Italian Journal of Zoology

ISSN: 1125-0003 (Print) 1748-5851 (Online) Journal homepage: http://www.tandfonline.com/loi/tizo20

Positive selection effects on the biochemical properties of fish pyroglutamylated RFamide peptide receptor (QRFPR)

R. Bakiu, A. M. Tolomeo & G. Santovito

To cite this article: R. Bakiu, A. M. Tolomeo & G. Santovito (2015) Positive selection effects on the biochemical properties of fish pyroglutamylated RFamide peptide receptor (QRFPR), Italian Journal of Zoology, 82:4, 460-472, DOI: 10.1080/11250003.2015.1071437

To link to this article: http://dx.doi.org/10.1080/11250003.2015.1071437

Italian Journal of Zoology, 2015, 460–472
Vol. 82, No. 4, http://dx.doi.org/10.1080/11250003.2015.1071437

Taylor & Francis
Taylor & Francis Group

Positive selection effects on the biochemical properties of fish pyroglutamylated RFamide peptide receptor (QRFPR)

R. BAKIU[1]*, A. M. TOLOMEO[2], & G. SANTOVITO[2]

[1]*Department of Aquaculture and Fisheries, Agricultural University of Tirana, Albania, and* [2]*Department of Biology, University of Padova, Italy*

(Received 16 April 2015; accepted 2 July 2015)

Abstract
Orphan receptor GPR103, a pyroglutamylated RFamide peptide receptor (QRFPR), is a class-A G protein-coupled receptor (GPCR) and it is coupled to a Gi alpha subunit (Gi/0) and/or to a Gq protein. Synteny analysis revealed the existence of *qrfpr* paralogous genes in mouse, zebrafish and coelacanth. These paralogous genes emerged along with the species-specific gene or genome duplications that occurred during vertebrate evolution. Neuropeptide 26RFa (also termed QRFP) is the latest member of the RFamide peptide family to be discovered in the hypothalamus of vertebrates. 26RFa/QRFP is a 26-amino acid residue peptide that was originally identified from the frog brain. It has been shown to exert orexigenic activity in mammals and to be a ligand of the previously identified orphan G protein-coupled receptor, QRFPR. The structure, tissue-specific expression and biochemical activity of the 26RFa/QRFP–QRFPR system are conserved across the Chordata phylum, from fish to mammals. In order to study the molecular evolution of fish QRFPRs, we investigated the presence of natural selection on the QRFPR family using a bioinformatic approach. Overall, the obtained results clearly indicate that fish QRFPRs are under positive selection, but the positively selected amino acids did not significantly alter the biochemical properties of these proteins.

Keywords: *Fish, molecular evolution, positive selection, QRFPR*

Introduction

Peptides, defined by their carboxy-terminal arginine (R) and amidated phenylalanine (F) residues (RFamide), have been identified in the nervous systems of animals within all major phyla. The first recognised member of the RFamide neuropeptide family was the cardioexcitatory peptide, Phe-Met-Arg-Phe-amide (FMRFamide), isolated from ganglia of the clam *Macrocallista nimbosa* (Lightfoot, 1786) (Price & Greenberg 1977). Vertebrates and more especially invertebrates can each express an array of RFamide peptides, owing to the fact that multiple genes encoding RFamides are often present in a single species, and multiple mature RFamide peptides can be generated by a single polypeptide precursor. These peptides seem to act as neurotransmitters and neuromodulators (Walker et al. 2009). Immunohistochemical studies that used antisera against FMRFamide suggested that the nervous system of vertebrates also contain neuropeptides immunologically related to FMRFamide (Raffa 1988; Rastogi et al. 2001). In fact, several neuropeptides harbouring the RFamide sequence at their C-terminus have been characterised in the brain of various vertebrates. In the past, the existence of five groups within the RFamide peptide family has been recognised in vertebrates, namely the neuropeptide FF (NPFF) group, the prolactin-releasing peptide (PrRP) group, the gonadotropin-inhibitory hormone (GnIH) group, the kisspeptin group, and the 26RFa/QRFP group (Chartrel et al. 2011; Leprince et al. 2013). These RFamide peptides have been shown to exert important neuroendocrine, behavioural, sensory and autonomic functions (Ukena & Tsutsui 2005; Tsutsui & Ukena 2006).

In humans, 26RFa/QRFP has been found to be an endogenous ligand for the orphan receptor, GPR103-QRFPR, which is a class-A G protein-coupled receptor

*Correspondence: R. Bakiu, Department of Aquaculture and Fisheries, Agricultural University of Tirana, Tirane 1000, Albania. Tel: +35 5694769532. Fax: +35 54200874. Email: bakiurigers@gmail.com

(GPCR; Jiang et al. 2003). Pyroglutamylated RFamide peptide receptor (QRFPR) shares relatively high sequence similarity with other RFamide receptors, notably those for NPFF, PrRP, kisspeptin and GnIH, and to a lesser extent with other peptidergic receptors for neuropeptide Y (NPY), galanin, orexin and cholecystokinin (Jiang et al. 2003). In addition, QRFPR possesses several features characteristic of class-A GPCRs, such as: (1) a disulphide bridge between the two Cys residues located in the first and second extracellular loops (EL1 and EL2); (2) the existence of an Asp residue within the second transmembrane domain (TM2), that seems to play a pivotal role in G protein coupling; (3) a conserved Glutamic acid-Arginine (Glu-Arg) doublet sequence at the N-terminus of the second intracellular loop (IL2); (4) three conserved residues, i.e. Phe, Pro and Asn, within TM6 and TM7, which are crucial for receptor activation (Ukena et al. 2014). Previous results suggested that QRFPR is coupled to a Gi/0 and/or to a Gq protein (Fukusumi et al. 2003). Furthermore, it has been reported that 26RFa/QRFP enhances corticosteroid secretion in human adrenocortical cells by regulating key steroidogenic enzymes involving Mitogen-Activated Protein Kinase/Protein Kinase C (MAPK/PKC) and calcium ion (Ca^{2+}) signalling pathways via QRFPR (Ramanjaneya et al. 2013).

In mammals, 26RFa/QRFP has been found to be a high-affinity endogenous ligand for the previously identified QRFPR. In rodents and monkeys, 26RFa/QRFP plays diverse biological roles, including regulation of food intake and energy homeostasis, hormone secretion, nociception and bone formation. Recently, the mature sequences of 26RFa/QRFP have been identified by structural analysis in quail and zebra finch. In birds, as well as in mammals, 26RFa/QRFP-producing neurons are only located in the hypothalamus, while QRFPR is widely distributed throughout the brain. In birds, 26RFa/QRFP also exerts an orexigenic action, as it does in rodents, and a similar effect of 26RFa/QRFP has been suggested in fish, because of upregulation of 26RFa/qrfp mRNA by a negative energy state (Ukena et al. 2014).

Synteny analysis revealed the existence of paralogous *qrfpr* genes in mouse, zebrafish and coelacanth (Ukena et al. 2014). These genes could have emerged along with species-specific gene or genome duplications that occurred during the evolution of vertebrates. Phylogenetic analysis data are consistent with synteny analysis (Ukena et al. 2014). Although in the genome database there are homologous QRFPR sequences of *Xenopus*, zebrafish, coelacanth and lamprey (Ukena et al. 2014), *qrfpr* genes have been studied only in mammals and birds. However, the structure, distribution pattern and biological actions of the 26RFa/QRFP–QRFPR system have been conserved across the vertebrate phylum, from fish to mammals (Ukena et al. 2014). Therefore, in order to study the molecular evolution of these proteins and have more information on functional significance of 26RFa/QRFP–QRFPR pair in vertebrates, we verified the presence of natural selection on QRFPR from a heterogeneous taxonomic group: fish.

Materials and methods

Coding region and amino acid sequences of fish QRFPRs were available in the European Molecular Biology Laboratory (EMBL)/GenBank database (Table I).

The T-Coffee multiple sequence alignment software package was used to obtain multiple sequence alignment (Notredame et al. 2000). We decided to use theT-Coffee program in order to align amino acid and nucleotide sequences, because even though the method is based on the popular progressive approach to multiple alignment, it is characterised by a dramatic improvement in accuracy with a modest sacrifice in speed as compared to the most commonly used alternatives (Notredame et al. 2000).

jModelTest 2.0 (Darriba et al. 2012) was used to carry out statistical selection of the best-fit model of nucleotide substitution. Analyses were performed using 88 candidate models and three types of criteria [Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (cAIC) and Bayesian Information Criterion (BIC)].

To select the best-fit model for the analysis of protein evolution, ProtTest 3 was used (Darriba et al. 2011). One hundred and twenty-two candidate models and the three previously mentioned criteria were used in these statistical analyses.

Phylogenetic trees were built using the Bayesian inference (BI) method implemented in Mr. Bayes 3.2 (Ronquist et al. 2012), and the Maximum Likelihood (ML) method implemented in PhyML 3.0 (Guindon et al. 2010). In the BI analyses, four independent runs, each with four simultaneous Markov Chain Monte Carlo (MCMC) chains, were performed for 1,000,000 generations sampled every 1000 generations. Using PhyML 3.0, bootstrap analyses were performed on 100,000 trees using two algorithms of tree topology improvement, nearest neighbour interchange (NNI) and subtree pruning and regrafting (SPR). FigTree (http://tree.bio.ed.ac.uk/software/figtree/) software was used to display the annotated phylogenetic trees.

Table I. QRFPR sequences used for phylogeny, and their EMBL/GenBank accession numbers.

Sequence name	Nucleotide accession number	Protein accession number
Poecilia formosa QRFPR3	XM_007558305.1	XP_007558367.1
Danio rerio QRFPR 3a	ENSDART00000098961	ENSDARP00000089731
Danio rerio QRFPR 3b	ENSDART00000041026	ENSDARP00000041025
Danio rerio QRFPR 4	ENSDART00000135119	ENSDARP00000113533
Dicentrarchus labrax QRFPR 3	FQ310507.3	CBN81515.1
Poecilia reticulata QRFPR 3	XM_008436920.1	XP_008435142.1
Cynoglossus semilaevis QRFPR 3	XM_008314519.1	XP_008312741.1
Stegastes partitus QRFPR 3	XM_008294887.1	XP_008293109.1
Oryzias latipes QRFPR 3	XM_004080411.1	XP_004080459.1
Astyanax mexicanus QRFPR 3	XM_007235442.1	XP_007235504.1
Astyanax mexicanus QRFPR 4	XM_007240638.1	XP_007240700.1
Neolamprologus brichardi QRFPR 3	XM_006791111.1	XP_006791174.1
Lepisosteus oculatus QRFPR 2	XM_006630338.1	XP_006630401.1
Lepisosteus oculatus QRFPR 3	XM_006625602.1	XP_006625665.1
Lepisosteus oculatus QRFPR 4	XM_006625648.1	XP_006625711.1
Haplochromis burtoni QRFPR 3	XM_005919582.1	XP_005919644.1
Xiphophorus maculatus QRFPR 3	XM_005794588.1	XP_005794645.1
Pundamilia nyererei QRFPR 3	XM_005736424.1	XP_005736481.1
Oreochromis niloticus QRFPR 3x1	XM_003441907.2	XP_003441955.2
Oreochromis niloticus QRFPR 3x2	XM_005471374.1	XP_005471431.1
Maylandia zebra QRFPR 3	XM_004561252.1	XP_004561309.1
Takifugu rubripes QRFPR 3	XM_003961238.1	XP_003961287.1
Latimeria chalumnae QRFPR 2	XM_005988242.1	XP_005988304.1
Homo sapiens QRFPR 1	ENST00000394427	ENSP00000377948
Mus musculus QRFPR1-1	ENSMUST00000091227	ENSMUSP00000088768
Mus musculus QRFPR1-2	ENSMUST00000170608	ENSMUSP00000130225
Monodelphis domestica QRFPR1	ENSMODT00000024050	ENSMODP00000023630
Ornithorhynchus anatinus QRFPR 1	ENSOANT00000023802	ENSOANP00000023798
Pelodiscus sinensis QRFPR 1	ENSPSIT00000016356	ENSPSIP00000016280
Pelodiscus sinensis QRFPR 2	ENSPSIT00000007026	ENSPSIP00000006986
Anolis carolinensis QRFPR 1	ENSACAT00000012987	ENSACAP00000012731
Xenopus tropicalis QRFPR 1	ENSXETT00000004500	ENSXETP00000004500

Genetic algorithm recombination detection (GARD) and single breakpoint recombination (SBP) programs (Kosakovsky Pond et al. 2006) were used to identify possible breakpoints in the *qrfpr* codon sequences.

In order to detect the presence of positive selection in QRFPR molecular evolution, we used statistical methods implemented in the HyPhy package (Kosakovsky Pond et al. 2005); single likelihood ancestor (SLAC), fixed effects likelihood (FEL), random effect likelihood (REL) and fast unconstrained Bayesian approximation (FUBAR) software is useful to detect the presence of possible positive selection, and mixed effects model evolution (MEME) is an excellent program to detect sites under episodic diversifying selection (Murrell et al. 2012).

The property-informed models of evolution (PRIME) program (http://hyphy.org/w/index.php/PRIME) was used to detect the variation of QRFPR biochemical properties which are driving the amino acid substitutions.

Results

Molecular clock tests

We performed a Bayes factor comparison using Mr. Bayes 3.2 to test the strict clock model against the *non*-clock (unconstrained) model using *qrfpr* cDNA coding region sequences. Furthermore, we used the same approach to test all three relaxed clock models implemented in MrBayes 3.2. The models are the Thorne–Kishino 2002 (TK02) model (Thorne & Kishino 2002), the compound Poisson process (CPP) model (Huelsenbeck et al. 2000) and the independent gamma rates (IGR) model (Lepage et al. 2007). We generated an accurate assessment of the marginal model likelihoods using the stepping-stone method. It estimates the model likelihood by sampling a series of distributions that represent different mixtures of the posterior distribution and the prior distribution (Xie et al. 2011). The stepping-stone method was applied to the *qrfpr* data set using 510,000 generations with a diagnostic frequency of 2500 in two independent runs

Table II. Marginal likelihood values in each of the two independent runs, and the resulting mean values for each of the tested models using the stepping-stone method. Compound Poisson Process (CPP), Thorne–Kishino 2002 (TK02) and Independent Gamma Rates (IGR) model.

Run	Unconstrained	Strict clock	Relaxed clock (CPP)	Relaxed clock (TK02)	Relaxed clock (IGR)
1	−15,365.05	−15,352.37	−15,318.84	−15,337.07	−15,339.37
2	−15,361.08	−15,351.63	−15,317.76	−15,335.87	−15,344.60
Mean of Marginal Likelihood	−15,361.75	−15,351.94	−15,318.16	−15,336.30	−15,340.23

for each of the tested models. The marginal likelihood values are shown in Table II. The CPP model (−15,318.16) is 34 log likelihood units better than the strict-clock model (−15,351.94) and nearly 18 log and 12 log likelihood units better than the other two local clock TK02 (−15,336.3) and IGR (−15,340.23) models, respectively.

Fish QRFPR sequences analyses

The multiple alignment of fish QRFPR amino acid sequences is reported in Figure 1. The total score value was 97, indicating that the multiple alignment was characterised by a high level of similarity among the different sequences. Furthermore all fish QRFPRs have the features characteristic of class-A GPCRs, represented by the putative Transmembrane Domains (TMDs), a disulphide bridge between two Cys residues located in the first and second extracellular loops, the Asp residue in TMD2 involved in G protein coupling, the conserved Glu-Arg residues in the second intracellular loop and the conserved Phe, Pro and Asn residues in TMD6 and TMD7.

Phylogeny of QRFPR

The CPP molecular clock model was used as evolution model in cDNA and amino acid phylogenetic tree building. All cDNA coding region sequences were aligned using T-Coffee in combined libraries of local and multiple alignments, which are known to induce high accuracy and performance in sequence alignments. jModelTest 2.0 software indicated that the GTR+G model is the best-fit model to analyse the evolution of *qrfpr* coding sequences, with a gamma shape value (four rate categories) of 0.607 using all statistical criterion: AIC, cAIC and BIC (−lnL = 15,162.51). BI and ML analyses generate phylogenies with the same topology (Figure 2).

ProtTest3 statistical results determined the JTT +G model as the best one to apply for the phylogenetic analysis of QRFPR amino acid sequences, with a gamma shape value (four rate categories) of 0.624 using all statistical criteria (−lnL = −7557.55). Figure 3 shows the common topology generated by the application of the BI and ML methods. The sequence used as out-group to root the tree, the amphioxus QRFP receptor, always branches off basally to all vertebrate QRFPR sequences, according to the results of Larhammar et al. (2014).

The cladogram based on amino acid sequences (Figure 3) is better resolved with respect to that obtained with open reading frames (ORFs), and all of them are supported by high posterior probabilities and bootstrap values of the BI and ML analyses, respectively. However, in the cladogram based on amino acid sequences, fewer nodes were supported by bootstrap values higher than 50%, even if all nodes were supported by posterior probability values higher than 50%. In both phylogenetic trees, QRFPR1, QRFPR2 and QRFPR4 are clearly grouped in the corresponding clade, but all of them are included in a "big" clade, which is separated from the clade containing all the QRFPR3 sequences. However, only in the cladogram based on amino acid sequences did a clade emerge containing all QRFPR3 sequences except spotted gar (*Lepisosteus oculatus* Winchell, 1864) QRFPR3. Another important difference concerns the respective positions of European seabass [*Dicentrarchus labrax* (Linnaeus, 1758)] and bicolour damselfish [*Stegastes partitus* (Poey, 1868)] QRFPR3s: in the ORF phylogenetic tree, *D. labrax qrfpr* is grouped together with Nile tilapia [*Oreochromis niloticus* (Linnaeus, 1758)], *Neolamprologus brichardi* (Poll, 1974), *Haplochromis burtoni* (Gunther, 1894), *Pundamilia nyererei* (Witte-Maas & Witte, 1985) and zebra mbuna [*Maylandia zebra* (Boulenger, 1899)] *qrfpr*s, while in the amino acid phylogenetic tree this relationship is verified for *S. partitus* QRFPR. Furthermore, there is a great difference between the two phylogenetic trees in branch lengths, as indicated by the respective scale bars.

Positively selected sites and conserved/altered biochemical property identifications

We used existing computational techniques (FUBAR, SLAC, FEL and REL) implemented in the HyPhy package (Kosakovsky Pond et al. 2005),

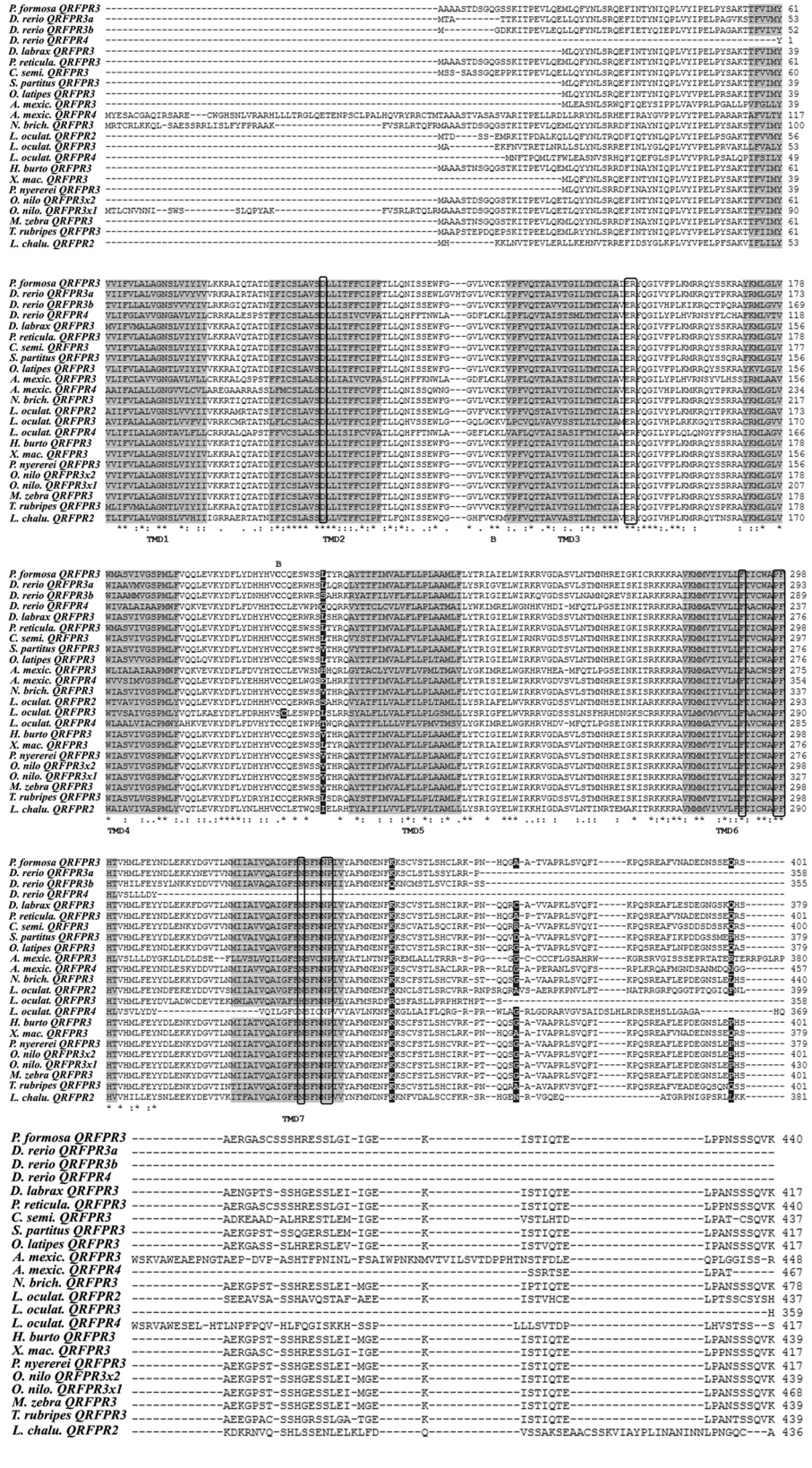

Figure 1. Multiple alignment of QRFPR amino acid sequences from different fish species. The grey background alignment regions refer to the TMDs (Transmembrane Domains). Boxed letters refer to the amino acids that play a crucial role in G protein coupling and receptor activation. The disulphide bridge between the two C (Cysteine) residues located in the first and second extracellular loops is indicated by "B". The letters marked by a black background refer to the positively selected amino acid sites. The symbols at the bottom of the QRFPR sequences correspond to the definitions of the T-coffee program: (*) fully conserved; (:) highly conserved; (.) conserved substitution.

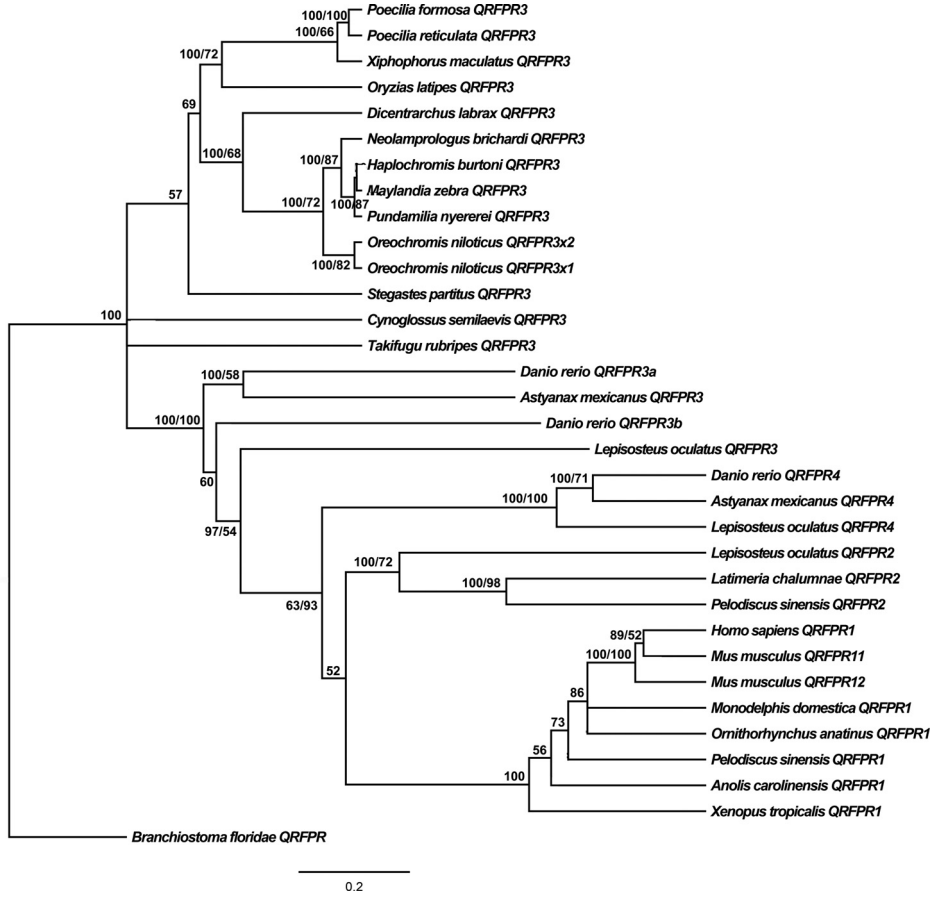

Figure 2. Phylogenetic relationships among different organisms QRFPRs reconstructed on the basis of cDNA coding region sequences and using both methods: BI (Bayesian Inference) (arithmetic mean = −22,234.94; harmonic mean = −22,263.61) and ML (Maximum Likelihood) (arithmetic mean = −22,133.8). Bayesian posterior probability (first number value) and bootstrap values higher than 50% are indicated on each node, respectively. The scale for branch length (0.2 substitution/site) is shown below the tree.

which are designed to identify the sites subject to pervasive selection (a large proportion of positively selected sites). These techniques may fail to recognise the sites where selection is episodic (Messier & Stewart 1997). For this reason, we used the MEME program, which is able to identify instances of both episodic and pervasive positive selections at the individual site level (Murrell et al. 2012). In order to investigate the presence of positive selection, we applied all the previously mentioned bioinformatics methods to the fish *qrfpr* codon sequence alignment. In Table III, the identified positively selected codon sites are presented. FUBAR and MEME results were statistically significant ($p < 0.05$). All the other program results (REL, SLAC and FEL) were not statistically significant. Kosakovsky Pond and Frost (2005) indicated that selection analyses of alignments with recombinants in them using a single tree could generate misleading results, if all *qrfpr* codon sequences had not been screened for recombination. Thus, we used the GARD and SBP programs (Kosakovsky Pond et al. 2006) to identify possible breakpoints in the *qrfpr* consensus codon sequence. The GARD and SBP applications did not find evidence of breakpoints in the *qrfpr* consensus codon sequence. In order to verify which biochemical properties are driving substitutions at different sites in the *qrfpr* codon sequence alignment and which properties are being selected for the advantageous changes in our previously identified positively selected sites, we performed other bioinformatics analyses using the PRIME program. In these analyses, both predefined sets of five amino-acid properties were used, being the five empirically measured properties used by Conant et al. (2007) and the five composite properties proposed by Atchley et al. (2005). PRIME builds on the same conceptual frameworks as FEL (Kosakovsky Pond & Frost 2005) and MEME (Murrell et al. 2012), but allows the nonsynonymous substitution rate (dN) to be depend not only on the site in object (like FEL and MEME), but also on which residues are being

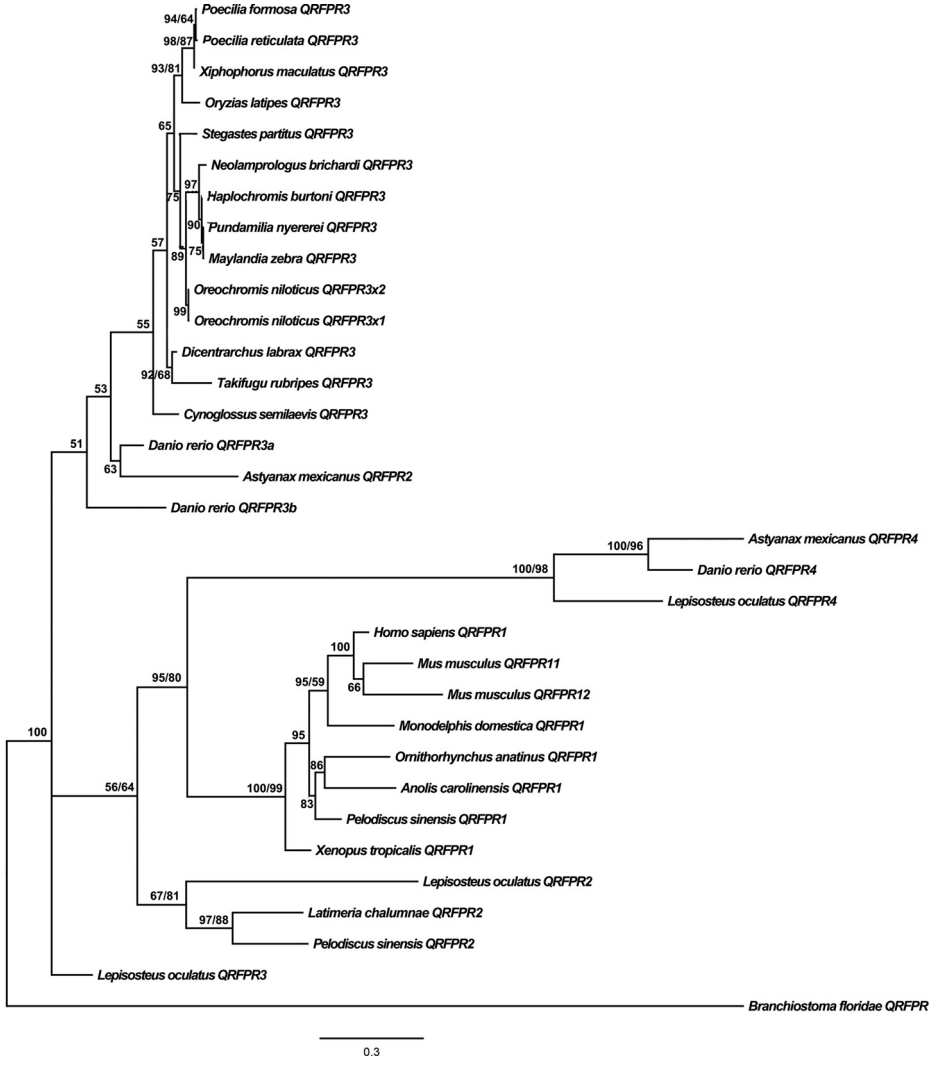

Figure 3. Phylogenetic relationships among different organisms' QRFPRs reconstructed on the basis of amino acid sequences and using both methods: BI (Bayesian Inference) (arithmetic mean = −10,508.75; harmonic mean = −10,545.58) and ML (Maximum Likelihood) (arithmetic mean = −10,438.4). Bayesian posterior probability (first number value) and bootstrap values higher than 50% are indicated on each node, respectively. The scale for branch length (0.3 substitution/site) is shown below the tree.

Table III. Positively selected codon sites identified by using Mixed Effects Model of Evolution (MEME) and Fast Unconstrained Bayesian AppRoximation (FUBAR), and their calculated statistics. dN – number of nonsynonymous substitutions for site and dS – number of synonymous substitutions for site, their difference must be a positive value in order to have positive selection at a given site; Posterior Pr. – posterior probability is a measure of the probability of having ω = dN/dS > 1 at a given site; p-value – the number of false positive tests.

Codon	MEME ω	MEME p-value	FUBAR dN–dS	FUBAR Post. Pr.
276	-	-	0.149	0.710
408	> 100	0.016	-	-
444	-	-	0.117	0.805
485	> 100	0.044	0.107	0.774

exchanged (e.g. I-V would be different from K-R). In Tables IV and V, conserved and altered biochemical properties, which are driving amino acid substitutions in fish QRFPR molecular evolution, are shown, based on Conant–Stadler and Atchley properties, respectively. Only one of the amino acids, corresponding to the positively selected codon sites previously identified by our analyses, seems to be significantly altered in at least one of the five considered biochemical properties. It is represented by the positively selected codon 485, whose refractivity or heat capacity was significantly altered at the corresponding amino acid site, according to the analyses based on Atchley properties. On the contrary, at the positively selected codon 276, the analysis based on

Table IV. Conserved and altered biochemical properties, which are driving amino acid substitutions in the fish QRFPR molecular evolution. These PRIME (property-informed models of evolution) analyses were based on Conant–Stadler properties and the results were statistically significant for p-values < 0.05. α is a measure of the importance of the alteration or conservation; for lower (negative) α values, the level of conservation of the corresponding property is higher at a given site and for higher (positive) values of α value, the alteration of the corresponding property is more radical at a given site. Altered properties results are shown on a dark grey background, while conserved properties results are shown on a light grey background.

Codon	Chemical composition		Polarity		Volume		Iso-electric point		Hydropathy	
	α	p-value	α	p-value	α	p-value	α	p-value	α	p-value
68	0.009	1	−2.378	0.293	15.083	0.036	−4.053	0.218	−0.474	1
86	1.435	1	0.38	0.917	1.606	1	−7.973	1	20	0.007
146	8.4	0.049	−3.924	1	8.869	0.432	2.61	1	−3.382	1
160	−6.03	0.026	9.828	0.45	3.551	0.536	17.257	0.533	0.079	1
173	1.312	1	−0.378	1	0.173	1	20	0.046	−0.68	1
217	4.434	0.29	−5.122	0.003	3.29	0.163	1.245	0.564	2.437	0.173
231	−0.247	1	17.455	0.029	2.052	1	−0.09	1	−5.004	0.188
246	1.494	1	−9.55	0.07	10.944	0.036	20	0.094	−1.315	0.76
265	13.595	1	−15.314	0.003	2.078	1	2.261	0.73	19.957	0.252
275	4.241	0.014	−8.717	0.011	1.28	0.011	4.605	0.061	5.84	0.012
303	0.181	1	−3.791	1	−1.101	1	0.244	1	20	0
306	−0.305	1	2.086	1	−0.219	1	16.7	0.013	0.38	1
318	−1.552	0.151	−0.297	0.82	2.493	0.045	20	0.02	−2.188	0.238
325	2.541	0.193	−1.068	0.452	−8.419	0.086	3.354	0.059	20	0.029
332	−6.477	0.025	0.573	1	2.248	0.639	20	1	3.485	1
375	3.144	0.181	−1.635	0.199	−4.035	0.161	2.24	0.008	6.553	0.059
376	−0.201	1	0.071	1	20	0.048	1.466	1	−1.256	1
435	−0.653	0.244	−6.01	0.362	2.196	0.071	−1.03	0.547	7.976	0.025
445	−1.037	0.447	7.384	0.074	2.335	0.351	−5.489	0.018	1.379	0.49
446	1.285	0.238	−1.944	0.366	4.659	0.265	12.636	0.001	−5.443	0.057
465	6.817	0.04	−0.402	0.231	−6.33	0.059	2.675	0.051	5.242	0.458
486	−1.247	1	5.388	0.048	−0.944	1	−0.086	0.936	−0.568	1
495	−1.218	1	0.903	1	−0.634	1	−0.703	1	4.329	0.023
498	0.968	1	2.119	1	0.587	1	3.128	0.02	−2.225	1
524	−0.528	1	−7.755	0.203	−0.545	1	3.148	0.231	15.737	0.01

Atchley properties indicated that the corresponding amino acid refractivity or its heat capacity was significantly conserved. Figures 4 and 5 show the property importance plots of Conant–Stadler and Atchley properties, respectively. As is shown by these graphics, although one of the properties of the amino acid site 276 was significantly conserved, the amino acid volume was altered (Figure 5), but it was not statistically supported.

Discussion

QRPFR molecular evolution

Although cDNA coding region and amino acid sequence alignments are characterised by a high level of similarity among the respective sequences, the amino acid sequence alignment turned out to be better than the coding sequence one, because its quality score value (98) was higher than the corresponding value (92) of the coding sequence alignment. Even considering only fish QRFPRs, the amino acid sequence alignment score value (98) was higher than the corresponding value (93) of the coding sequence alignment. The high score values obtained by sequence alignment analyses suggested a high level of homology among the various fish sequences, also supported by the high conservation level of the putative TMDs (Figure 1). A similar level of conservation was observed in the QRFPR amino acid alignment of *Homo sapiens, Rattus norvegicus, Gallus domesticus* (Linnaeus, 1758), *Taeniopygia guttata* (Vieillot, 1817), *Xenopus laevis* (Daudin, 1802) and *Danio rerio* (Hamilton, 1822), reported by Ukena et al. (2014). Fish proteins have the features characteristic of class-A GPCRs. *L. oculatous* QRFPR2 is an exception, because the second Cys, located in the second extracellular loop, was substituted by a Ser residue. However, both these amino acids are uncharged, despite the fact that Ser is a polar residue. Balaji et al. (2003) suggested that there is tolerance to the substitution of buried apolar residues by charged residues in the homologous protein structures, and in some cases it depends on the nature of the substitutions and the presence of an appropriate amino acid in the proximity of the substituting position. Thus, we

Table V. Conserved and altered biochemical properties which are driving amino acid substitutions in the fish QRFPR molecular evolution. These PRIME (property-informed models of evolution) analyses were based on Atchley properties and the results were statistically significant for p-values < 0.05. Altered properties results are shown on a dark grey background, while conserved properties results are shown on a light grey background.

Codon	Polarity		Secondary structure		Volume		Refractivity/heat capacity		Charge/iso-electric point	
	α	p-value	α	p-value	α	p-value	α	p-value	α	p-value
36	−0.888	0.666	−0.526	0.057	6.041	0.497	6.285	0.045	−0.456	0.635
79	13.305	0.056	−0.461	1	0.178	1	−4.25	0.043	1.811	0.881
83	13.243	0.006	4.668	0.206	−0.993	1	−1.79	1	0.859	0.7
88	1.018	1	8.315	0.029	0.239	1	−1.486	1	−0.333	1
95	2.561	0.048	1.521	0.138	−1.828	0.174	1.283	1	2.053	0.148
146	−0.778	0.467	3.584	0.007	−1.623	0.619	3.978	0.578	1.943	0.769
154	4.563	0.346	−2.577	0.349	−3.624	0.029	1.434	0.185	7.187	0.029
177	−3.076	0.026	5.14	0.011	−1.769	0.077	0.256	1	4.508	0.064
183	−0.497	1	1.623	0.728	−2.521	1	2.577	0.544	3.051	0.046
217	3.215	0.122	−5.621	0.147	−0.381	1	10.617	0.008	0.982	1
229	−1.566	0.505	9.917	0.017	−3.885	0.002	−1.111	0.547	4.58	0.016
232	18.826	0.003	−1.025	1	−0.389	1	−6.538	0.136	−0.396	1
241	9.905	0.001	−0.956	0.272	−1.102	0.385	−0.46	0.695	2.434	0.391
272	−0.711	0.855	1.562	0.422	0.407	0.44	3.209	0.021	−0.64	0.83
276	−0.647	0.289	−0.334	0.466	−3.229	0.142	5.019	0.035	3.555	0.375
281	−1.31	0.523	1.524	0.289	0.905	0.574	6.277	0.04	−1.637	0.45
287	−0.103	0.906	5.858	0.054	−3.603	0.028	−1.467	0.299	6.674	0.049
299	5.419	0.248	−0.437	0.689	2.056	0.24	7.32	0.218	−3.606	0.04
303	20	0	−0.052	1	−0.485	1	−1.698	0.311	−0.048	1
312	7.677	0.044	−6.147	0.143	1.37	0.51	−2.002	0.336	0.965	0.655
325	5.714	0.014	−0.095	0.646	−0.793	0.568	−3.365	0.175	3.846	0.264
332	1.729	0.753	−3.216	1	−0.3	1	15.442	0.007	1.186	1
333	5.035	0.048	−0.154	1	0.254	1	−4.676	0.046	0.57	1
367	−0.548	1	5.455	0.011	−0.205	1	−0.473	1	−0.169	1
371	4.338	0.617	−5.818	0.029	−1.554	1	5.669	0.083	1.66	1
372	2.641	0.393	11.944	0.031	3.347	0.064	−1.79	0.572	−4.894	0.06
435	−4.978	0.602	0.175	1	−1.657	0.089	13.537	0.024	2.563	0.162
449	10.015	0.014	−4.877	0.015	2.22	0.014	7.65	0.011	−0.464	0.01
450	5.898	0.022	−3.297	0.618	1.056	0.361	−3.094	0.523	1.106	0.408
453	−0.458	1	8.634	0.052	2.416	0.665	0.746	0.828	−4.946	0.038
475	0.093	1	0.024	1	2.468	0.032	6.584	0.466	−2.409	0.464
478	9.426	0.05	−4.535	0.104	0.075	1	0.028	1	1.787	1
481	−0.89	0.608	1.07	0.75	−1.958	0.059	4.331	0.655	2.391	0.048
484	−1.327	0.077	0.283	1	1.185	0.052	5.93	0.087	−1.748	0.044
485	1.899	0.235	−1.073	0.154	−0.43	1	−4.274	0.018	4.514	0.179
488	1.154	0.692	−0.486	1	−2.941	0.016	5.912	0.053	4.385	0.067
493	5.782	0.046	−1.059	1	−0.241	0.781	0.709	1	0.597	1
495	1.073	1	4.5	0.005	−0.587	1	0.286	1	0.007	1
499	−1.695	0.351	−4.556	0.117	0.369	1	9.694	0.125	3.029	0.011
508	19.514	1	0.56	1	0.119	1	−13.88	0.043	0.295	1

can hypothesise that in this apolar-to-polar substitution, an important role is played by the Cys residue located next to the Ser, which might have been constrained to take possession of the necessary characteristic in order to form the disulphide bridge with the other C residue located in the first extracellular loop.

The coelacanth *Latimeria* has been designated as a "living fossil" because the lineage disappeared from the fossil record in the Cretaceous period, about 80 million years ago (Pyron 2010). Together with the lungfishes, the coelacanth is considered the closest living relative of the tetrapods (Shan & Gras 2011). Thus, *Latimeria* harbours a key position in the evolution of vertebrates, including mammals. *L. oculatus* is a member of the Holostei infraclass, which diverged from teleosts (the Teleostei infraclass) before the teleost genome duplication (Amores et al. 2011). In the phylogenesis proposed by Braasch et al. (2014), based on paired-related homeobox (Prrx) gene evolution, and by Near et al. (2012), based on nine nuclear genes, *L. oculatus* emerges early in the Actinopterygii clade, ending up nearer to *L.*

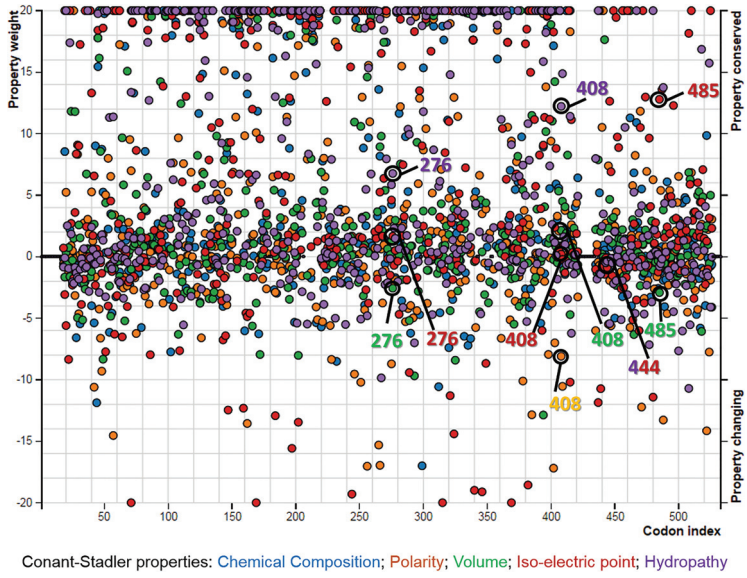

Figure 4. Importance plot of the altered and conserved properties at different codon sites, identified by PRIME (property-informed models of evolution) analyses based on Conant–Stadler properties. The colour of each of the circles refers to the colour of a specific biochemical property shown on the bottom of the graphic. For higher vertical distance from 0, the property alterations or conservations become more radical. The altered properties of the positively selected sites are indicated by empty circles. In the online document you can find the colored version of this figure.

Figure 5. Importance plot of the altered and conserved properties at different codon sites, identified by PRIME (property-informed models of evolution) analyses based on Atchley properties. The altered and conserved properties of the positively selected sites are indicated by empty circles. In the online document you can find the colored version of this figure.

chalumnea Smith, 1939 than other teleosts. This could indicate that Actinopterygii and Holostei may have retained some ancestral characteristics; one of them could be the QRFPR2 isoform.

The phylogenetic relationships among QRFPRs of other fish species are generally compatible with the phylogenies reported by Near et al. (2012) and Braasch et al. (2014). For instance, QRFPR sequences of Ostariophysi members, such as Mexican tetra [*Astyanax mexicanus* (De Filippi, 1853)] and *D. rerio*, are grouped together in QRFPR3 and QRFPR4 clades, like in the species

phylogeny reconstruction. The remaining teleost QRFPRs, belonging to Acanthomorphata, are positioned into the fish QRFPR phylogenetic tree in the same position, as in the species phylogenetic tree reported by Braasch et al. (2014). For instance, in amino acid base tree topology, the QRFPR of the Tetraodontiformes, Japanese pufferfish [*Takifugu rubripes* (Temminck & Schlegel, 1850)] was grouped together with the QRFPR of another member of Percomorpharia (*D. labrax*), similarly to the results reported by Braasch et al. (2014), where *T. rubripes* is grouped together with three-spined stickleback (*Gasterosteus aculeatus* Linnaeus, 1758; member of Percomorpharia).

Some discordances emerged also from the comparison between the topologies obtained from the analyses performed with coding regions and amino acids, especially in relation to the positions of *D. labrax, S. partitus* and *L. oculatus* QRFPR3s. We hypothesise that they might be the result of differences in substitution rates, supported by the great difference in the branch lengths of the two phylogenetic trees. The estimated mean amino acid substitution rate (0.8 substitution/site) was bigger than the average nucleotide substitution rate (0.57 substitution/site), estimated by MEGA 6 (Tamura et al. 2013). However, the estimated mean amino acid substitution rate of the fish data set (0.49 substitution/site) is comparable to the mean base substitution rate (0.56 substitution/site), but generally the estimated evolutionary divergence among the nucleotide sequences is much higher than the evolutionary divergence among amino acid sequences. Probably this could be explained only by a situation where the nonsynonymous substitution rates (dN) were higher than synonymous substitution rates (dS). A statistically significant excess of nonsynonymous substitution (dN > dS) could be interpreted as positive selection (Kosakovsky Pond & Frost 2005).

Another point to highlight is the phenomenon of gene duplication that seems to be related to the evolutionary history of these proteins. This is evident in some species having multiple isoforms (*D. rerio, L. oculatus* and *A. mexicanus*), but it does not seem a general phenomenon, although we can not exclude that other species may have isoforms not yet identified. However, the gene duplication of *qrfpr*s does not seem to be a phenomenon that occurred at an ancestral level, before the differentiation of fish taxa, but rather occurred several times during their evolution. The presence of four *qrfpr* clades (Larhammar et al. 2014) suggests an origin by duplications resulting from the two basal tetraploidisations, 1R and 2R, thus forming a paralogon (a set of paralogous chromosomal regions containing members of the same gene families as a result of the duplication of a large block or an entire chromosome). However, sequence-based analyses may be skewed due to uneven selection pressures, due to the many possible losses of *qrfpr* genes in the various species or lineages. Therefore, some additional information should be considered.

As result of our phylogenetic analyses, there is probably the need to rename some fish QRFPR isoforms. In particular, according to the isoform nomenclature used by Larhammar and colleagues (2014), *D. rerio, L. oculatus* and *A. mexicanus* QRFPRs, emerging together and clearly separated from the other isoforms, can be indicated as belonging to the same isoform, 4. Since the two sequences of *O. niloticus* are the product of a single-gene alternative splicing, they should be considered two variants of isoform 3, and we propose to indicate them as QRFPR3x1 and QRFPR3x2.

Positive selection in fish QRFPR molecular evolution

Evolutionary biologists have typically invoked two types of selective forces that shape the evolution of species. One is purifying selection, which favours the conservation of existing phenotypes. The other is positive selection (also known as Darwinian selection), which promotes the emergence of new phenotypes. Positive selection can leave a set of telltale signatures in the genes under its influence, such as the rapid divergence of functional sites between species and the depression of polymorphism within species (Bamshad & Wooding 2003). The imprint of natural selection (positive selection) on protein coding genes is often difficult to identify, because selection is frequently transient or episodic, i.e. it affects only a subset of lineages. Our results support the Murrell et al. (2012) hypothesis, that natural selection is predominantly episodic, with transient periods of adaptive evolution masked by the prevalence of purifying or neutral selection on other branches. Previous work on calreticulin molecular evolution indicated that a significant number of codon sites involved in positive and purifying selection were functionally or structurally important, as demonstrated by wet-lab analyses (Bakiu 2014). However, our results must be confirmed experimentally, using a targeted experimental approach. For example, it could be extremely interesting to perform site-specific mutagenesis experiments, in order to determine the functional and/or structural importance of the positively selected sites.

Conserved/altered amino acid biochemical properties in fish QRFPR molecular evolution

Recent studies have shown that amino acid exchangeability varies across organisms (Dunn et al. 2013) and across genes (Conant et al. 2007), depending on physico-chemical properties, so that the same substitution may sometimes be radical (having a large effect on protein structure and/or function) or conservative (having little effect on structure or function). Tourasse and Li (2000) found that observed substitution patterns resulted from the unique functional characteristics of individual protein families. Variation can be expected from site to site within a protein. For instance, amino acids with different hydrophobicity may be unchangeable at sites where the protein fold is sensitive to hydrophobicity, but exchangeable at sites where it is insensitive to hydrophobicity.

PRIME analyses suggest that there are more conserved than altered amino acid biochemical properties in fish QRFPR molecular evolution. However, at least one statistically significant alteration was observed at one of the positively selected amino acid sites. Atchley property results (Figure 5) indicate that the polarity index property was not altered in any of the positively selected sites. A strong evolutionary basis exists for a complex pattern of covariation involving all the polarity index-related attributes. Atchley et al. (2005) described in considerable detail the patterns of variability in buried hydrophobic versus accessible hydrophilic amino acids in the dimerisation domain of basic helix–loop–helix (bHLH) proteins. These observed patterns were related to natural selection, evolutionary change and phylogenetic divergence. Refractivity and heat capacity, which reflect codon and amino acid diversity, are properties that exhibit significant correlation with positive selection and evolutionary change (Atchley et al. 2005). In our analyses, the positively selected site 485 exhibited a statistically significant ($p < 0.05$) altered refractivity/heat capacity.

Atchley et al. (2005) demonstrated that the remaining three properties do not indicate significant association between physiochemical attribute variation and evolutionary patterns of amino acid substitution, and they cannot be ascribed to evolutionary divergence, but rather to nonevolutionary changes in structure and function. Although there is a significant altered refractivity or heat capacity, as mentioned, due to a polarity alteration at amino acid site 485, like the secondary structure and volume properties alteration, it is not statistically significant. In conclusion, our results suggest that in the molecular evolution of fish QRFPRs, positive selection did not significantly alter the amino acid biochemical properties of any positively selected sites.

References

Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH. 2011. Genome evolution and meiotic maps by massively parallel DNA sequencing: Spotted gar, an outgroup for the teleost genome duplication. Genetics 188:799–808.

Atchley WR, Zhao J, Fernandes AD, Drüke T. 2005. Solving the protein sequence metric problem. Proceedings of the National Academy of Science USA 102:6395–6400. doi:10.1073/pnas.0408677102.

Bakiu R. 2014. Calreticulin molecular evolution: A strong purifying and episodic diversifying selection result. Biologia 69:270–280. doi:10.2478/s11756-013-0327-7.

Balaji S, Aruna S, Srinivasan N. 2003. Tolerance to the substitution of buried apolar residues by charged residues in the homologous protein structures. Proteins 53:783–791. doi:10.1002/prot.v53:4.

Bamshad M, Wooding SP. 2003. Signatures of natural selection in the human genome. Nature Reviews Genetics 4:99–111. doi:10.1038/nrg999.

Braasch I, Guiguen Y, Loker R, Letawa JH, Ferrara A, Bobe J, Postlethwait JH. 2014. Connectivity of vertebrate genomes: Paired-related homeobox (Prrx) genes in spotted gar, basal teleosts, and tetrapods. Comparative Biochemistry and Physiology C 163:24–36.

Chartrel N, Alonzeau J, Alexandre D, Jeandel L, Alvear-Perez R, Leprince J, Boutin J, Vaudry H, Anouar Y, Llorens-Cortes C. 2011. The RFamide neuropeptide 26RFa and its role in the control of neuroendocrine functions. Frontiers in Neuroendocrinology 32:387–397. doi:10.1016/j.yfrne.2011.04.001.

Conant GC, Wagner GP, Stadler PF. 2007. Modeling amino acid substitution patterns in orthologous and paralogous genes. Molecular Phylogenetics and Evolution 42:298–307. doi:10.1016/j.ympev.2006.07.006.

Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: Fast selection of best-fit models of protein evolution. Bioinformatics 27:1164–1165. doi:10.1093/bioinformatics/btr088.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: More models, new heuristics and parallel computing. Nature Methods 9:772.

Dunn KA, Jiang W, Field C, Bielawski JP. 2013. Improving evolutionary models for mitochondrial protein data with site-class specific amino acid exchangeability matrices. PLoS One 8:1e55816. doi:10.1371/journal.pone.0055816.

Fukusumi S, Yoshida H, Fujii R, Murayama M, Komatsu H, Habata Y, Shintani Y, Hinuma S, Fujino M. 2003. A new peptidic ligand and its receptor regulating adrenal function in rats. Journal of Biological Chemistry 278:46387–46395. doi:10.1074/jbc.M305270200.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Systematic Biology 59:307–321. doi:10.1093/sysbio/syq010.

Huelsenbeck JP, Larget B, Swofford D. 2000. A compound Poisson process for relaxing the molecular clock. Genetics 154:1879–1892.

Jiang Y, Luo L, Gustafson EL, Yadav D, Laverty M, Murgolo N, Vassileva G, Zeng M, Laz TM, Behan J, Qiu P, Wang L,

Wang S, Bayne M, Greene J, Monsma F, Zhang FL. 2003. Identification and characterization of a novel RF-amide peptide ligand for orphan G-protein-coupled receptor SP9155. Journal of Biological Chemistry 278:27652–27657. doi:10.1074/jbc.M302945200.

Kosakovsky Pond SL, Frost SDV. 2005. Not so different after all: A comparison of methods for detecting amino acid sites under selection. Molecular Biology and Evolution 22:1208–1222. doi:10.1093/molbev/msi105.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Molecular Biology and Evolution 23:1891–1901. doi:10.1080/10635150290102456.

Larhammar D, Xu B, Bergqvist CA. 2014. Unexpected multiplicity of QRFP receptors in early vertebrate evolution. Frontiers in Neuroscience 8:337. doi:10.3389/fnins.2014.00337. Oct 24; eCollection.

Lepage T, Bryant D, Philippe H, Lartillot N. 2007. A general comparison of relaxed molecular clock models. Molecular Biology and Evolution 24:2669–2680. doi:10.1093/molbev/msm193.

Leprince J, Neveu C, Lefranc B, Guilhaudis L, Milazzo-Segalas I, Do Rego JC, Tena-Sempere M, Tsutsui K, Vaudry H. 2013. 26RFa. In: Kastin AJ, editor. Handbook of biologically active peptides. 2nd ed. Amsterdam, The Netherlands: Elsevier. pp. 917–923.

Messier W, Stewart CB. 1997. Episodic adaptive evolution of primate lysozymes. Nature 385:151–154. doi:10.1038/385151a0.

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K. 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genetics 8:e1002764. doi:10.1371/journal.pgen.1002764.

Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Smith WL. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proceedings of the National Academy of Science US 109:13698–13703. doi:10.1073/pnas.1206625109.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of Molecular Biology 302:205–217. doi:10.1006/jmbi.2000.4042.

Pond SL, Frost SD, Muse SV. 2005. HyPhy: Hypothesis testing using phylogenies. Bioinformatics 21:676–679. doi:10.1093/bioinformatics/bti079.

Price DA, Greenberg MJ. 1977. Structure of a molluscan cardioexcitatory neuropeptide. Science 197:670–671. doi:10.1126/science.877582.

Pyron RA. 2010. A likelihood method for assessing molecular divergence time estimates and the placement of fossil calibrations. Systematic Biology 59:185–194. doi:10.1093/sysbio/syp090.

Raffa RB. 1988. The action of FMRFamide (Phe-Met-Arg-Phe-NH2) and related peptides on mammals. Peptides 9:915–922. doi:10.1016/0196-9781(88)90141-6.

Ramanjaneya M, Karteris E, Chen J, Rucinski M, Ziolkowska A, Ahmed N, Kagerer S, Jöhren O, Lehnert H, Malendowicz LK. 2013. QRFP induces aldosterone production via PKC and T-type calcium channel-mediated pathways in human adrenocortical cells: Evidence for a novel role of GPR103. American Journal of Physiology-Endocrinology and Metabolism 305:E1049–E1058. doi:10.1152/ajpendo.00191.2013.

Rastogi RK, D'Aniello B, Pinelli C, Fiorentino M, Di Fiore MM, Di Meglio M, Iela L. 2001. FMRFamide in the amphibian brain: A comprehensive survey. Microscopy Research and Technique 54:158–172. doi:10.1002/jemt.1130.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, L L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61:539–542. doi:10.1093/sysbio/sys029.

Shan Y, Gras R. 2011. 43 genes support the lungfish-coelacanth grouping related to the closest living relative of tetrapods with the Bayesian method under the coalescence model. BMC Research Notes 4:49.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, Tamura K, Stecher G, Peterson D, Filipski A. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30:2725–2729. doi:10.1093/molbev/mst197.

Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. Systematic Biology 51:689–702. doi:10.1080/10635150290102456.

Tourasse NJ, Li WH. 2000. Selective constraints, amino acid composition, and the rate of protein evolution. Molecular Biology and Evolution 17:656–664.

Tsutsui K, Ukena K. 2006. Hypothalamic LPXRF-amide peptides in vertebrates: Identification, localization and hypophysiotropic activity. Peptides 27:1121–1129.

Ukena K, Osugi T, Leprince J, Vaudry H, Tsutsui K. 2014. Molecular evolution of GPCRs: 26Rfa/GPR103. Journal of Molecular Endocrinology 52:T119–131. doi:10.1530/JME-13-0207.

Ukena K, Tsutsui K. 2005. A new member of the hypothalamic RF-amide peptide family, LPXRF-amide peptides: Structure, localization, and function. Mass Spectrometry Reviews 24:469–486.

Walker RJ, Papaioannou S, Holden-Dye L. 2009. A review of FMRFamide- and RFamide-like peptides in metazoa. Invertebrate Neuroscience 9:111–153. doi:10.1007/s10158-010-0097-7.

Xie W, Lewis PO, Fan Y, Kuo L, Chen MH. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Systematic Biology 60:150–160. doi:10.1093/sysbio/syq085.