Gene 591 (2016) 245–254

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Evolutionary history of the reprimo tumor suppressor gene family in vertebrates with a description of a new reprimo gene lineage

Ignacio A. Wichmann [a,e,1], Kattina Zavala [b,1], Federico G. Hoffmann [c,d], Michael W. Vandewege [c], Alejandro H. Corvalán [a,e,g], Julio D. Amigo [f], Gareth I. Owen [a,f,g], Juan C. Opazo [b,*]

[a] Advanced Center for Chronic Diseases (ACCDiS), Pontificia Universidad Católica de Chile, Santiago, Chile
[b] Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile
[c] Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, MS, USA
[d] Institute for Genomics, Biocomputing, and Biotechnology, Mississippi State University, MS, USA
[e] Departamento de Oncología y Hematología, Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile
[f] Departamento de Fisiología, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Santiago, Chile
[g] Center UC for Investigation in Oncology (CITO), Pontificia Universidad Católica de Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 28 March 2016
Received in revised form 12 July 2016
Accepted 14 July 2016
Available online 16 July 2016

Keywords:
Whole genome duplication
Gene family evolution
Cancer
Gastric cancer
Evolutionary medicine

ABSTRACT

Genes related to human diseases should be natural targets for evolutionary studies, since they could provide clues regarding the genetic bases of pathologies and potential treatments. Here we studied the evolution of the reprimo gene family, a group of tumor-suppressor genes that are implicated in p53-mediated cell cycle arrest. These genes, especially the reprimo duplicate located on human chromosome 2, have been associated with epigenetic modifications correlated with transcriptional silencing and cancer progression. We demonstrate the presence of a third reprimo lineage that, together with the reprimo and reprimo-like genes, appears to have been differentially retained during the evolutionary history of vertebrates. We present evidence that these reprimo lineages originated early in vertebrate evolution and expanded as a result of the two rounds of whole genome duplications that occurred in the last common ancestor of vertebrates. The reprimo gene has been lost in birds, and the third reprimo gene lineage has been retained in only a few distantly related species, such as coelacanth and gar. Expression analyses revealed that the reprimo paralogs are mainly expressed in the nervous system. Different vertebrate lineages have retained different reprimo paralogs, and even in species that have retained multiple copies, only one of them is heavily expressed.

1. Introduction

Genes related to human diseases (e.g. cancer) are natural targets for evolutionary studies, as this research could reveal the genetic bases of pathologies and suggest potential treatments. Comparative studies involving genes related to human pathology have gained a medical relevance with the discovery that several non-model organisms can help us better understand certain pathologies. An example is the discovery that some subterranean rodents (e.g. *Heterocephalus glaber*, *Fukomys anselli*, *Spalax* sp.) are resistant to the aging processes and age-related diseases, such as cancer (Yu et al., 2011; Edrey et al., 2012; Gorbunova et al., 2012; Manov et al., 2013; Novikov and Burda, 2013; Henning et al., 2014; Faulkes et al., 2015), or the finding that some wild animals (e.g. *Octodon degus*) develop pathologies (such as Alzheimer and diabetes) in a similar way as humans (Castro-Fuentes and Socas-Pérez, 2013; Tarragon et al., 2013; Braidy et al., 2015; Inestrosa et al., 2015), making them promising biomedical models.

Reprimo (*RPRM*) is a poorly studied (Ohki et al., 2000) single exon gene that appears to be involved in tumor suppression based on its induction after X-ray irradiation, as well as, its identification as a downstream target of p53. The *RPRM* protein is believed to mediate, at least in part, the p53-dependent cell cycle arrest at G2 by the inhibition of Cdc2 activity and nuclear translocation of cyclin B1 (Ohki et al., 2000; Taylor and Stark, 2001). It has also been suggested that *RPRM* may also be regulated by p73 in an independent manner from p53 in gastric carcinogenesis (Saavedra et al., 2015). The loss of *RPRM* expression in cancer through abnormal methylation patterns suggests that this gene and its corresponding biochemical pathway might play a role in cancer metabolism and development, and suggests a potential target for treatment against malignancies (Ohki et al., 2000; Wong et al., 2005; Hamilton et al., 2006; Bernal et al., 2008; Luo et al., 2011; Saavedra et al., 2015). To date, efforts to find biomarkers for early stages of cancer have found limited success, and the gene *RPRM* has emerged as a promising candidate in this area (Bernal et al., 2008).

Abbreviations: RPRM, reprimo; RPRM-like, reprimo like; Mya, million of years ago.

* Corresponding author.
E-mail address: jopazo@gmail.com (J.C. Opazo).

[1] *Equal contribution

http://dx.doi.org/10.1016/j.gene.2016.07.036
0378-1119/

The inverse relationship between methylation and transcription has identified *RPRM* as a target for aberrant methylation in various cancers, including prostate cancer (Ellinger et al., 2008), lung cancer (Nakajima et al., 2009), hepatocellular carcinoma (Nishida et al., 2008), pancreatic cancer (Sato et al., 2008), head and neck carcinomas (Wong et al., 2005), esophageal cancer (Hamilton et al., 2006) and gastric adenocarcinoma (Bernal et al., 2008; Luo et al., 2011; Saavedra et al., 2015). The abnormal regulation of the genome during cancer is thought to be due to atypical methylation processes rather than the presence of somatic mutations. In recent years, this hypothesis has driven attempts to explore correlations between methylation profiles and early cancer detection. Thus, assessment of the methylation status of *RPRM* cell-free circulating DNA has emerged as a promising biomarker for the early diagnosis of stomach cancer.

The *RPRM* gene family also includes a second paralog, named reprimo-like (*RPRML*), which in humans retains 52.5% identity with *RPRM* at the amino acid level. In humans, the C-terminal portion of these genes demonstrates a strikingly high degree of conservation, whereas in the N-terminal the situation is the opposite (Fig. 1). There is little information on the emergence of the RPRM gene family and the functional role of its members. The evidence so far suggests that RPRML plays a role in cell cycle dynamics that is distinct from that of RPRM, and that *RPRML* expression is inducible by gamma-irradiation (Lin, 2010).

The goal of this study was to provide an initial assessment of the evolutionary history of the reprimo gene family, focused on assessing the diversity of reprimo genes in the major groups of vertebrates; and examine gene expression patterns among different tissues in representative species of vertebrates. Our results uncover the presence of at least three reprimo lineages that were differentially retained during the evolutionary history of vertebrates, and also suggest that they originated and diversified as a product of the two whole genome duplications in the last common ancestor of vertebrates. Expression analyses revealed that reprimo genes are mostly expressed in the nervous system and, when multiple gene copies are present in the genome, only one of them is predominantly expressed.

2. Material and methods

2.1. Data collection and phylogenetic analyses

We used bioinformatic tools to annotate reprimo genes in representative species of the major groups of vertebrates. Our sampling included representatives from placental mammals, marsupials, monotremes, birds, reptiles, amphibians, lobe-finned fish, teleost fish, non-teleost ray-finned fish, cartilaginous fish, and cyclostomes (Supplementary Table S1). To investigate the origin of the reprimo gene family we also searched reprimo sequences in non-vertebrate deuterostomes using tblastn strategy in the non-redundant and whole-genome shotgun contig databases (Tatusova and Madden, 1999). Reprimo amino acid sequences were aligned using the L-INS-i strategy from MAFFT v.7 (Katoh and Standley, 2013). Phylogenetic relationships among reprimo genes were estimated using maximum likelihood and Bayesian approaches. The program MEGA v6.06 (Tamura et al., 2013) was used to perform a maximum likelihood analysis to obtain the best tree. We used the propose model tool of MEGA v6.06 (Tamura et al., 2013) to select the best-fitting model of amino acid substitution (JTT + G). Bayesian searches were conducted in MrBayes v.3.1.2 (Ronquist and Huelsenbeck, 2003), setting two independent runs of six simultaneous chains for 30×10^6 generations, sampling every 2500 generations, and using default priors. The run was considered to have reached convergence once the likelihood scores reached an asymptotic value and the average standard deviation of split frequencies remained <0.01. We discarded all trees that were sampled before convergence, and we evaluated support for the nodes and parameter estimates from a majority rule consensus of the last 6000 trees.

```
Human RPRM          MNPALGNQTDV---------AGLFLANSS-----EAL-ERAVRCCTQAS--VVTDD--GFAEGGP
Coelacanth RPRM     MNSTIINQTD----------TGLLFSNRS-----DSL-ETVFRCCNQTS--VVTDD--GFVVTAP
Zebrafish RPRMb     MNSTGFNDTD----------SA-LFSNTS----------FLSCCNVSS--VVTDS--GFVSAAL
Zebrafish RPRMa     MNST-FNQTD----------SG-IFSNRT---------EENLLCCNFSS--VVTDN--GFAAAAP

Human RPRML         MNATFLNHSGLEEVDGVGGGAGAALGNRT-----HGL-GTWLGCCPGGAP-LAASD-GVPAGLAP
Coelacanth RPRML    MNGTFFNQTMLF--------QG-IYSNKT-----QSV-GTLVGCCNGSSP-VVTSD-GESLAHVP
Zebrafish RPRML     MNGTFFNHTVFT--------QG-VLLNRS-----QELAGTLVDCCTGNGSEVTANDGGGSLVLAQ

Coelacanth RPRM3    MNGT-------------------VLGRQA---------DSLLGCCNFSSA-VVWDE--RLGVSGA
Zebrafish RPRM3     MNVS----------------HG--LLNRTISTVKHPL-DAGRGCCNFS---VITSD--GFGSPIQ


Human RPRM          DERSLYIMRVVQIAVMCVLSLTVVFGIFFLGCNLLIKSEGMINFLVKDRRPSKEVEAVVVGPY
Coelacanth RPRM     DERSLYIMRVVQIAVMCVLSLTVVFGIFFLGCNLLIKSEGMINFLVKDRRPSKEVEAVIVGPY
Zebrafish RPRMb     DERSVFIMRTVQIAVMCVLALTVVFGIFFLGCNLLIKSEGMINFLVTDRRPSKDVEAVIVGSY
Zebrafish RPRMa     DERSLFIMRIVQIAVMCVLSLTVVFGIFFLGCNLLIKSEGMINFLVTDRRPSKEVEAVIVGAY

Human RPRML         DERSLWVSRVAQIAVLCVLSLTVVFGVFFLGCNLLIKSESMINFLVQERRPSKDVGAAILGLY
Coelacanth RPRML    DERNLFIMRVVQIAVLCVLSLTVVFGIFFLGCNLLIKSESMINFLVKDRRPSKDVEAVIVGLY
Zebrafish RPRML     DERKLFVTRVVQIAVLCVLSLTVMFGIFFLGCNLMIKSESMINFLVKDRRSSKDVEAVMIGLS

Coelacanth RPRM3    DERSLYIMRVVQIAVLCVLSITVIFGIFFLGCNLLLKSESMINLLVKERRPSNEVEAIIIGSY
Zebrafish RPRM3     DEREQFITRLVQIAVLCVLSLTVIFGIFFLGCNLLIKSESMINLLVEDRRPSKDAEIIMIAA-
```

Fig. 1. Alignment of reprimo amino acid sequences from human, coelacanth and zebrafish. Amino acids that delimits the identity the *RPRM* gene are in blue, amino acids that delimits the identity the *RPRML* gene are in red, and amino acids delimits the identity the third reprimo lineage are in green.

2.2. Assessments of conserved synteny

To examine patterns of conserved synteny, we examined genes found upstream and downstream of the reprimo genes on species representative of all main groups of vertebrates. In the case of the *RPRM* and *RPRML* genes, that were present in a wide variety of vertebrates, synteny analyses was performed in humans, opossum, platypus, Chinese softshell turtle, chicken, zebrafinch, anole lizard, clawed frog, coelacanth, spotted gar and elephant shark. For the third reprimo lineage that was only present in the coelacanth, spotted gar and elephant shark, we assessed synteny in these three species. In the case of lampreys, both species included in this study were investigated. Initial orthologue predictions were derived from the Ensembl Compara database (Vilella et al., 2009) and were visualized using the program Genomicus (Muffato et al., 2010). In the case of the Japanese lamprey (http://jlampreygenome.imcb.a-star.edu.sg/) and the elephant shark (http://esharkgenome.imcb.a-star.edu.sg/) the genomic pieces containing reprimo genes were annotated using the program Augustus (Stanke et al., 2008). The predicted genes were then compared with the non-redundant protein database using Basic Local Alignment Search Tool (BLAST) (Tatusova and Madden, 1999).

2.3. Gene expression assessment

Reprimo expression was measured in a representative sample of vertebrate species. Our sampling included the elephant shark, spotted gar, chicken, and human. For each species, reprimo expression was assessed in six different tissues (brain, heart, kidney, liver, muscle and testis). RNASeq data was gathered from GenBank's SRA database (Supplementary Table S2). Reference cDNA sequences of predicted genes from each genome were collected from Ensembl, for each gene only the longest cDNA sequence was included. Elephant shark cDNA sequences were collected from its own web site (http://esharkgenome.imcb.a-star.edu.sg/). Gene expression levels were estimated using RSEM v1.2.3 (Li and Dewey, 2011) which uses Bowtie v. 0.12.9 (Langmead et al., 2009) to map reads to the proper set of coding sequences. Default settings were used, and expression was measured in transcripts per million (TPM).

3. Results and discussion

By combining phylogenetic, synteny and gene expression analyses, we reconstructed the evolutionary history of the reprimo gene family in representative species of vertebrates. Our results demonstrate the presence of three separate lineages in the reprimo gene family that emerged early in vertebrate evolution, and have been differentially retained among the different vertebrate lineages. As a first step to investigate the origin of this gene family, we searched for reprimo genes in non-vertebrate deuterostomes available in the non-redundant and whole-genome shotgun contig databases. According to our analysis, there are no reprimo sequences outside vertebrates, suggesting that this gene family is an evolutionary innovation of this group.

3.1. Gene phylogenies and synteny resolve orthology for RPRM and RPRML

Our phylogenetic analyses revealed the presence of two well-supported reprimo lineages among gnathostomes (Fig. 2). The *RPRM* lineage (Fig. 2, blue clade), represented by the gene located on human chromosome 2, and the RPRML lineage (Fig. 2, red clade), represented by gene located on human chromosome 17, were recovered as a well-supported clades providing a clear definition of orthology. The RPRM gene lineage has been retained as a single copy in all major groups of gnathostomes other than birds, where it has been lost, and pig (*Sus scrofa*), where two tandem copies are found on chromosome 15. The *RPRML* gene lineage was also present as a single copy gene in all surveyed species with the exception of monotremes, the group that includes platypuses and echidnas, where the gene was not found. In this case, we cannot rule out that the *RPRML* gene is present in their genomes but not in the current genome assembly.

Synteny analysis provides additional support for the identity of the *RPRM* and *RPRML* gene lineages (Figs. 3 and 4). Genes found up- and downstream of the *RPRM* gene are well conserved in all main groups of gnathostome vertebrates (Fig. 3, upper panel). In most surveyed species there are four upstream genes (*ARL6IP6*, *PRPF40A*, *FMNL2 and STAM2*) and four downstream genes (*GALNT13*, *KCNJ3*, *NR4A2*, *GPD2*) that define the identity of this genomic region (Fig. 3, upper panel). Synteny for the *RPRML* gene is not as conserved as for *RPRM* (Fig. 4), because genes defining this genomic region appear to have been rearranged on the same chromosome. Even so, we were able to identify 12 genes commonly found in the proximity of the *RPRML* gene (e.g. *GOSR2*, *WNT9B*, *WNT3*, *NSF*, *ITGB3*, *FMNL1*) that defines this genomic region (Fig. 4). We also found that in humans and Old World monkeys, the *RPRML* gene is located between exon 7 and 8 of the Golgi SNAP receptor complex member 2 gene (*GOSR2*, Fig. 5). The transcription of *RPRML* appears to be independent to that of the *GOSR2* gene, as in contrast to the *RPRML* expression pattern, the host gene is expressed in five (brain, heart, kidney, liver and testis) out of the six tissues included in this study.

3.2. Presence of a third reprimo gene lineage in gar, elephant shark, coelacanth

Our study uncovered the presence of a third reprimo gene lineage in gnathostomes, which was identified in three distantly related species: elephant shark, spotted gar and coelacanth (Fig. 2, green clade). This gene lineage is recovered as a monophyletic group in our phylogenies, but with no support (Fig. 2, green clade). However, the orthology of this lineage is supported by synteny, as we found five genes that consistently are in its proximity in all three species (*NR4A1*, *GRASP*, *ACVR1B*, *ACVRL1*, *FMNL3*, Fig. 3). The phyletic distribution of this gene lineage allows us to trace back its origin to the last common ancestor of gnathostomes, and synteny revealed that genes found in its proximity in the spotted gar, coelacanth and elephant shark mapped unequivocally to chromosome 12 of humans (Fig. 3). We infer that this chromosome would be the putative genomic location of the third reprimo paralogon in humans, and that the gene copy was lost in the ancestor of tetrapods.

3.3. The reprimo gene repertoire in lampreys

In addition to the reprimo genes found in gnathostomes, we also annotated four reprimo sequences in the two species of lampreys. These genes were recovered as two distinct sister clades (Fig. 2, upper and lower clades). The sister group relationship of the two clades of lamprey reprimo genes is not well resolved. In the case of the lower clade we found three genes (FMNL2a, *MYLKb* and *SLC12A8*) downstream the reprimo sequence that define the orthology in both lamprey species. Two of these genes do not map to any of the already identified reprimo paralogons, suggesting that this reprimo gene could be either the fourth paralogon or the product of a duplication restricted to the lamprey lineage. In support of the former option we found that two flanking genes (*MYLKb* and *SLC12A8*) identified in both lamprey species consistently map on chromosome 3 of humans. In the case of the upper clade we did not find shared synteny between the two lampreys, probably due to the state of the current lamprey genome assemblies. Finally, the difficulty in resolving orthology using gene phylogenies between cyclostome and gnathostome genes is apparently related to particular features of the cyclostome genomes, such as GC and amino acid bias, that set them apart from all other vertebrate genomes (Qiu et al., 2011; Mehta et al., 2013; Smith et al., 2013; Campanini et al., 2015).

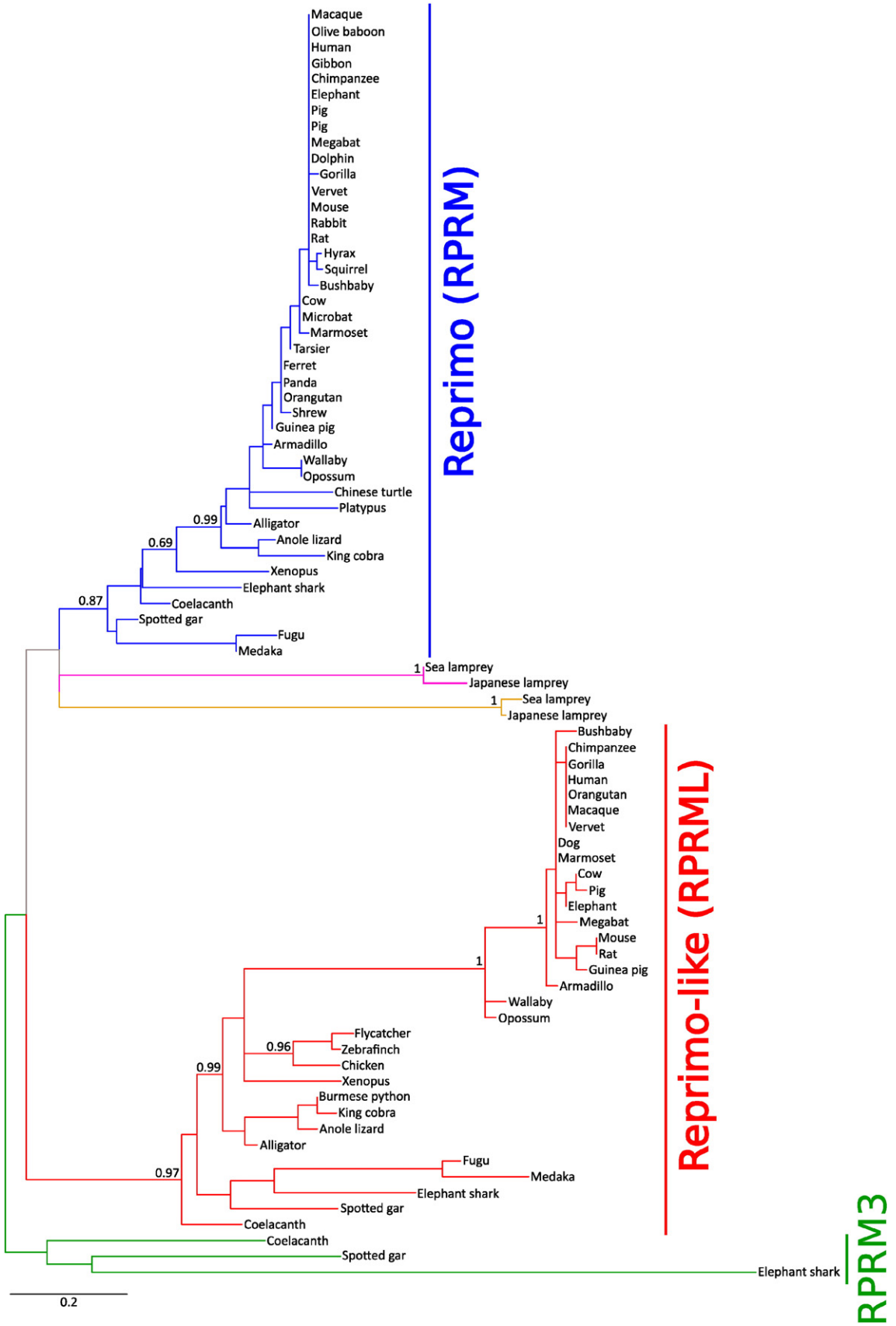

Fig. 2. Unrooted maximum likelihood phylogram depicting relationships among vertebrate reprimo genes based on amino acid sequence data. The number on the nodes corresponds to the posterior probability using the Bayesian approach.

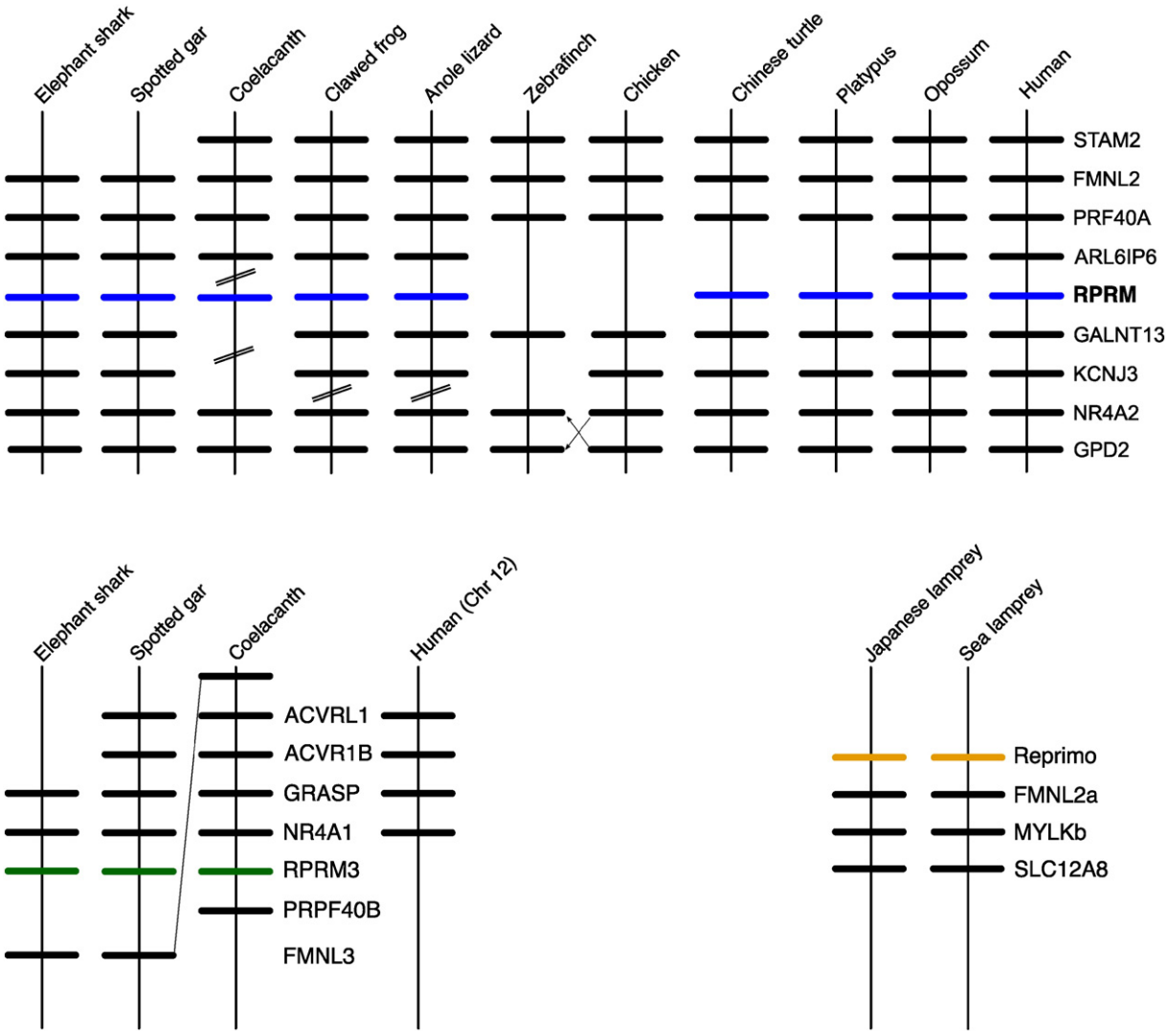

Fig. 3. Patterns of conserved synteny in the genomic regions that harbor reprimo genes in gnathostome vertebrates. Upper panel, synteny comparison of the RPRM gene region in representative species in gnathostome vertebrates. Lower panel (left), synteny comparison of the region that harbors the RPRM gene that belongs to the third RPRM lineage in the elephant shark, spotted gar and coelacanth. Syntenic genes mapped on chromosome 12 of humans suggesting that this chromosome would be the putative genomic location of the third reprimo paralogon in humans. Lower panel (right), synteny comparison of the region that harbors the RPRM gene of the orange clade in lampreys.

3.4. Additional duplications of the reprimo gene family in teleost fish

Given that teleost fish experienced an extra round of whole genome duplication (Meyer and Van de Peer, 2005; Kasahara, 2007; Sato and Nishida, 2010), we performed an additional phylogenetic analysis with an increased taxonomic sampling, including all available fish species in genome databases, as well as individual records. For these analyses we also included sequences from the elephant shark and coelacanth as they possess the three reprimo lineages. In this case the phylogenetic prediction is to find duplicated copies in all reprimo gene lineages in teleost fish relative to gar, a species that did not experienced the extra round of whole genome duplication (Braasch et al., 2016). Our phylogenies revealed the presence of the three already identified reprimo gene lineages (Fig. 6). In the case of the reprimo clade (Fig. 6, blue clade) traces of the two gene lineages derived from the teleost-specific genome duplication are present. In support of this claim we found that that there is at least three pair of genes on chromosome 9 and 6 (FMNL2a and FMNL2b, MYLKa and MYLKb, KCNJ3a and KCNJ3b) that support that RPRMa and RPRMb genes of zebrafish were originated as the product of the teleost specific genome duplication. For the reprimo-like clade, teleost species retained only one of the lineages derived from the teleost-specific genome duplication (Fig. 6, red clade). Finally, the third reprimo lineage has only been retained in cavefish and zebrafish (Fig. 6, green clade).

3.5. Expression of the reprimo genes in vertebrates

We also examined expression of the reprimo paralogs in a representative sample of gnathostomes. Our results revealed that in most examined species reprimo genes are mostly expressed in the brain (Table 1), and that there is not a clear pattern of which reprimo gene is the most abundantly expressed among species (Table 1). In humans, which possess *RPRM* and *RPRML* genes, the former gene is expressed more than three times relative to *RPRML* (Table 1), whereas in the spotted gar, the *RPRML* gene is the most abundantly expressed gene, almost seven times higher than the reprimo gene belonging to the third reprimo lineage. Interestingly, we did not detect *RPRM* expression in any tissue on this species. As in gar, the elephant shark reprimo gene belonging to the third reprimo lineage is the most abundantly expressed; followed by *RPRM*, while the *RPRML* gene is the least expressed (Table 1). In chicken, where *RPRML* is the only reprimo gene present in its genome, this gene is only expressed in the brain among the tissues and developmental

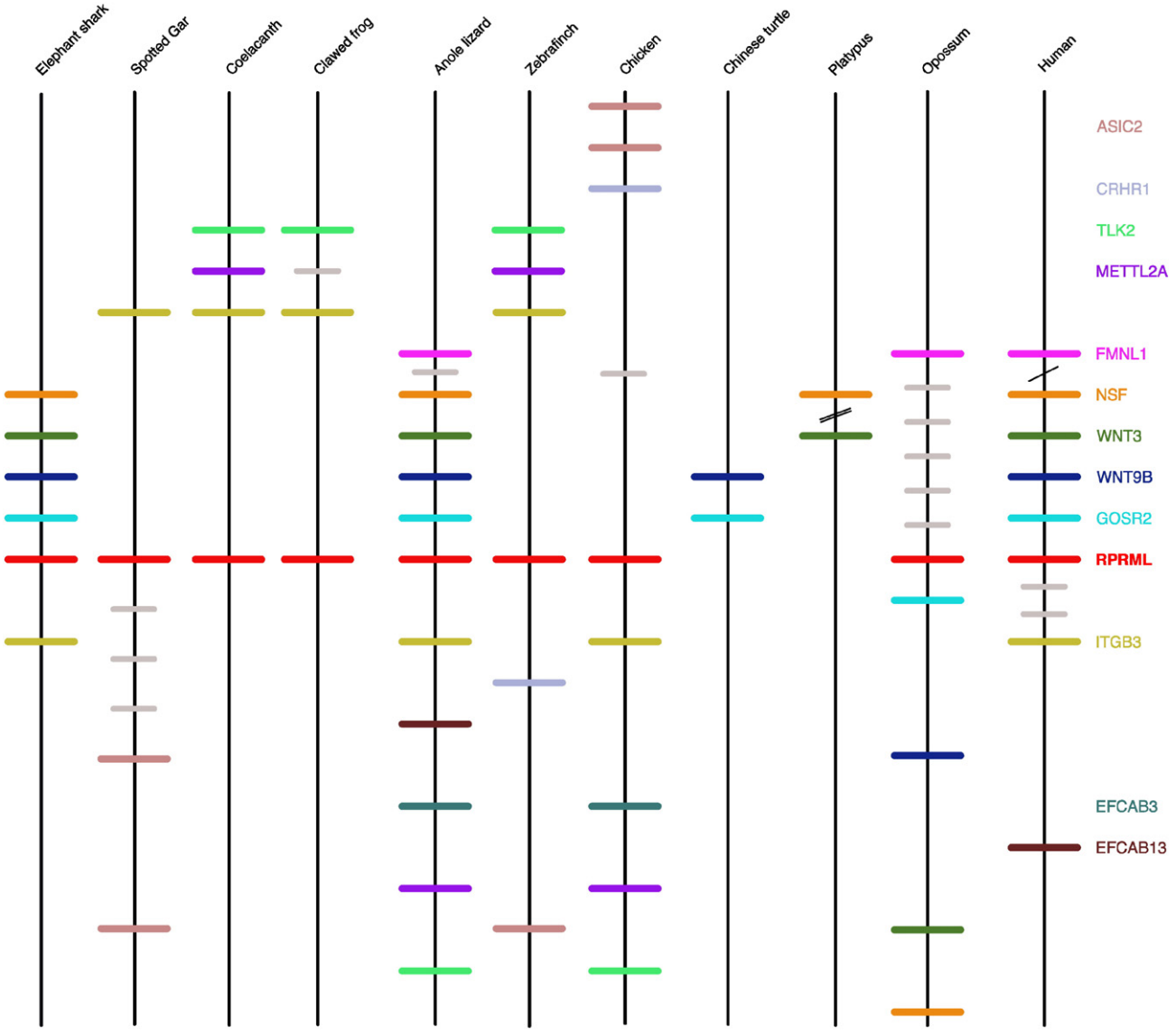

Fig. 4. Patterns of conserved synteny in the genomic region that harbors the RPRML gene in gnathostome vertebrates. The dotted box around the RPRML gene of humans and opossum denotes that this gene is found within another gene.

stages included in our analyses (Table 1). In zebrafish, all reprimo paralogs are expressed at low levels, where RPRMa and RPRMb are the most abundantly expressed genes in the brain and testis. The fact that reprimo genes are mainly expressed in the brain, a tissue that in general does not experience cell division, and that this pattern has been conserved for almost 500 million years presents an intriguing pattern. One possibility could be that reprimo expression plays a protector role in this organ during adulthood (Zohrabian et al., 2007).

3.6. Evolutionary diversification of the reprimo gene family

If the main lineages of reprimo genes represent the product of the two rounds of whole genome duplications occurred in the last common ancestor of vertebrates (Meyer and Schartl, 1999; McLysaght et al., 2002; Dehal and Boore, 2005; Hoegg and Meyer, 2005; Putnam et al., 2008), then representatives of the three reprimo gene lineages should be embedded in chromosomal regions that share similar arrangements of ohnologous genes (Dehal and Boore, 2005). To test this prediction, we searched for co-duplicated genes that are present on the chromosomes where the human species possess reprimo gene copies (chr 2 and 17), plus the chromosome where the third reprimo lineage it should be (chr 12). If the diversity of reprimo gene family is the product of the two rounds of whole genome duplications, ohnologous copies should be found on all three chromosomes. However, given the dynamics of the genomes after whole genome duplications only a fraction of gene families would be expected to retain all the resultant ohnologs (Dehal and Boore, 2005; Braasch et al., 2006; Braasch et al., 2007; Braasch et al., 2009). In our survey we found 17 ohnologous gene families (ASIC, ERBB, HDAC, IGFBP, IKZF, MYL, NEUROD, NFE, RAPGEF, SP, HOX, ATP5G, FMNL, FZD, ITGA, SCN, STAT) that were present on the three human chromosomes that suggest that the diversity of reprimo genes observed in vertebrates is the product of the two whole genome duplications occurred in the last common ancestor of the group. According to the ohnolog repository (http://ohnologs.curie.fr/) (Singh et al., 2015) all these genes also diversified as product of the whole genome duplications occurred in the last common ancestor of vertebrates, thus providing further support to our claim. All of these segments map to regions of the human genome that derive from the Ancestral Chordate Linkage group 16 (Putnam et al., 2008), providing further support for the role of whole genome duplications in the diversification of this

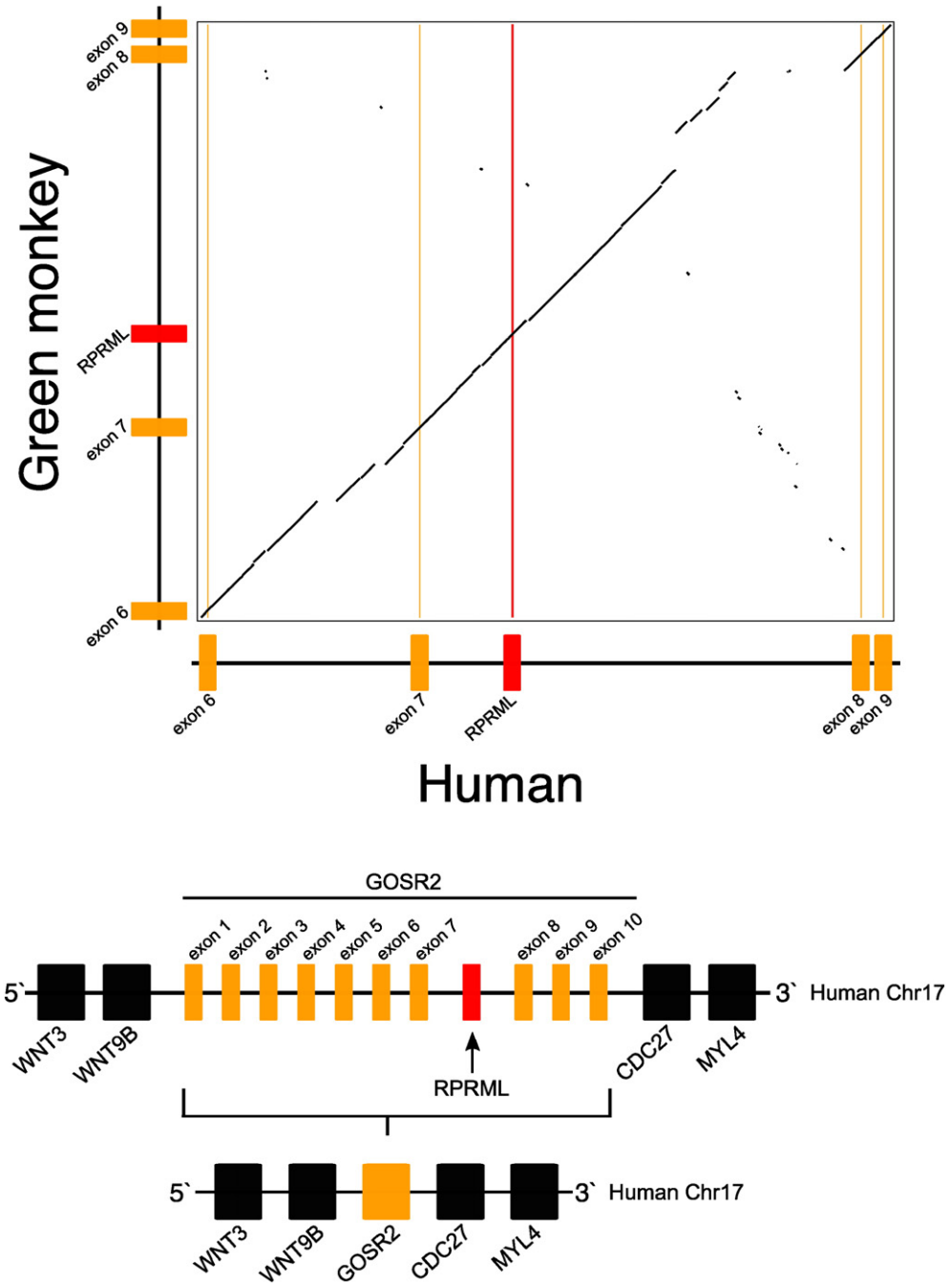

Fig. 5. Schematic representation of the RPRML gene location in humans. Upper panel, dot-plot comparison of the GOSR2 gene region containing the RPRML gene in humans and green monkey. Yellow vertical lines correspond to exons of the GOSR2 gene, whereas, the red vertical line correspond to the RPRML sequence. Lower panel, schematic representation of the syntenic region in which RPRML gene in humans is located. RPRML gene is in the reverse strand, whereas the GOSR2 gene is in the forward strand.

gene family. Pointing into the same direction, ohnologs that were lost in one of the chromosomes were also identified (e.g *BIN*, *GLI*, *NOS*, *PER*, *ATF*, *RAB5*).

3.7. Evolution of the p53 gene family a key regulator of reprimo genes

As stated above, *RPRM* is a downstream effector of p53, involved in cell cycle arrest at G2. Interestingly, p53-dependent cell cycle arrest may, to some extent, be restricted to vertebrates; as is the case for p53-dependent p21-mediated cell cycle arrest (Lu et al., 2009). Therefore, it is important to contextualize the evolutionary origins of the reprimo gene family, within the evolutionary history of the p53 gene family, upstream regulators for *RPRM*. The human p53 gene family consists of three members: p53, p63 and p73; all of which share a common "blueprint" consisting of an amino-terminal transactivation domain, a highly conserved central DNA-binding domain and a carboxy-terminal oligomerization domain (Lu et al., 2009). Its evolutionary origin apparently predates the emergence of the reprimo gene family as the presence of a p53-like gene has been reported in choanoflagellates, the sister group of animals, where two to three p53 paralogs have been described (Nedelcu and Tan, 2007; Lu et al., 2009; Belyi et al., 2010; Dötsch et al., 2010; Zmasek and Godzik, 2013). In insects and *Caenorhabditis elegans*, one homolog that is a hybrid p63/p73 gene has been described. In vertebrates the three members of the family (p53, p63 and p73) have

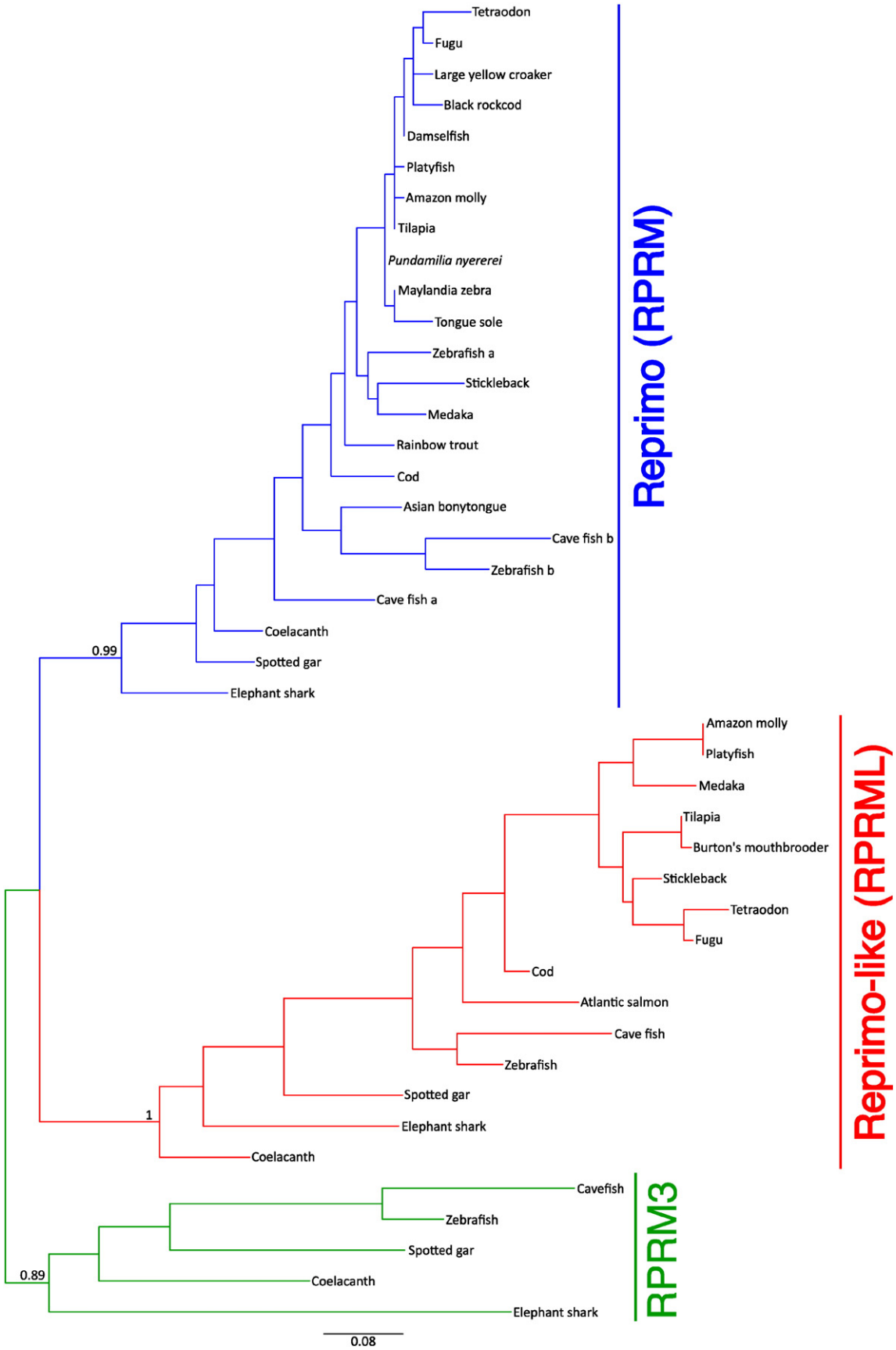

Fig. 6. Unrooted maximum likelihood phylogram depicting relationships among fish reprimo genes based on amino acid sequence data. The number on the nodes corresponds to the posterior probability using the Bayesian approach.

Table 1
Gene transcription profiles of the reprimo gene lineages in humans, chicken, spotted gar, zebrafish and elephant shark.

	Human (Two reprimo genes)			Chicken (One reprimo gene)			Spotted gar (Three reprimo genes)			Zebrafish (four reprimo genes)				Elephant shark (Three reprimo genes)		
	RPRM	*RPRML*	*RPRM3*	*RPRM*	*RPRML*	*RPRM3*	*RPRM*	*RPRML*	*RPRM3*	*RPRMa*	*RPRMb*	*RPRML*	*RPRM3*	*RPRM*	*RPRML*	*RPRM3*
Brain	53.9	16.4	NP	NP	311	NP	0	116.7	17	1.1	1.16	0.17	0.54	51.73	7.8	154.16
Heart	0	0	NP	NP	0	NP	0	0	0	0.56	0.12	0	0	1.81	5.65	0.14
Kidney	4.23	0.41	NP	NP	0	NP	0	0	8.62	0.79	0.06	0.1	0	0.25	2.81	0.15
Liver	0	0	NP	NP	0	NP	0	0	1.65	0	0	0	0	0.07	0.35	0
Muscle	0	0	NP	NP	0	NP	0	0	0	0	0	0.05	0.05	0.65	10.03	0
Testis	1.66	0.15	NP	NP	0	NP	0	0	0	1.32	1.8	0.3	0.25	9.12	0.86	1.5

NP denotes that the reprimo lineage is not present in the genome of the species. Gene expression was measured in transcript per million (TPM).

been identified (Lane et al., 2011), suggesting that the vertebrate specific whole genome duplications played a pivotal role in the diversification of the p53 gene family as well.

Taken together, our results recreate a story alongside RPRM, where genes involved in cell cycle dynamics in the p53 pathway diversified and acquired more complex and specific functions as a result of the gene family expansions derived from the vertebrate specific whole genome duplications. The diversification of regulatory genes in cell cycle dynamics (p53, p63, p73, and downstream effectors such as RPRM, RPRML and RPRM3) shows a more complex network guarding the integrity of cell replication in higher vertebrates, which may help provide further insight and understanding of deregulations leading to disease development, especially cancer.

4. Conclusions

In this study, we provided a comprehensive evolutionary analysis of the reprimo gene family in a sample of vertebrates including representative species of tetrapods, coelacanths, teleost fish, non-teleost ray-finned fish, cartilaginous fish, and cyclostomes. By tracking the evolution of the reprimo gene family among vertebrates and comparing patterns of expression in an evolutionary framework our study shows that 1- the reprimo gene family emerged early in vertebrate evolution, 2- even if multiple members of the family are present, a single representative is the most heavily expressed, 3- there is variation in which representative gene is the most heavily expressed, 4- gene expression occurs predominantly in the brain, and 5- there is preliminary evidence that the diversification of the reprimo gene family might have coincided with that of its signaling partner, p53. Our results are in agreement with previous reports that suggest that a relatively high proportion of human disease related genes originated in the last common ancestor of vertebrates (Makino and McLysaght, 2010; Singh et al., 2012; Maxwell et al., 2014).

Future studies should be directed towards obtaining basic information regarding the biology of the reprimo gene lineages and its relationship to cancer. This information will be also fundamental to understand the pattern of differential retention observed in this group of genes, specially the limited phyletic distribution of the *RPRM3* gene.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2016.07.036.

Declaration

The authors have no conflict of interests.

Acknowledgment

We thank two anonymous reviewers for their useful comments and suggestions on the manuscript. This work was funded by Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) grants 1160627 to JCO, 1151411 to AHC, 1140970 to GIO, CORFO 13IDL2-18608 & 13CTI-21526-P6 to GIO and CONICYT-FONDAP 15130011 to AHC and GIO, BMBF-CONICYT 20140027, CRP-ICGEB CH15-01 and CONICYT-ANILLO ACT1402 to JDA. FGH acknowledges support from grant EPS-0903787 from the National Science Foundation.

References

Belyi, V.A., Ak, P., Markert, E., Wang, H., Hu, W., Puzio-Kuter, A., Levine, A.J., 2010. The origins and evolution of the p53 family of genes. Cold Spring Harb. Perspect. Biol. 2, a001198.

Bernal, C., Aguayo, F., Villarroel, C., Vargas, M., Díaz, I., Ossandon, F.J., Santibáñez, E., Palma, M., Aravena, E., Barrientos, C., Corvalan, A.H., 2008. Reprimo as a potential biomarker for early detection in gastric cancer. Clin. Cancer Res. 14, 6264–6269.

Braasch, I., Salzburger, W., Meyer, A., 2006. Asymmetric evolution in two fish-specifically duplicated receptor tyrosine kinase paralogons involved in teleost coloration. Mol. Biol. Evol. 23, 1192–1202.

Braasch, I., Schartl, M., Volff, J.N., 2007. Evolution of pigment synthesis pathways by gene and genome duplication in fish. BMC Evol. Biol. 7, 74.

Braasch, I., Brunet, F., Volff, J.N., Schartl, M., 2009. Pigmentation pathway evolution after whole-genome duplication in fish. Genome Biol Evol 1, 479–493.

Braasch, I., Gehrke, A.R., Smith, J.J., Kawasaki, K., Manousaki, T., Pasquier, J., Amores, A., Desvignes, T., Batzel, P., Catchen, J., Berlin, A.M., Campbell, M.S., Barrell, D., Martin, K.J., Mulley, J.F., Ravi, V., Lee, A.P., Nakamura, T., Chalopin, D., Fan, S., Wcisel, D., Caestro, C., Sydes, J., Beaudry, F.E., Sun, Y., Hertel, J., Beam, M.J., Fasold, M., Ishiyama, M., Johnson, J., Kehr, S., Lara, M., Letaw, J.H., Litman, G.W., Litman, R.T., Mikami, M., Ota, T., Saha, N.R., Williams, L., Stadler, P.F., Wang, H., Taylor, J.S., Fontenot, Q., Ferrara, A., Searle, S.M., Aken, B., Yandell, M., Schneider, I., Yoder, J.A., Volff, J.N., Meyer, A., Amemiya, C.T., Venkatesh, B., Holland, P.W., Guiguen, Y., Bobe, J., Shubin, N.H., Di Palma, F., Alfoldi, J., Lindblad-Toh, K., Postlethwait, J.H., 2016. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. Nat. Genet. 48, 427–437.

Braidy, N., Poljak, A., Jayasena, T., Mansour, H., Inestrosa, N.C., Sachdev, P.S., 2015. Accelerating Alzheimer's research through 'natural' animal models. Curr Opin Psychiatry 28, 155–164.

Campanini, E.B., Vandewege, M.W., Pillai, N.E., Tay, B.H., Jones, J.L., Venkatesh, B., Hoffmann, F.G., 2015. Early evolution of vertebrate Mybs: an integrative perspective combining synteny, phylogenetic, and gene expression analyses. Genome Biol Evol 7, 3009–3021.

Castro-Fuentes, R., Socas-Pérez, R., 2013. Octodon degus: a strong attractor for Alzheimer research. Basic Clin Neurosci 4, 91–96.

Dehal, P., Boore, J.L., 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3, e314.

Dötsch, V., Bernassola, F., Coutandin, D., Candi, E., Melino, G., 2010. p63 and p73, the ancestors of p53. Cold Spring Harb. Perspect. Biol. 2, a004887.

Edrey, Y.H., Casper, D., Huchon, D., Mele, J., Gelfond, J.A., Kristan, D.M., Nevo, E., Buffenstein, R., 2012. Sustained high levels of neuregulin-1 in the longest-lived rodents; a key determinant of rodent longevity. Aging Cell 11, 213–222.

Ellinger, J., Bastian, P.J., Jurgan, T., Biermann, K., Kahl, P., Heukamp, L.C., Wernert, N., Müller, S.C., von Ruecker, A., 2008. CpG island hypermethylation at multiple gene sites in diagnosis and prognosis of prostate cancer. Urology 71, 161–167.

Faulkes, C.G., Davies, K.T., Rossiter, S.J., Bennett, N.C., 2015. Molecular evolution of the hyaluronan synthase 2 gene in mammals: implications for adaptations to the subterranean niche and cancer resistance. Biol. Lett. 11, 20150185.

Gorbunova, V., Hine, C., Tian, X., Ablaeva, J., Gudkov, A.V., Nevo, E., Seluanov, A., 2012. Cancer resistance in the blind mole rat is mediated by concerted necrotic cell death mechanism. Proc. Natl. Acad. Sci. U. S. A. 109, 19392–19396.

Hamilton, J.P., Sato, F., Jin, Z., Greenwald, B.D., Ito, T., Mori, Y., Paun, B.C., Kan, T., Cheng, Y., Wang, S., Yang, J., Abraham, J.M., Meltzer, S.J., 2006. Reprimo methylation is a potential biomarker of Barrett's-associated esophageal neoplastic progression. Clin. Cancer Res. 12, 6637–6642.

Henning, Y., Vole, C., Begall, S., Bens, M., Broecker-Preuss, M., Sahm, A., Szafranski, K., Burda, H., Dammann, P., 2014. Unusual ratio between free thyroxine and free triiodothyronine in a long-lived mole-rat species with bimodal ageing. PLoS One 9, e113698.

Hoegg, S., Meyer, A., 2005. Hox clusters as models for vertebrate genome evolution. Trends Genet. 21, 421–424.

Inestrosa, N.C., Ríos, J.A., Cisternas, P., Tapia-Rojas, C., Rivera, S., Braidy, D., Zolezzi, N., Godoy, J.M., Carvajal, J.A., Ardiles, F.J., Bozinovic, A.O., Palacios, A.G., F., Sachdev, P.S., 2015. Age progression of Neuropathological markers in the brain of the Chilean rodent Octodon Degus, a natural model of Alzheimer's disease. Brain Pathol. http://dx.doi.org/10.1111/bpa.12226.

Kasahara, M., 2007. The 2R hypothesis: an update. Curr. Opin. Immunol. 19, 547–552.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Lane, D.P., Madhumalar, A., Lee, A.P., Tay, B.H., Verma, C., Brenner, S., Venkatesh, B., 2011. Conservation of all three p53 family members and Mdm2 and Mdm4 in the cartilaginous fish. Cell Cycle 10, 4272–4279.
Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.
Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Biochem. 12, 323.
Lin, S., 2010. The role of reprimo-like gene methylation in radiation responsiveness and prognosis in pancreatic cancer. J. Clin. Oncol. 28, e14593.
Lu, W.J., Amatruda, J.F., Abrams, J.M., 2009. p53 ancestry: gazing through an evolutionary lens. Nat. Rev. Cancer 9, 758–762.
Luo, J., Zhu, Y., Yang, G., Gong, L., Wang, B., Liu, H., 2011. Loss of reprimo and S100A2 expression in human gastric adenocarcinoma. Diagn. Cytopathol. 39, 752–757.
Makino, T., McLysaght, A., 2010. Ohnologs in the human genome are dosage balanced and frequently associated with disease. Proc. Natl. Acad. Sci. U. S. A. 107, 9270–9274.
Manov, I., Hirsh, M., Iancu, T.C., Malik, A., Sotnichenko, N., Band, M., Avivi, A., Shams, I., 2013. Pronounced cancer resistance in a subterranean rodent, the blind mole-rat, Spalax: in vivo and in vitro evidence. BMC Biol. 11, 91.
Maxwell, E.K., Schnitzler, C.E., Havlak, P., Putnam, N.H., Nguyen, A.D., Moreland, R.T., Baxevanis, A.D., 2014. Evolutionary profiling reveals the heterogeneous origins of classes of human disease genes: implications for modeling disease genetics in animals. BMC Evol. Biol. 14, 212.
McLysaght, A., Hokamp, K., Wolfe, K.H., 2002. Extensive genomic duplication during early chordate evolution. Nat. Genet. 31, 200–204.
Mehta, T.K., Ravi, V., Yamasaki, S., Lee, A.P., Lian, M.M., Tay, B.H., Tohari, S., Yanai, S., Tay, A., Brenner, S., Venkatesh, B., 2013. Evidence for at least six Hox clusters in the Japanese lamprey (*Lethenteron japonicum*). Proc. Natl. Acad. Sci. U. S. A. 110, 16044–16049.
Meyer, A., Schartl, M., 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. 11, 699–704.
Meyer, A., Van de Peer, Y., 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). BioEssays 27, 937–945.
Muffato, M., Louis, A., Poisnel, C.E., Roest Crollius, H., 2010. Genomicus: a database and a browser to study gene synteny in modern and ancestral genomes. Bioinformatics 26, 1119–1121.
Nakajima, T., Yasufuku, K., Suzuki, M., Fujiwara, T., Shibuya, K., Takiguchi, Y., Hiroshima, K., Kimura, H., Yoshino, I., 2009. Assessment of Chemosensitivity-related aberrant methylation of nonsmall cell lung cancer by EBUS-TBNA. J Bronchology Interv Pulmonol 16, 10–14.
Nedelcu, A.M., Tan, C., 2007. Early diversification and complex evolutionary history of the p53 tumor suppressor gene family. Dev. Genes Evol. 217, 801–806.
Nishida, N., Nagasaka, T., Nishimura, T., Ikai, I., Boland, C.R., Goel, A., 2008. Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. Hepatology 47, 908–918.
Novikov, E., Burda, G., 2013. Ecological and evolutionary preconditions of extended longevity in subterranean rodents. Biol. Bull. Rev. 3, 325–333.
Ohki, R., Nemoto, J., Murasawa, H., Oda, E., Inazawa, J., Tanaka, N., Taniguchi, T., 2000. Reprimo, a new candidate mediator of the p53-mediated cell cycle arrest at the G2 phase. J. Biol. Chem. 275, 22627–22630.
Putnam, N.H., Butts, T., Ferrier, D.E., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.K., Benito-Gutiérrez, E.L., Dubchak, I., Garcia-Fernàndez, J., Gibson-Brown, J.J., Grigoriev, I.V., Horton, A.C., de Jong, P.J., Jurka, J., Kapitonov, V.V., Kohara, Y., Kuroki, Y., Lindquist, E., Lucas, S., Osoegawa, K., Pennacchio, L.A., Salamov, A.A., Satou, Y., Sauka-Spengler, T., Schmutz, J., Shin-I, T., Toyoda, A., Bronner-Fraser, M., Fujiyama, A., Holland, L.Z., Holland, P.W., Satoh, N., Rokhsar, D.S., 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453, 1064–1071.
Qiu, H., Hildebrand, F., Kuraku, S., Meyer, A., 2011. Unresolved orthology and peculiar coding sequence properties of lamprey genes: the KCNA gene family as test case. BMC Genomics 12, 325.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Saavedra, K., Valbuena, J., Olivares, W., Marchant, M.J., Rodríguez, A., Torres-Estay, V., Carrasco-Avino, G., Guzmán, L., Aguayo, F., Roa, J.C., Corvalán, A.H., 2015. Loss of expression of reprimo, a p53-induced cell cycle arrest gene, correlates with invasive stage of tumor progression and p73 expression in gastric cancer. PLoS One 10, e0125834.
Sato, Y., Nishida, M., 2010. Teleost fish with specific genome duplication as unique models of vertebrate evolution. Environ. Biol. Fish 88, 169–188.
Sato, N., Fukushima, N., Hruban, R.H., Goggins, M., 2008. CpG island methylation profile of pancreatic intraepithelial neoplasia. Mod. Pathol. 21, 238–244.
Singh, P.P., Affeldt, S., Cascone, I., Selimoglu, R., Camonis, J., Isambert, H., 2012. On the expansion of "dangerous" gene repertoires by whole-genome duplications in early vertebrates. Cell Rep. 2, 1387–1398.
Singh, P.P., Arora, J., Isambert, H., 2015. Identification of ohnolog genes originating from whole genome duplication in early vertebrates, based on synteny comparison across multiple genomes. PLoS Comput. Biol. 11, e1004394.
Smith, J.J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M.S., Yandell, M.D., Manousaki, T., Meyer, A., Bloom, O.E., Morgan, J.R., Buxbaum, J.D., Sachidanandam, R., Sims, C., Garruss, A.S., Cook, M., Krumlauf, R., Wiedemann, L.M., Sower, S.A., Decatur, W.A., Hall, J.A., Amemiya, C.T., Saha, N.R., Buckley, K.M., Rast, J.P., Das, S., Hirano, M., McCurley, N., Guo, P., Rohner, N., Tabin, C.J., Piccinelli, P., Elgar, G., Ruffier, M., Aken, B.L., Searle, S.M., Muffato, M., Pignatelli, M., Herrero, J., Jones, M., Brown, C.T., Chung-Davidson, Y.W., Nanlohy, K.G., Libants, S.V., Yeh, C.Y., McCauley, D.W., Langeland, J.A., Pancer, Z., Fritzsch, B., de Jong, P.J., Zhu, B., Fulton, L.L., Theising, B., Flicek, P., Bronner, M.E., Warren, W.C., Clifton, S.W., Wilson, R.K., Li, W., 2013. Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. Nat. Genet. 45 (415–21), 421e1–421e2.
Stanke, M., Diekhans, M., Baertsch, R., Haussler, D., 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24, 637–644.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tarragon, E., Lopez, D., Estrada, C., Ana, G.C., Schenker, E., Pifferi, F., Bordet, R., Richardson, J.C., Herrero, M.T., 2013. *Octodon degus*: a model for the cognitive impairment associated with Alzheimer's disease. CNS Neurosci Ther 19, 643–648.
Tatusova, T.A., Madden, T.L., 1999. BLAST 2 sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. 174, 247–250.
Taylor, W.R., Stark, G.R., 2001. Regulation of the G2/M transition by p53. Oncogene 20, 1803–1815.
Vilella, A.J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., Birney, E., 2009. EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. Genome Res. 19, 327–335.
Wong, T.S., Kwong, D.L., Sham, J.S., Wei, W.I., Yuen, A.P., 2005. Methylation status of Reprimo in head and neck carcinomas. Int. J. Cancer 117, 697.
Yu, C., Li, Y., Holmes, A., Szafranski, K., Faulkes, C.G., Coen, C.W., Buffenstein, R., Platzer, M., de Magalhães, J.P., Church, G.M., 2011. RNA sequencing reveals differential expression of mitochondrial and oxidation reduction genes in the long-lived naked mole-rat when compared to mice. PLoS One 6, e26729.
Zmasek, C.M., Godzik, A., 2013. Evolution of the animal apoptosis network. Cold Spring Harb. Perspect. Biol. 5, a008649.
Zohrabian, V.M., Nandu, H., Gulati, N., Khitrov, G., Zhao, C., Mohan, A., Demattia, J., Braun, A., Das, K., Murali, R., Jhanwar-Uniyal, M., 2007. Gene expression profiling of metastatic brain cancer. Oncol. Rep. 18, 321–328.