 2009. Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts. Proc Natl Acad Sci USA 106:4313-4318.

Ishimaru Y, Okada S, Naito H, Nagai T, Yasuoka A, Matsumoto I, Abe K, others. 2005. Two families of candidate taste receptors in fishes. Mech Dev 122:1310-1321.

Jeon TI, Seo YK, Osborne TF. 2011. Gut bitter taste receptor signaling induces ABCB1 through a mechanism involving CCK. Biochem J 438:33-7.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567-580.

Kuang D, Yao Y, Lam J, Tsushima RG, Hampson DR. 2005. Cloning and characterization of a family C orphan G-protein coupled receptor. J Neurochem 93:383-391.

Libants S, Carr K, Wu H, Teeter JH, Chung-Davidson YW, Zhang Z, Wilkerson C, Li W. 2009. The sea lamprey Petromyzon marinus genome reveals the early origin of several chemosensory receptor families in the vertebrate lineage. BMC Evol Biol 9:180-194.

Liberles SD, Buck LB. 2006. A second class of chemosensory receptors in the olfactory epithelium. Nature 442:645-650.

Mombaerts P. 2004. Genes and ligands for odorant, vomeronasal and taste receptors. Nature Reviews Neuroscience 5:263-278.

Nakamuta S, Nakamuta N, Taniguchi K, Taniguchi K. 2012. Histological and Ultrastructural Characteristics of the Primordial Vomeronasal Organ in Lungfish. The Anatomical Record 295:481-491.

Niimura Y. 2009. On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. Genome Biol Evol 1:34-44.

Niimura Y, Nei M. 2005. Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. Proc Natl Acad Sci USA 102:6039-6044.

Notredame C, Higgins DG, Heringa J. 2000. T-coffee: A novel method for multiple sequence alignments. J Mol Biol 302:205-217.

Ohmoto M, Okada S, Nakamura S, Abe K, Matsumoto I. 2011. Mutually exclusive expression of Gαia and Gα14 reveals diversification of taste receptor cells in zebrafish. J Comp Neurol 519: 1616-1629.

Oike H, Nagai T, Furuyama A, Okada S, Aihara Y, Ishimaru Y, Marui T, Matsumoto I, Misaka T, Abe K. 2007. Characterization of ligands for fish taste receptors. J Neurosci 27:5584-5592.

Pfister P, Rodriguez I. 2005. Olfactory expression of a single and highly variable V1r pheromone receptor-like gene in fish species. Proc Natl Acad Sci USA 102:5489-5494.

Rolen SH, Sorensen PW, Mattson D, Caprio J. 2003. Polyamines as olfactory stimuli in the goldfish Carassius auratus. J Exp Biol 206:1683-1696.

Saraiva LR, Korsching SI. 2007. A novel olfactory receptor gene family in teleost fish. Genome Res 17:1448-1457.

Shah AS, Ben-Shahar Y, Moninger TO, Kline JN, Welsh MJ. 2009. Motile cilia of human airway epithelia are chemosensory. Science 325:1131-1134.

Shi P, Zhang J. 2006. Contrasting modes of evolution between vertebrate sweet/umami receptor genes and bitter receptor genes. Mol Biol Evol 23:292-300.

Shi P, Zhang J. 2007. Comparative genomic analysis identifies an evolutionary shift of vomeronasal receptor gene repertoires in the vertebrate transition from water to land. Genome Res 17:166-174.

Tirindelli R, Dibattista M, Pifferi S, Menini A. 2009. From pheromones to behavior. Physiol Rev 89:921-956.

Tizzano M, Gulbransen BD, Vandenbeuch A, Clapp TR, Herman JP, Sibhatu HM, Churchill MEA, Silver WL, Kinnamon SC, Finger TE. 2010. Nasal chemosensory cells use bitter taste signaling to detect irritants and bacterial signals. Proc Natl Acad Sci USA 107:3210-3215.

Uyeno T, Tsutsumi T. 1991. Stomach contents of Latimeria chalumnae and further notes on its feeding habits. Enviro Biol Fishes 32:275-279.

Vroling B, Sanders M, Baakman C, Borrmann A, Verhoeven S, Klomp J, Oliveira L, de Vlieg J, Vriend G. 2011. GPCRDB: information system for G protein-coupled receptors. Nucleic Acids Res 39:309-319.

Wellendorph P, Bräuner-Osborne H. 2004. Molecular cloning, expression, and sequence analysis of GPRC6A, a novel family C G-protein-coupled receptor. Gene 335:37-46.

Wright PA. 1995. Nitrogen excretion: three end products, many physiological roles. J Exp Biol 198:273-281.

Wu SV, Chen MC, Rozengurt E. 2005. Genomic organization, expression, and function of bitter taste receptors. T2R. in mouse and rat. Physiol Genomics 22:139 -149.

Yang H, Shi P, Zhang Y, Zhang J. 2005. Composition and evolution of the V2R vomeronasal

receptor gene repertoire in mice and rats. Genomics 86:306 - 315.

Zdobnov EM, Apweiler R. 2001. InterProScan - an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847-848.

Figure Legends

Figure 1. Phylogenetic tree of the T1R genes. The tree was constructed using Maximum Likelihood (ML) and a WAG substitution model as the best fit for the protein dataset. Zebrafish V2Rs served as outgroup (black); Blue = teleost T1Rs (zebrafish, fugu, pufferfish), green = tetrapod T1Rs (lizard, human), red = coelacanth T1Rs.

Figure 2. Phylogenetic tree of the T2R genes. The unrooted tree was constructed using Maximum Likelihood (ML) and a JTT substitution model as the best fit for the protein dataset. Blue = teleost T2Rs (zebrafish, fugu, pufferfish), green = tetrapod T2Rs (frog, chicken, human), red = coelacanth T2Rs.

Figure 3. Phylogenetic tree of the OR genes. The tree was constructed using Maximum Likelihood (ML) and an LG substitution model as the best fit for the protein dataset. Six non-OR GPCR genes were used as outgroup (black). OR classification adopted from Niimura and Nei (2005). Blue = zebrafish ORs, green = human ORs, red = coelacanth ORs.

Figure 4. Phylogenetic tree of the TAAR genes. The tree was constructed using Maximum Likelihood (ML) and an LG substitution model as the best fit for the protein dataset. TAAR classes and subclasses are adopted from Hussein et al. (2009). The sea lamprey genes as outgroup (black); Violet = elephant shark TAARs; blue = teleost TAARs (zebrafish, fugu), green = tetrapod TAARs (frog, lizard, mouse, human), red = coelacanth TAARs.

Figure 5. Phylogenetic tree of the V1R/ORA genes. The tree was constructed using Maximum Likelihood (ML) and a JTT substitution model as the best fit for the protein dataset. Zebrafish T2Rs served as outgroup (black). Violet = elephant shark ORAs; blue = teleost ORAs (zebrafish), green = tetrapod V1Rs (frog, lizard, human), red = coelacanth V1Rs. The two frog and the lizard V1Rs that grouped separately from other tetrapods were identified by Dong et al. (2012).

Figure 6. Phylogenetic tree of the V2R and other OlfC genes. The tree was constructed using Maximum Likelihood (ML) and a JTT substitution model as the best fit for the protein dataset. The OlfCs classification system is adopted from Alioto and Ngai (2006), see Discussion. To show the phylogenetic position of OlfC Group III, we included CaSR, T1R, and mGluR genes from zebrafish and mouse (black). Violet = elephant shark V2Rs; blue = teleost V2Rs (zebrafish, fugu), green = tetrapod V2Rs (frog, mouse), red = coelacanth V2Rs.

Table Legend

Table 1. Number of potentially functional GPCR-CRs in lamprey, fish and tetrapod species

Supplementary Material

Figure S1. Phylogenetic tree of the OR genes with all gene names and bootstrap values. The tree was constructed using Maximum Likelihood (ML) and an LG substitution model as the best fit for the protein dataset. Six non-OR GPCR genes were used as outgroup (black). Blue = zebrafish ORs, green = human ORs, red = coelacanth ORs.

Figure S2. Conservation of amino acids in the 15 putative *L. chalumnae* V1Rs as compared to conserved amino acids in the ORAs of teleost fish (Saraiva and Korsching, 2007) and V1Rs in mouse (Pfister and Rodriguez, 2005). The sequence logo was constructed using WebLogo (Crooks et al., 2004).

Figure S3. Phylogenetic tree of the V1R/ORA genes including five putative lamprey ORAs. The tree was constructed using Maximum Likelihood (ML) and a JTT substitution model as the best fit for the protein dataset. Zebrafish T2Rs served as outgroup (black). Grey = sea lamprey ORAs, violet = elephant shark ORAs; blue = teleost ORAs (zebrafish), green = tetrapod V1Rs (frog, lizard, human), red = coelacanth V1Rs.

Table S1. Annotations for all *L. chalumnae* Chemosensory Receptors identified in this study.

Table S2. Mapping information for the protein identifiers used in this study to the respective gene/protein identifiers from the original sources for all GPCR-CR from species other than *L. chalumnae*.

Table S3. Statistical verification of the T1R tree topology.