The vertebrate muscle-specific RING finger protein family includes MuRF4 – a novel and widely conserved E3-ubiquitin ligase

Daniel J. Macqueen [1] *, Eduardo N. Fuentes [1,2,3], Juan Antonio Valdés [2,3], Alfredo Molina [2,3], Samuel A.M. Martin [1]

[1] *Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK*

[2] *Laboratorio de Biotecnología Molecular, Facultad de Ciencias Biológicas, Universidad Andres Bello, Santiago, Chile*

[3] *Interdisciplinary Center for Aquaculture Research (INCAR), Víctor Lamas 1290, PO Box 160-C, Concepción, Chile*

*Corresponding Author

Abstract

Muscle-specific RING finger (MuRF) proteins are E3-ubiquitin ligases and key regulators of muscle growth and turnover. Here, using a range of phylogenomic approaches, we established the complete-definitive MuRF family of vertebrates. Adding to recognized MuRF1, 2 and 3, we describe a novel family member, hereafter MuRF4, which was independently lost during placental mammal and bird evolution, but is otherwise conserved. *MuRF4* transcripts were expressed in heart and skeletal muscles of zebrafish, but were barely detectable in striated muscles of adult anole lizards. We also demonstrate that *MuRF1* underwent retrotransposition in the teleost fish ancestor, before the retrogene fully replaced the original gene and muscle-specific function.

Key Words: Muscle-specific RING finger family, Novel MuRF: MuRF4, Striated muscle, E3-Ubiquitin ligase, Evolution, Retrotransposition

Abbreviations: AA: amino acid; cDNA: first strand complementary cDNA; FoXO: Forkhead box protein class O; GR: Glucocorticoid receptor; JTT: Jones-Taylor-Thornton AA substitution model; ML: Maximum likelihood; MuRF: Muscle-specific RING finger; NCBI: National Center for Biotechnology Information; NF-κB: nuclear factor kappa-light-chain-enhancer of activated B cells; qPCR: quantitative polymerase chain replication.

Highlights:

- The complete-definitive vertebrate MuRF family was characterized
- The ancestral jawed vertebrate genome had four MuRF family members
- MuRF4 is uncharacterised yet conserved in many major vertebrate lineages
- Expression of all MuRF family member genes was studied in distant vertebrate taxa
- *MuRF1* gene paralogues of teleost fish are functional retrogenes

1. Introduction

The MuRFs have been recognised for around fourteen years as a small group of related proteins within the TRIM/RBCC superfamily [1-4]. Like other TRIM/RBCC proteins, they are characterized by a conserved tripartite domain, which is split into an N-terminal RING-finger motif, sequentially followed by a MuRF-family specific conserved box (MFC), a zinc-binding B-box motif and two coiled-coil dimerization boxes [1-4]. The C-terminal of MuRFs is less well conserved but contains the acidic region (AR), a tail domain rich in acidic residues [2-4]. Until now, three MuRF family members have been characterized in vertebrates called MuRF1, 2 and 3 - also called TRIM63, 55 and 54, respectively. Each of these MuRFs is largely restricted to cardiac and skeletal muscle [3-6], where, in mammals, they localize to sarcomeres [4].

The importance of MuRFs as regulators of mammalian striated muscle turnover has been demonstrated through genetic, pharmacological and biochemical approaches, with MuRF1 receiving particular attention. In model mammal species, this molecule is a robust marker for muscle atrophy that is transcriptionally up-regulated in response to denervation, injury, joint immobilization, glucocorticoid treatment, sepsis, cancer, and aging [7-8]. The knockout of each *MuRF* gene has been achieved in mice, both individually [8, 9-10] or in combination for MuRF1 and 2 [9], demonstrating the importance of MuRF1 as a regulator of atrophy under catabolic contexts [8] and the individual or combined importance of MuRFs in the normal development [9] and protection of heart muscle [10]. In terms of their roles as E3-ubiquitin ligases, characterized mammalian MuRF targets include major sarcomeric proteins such as myosin heavy chain, myosin light chain and troponin-I, [11-13], while additional binding partners are known [2, 10, 14] that may or may not be targeted for degradation [2].

Progress made in understanding the functions and regulation of mammalian MuRFs is not mirrored at the evolutionary level. The major focus of MuRF research has been on MuRF1 in human and mouse, particularly in the context of muscle atrophy. Due to a general lack of work with other vertebrate taxa, it still remains unknown whether the mammalian MuRF repertoire is even representative of the remaining ninety percent of vertebrate species, where muscle turnover is just as crucial for survival. It is currently thought that many teleost fish have orthologues of mammalian MuRF1, 2 and 3 [5-6, 15-19], suggesting the family arose

during early vertebrate evolution. In addition, *MuRF1* and/or *2* are transcriptionally induced in the skeletal muscle of teleosts under a range of conditions promoting muscle remodelling, including fasting [5-6, 15-16], spawning [6] and treatment with lipopolysaccharide [5] and 17β-estradiol [20]. Therefore, the function of MuRFs in muscle atrophy is thought to be conserved. However, current data on MuRF evolution and conservation is not comprehensive, especially considering that a restricted taxonomic focus during gene family characterization studies may limit discovery of ancestral vertebrate gene family repertoires [21]. Motivated by such issues, this studies aim was to characterize the complete vertebrate MuRF family. Our findings reveal a hitherto unrecognized MuRF family member and the existence of functional *MuRF1* retrogenes in all teleosts - the single largest vertebrate group.

2. Materials and Methods

2.1. Sequence searches

We searched for and downloaded vertebrate *MURF* genes from genome assemblies available in Ensembl (http://www.ensembl.org/). Details of the species studied, including assembly versions, used is provided in Table S1. Searches were facilitated by the EnsemblCompara GeneTrees paralogy function [22]. We also searched for *MuRF* sequences using BLAST [23] against a range of NCBI databases, including non-redundant proteins and shotgun-transcriptome assemblies.

2.2. Phylogenetic Analyses

58 putative MuRF sequences were aligned at the AA-level using MAFFT [24] and the GUIDANCE algorithm [25-26] to gain statistical confidence at each aligned site. After filtering sites below the recommended cut-off [25-26], sequences were uploaded to Mega 5.0 [27], where the best-fitting AA substitution model was identified by ML. According to Bayesian model selection, this was JTT [28] with estimation of the gamma parameter to account for among site rate variation. The same sequence data and substitution model was used in a Bayesian phylogenetic analysis performed in BEAST v.1.7 [29], employing an uncorrelated lognormal relaxed molecular clock model [30], a Yule speciation prior [31] and a UPGMA starting tree. The BEAST analysis was run twice using a Markov chain Monte Carlo (MCMC) chain of 10,000,000 steps, sampling every 500 steps. Convergence of the MCMC chains was confirmed using Tracer v.1.5 (http://tree.bio.ed.ac.uk/software/tracer/). A maximum clade credibility tree from one run was produced with TreeAnnotator [29] after removing the first 10% of MCMC samples. We used the same data for ML phylogenetic analyses performed in PhyML [32] via an online service [33], employing the same substitution model and an approximate likelihood ratio test (aLRT) [34] to gain support for branching patterns.

The same approaches were used to create further Bayesian and ML phylogenies either for more limited sets of MuRF sequences, or the original complete set plus additional TRIM/RBCC family members from human (*Homo sapiens*) and zebrafish (*Danio rerio*). Ensembl identifier numbers for all MuRF and TRIM family member sequences used are provided within figures. All sequence alignments used in phylogenetic analysis are provided in the supporting information (Fig. S1A-C).

2.3. Comparative genomics and sequence analyses

We established MuRF protein domain organization with respect to gene intron-exon structures based on Ensembl gene model predictions. All protein domain annotations were made by comparison to the MuRF1 reference sequence from *Homo sapiens* (NCBI RefSeq: NP_115977). We used the NCBI tool Open Reading Frame Finder to identify putative *MuRF1* retrogenes from genomic sequences downloaded from Ensembl for various teleost species.

2.4. Transcript expression analyses

We used qPCR to estimate the relative mRNA expression of the full repertoire of *MuRF* genes in *Anolis carolinensis* (anole lizard, Tetrapoda) and *Danio rerio* (zebrafish, Teleostei). Accordingly, four and six respective primer pairs were designed to *MuRF* gene exons (Table S2). Primers were positioned in highly differing regions between *MuRF* genes and, when possible, in different exons or spanning exon-boundaries. The cDNA samples used for each species have been described elsewhere, along with detailed methods of the qPCR study design [21, 35]. Briefly, the cDNAs were reverse transcribed from total RNA pooled for four and six adult individuals of anole lizard and zebrafish respectively. The reverse transcriptions included a genomic DNA removal step. The cDNA samples were run in duplicate qPCR assays including gene-specific primers and Brilliant III ultra-fast SYBRgreen (Agilent Technologies) on an Mx3005P system (Agilent Technologies). For each *MuRF* assay, all the samples were run within single plates that always included minus-cDNA controls and equivalent assays for *rps13,* a reference gene used for normalization [21, 35]. Dissociation curves were used to ensure a single product was amplified in all final qPCR assays. The data was analysed in Genex v.5. (MultiD Analyses AB) with each gene placed on a relative scale following normalization to *rps13* before final presentation in the style of a Northern dot blot [36]. Despite using qPCR, this method should be considered semi-quantitative, given the lack of biological replication, and the fact that it does not account for differences in primer efficiency/lacks a robust normalization strategy.

3. Results and Discussion

3.1. Identification of MuRF genes in vertebrate genomes

We searched for *MuRF* genes in taxa broadly spanning the vertebrate phylogeny. MuRF1, 2 and 3 share around 50% AA identity and can therefore be easily distinguished from the next closest TRIM/RBCC members, which share less than 25% identity with any MuRF. In several distantly related species, including coelacanth, spotted gar, anole lizard, platypus and Tasmanian devil, we identified four distinct putative *MuRF* genes. However, unlike teleost fish, where *MuRF* family gene duplicates are recognized [5-6], which may result from teleost-specific whole genome duplication (WGD) events [e.g. 37-38], these species have not undergone WGDs beyond those common to all vertebrates (i.e. two WGD rounds [39]).

3.2. Phylogenetic analysis of MuRF proteins

Phylogenetic analyses were used to determine evolutionary relationships among putative MuRF sequences. Initially, we built trees based on MuRF sequences alone, i.e. excluding other TRIM family member sequences. We used a Bayesian method that incorporates a relaxed molecular clock model allowing statistical inference of the tree's root [29-30] without enforcing distant outgroup sequences as the root, which can lead to branching artefacts during analyses of vertebrate gene families [e.g. 40-41]. The Bayesian tree, along with supporting data from an independent ML reconstruction, is presented in Fig. 1, where the sequences split into four strongly supported clades. Each of these clades contains a range of vertebrate species that last shared an ancestor before the divide of the lobe-finned fish (i.e. the group containing tetrapods such as humans) and ray-finned fish (i.e. the group containing teleost fish) (Fig. 1). Within each clade, the branching patterns were largely congruent with expected phylogenetic relationships and most of the major taxonomic groups were represented (Fig. 1).

The recognized mammalian MuRFs each fell into one of the four vertebrate clades (Fig. 1), providing strong support for the existence of true MuRF1, 2 and 3 orthologues in a wide range of jawed vertebrates. The fourth MuRF clade contains a zebrafish sequence previously identified in a study of teleost TRIM family genes [17], where it was tentatively called TRIM101, but was not linked to the MuRF family. As for MuRF1, 2 and 3, a broad range of vertebrates are represented in this clade, which thus represents a grouping of novel vertebrate orthologues, hereafter called MuRF4. Under the Bayesian method, MuRF4 received maximal support as being ancestral to MuRF1, 2 and 3 (Fig. 1). However, this arrangement was not recaptured in ML analysis (Fig. 1).

We repeated the Bayesian and ML phylogenetic analyses including sequences for human and zebrafish TRIM/RBCC superfamily members that are most closely related to MuRFs in terms of sequence identity (TRIM9/13/46/59/67). Using both methods, the vertebrate MuRF sequences formed a single grouping with maximal statistical support (Fig. 2A). This provides evidence that MuRF4 is a new member of the vertebrate MuRF family. This is independently

supported by comparison of N-terminal RING-finger and MFS domains, where MuRF4 shares a similar level of identity with MuRF1, 2 and 3 as these proteins do to one another (Fig. 2B).

The inclusion of TRIM/RBCC outgroup sequences recovered a distinct branching of the four MuRF clades when compared to their exclusion (compare Fig. 1 and 2A). In the Bayesian outgroup tree, MuRF1 and 3 share a sister relationship, as do MuRF2 and 4, although the support for these groupings is weak, and were not recovered with ML reconstruction (Fig. 2A). Thus, phylogenetic analysis alone cannot resolve evolutionary relationships among the different MuRF family members.

3.3. Duplications of teleost MuRFs

There was evidence for the presence of teleost-specific gene duplicates for MuRF1 and 2, but not MuRF3 and 4 (Fig. 1). MuRF1 and 2 sequences split into two sister clades, each represented by distant teleost species, often including the same species (Fig. 1; branches coloured red and blue to show paralogous groups). We propose that future studies of teleost MuRFs employ a nomenclature using 'A' and 'B' to distinguish the two teleost paralogues (highlighted in Fig. 1).

3.4. Distribution and losses of MuRFs in vertebrate genomes

The number of *MuRF* family genes identified in Ensembl vertebrate genomes ranged from two to four. *MuRF1* and *2* were represented among all the major vertebrate lineages (Fig. 3), while *MuRF3* was not identified in reptile (n=2) or amphibian genomes (n=1). However, BLAST searches revealed true *MuRF3* orthologues in the amphibians *Xenopus leavis* and *Hynobius chinensis* as well as the reptile *A. carolinensis* (Fig. S2). Thus, we conclude that MuRF1, 2 and 3 are represented in all the major vertebrate lineages (Fig. 3).

There was no evidence for a *MuRF4* orthologue in any placental mammal genome in Ensembl. As this represents over thirty genome assemblies spanning the entire evolution of this group, it is parsimonious to conclude a true loss of *MuRF4* in a stem placental mammal. There was also no evidence for a MuRF4 orthologue in any Ensembl avian genome (n=5). We performed BLAST searches against the complete predicted protein complements of 13 avian genomes (8 additional to Ensembl), where the top hits were invariably MuRF2. As for mammals, the species searched broadly span the avian phylogeny. Thus, we conclude that MuRF4 was lost during an early point of avian evolution, independent from placental mammals (Fig. 3). Otherwise, *MURF4* is found in species representing all major remaining vertebrate groups, which account for around three-quarters of known species (Fig. 3).

3.5. Conservation of MuRF gene and protein structures

The protein domain structure of each MuRF family member was characterized with respect to genomic organization in representative mammal and teleost species (Fig. 4). We observed that *MuRF* family member genes (barring zebrafish *MURF1* paralogues; see next section) share genomic features, including conservation of exon length leading up to the AR domain, along with positional conservation of protein domains with respect to exon boundaries. The length of the MFC, BBOX and BBC domains is almost invariant among MuRF family members, while the RING and AR show greater length variation (Fig. 4). We also observed putative intron gain events in *MuRF2A* of zebrafish and *MuRF4* of Tasmanian devil (black arrows in Fig. 4). These data clearly highlight conservation of genomic organization among *MuRF* genes and add another level of support to the status of *MuRF4* as a true MuRF family member.

3.6. Teleost MuRF1 paralogues are functional retrogenes

Zebrafish *MuRF1A* and *1B* are intronless genes (Figs. 4 and 5) and there is no evidence for zebrafish *MuRF1* copies containing introns. This finding might be explained if, during the evolution of zebrafish, *MuRF1A* and *1B* mRNAs (transcribed from intron-containing genes) were independently reverse-transcribed to cDNAs then reinserted into the genome by retrotransposition [42], thereafter replacing the original genes. However, we consider this hypothesis implausible, given that retrotransposition followed by functional replacement of the original intron-containing gene was recently quantified systemically in humans and represents an extremely rare evolutionary event [43], unlikely to affect two related genes by chance.

A more parsimonious model is that a *MURF1* retrogene functionally replaced a single 'mother' *MURF1* gene in an ancestor to teleost fish, with the retrogene then being duplicated during the basal teleost WGD [37, 39] and the resultant paralogues subsequently descended during evolution. This model predicts the presence of two teleost MURF1 sister clades in phylogenetic analysis (as observed in Fig. 1; corroborated in Fig 5), and that all teleost *MURF1* genes are intronless. Indeed this second prediction is strongly supported, as all identified teleost *MURF1* genes code a complete MuRF1 protein with all MuRF domains within a single uninterrupted open reading frame (Fig. 5; see Fig. S1D). The spotted gar, a ray-finned fish that split from teleosts over 350 million years ago (before the teleost-specific WGD [44]), contains the same *MURF1* genomic organization as a range of lobe-finned fish lineages, including tetrapods (Fig. 5). Thus, we hypothesize that a putative *MURF1* retrotransposition event occurred specifically within the teleost lineage, sometime before the basal WGD event [39, 41] (Fig. 5).

3.7. Expression of MURF genes in vertebrates

To examine the conservation of *MuRF* gene family expression in distant vertebrate taxa, we performed qPCR assays to estimate tissue transcript levels of *MuRF* genes in adult anole lizards and zebrafish, species separated by around 420 million years evolution [45]. We used the housekeeping gene *rps13* as a reference gene to normalize the data, which was expressed abundantly in all tested tissues. In lizards, *MuRF1, 2* and *3* transcripts were much more abundantly expressed in heart and skeletal muscle than other tissues (Fig. 6A), as observed previously in mammals [4]. However, *MuRF4* transcripts were barely detected in striated muscles, despite being detected at in brain (Fig. 6A). We also detected low levels of *MuRF1, 2* and *3* transcripts in the lizard brain, at comparably lower levels than *MuRF4* (Fig. 5A). As observed in mammals [4], *MuRF3* was present to some extent in all the examined tissues, but at relatively low levels outside striated muscles (Fig. 5A).

In adult zebrafish, we detected transcripts for all *MuRF* genes except *MuRF3*, observing predominant striated muscle expression in each case, including for *MuRF4* (Fig. 6B). However, low levels of *MuRF* family member transcripts were also detected in zebrafish tissues outside striated muscle (Fig. 6B). While we failed to detect *MuRF3*, a previous study used qPCR to quantify *MuRF3* transcript expression in zebrafish tissues, revealing highest expression in skeletal muscle [17].

These data confirm that zebrafish *MURF1* retrogenes have expression consistent with striated-muscle specific functions. Past studies with zebrafish [46-47] and other teleost species [see introduction] have also clearly shown that *MuRF1* retrogenes are highly transcriptionally-responsive to conditions favouring muscle atrophy. Therefore, teleost *MuRF1* retrogenes and intron-containing orthologues from mammals evidently conserve similar roles. This in turn suggests that a *MuRF1* retrogene 'perfectly' replaced the function of an ancestral *MuRF1* gene during teleost evolution. This is notable, considering that when a retrogene is inserted into a genome, it will normally lack regulatory sequences required for transcription (e.g. a promoter) [48], let alone sequence elements regulating transcription under different biological contexts. Transcriptional regulation of mammalian *MuRF1* in response to atrophy depends on promoter transcription factor binding sites for NF-κB, FoxO and GR transcription factors, which may act separately or in synergy depending on the context [49-50]. Such inherent complexity in the *MuRF1* promoter makes it unlikely that a *MuRF1* retrogene instantly replaced the expression of the ancestral *MuRF1* gene. Therefore, teleost *MuRF1* may provide a useful model system to study evolutionary mechanisms that lead to complete functional replacement of genes by retrogenes, a topic that was recently discussed elsewhere [43].

3.8. Conclusions

This study establishes that the vertebrate MuRF family is comprised of four genes that were present in a basal ancestor to jawed vertebrates. The results demonstrate that *MuRF* genes, including *MuRF4* in zebrafish, are predominantly transcribed in heart and skeletal muscle in distant vertebrate taxa. This suggests that the main ancestral role of MuRFs was in striated muscle. Equally, our data accommodates the possibility that *MuRF* functions within striated muscles can be secondarily lost, at least at certain life stages, as observed for *MuRF4* in adult lizards. The observation of *MuRF* expression outside striated muscles in both zebrafish and lizard, albeit at a relatively low level, also points to the possibility of functions outside muscle. Finally, future work might consider further characterizing the roles of *MuRF4*, a gene that was dispensable in placental mammal and bird evolution, but has otherwise been maintained in vertebrates.

Acknowledgements

This work was supported by a visiting fellowship funded by the Marine Alliance for Science and Technology for Scotland (grant number: VF20) and by start-up funds to D.J.M from the University of Aberdeen. E.N.F. also receives support from a FONDAP project (grant number: 15110027) granted by CONICYT-Chile. Samples used for transcript expression studies were a gift from Professor Ian A. Johnston (University of St Andrews, UK).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version.

References

[1] Meroni, G. and Diez-Roux, G. (2005) TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. Bioessays. 27, 1147-1157.

[2] Mayans, O. and Labeit, S. (2012) MuRFs: specialized members of the TRIM/RBCC family with roles in the regulation of the trophic state of muscle and its metabolism. Adv. Exp. Med. Biol. 770, 119-129.

[3] Spencer, J. A., Eliazer, S., Ilaria, R. L. Jr, Richardson, J. A. and Olson, E. N. (2000) Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein. J. Cell Biol. 150, 771-784.

[4] Centner, T. et al. (2001) Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain. J. Mol. Biol. 306, 717-726.

[5] Tacchi, L., Bickerdike, R., Secombes, C. J and Martin, S. A. (2012) Muscle-specific RING finger (MuRF) cDNAs in Atlantic salmon (Salmo salar) and their role as regulators of muscle protein degradation. Mar. Biotechnol. (NY). 14, 35-45.

[6] Wang, J., Salem, M., Qi, N., Kenney, P. B., Rexroad, C. E. 3rd and Yao, J. (2011) Molecular characterization of the MuRF genes in rainbow trout: Potential role in muscle degradation. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 158, 208-215.

[7] Glass, D.J. (2005) Skeletal muscle hypertrophy and atrophy signaling pathways. Int. J. Biochem. Cell. Biol. 37, 1974-1984.

[8] Bodine, S.C. et al. (2001) Identification of ubiquitin ligases required for skeletal muscle atrophy. Science. 294, 1704-1708.

[9] Witt, C. C., Witt, S. H., Lerche, S., Labeit, D., Back, W. and Labeit, S. (2008) Cooperative control of striated muscle mass and metabolism by MuRF1 and MuRF2. EMBO J. 27, 350-360.

[10] Fielitz, J. et al. (2007). Loss of muscle-specific RING-finger 3 predisposes the heart to cardiac rupture after myocardial infarction. Proc. Natl. Acad. Sci. USA. 104, 4377-4382.

[11] Kedar, V., McDonough, H., Arya, R., Li, H.H., Rockman, H.A. and Patterson, C. (2004). Muscle-specific RING finger 1 is a bona fide ubiquitin ligase that degrades cardiac troponin I. Proc. Natl. Acad. Sci. USA. 101, 18135-18140.

[12] Clarke, B. A. et al. (2007). The E3 Ligase MuRF1 degrades myosin heavy chain protein in dexamethasone-treated skeletal muscle. Cell Metab. 6, 376-385.

[13] Cohen, S., Brault, J. J., Gygi, S.P., Glass, D. J., Valenzuela, D. M., Gartner, C., Latres, E. and Goldberg, A.L. (2009) During muscle atrophy, thick, but not thin, filament components are degraded by MuRF1-dependent ubiquitylation. J. Cell Biol. 185, 1083-1095.

[14] Witt, S. H., Granzier, H., Witt, C. C. and Labeit, S. (2005) MURF-1 and MURF-2 target a specific subset of myofibrillar proteins redundantly: towards understanding MURF-dependent muscle ubiquitination. J. Mol. Biol. 350, 713-722.

[15] Macqueen, D. J., Kristjánsson, B. K., Paxton, C. G., Vieira, V. L. and Johnston I. A. (2011) The parallel evolution of dwarfism in Arctic charr is accompanied by adaptive divergence in mTOR-pathway gene expression. Mol. Ecol. 20, 3167-3184.

[16] Fuentes, E. N., Ruiz, P., Valdes, J. A. and Molina, A. (2012) Catabolic signaling pathways, atrogenes, and ubiquitinated proteins are regulated by the nutritional status in the muscle of the fine flounder. PLoS One. 7, e44256.

[17] Boudinot, P., van der Aa, L.M., Jouneau, L., Du Pasquier, L., Pontarotti, P., Briolat, V., Benmansour, A. and Levraud, J.P. 2011. Origin and evolution of TRIM proteins: new insights from the complete TRIM repertoire of zebrafish and pufferfish. PLoS One. 6, e22022.

[18] Johnston, I.A., Bower, N.I. and Macqueen, D.J. (2011) Growth and the regulation of myotomal muscle mass in teleost fish. J. Exp. Biol. 214, 1617-1628.

[19] Johnston, I.A, Macqueen, D.J. and Watabe, S. (2008) Molecular biotechnology of development and growth in fish muscle. In: Fisheries for Global Welfare and Environment. Memorial book of the 5th World Fisheries Congress 2008. Edited by: K. Tsukamoto, T. Kawamura, T. Takeuchi, T. D. Beard, Jr and M. J. Kaiser. Terrapub. pp 241-262.

[20] Cleveland, B.M. and Weber, G.M. (2011) Effects of sex steroids on indices of protein turnover in rainbow trout (*Oncorhynchus mykiss*) white muscle. Gen. Comp. Endocrinol. 174, 132-142.

[21] Macqueen, D.J. and Wilcox, A.H. (2014) Characterization of the definitive classical calpain family of vertebrates using phylogenetic, evolutionary and expression analyses. Open Biol. 4, 130219.

[22] Vilella, A.J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R. and Birney E. (2009) EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. Genome Res. 19, 327–335.

[23] Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

[24] Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

[25] Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D.and Pupko, T. (2010) GUIDANCE: a web server for assessing alignment confidence scores. Nucleic Acids Res. 38, W23–W28.

[26] Penn, O., Privman, E., Landan, G., Graur, D. and Pupko T. (2010). An alignment confidence score capturing robustness to guide tree uncertainty. Mol. Biol. Evol. 27, 1759–1767.

[27] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

[28] Jones, D. T., Taylor, W. R. and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

[29] Drummond, A. J., Suchard, M. A., Xie, D. and Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

[30] Drummond, A.J., Ho, S.Y., Phillips, M.J. and Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

[31] Gernhard, T. (2008) The conditioned reconstructed process. J. Theor. Biol. 253, 769–778.

[32] Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696-704.

[33] Dereeper, A., et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465-W469.

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

[35] Macqueen, D. J., Delbridge, M. L., Manthri, S. and Johnston, I. A. (2010) A newly classified vertebrate calpain protease, directly ancestral to CAPN1 and 2, episodically evolved a restricted physiological function in placental mammals. Mol. Biol. Evol. 27, 1886-1902.

[36] Macqueen, D. J., Garcia de la Serrana, D., Johnston, I. A. (2013) Evolution of ancient functions in the vertebrate insulin-like growth factor system uncovered by study of duplicated salmonid fish genomes. Mol. Biol. Evol. 30, 1060-1076.

[37] Jaillon, O. (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature. 431, 946-57.

[38]. Macqueen, D. J. and Johnston, I. A. 2014. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. Proc. R. Soc. B. 281, 20132881

[39] Van de Peer, Y., Maere, S. and Meyer, A. (2009) The evolutionary significance of ancient genome duplications. Nat. Rev. Genet. 10, 725-732.

[40] Macqueen, D. J., Johnston, I. A. (2006) A novel salmonid myoD gene is distinctly regulated during development and probably arose by duplication after the genome tetraploidization. FEBS Lett. 580, 4996-5002.

[41] Macqueen, D. J. and Johnston, I. A. (2008) An update on MyoD evolution in teleosts and a proposed consensus nomenclature to accommodate the tetraploidization of different vertebrate genomes. PLoS One. 3, e1567.

[42] Weiner, A.M., Deininger, P.L. and Efstratiadis, A. (1986) Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. 55, 631-661.

[43] Ciomborowska, J., Rosikiewicz, W., Szklarczyk, D., Makałowski, W. and Makałowska, I. (2013) “Orphan” Retrogenes in the Human Genome. Mol. Biol. Evol. 30, 384-396.

[44] Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M. and Smith, W.L. (2012) Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. USA. 109, 13698-13703.

[45] Benton, M.J. and Donoghue, P.C. (2007) Paleontological evidence to date the tree of life. Mol. Biol. Evol. 24, 26-53.

[46] Amaral, I.P. and Johnston, I.A. 2009. Insulin-like growth factor (IGF) signalling and genome-wide transcriptional regulation in fast muscle of zebrafish following a single-satiating meal. J. Exp. Biol. 214, 2125-2139.

[47] Seiliez, I., Médale, F., Aguirre, P., Larquier, M., Lanneretonne, L., Alami-Durante, H., Panserat, S. and Skiba-Cassy, S. (2013) Postprandial regulation of growth- and metabolism-related factors in zebrafish. Zebrafish. 10, 237-248.

[48] Mighell, A.J., Smith, N.R., Robinson, P.A. and Markham, A.F. (2000) Vertebrate pseudogenes. FEBS Lett. 468, 109-114

[49] Waddell. D. S, Baehr. L. M., van den Brandt. J., Johnsen, S. A., Reichardt, H. M., Furlow, J. D. and Bodine, S.C. (2008) The glucocorticoid receptor and FOXO1 synergistically activate the skeletal muscle atrophy-associated MuRF1 gene. Am. J. Physiol. Endocrinol. Metab. 295, E785-797.

[50] Wu, C.L., Cornwell, E.W., Jackman, R.W. and Kandarian, S.C. (2014) NF-κB but not FoxO sites in the MuRF1 promoter are required for transcriptional activation in disuse muscle atrophy. Am. J. Physiol. Cell Physiol. 306, C762-C767.

Figures and Legends

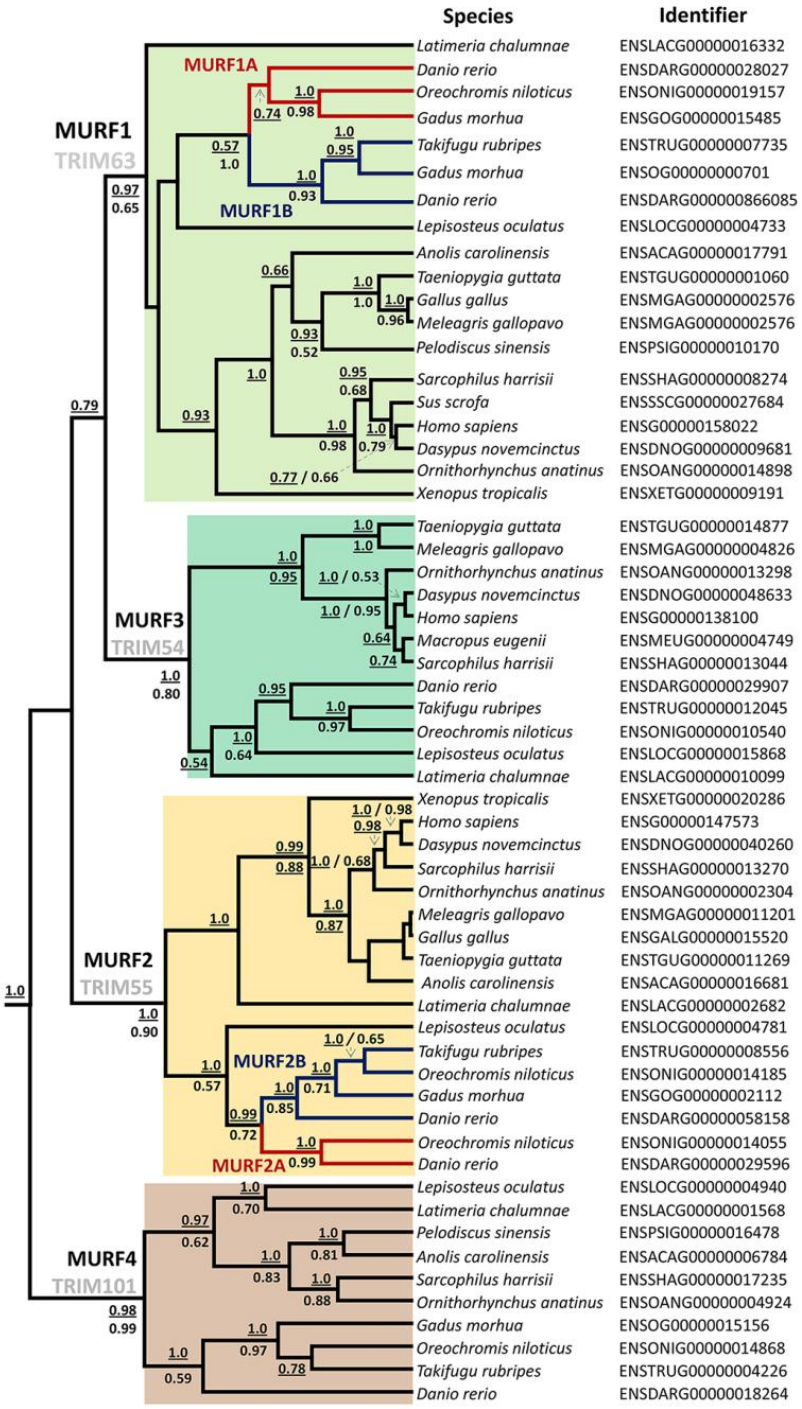
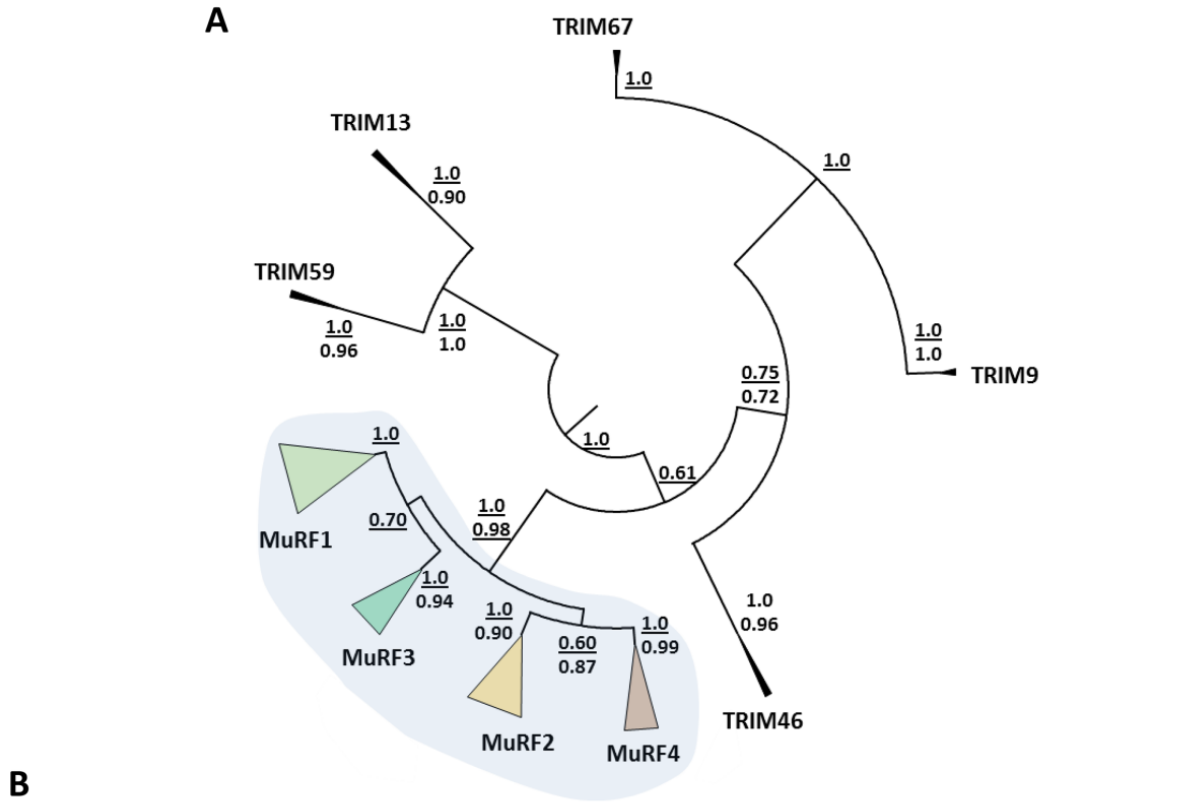


Fig. 1. Bayesian phylogenetic tree of fifty-eight MuRF sequences spanning vertebrate evolution. The analysis was based on a high-confidence alignment of 290 AA sites (Fig. S1A). The length of branches is proportionate to an uncalibrated timescale. Posterior probability branch support values from the Bayesian analysis, along with proportionate bootstrap support values from a supporting ML analysis are provided at each node (given as underlined and non-underlined numbers, respectively; values greater than 0.5 shown). Monophyletic clades that support jawed-vertebrate wide MuRF family members are shaded in different colours. The position of teleost-specific MuRF duplications on the tree are highlighted with black stars.



B

```
Human       MuRF1      EKQLICPICLEMFTKPVVILPCQHNLCRKCANDIFQAANPYWTSRGSSVSMSGGRFRCPTCRHEVIMDRHGVYGLQRNLLVENIIDIYKQ
Coelocanth  MuRF1      ................................V........Q...T.I--........S......L.....................E
Zebrafish   MuRF1A     ....S.....D.................S..SELYDSK...--------HF...I.......F..VL.....................
Zebrafish   MuRF1B     ....S.....D.................G..S.LYDSR...--------RF...V.......F..VL.....H...............
Human       MuRF2      ................................S.....S...LPT..GTTMA........S.....VL.....................
Coelocanth  MuRF2      .....................................S...LPT..GTTVA........S......L.....................
Zebrafish   MuRF2A     .R.........I.........................S...LPT..G.SLG........S.....VL..................M...
Zebrafish   MuRF2B     ................................SS...LPI..G-SVT........S.....VL..................M...
Human       MuRF3      ............S...................V...S..L.Q....TTVS........S.....VL.....................
Coelocanth  MurF3      .....................................S..L.Q...T.TMS........S.....VL.....................
Zebrafish   MuRF3      .R.........I.S.....................V...S..L.QN.A.TTVST.......S.....VL.....................S
Tas. devil  MuRF4      .R...........S.....................SS.RCLKTLPGTTLG.S......S.....VL.....................
Coelocanth  MuRF4      .R.........I.......................L..GC.V.CESRGATLG........S.....VL.....................
Zebrafish   MuRF4      .........MDV..........L.........ELY.---.TL-----FQVGIE.S....S..R..VLG....F.............V...
                       * ** ****   * ******** ***** **        *             * **** ** **   **** ************ ****

                                                                RING-finger domain                MuRF-family specific box

Human       TRIM46     ..E.L..V.Q..YKQ.L.-...T..V.QA..REVLPSSE.TSPASTP.TRPQVIM.P..A.QGD.ELGER.LA..F...TL.RVVER.R.
Human       TRIM9      .EE.K..V.GSFYRE.-I....S....QA..RN.LESES.---------QPRNSCIT..Q.HRSL.L.DR.LR.FPK.RVL.GV..R.Q.
Human       TRIM67     .EE.K..V.GSL.RE.-I....S..V.LP..RT.ADGEQ.LLL....GLQS.SSSIT..Q.HRSASL.HR.LR.F...R.L.A.VQR.Q.
                       *  * ** *          *** ** *  **        *                 ** *         *  *   *   *     *
```

Fig. 2. (**A**) Circular cladogram depicting Bayesian/supporting ML phylogenetic analyses of sixty-eight sequences including the MuRFs presented in Fig. 1, along with additional out-group sequences from the TRIM/RBCC family. The analysis was based on a high-confidence alignment of 208 AA sites (Fig. S1B). Other details are as described in the Fig. 1 legend. (**B**) Sequence alignment highlighting the conservation of the RING and MFC domains in MuRF proteins with respect to the next most-related proteins of the TRIM/RBCC superfamily.

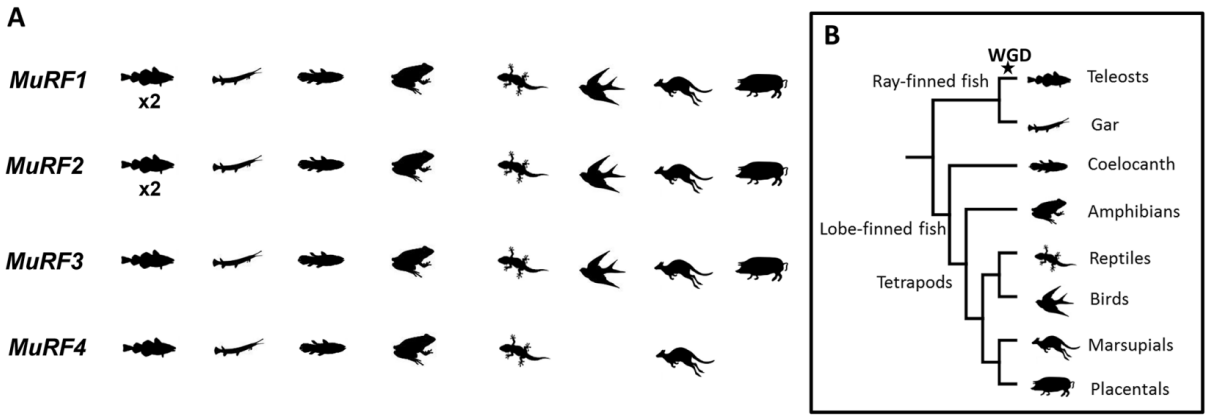


Fig. 3. (**A**) Diagram summarizing the evolutionary conservation of different MuRF family members in major vertebrate lineages according to the results of this study. (**B**) Key for silhouette diagrams highlighting vertebrate lineages in part **A**. Established evolutionary relationships are shown in the form of a cladogram that is not scaled.

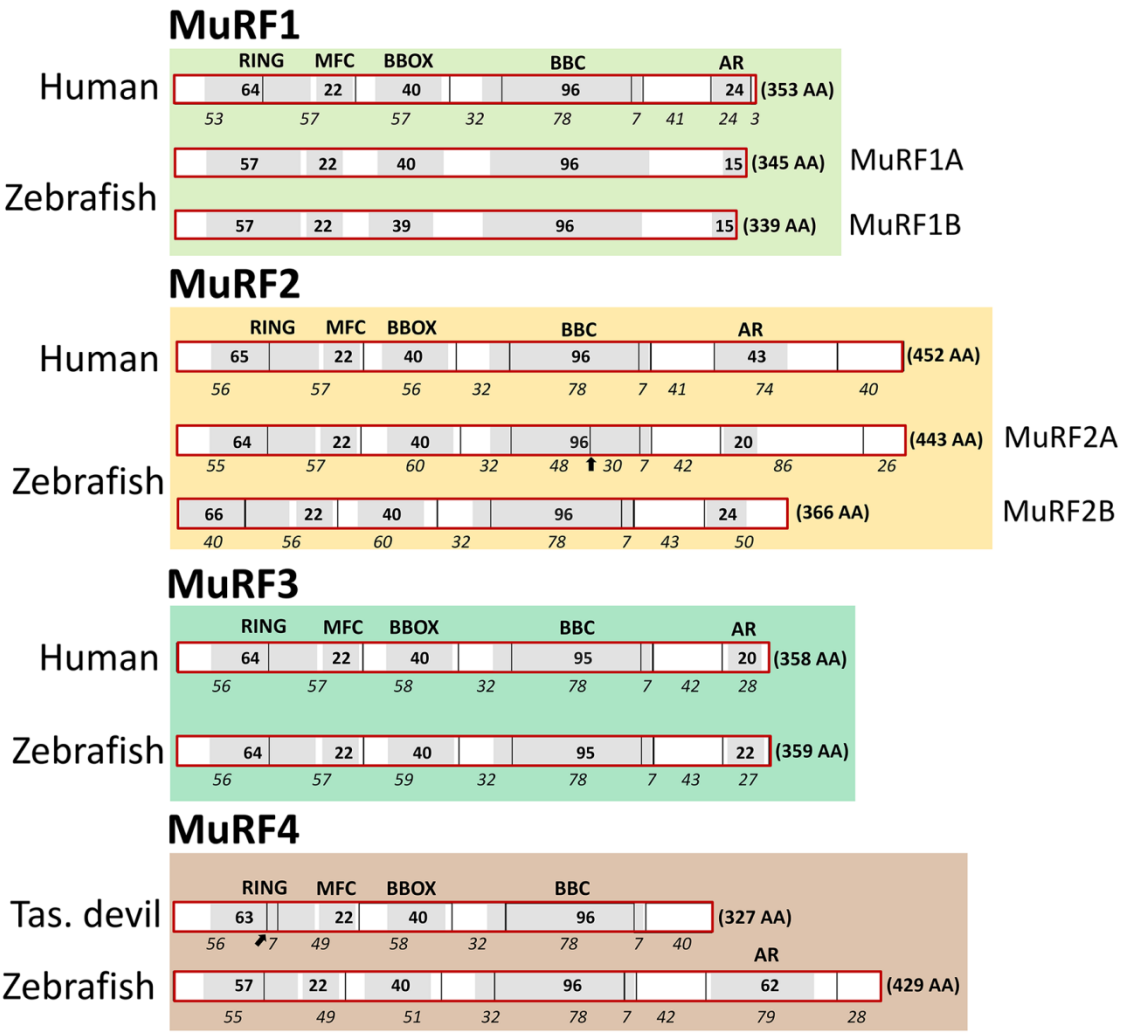


Fig. 4. Protein domain organization of the vertebrate MuRF family for representative mammalian and teleost species, presented with respect to gene structure. Exon boundaries are shown as solid black vertical lines and conserved MuRF domains are shaded grey and indicated by text above mammalian orthologues of each MuRF family member. Also shown for each MuRF protein is the number of AAs comprising different exons and MuRF domains (numbers in italic and bold text, respectively)

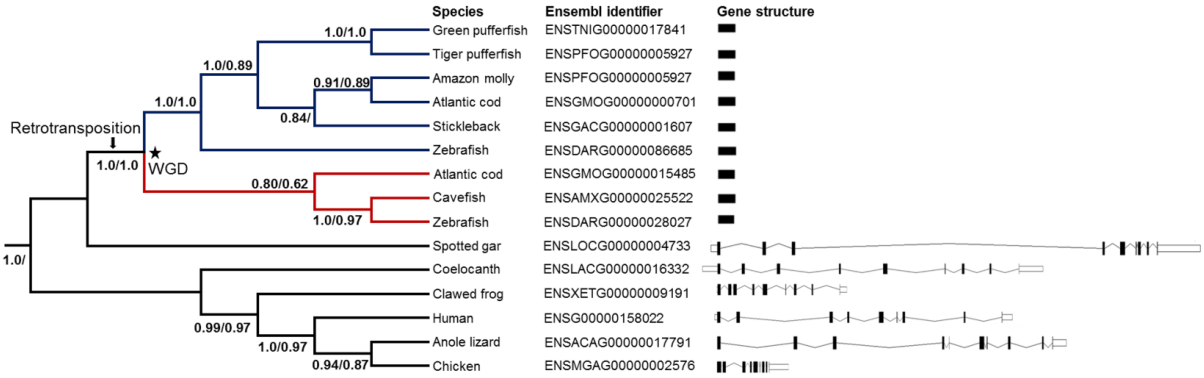


Fig. 5. Evidence for an ancestral retrotransposition of *MuRF1* during teleost evolution. On the left side of the figure, an empirical Bayesian/supporting ML phylogenetic tree is shown built from fifteen MuRF1 sequences, including a range of teleost species not included in Fig. 1 or 2. This tree was based on a high-confidence alignment of 345 AA sites (Fig. S1C). Other details about the phylogenetic analysis are as described in the Fig. 1 legend. On the right side of the figure, *MuRF1* gene structures are shown to scale, including protein coding exons (solid black rectangle), untranslated exons (unfilled rectangles) and introns (lines between exons). Many teleost *MuRF1* genes are incorrectly predicted to have one or a small number of short introns in Ensembl (example provided in Fig. S1D). However, these predictions are spurious, because the associated protein models lack a complete MuRF domain structure (i.e. as described Fig. 4; see Fig. S1D). Conversely, if the genomic region containing *MuRF1* genes is translated as a single ORF (as predicted correctly in some teleost species) the resultant proteins contain all recognized MuRF domains (see Fig. S1D)

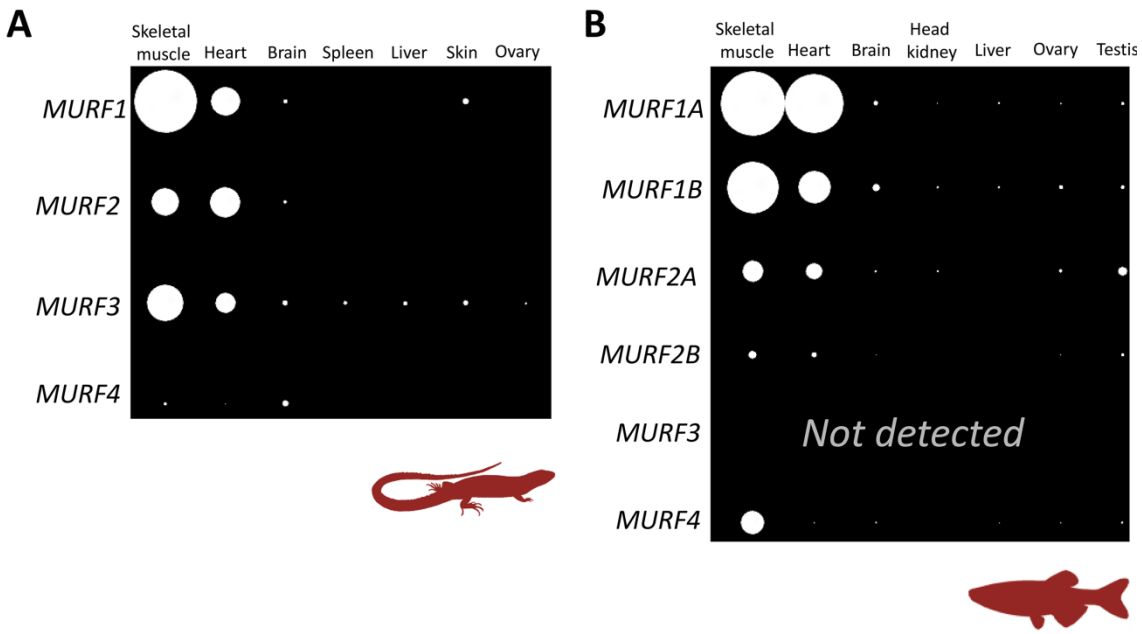


Fig. 6. qPCR estimated transcript expression of *MuRF* gene repetoires across tissues of (**A**) adult anole lizards, and (**B**) adult zebrafish. White bubbles are scaled to show relative transcript levels within each species, normalized to the reference gene *rps13*.