Repeated evolutionary turnover of vertebrate skeletal muscle myosins

Christina M. Harvey[1,*]
Eric R. Schuppe[2]
Michael S. Brainard[2]
Matthew J. Fuxjager[3]
James B. Pease[4,*]

[1]Department of Biology, Wake Forest University, 1834 Wake Forest Road, Winston-Salem, North Carolina, 27109, USA

[2]Howard Hughes Medical Institute and Center for Integrative Neuroscience, University of California San Francisco, San Francisco, California, 94143, USA

[3]Department of Ecology, Evolution, and Organismal Biology, Brown University, Providence, Rhode Island, 02912, USA

[4]Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, Ohio, 43210, USA

[*]Corresponding Authors: keatcm21@wfu.edu, pease.25@osu.edu

Short Title: Evolution of vertebrate skeletal muscle myosins

Keywords: myosin, vertebrates, evolution, molecular evolution, skeletal muscle, gene family evolution, RNA-Seq, protein structure

Abstract

Myosin heavy chain proteins are essential for muscle contraction and nearly every physiological function in animals, but their diversity and evolution outside mammals is largely unknown. Here, we analyze over 1,000 myosin protein sequences and dozens of skeletal muscle expression profiles to comprehensively model the evolutionary history of heavy-chain myosins and understand common themes of their tissue-specific expression. We show that even though skeletal muscle myosins are located in the same tandem gene cluster across vertebrate species, repeated gene duplication-loss turnover has resulted in each major vertebrate group now having an independently evolved set of core skeletal muscle myosins. Despite these separate derivations of skeletal myosins, each major vertebrate group exhibits tissue-specific patterns of expression, includes diverse myosin molecular properties, and employs different myosins in extreme muscles. Our results suggest that muscle evolution across vertebrates involves a repeated and multifaceted process of gene family diversification of sarcomeric myosin proteins, with shifting expression of myosins in different muscle types as a general theme of their evolution across vertebrates.

Physical movement is fundamental for animal survival and reproduction, helping individuals find food, evade predators, mate successfully, and parent offspring. Myosin heavy chain (MYH) proteins actuate much of this movement by endowing skeletal muscles with contractile properties. Physiological studies have repeatedly suggested that MYH proteins are conserved across taxa suggesting minimal evolutionary change in these core molecules (Goodson, et al. 1999, Weiss, et al. 1999a, Weiss et al. 1999b, Schiaffino & Reggiani 2011, Lee, et al. 2019). In contrast, molecular genetic studies have found extraordinary diversity in MYH proteins broadly, and have long hinted at a more complex underlying process for the adaptive evolution of skeletal muscle MYHs involving repeated evolution of distinct sets of MYHs (Moore, et al. 1992, Moore, et al. 1993, McGuigan, et al. 2004, Foth, et al. 2006, Ikeda, et al. 2007, Sebé-Pedrós, et al. 2014, Wagoner & Dill 2021, Schiaffino et al. 2024a, Schiaffino, et al. 2024b). The fundamental tension between diametric perspectives of genetic conservation and repeated divergence can only be resolved by comparing MYH gene sequences and expression across a large group of unrelated taxa. Such work promises to reveal how MYH proteins have diversified historically, how they contribute to contemporary performance traits, and the central question of whether myosin evolution is dominated by slow conservative evolution or repeated divergence and innovation.

Vertebrate skeletal muscle MYH proteins pull against actin filaments to generate a power stroke that creates force by shortening or lengthening the muscle cell. MYH proteins initiate this process by hydrolyzing ATP into ADP to power actin-myosin binding kinetics (Rayment et al. 1993, Korn 2000, Piazzesi, et al. 2002, Piazzesi, et al. 2007). Even subtle amino acid sequence variation can confer major alterations to MYH molecular properties (Larsson & Moss 1993, He, et al. 2000, Resnicow, et al. 2010, Johnson, et al. 2021). The presence of impactful variation creates ample opportunity for selection to drive evolutionary changes in muscle performance.

Mammals have a set of sarcomeric MYH proteins known to have distinct expression patterns and molecular properties that have contributed to the past categorization of “fast” and “slow” contractile speeds (Bárány 1967, Whalen et al. 1981, Ennion, et al. 1995, Schiaffino & Reggiani 2011). MYH3 and MYH8 are primarily expressed in embryonic and neonatal tissues, with only minimal expression in some specialized adult muscles (Schiaffino et al. 2015, Mascarello et al. 2016). Adult skeletal muscles typically express MYH1, MYH2, MYH4 that are exclusively expressed in skeletal muscles, along with MYH7 that has roles in both cardiac and skeletal muscles (Schiaffino et al. 2015). Larger mammals tend to use a combination of MYH1, MYH2, and MYH7 in their skeletal muscles (Fig. 1A), but smaller-bodied rodents have been shown to primarily express MYH4 (Lucas et al. 2000, Hoh 2023). The molecular properties of these MYH proteins have been shown to be distinct with regard to ATP efficiency and mechanical movement, which connects the different MYH molecular properties to the performance properties of the muscles (Resnicow et al. 2009).

Current conventional wisdom holds that other vertebrate genomes contain MYHs that are orthologous in origin and function to those observed in mammalian genomes. This belief most likely stems from the observation that distinct muscle fiber types with notable performance differences have been observed in a variety of nonmammals (Bonine et al. 2005, Luna et al. 2015). For example, “superfast” muscles in vertebrates support this idea, where extremely rapid contraction-relaxation speeds are associated with distinct MYH compositions (Mead et al. 2017, Schiaffino et al. 2024b). The presence of distinct MYHs in these tissues has been noted, though their genomic origins and relatedness to those observed in

mammalian fibers have not been assessed in detail (Listrat et al. 2016, Robinson et al. 2021). MYHs across vertebrates may be genetically and functionally distinct in ways that have yet to be fully described (Fig. 1A; McGuigan et al. 2004). This suggests a different relationship between physiological evolution of vertebrate muscles and the molecular evolution of MYHs where evolution of individual MYH sequences and evolution of the relative expression of MYHs within muscles are joint, interrelated processes.

Here, we conduct a comprehensive study of the origins and expression of MYH genes across chordates by analyzing over a thousand MYH sequences and dozens of skeletal muscle expression profiles. We use our results to test whether MYH evolution is more generally conservative or diversifying. We find that vertebrate genomes encode sets of MYH proteins that are more diverse in gene number, genomic arrangement, and molecular structure than has been described previously. Our results conclusively demonstrate that primary skeletal MYH proteins in each major group of vertebrates evolved separately with respect to each other and lack one-to-one evolutionary relationships. These separate MYH diversifications precludes the possibility of homologous functional roles and suggests convergent pressures where similar muscle properties and functions occur in separate groups. Finally, the tissue-specific expression of MYHs in both ordinary and extreme-performance muscles points to a strong and generalized connection across vertebrates between MYH gene functional properties and muscle performance. Most broadly, this expanded evolutionary history of sarcomeric MYHs shows that skeletal myosins show lability despite being fundamental to animal reproduction and fitness, demonstrating a dramatic example of how core genes simultaneously maintain homeostasis while adapting new innovations.

Core skeletal muscle MYH proteins are not one-to-one related among vertebrate classes.

To create a comprehensive model of sarcomeric vertebrate MYH evolution, we modeled the molecular relationships of 1137 sarcomeric heavy-chain myosin (MYH) amino acid sequences from 113 species from all classes and most major orders of chordates (Fig. 1B). Analysis of the evolutionary change in *MYH* sequences and genomic locations reveals broader patterns of vertebrate evolution and many specific patterns within vertebrate subgroups (see Supplementary Text). We find that all vertebrate sarcomeric MYH proteins are related to a single common ancestor, with cytoskeletal *MYH11* as an outgroup. Across vertebrates, *MYH* gene copy numbers and genomic arrangements are far more variable than have been reported previously (Figs. 1B and 1C).

Present across all vertebrates is a tandem cluster of 1–12 *MYH* genes, which has maintained the same syntenic context throughout vertebrate evolution. This cluster generally includes the core skeletal muscle *MYH* genes in each major group of vertebrates (Fig. 1C). While phylogenetic, pairwise distance, and MDS analyses show that embryonic MYH3 and specialized/amphibian MYH13 are related across major groups, core skeletal MYHs do not have one-to-one relationships between any two major vertebrate groups (Fig 1B; Supplementary Figs. S1 and S2). Instead, each group's present core MYHs descend from a group-specific common ancestor (Figs. 1B, 1D, and 2). As a consequence, each major group of vertebrates has a set of *MYH* proteins that have evolved and diversified separately. Therefore, each vertebrate group's set of MYHs would change in response to their separate evolutionary pressures. Within each group, we also identified clearly conserved and previously unidentified *MYH* subtypes. This discovery prompted us to establish a new classification for non-mammalian subtypes *MYH20–MYH60* to

promote discussion of their distinct functional roles and evolutionary histories (Fig. 1B, 1C; Tables S1 and S2).

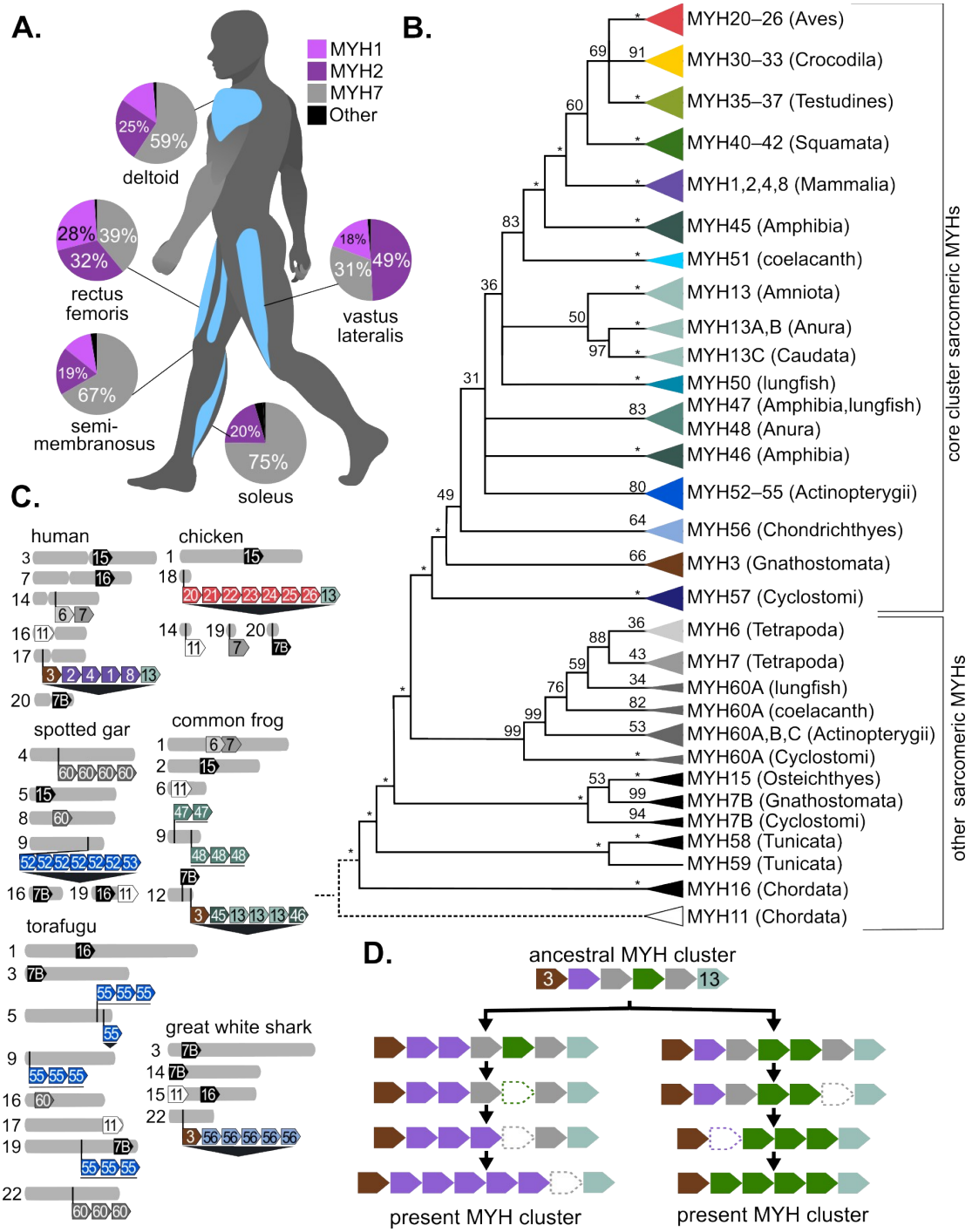

Figure 1. Evolution of skeletal muscle MYH genes and subfamilies. (A) MYH average relative expression differs in human large skeletal muscles (Art credit: C. Harvey). (B) Cladogram of MYH subfamily phylogenetic relationships based on a phylogeny of 1137 sarcomeric *MYH* from vertebrates and tunicates, showing the non-orthologous relationships of core skeletal MYHs across major vertebrate groups. Non-parametric bootstraps are shown (* for 100). Nodes are collapsed when non-parametric

122 bootstrap < 30. MYH symbols numbered *MYH1–MYH16* follow established mammalian nomenclature,
123 while numbers *MYH20–60* are new non-mammalian symbols established by this study. Each major
124 group's core cluster skeletal myosin subfamilies are indicated by different color groups. Other sarcomeric
125 myosins shared across groups are grayscale. For additional details on the myosin labels, see the
126 Supplementary Text. (C) Chromosomal positions of sarcomeric *MYH* genes emphasizing that each
127 species has a core tandem cluster skeletal muscle myosins (closed triangle underline) and sometimes
128 other derived tandem clusters (open triangle underline) in human, chicken (*Gallus gallus*), common frog
129 (*Rana temporaria*), spotted gar (*Lepisosteus oculatus*), torafugu (*Takifugu rubripes*), and great white
130 shark (*Carcharodon carcharias*). *MYH* subfamily numbers are shown inside chevrons with colors
131 following Figure 1B. (D) Diagram of the proposed model of evolution of the *MYH* tandem cluster through
132 gene duplications and losses via unequal crossing over and pseudogenization.

133 Core skeletal cluster *MYH* genes apparently evolve via tandem gene gain and loss, where each group's
134 present MYH proteins are the result of gene duplications and deletions that all occurred after divergence
135 from other groups (Figs. 1D and 2). The core skeletal cluster is a gene cassette consisting of an
136 uninterrupted tandem array of genes that are all encoded on the same strand and by the same number of
137 exons. We frequently observed instances of recent duplicates (>99% similar) and inactivated and partial
138 pseudogenes. The appearance of these features in several independent lineages indicates repeated
139 instances of gene gain and loss via unequal crossing over as the primary genetic mechanism. We further
140 observed sequence homogenization through gene conversion, biased nucleotide content shifts, and
141 intergenic recombination, but the evidence indicates these are lesser and rarer factors in driving cluster
142 evolution in the long term (Supplementary Figs. S3 and S4, Supplementary Text). Amphibians, lungfish,
143 and many ray-finned fishes have additional translocated individual MYHs or tandem clusters whose
144 sequence relationships indicate they are derived from the core skeletal cluster (*MYH47, MYH48, MYH55*
145 subtypes). These other tandem clusters also show expansion, contraction, and pseudogenization that
146 indicates evolution under the same process as the core cluster. We conclude that this core skeletal MYH
147 cluster as a whole has persisted throughout vertebrate history as a conserved gene cassette, but the
148 specific individual genes in each cluster have gradually been replaced through duplication and loss.
149 Additionally, translocated duplicates have established alternate clusters in certain vertebrate groups.

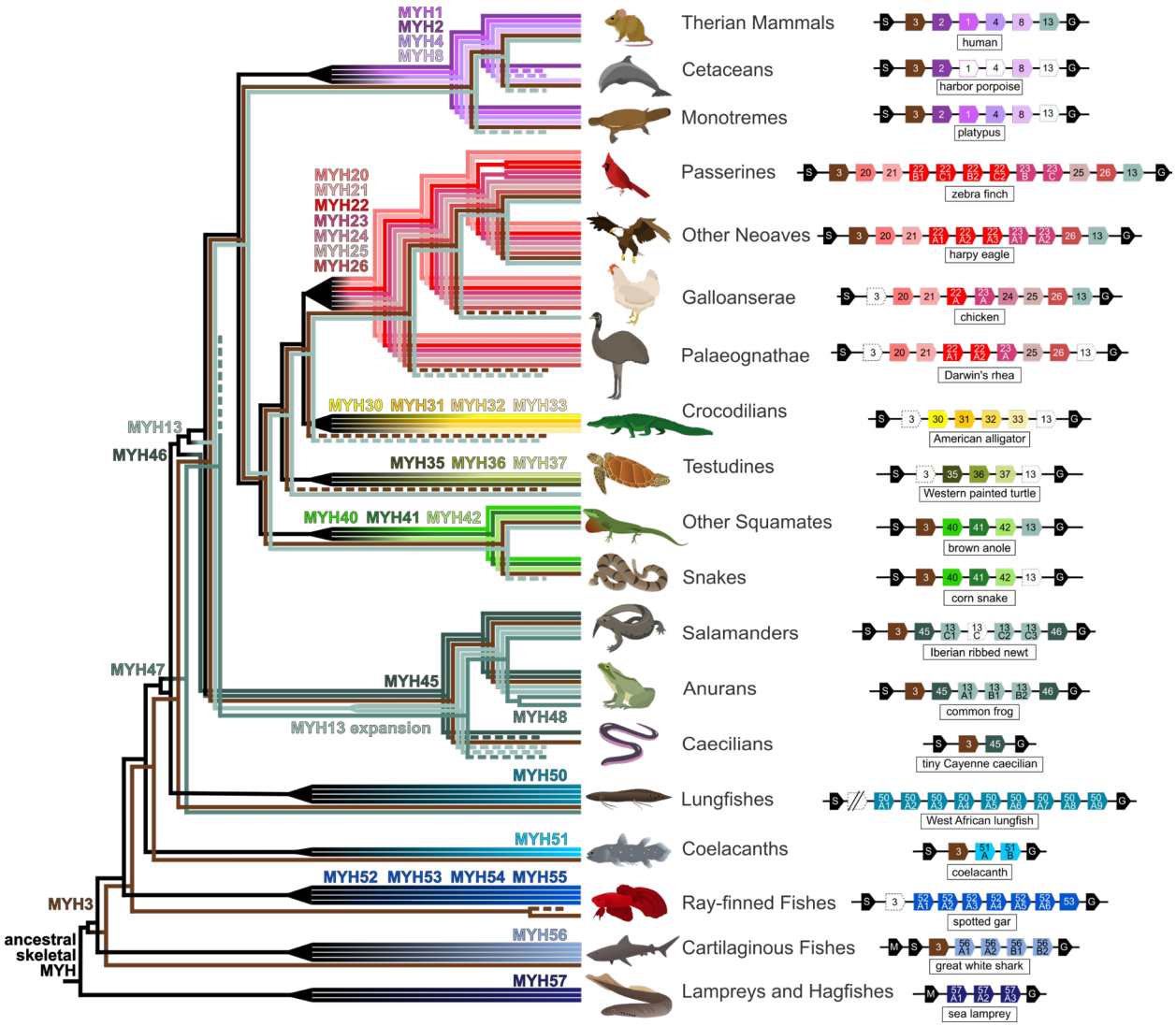

Figure 2. Core skeletal muscle MYH proteins are not one-to-one related among vertebrate classes. Vertebrates show repeated evidence of group-specific myosin subfamily expansion (black triangles) in their evolutionary history (left) and variation in the number of *MYH* genes in their core skeletal muscle MYH cluster (right). Specialized MYH proteins *MYH3* and *MYH13* are related across classes, but pseudogenization through mutational inactivation or partial deletion is also common (dotted outlines). Gene gain-loss in the core cluster occurs in a syntenically conserved locus near *GAS7* (G) and *SCO1* (S) or *MAP24K2* (M). Lungfish is not contiguously assembled and contains *MYH3* fragments on two scaffolds. (Art credit for all drawings: C. Harvey).

Vertebrates express their separate MYH protein sets in muscle-specific combinations. Expression patterns of sarcomeric MYHs have largely been studied in mammals, where correlations between MYH expression and muscle performance are commonly observed (Schiaffino & Reggiani 2011). However, the findings described above that other vertebrate groups have independently evolved sets of MYHs provokes the question: do other vertebrate groups exhibit patterns of muscle-specific expression of their various MYH protein sets? Knowledge of the individual and collective patterns of expression are a window into the potential performance properties of muscles in the present and the past

evolutionary regimes that have shaped them. To address this, we quantified whole RNA-Seq data sets from diverse vertebrate species where multiple muscles were available (Fig. 3). We selected cases where we could analyze locomotor muscles, specialized muscles, and extreme performance muscles. In mammals, both human and sika deer show expression of *MYH1*, *MYH2*, and *MYH7* in major locomotor skeletal muscles (Figs. 1A and 3A; Supplementary Fig. S5). *MYH* genes were also in the highest 1% of genes by expression level and showed moderate levels of tissue specificity in skeletal muscles. Bat breast muscles expressed *MYH1*, *MYH2*, and the usual cardiac-specific *MYH6*, but the "superfast" laryngeal muscles that generate echolocation signals expressed almost exclusively *MYH4* (Lee, et al. 2018, Usui, et al. 2024). A primary locomotor muscle in the harbor porpoise expressed only *MYH2* out of the core skeletal muscle myosins, corroborating our observation that all cetaceans appear to have pseudogenized *MYH1*, *MYH4*, and *MYH13* (Fig. 3C).

For birds, we focused on two passerine species, zebra finches and golden-collared manakins. Two important muscles that power flight and control wing posture, respectively, showed consistent and differentiated patterns of expression involving multiple MYHs (Fig. 3D, E). In zebra finches, the flight-powering muscle expressed seven of its twelve *MYH* genes at relative expression proportions $> 5\%$, while golden-collared manakins overwhelmingly expressed a single *MYH22*. Notably, this latter species is known for its rapid wing-snap courtship display behavior (Fuxjager, et al. 2016). The wing muscle that actuates this gestural signal showed expression of both *MYH22* and *MYH25*. Interestingly, this same gene (*MYH25*) is also 92% of the *MYH* expression in hummingbird flight-powering muscle (Fig. 3F). The presence of *MYH25* in two separately evolved extreme muscles strongly suggests the possibility of specialization of this *MYH25* protein. Meanwhile, "superfast" syringeal muscle from Bengalese finches showed strong expression of *MYH13*, as expected from previous evidence in zebra finches (Fig. 3G; Mead et al. 2017). However, we also found substantial syringeal expression of *MYH26*, also the most prevalent myosin in hummingbird tongues (Fig. 3F).

Other vertebrate groups show similar expression patterns. In reptiles, Western Diamondback rattlesnakes show distinct *MYH* expression patterns in three different muscle regions in its body, which govern different types of movement from locomotion to tail-shaking (Fig. 3H; Bothe et al. 2024). Across individuals, each segment consistently showed the highest expression of a different *MYH* gene. Notably, the "superfast" caudal rattle-controlling muscles expressed high amounts of *MYH42*, which was absent from the other segments (Martin and Bagby 1973, Bothe et al. 2024). In amphibians, Bornean rock frogs expressed multiple forms of the *MYH13A* subtype in hind limb musculature (Fig. 3I). Amniotes have no more than one copy of *MYH13* that is used primarily in specialized muscles like the larynx and syrinx. However, amphibians have many copies of related protein subtypes MYH13A, MYH13B, and MYH13C. This case study preliminarily confirms that frogs use proteins from the MYH13 subfamily as primary locomotor myosins. Finally, an analysis of yellowfin tuna shows expression of multiple MYH55 proteins in white muscle and several MYH60 proteins from a separate cluster in red muscle (Fig 3E, Supplementary Fig. S6).

Collectively, these expression data cases show that MYH proteins are not interchangeable, but instead show tissue specificity and consistent ratios among locomotor and specialized muscles. Tissue-specific MYH expression is a common theme across major vertebrate groups, which is notable given the independent evolution of their MYH gene sets. Intriguingly, this shifting balance of expression of MYH

genes broadly observed across vertebrates is complementary with recent evidence that cluster MYH genes are collectively activated by a superenhancer described in mice (Dos Santos, et al. 2022). In particular, we further show that various specialized and "superfast" muscles predominantly express different myosins, and not a universal "extreme myosin." The tissue-specific combinations of MYH strongly suggest that these proteins have diversified to have different molecular properties, and then combine in muscles to facilitate various muscle performance properties.

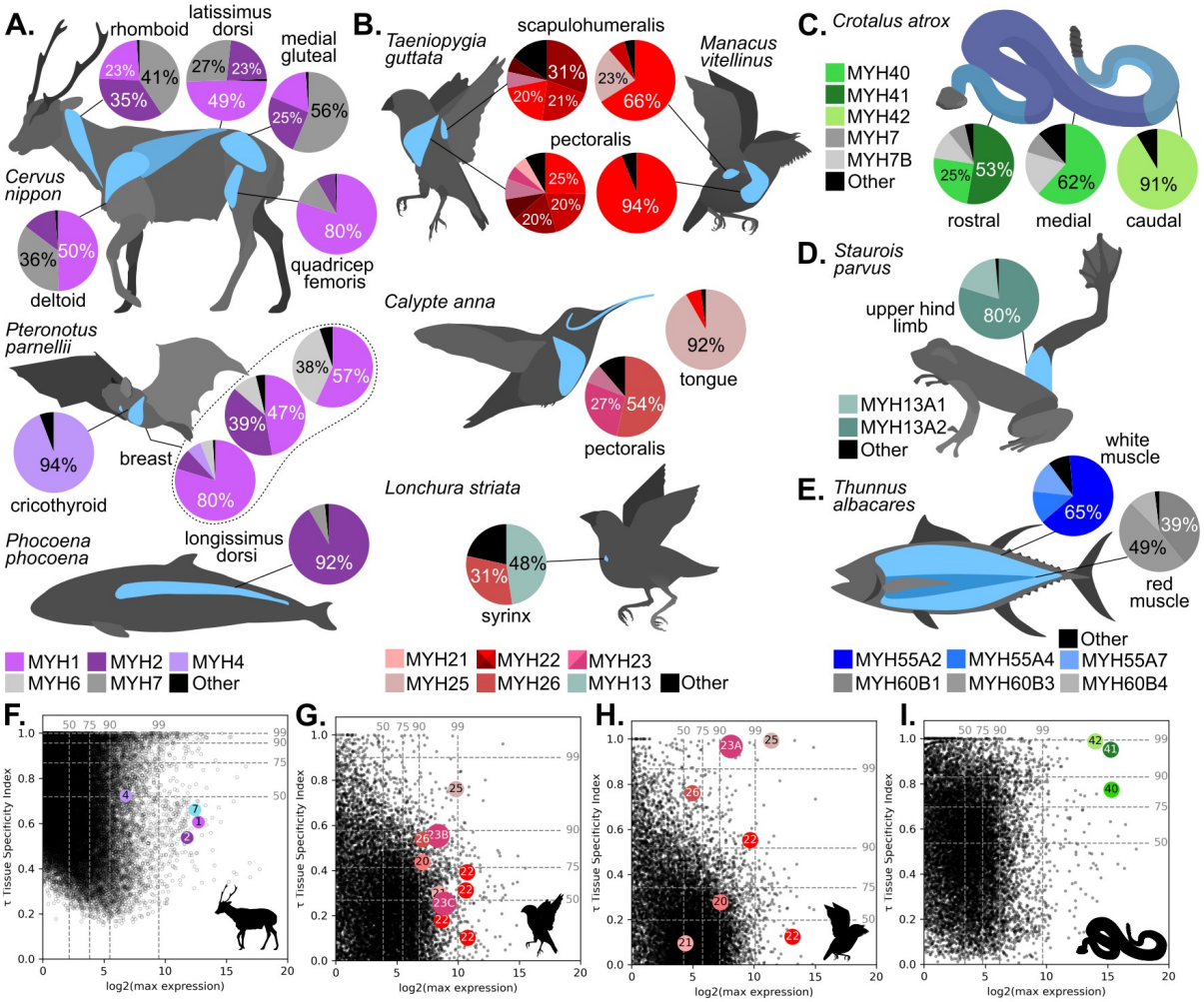

Figure 3. *MYH* genes show variable expression in muscles across vertebrates. The mean proportion of *MYH* expression in each of the highlighted muscles for (A) mammals (sika, Parnell's mustached bat, and harbor porpoise), (B) birds (zebra finch, golden-collared manakin, Anna's hummingbird, and Bengalese finch), (C) Western diamondback rattlesnake, (D) white-spotted foot-flagging frog, and (E) yellowfin tuna. *MYH* genes < 5% relative expression are included as "Other." Breast muscles of Parnell's mustached bat individual samples are depicted with their own pie charts. (F–I) Maximum $\log_2$-expression and tissue specificity index (τ) for all genes with *MYH* genes highlighted as color circles labeled with *MYH* numbers for (F) sika, (G) zebra finch, (H) golden-collared manakin, and (I) Western diamondback rattlesnake. Data distribution percentiles (dashed lines) are shown for each axis. (Art credit for all silhouettes: C. Harvey).

Vertebrate sarcomeric myosin diversity is concentrated in two ATP-interacting loops.
To explore the molecular variation of MYH proteins both within each species and across protein subfamilies, we analyzed sequence variation of core skeletal muscle MYHs, MYH3, MYH6, MYH7, and MYH13 in a structural context. MYH sequences are highly conserved within and among subfamilies across the protein, except two regions in the motor domain (Fig. 4A–C). These variable regions constitute the known loop structures "Loop 1" and "Loop 2," which are intermolecular binding sites for the crucial interaction with ATP. Both loops include several positively and negatively charged residues separated by multiple short amino acid segments of variable length (Fig. 4D). Across them, the overall length of the Loop 1 (11–20aa) and Loop 2 (20–31aa) vary by 2- and 1.5-fold, respectively. This length variation stands in sharp contrast to the rest of the MYH protein, where even a single amino acid insertion or deletions are virtually nonexistent except for minor 2–3aa variation in the termini. Furthermore, we found that each species has a set of MYHs with diverse Loop 1 and 2 lengths (Figs. 4E–H). Even recently duplicated subfamilies exhibit variable loop length, such as MYH22A1–MYH22A4 in pigeon (Fig. 4F). This means that, although each group's set of core skeletal MYHs are separately duplicated from a group-specific ancestral protein, their loop lengths have repeatedly evolved variation across all species and in all vertebrate groups (Fig. 4I).

ATP-binding loop structural variation helps determine MYH contractile properties, largely governing rate-limiting kinetics of ATP and ADP binding (Weiss et al. 1999b, Vilfan and Šarlah 2023). Thus, we hypothesize that the differences we observe in MYH proteins among muscles and species contribute to variation in muscle performance. Supporting this view, we find that muscles with markedly different performance attributes express different ratios and combinations of MYH proteins. Importantly, these data are consistent with a model where the molecular structure of individual MYH genes is, at minimum, not being homogenized in the ATP-binding loops toward a single optimum. This is especially intriguing in the context of MYH proteins expressed in "superfast" muscles. Despite having a relatively similar performance of rapid contractions with little force generation, each species that maintains these muscles does so by expressing different MYH genes that share no obvious loop configuration properties. Selection for "superfast" muscle performance therefore has likely followed lineage-specific biochemical and biophysical pathways. Furthermore, muscle performance variation is not likely determined strictly by a particular molecular structure or even a particular MYH protein, but instead through the combined molecular properties and expression ratios of several MYH proteins with diverse structures.

To characterize further protein variation, we calculated amino acid diversity for each fully aligned site for core MYHs and MYH6/7/60 subfamilies separately within mammals, birds, amphibians, and ray-finned fish (Fig. 4J; Grantham 1974). Out of 1893 sites without gaps, we found that 412 (22%) were completely invariant across all sequences from these four groups with 68% of invariant sites appearing in the head domain. In contrast, 645 sites (34%) exhibited repeated amino acid variation in all four groups, while the remaining 278 (15%), 258 (14%), and 288 (15%) fully aligned sites were variable in three, two, or a single group, respectively. Sites showing variation in all four groups and variation hotspots overall were found in the N-terminal domain, the light-chain myosin binding site on the neck, and midway down the tail (1230–1260). These intersected with several sites of known mammalian MYH7 functional impactful, while pathogenic variants were nearly always invariant across core myosins (Johnson, et al. 2021; Supplementary Table S3). No specific amino acid state was identified as being common to the myosins known to exhibit extreme muscle properties, concurring with previous hypotheses that the specific impact

of individual amino acid changes on MYH fibers or muscle performance as a whole is difficult to predict (Weiss 1999b, Robert-Paganin et al. 2019). This analysis highlights that, particularly in the actin- and ATP-interacting myosin head region, MYH is balancing sites of extreme biochemical and structural conservation with sites that are repeatedly mutating and show variation in several independent groups. Therefore, structural variation is maintained in this gene family across long evolutionary distances, but the same regions have repeatedly evolved higher variation, consistent with our characterization of a intermixed evolutionary pressures of homeostatic conservation and physiological innovation.

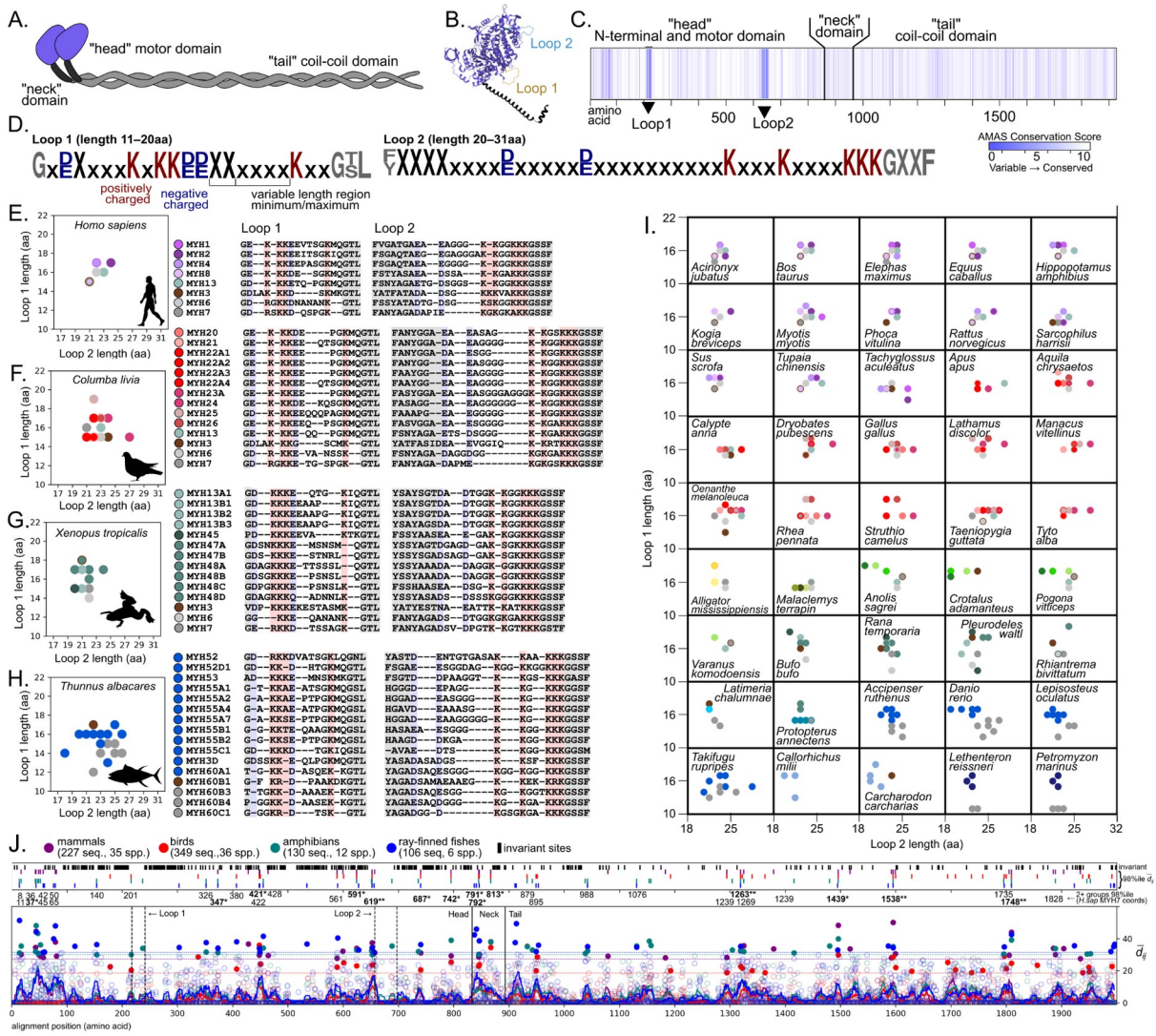

Figure 4. MYH loops and other sites are variable among and within vertebrate species. (A) Myosin protein structural domains shown for a dimerized protein pair. (B) Structure head and neck with Loop 1 and Loop 2 highlighted (PDB#P12882; Berman, et al. 2003, Sehnal, et al. 2021). (C) AMAS conservation scores for each position show that Loop 1 and Loop 2 have the highest rates of variation. (D) MYH Loop 1 and Loop 2 motifs, showing the positive and negative residues and the variable region minimum (X) and maximum (x) lengths observed. (E,F,G,H) Loop 1 and Loop 2 overall lengths create variable combinations for the core sarcomeric and slow sarcomeric MYH proteins in human, pigeon, and western clawed frog, and yellowfin tuna. (I) Loop 1 and 2 combinations for 45 additional vertebrate species show

variation in MYH loop lengths is present in species in each major group. Each point represents a single MYH gene, with colors matching subfamilies in Figures 1b and 2. (J) Mean Grantham's distance ($\bar{d}_{ij}$) for all pairs of amino acids at each aligned position in skeletal sarcomeric myosins separately for mammals, birds, amphibians, and ray-finned fish, show amino acid positions that are recurrently variably in multiple groups. Solid graph lines show the mean value for 10aa windows. Positions with more than one gap and MYH7B, MYH15, and MYH16 were excluded. Vertical dashed lines show the Loop 1 and Loop 2 regions, and vertical solid lines show the head/heck/tail sections. Horizontal dashed lines show the 98th percentiles for each group. High values ($\bar{d}_{ij} \geq$ 98%ile) are labeled with color dots. The upper track shows the location of completely invariant sites across all sequences (black boxes) and high $\bar{d}_{ij}$ values for each group (color boxes). Numbers shown below are amino acid positions relative to human MYH7 shown for sites with high $\bar{d}_{ij}$ values in at least two groups, three groups (*), or all four groups (**). (Art credit for all silhouettes: C. Harvey).

Skeletal muscle myosins evolve in a balance of functional maintenance and innovation.
Understanding the molecular evolution of skeletal muscles is crucial to deciphering their physiological evolution and the evolution of vertebrates as a whole. Our finding that major vertebrate groups have separately duplicated skeleton muscle myosin subfamilies substantially recontextualizes discussion of vertebrate muscle evolution. The lack of one-to-one relationships among individual core MYH proteins between any two major vertebrate groups rejects the extrapolation of mammalian MYH functional identities or fiber type classifications to other vertebrates. Our results point to an alternative perspective on skeletal muscle myosins to explain the surprising evolutionary lability of such a core genetic system. Unequal crossing over creates new myosin copies that gradually acquire diverse molecular properties (particularly in the hypervariable Loop 1 and Loop 2) leading to subfunctionalization or neofunctionalization. Different muscles express MYH copies in specific combinations to achieve differences in performance ability, with such differences likely ranging from subtle to extreme. At the same time, many performance attributes will also require the collective action of several muscles that use a range of expression ratios of various MYH proteins. We therefore propose that MYH evolution proceeds under layers of tandem gene cluster turnover, shifting expression, and changing musculoskeletal design. This dynamic presumably creates the opportunity for gradual evolution, with the entire set of myosins being under a collective coselection at the phenotypic level.

Our results also speak to an evolutionary paradox about balancing the necessity of maintaining core fitness with the observation of abundant innovation of muscles throughout vertebrates. MYH genes are individually gained, lost, or diversified as long as the overall functional set of MYHs is maintained. The flexibility of the multi-layer system creates an opportunity for selection to act on certain muscles while maintaining the functionality of others, balancing homeostasis and innovation. Several other cases of vertebrate gene family tandem cluster duplications (e.g., opsins, venom, immune factors, hormones) often involve antagonistic genetic systems whose diversification is stimulated by a shifting, exogenous factor (Ramirez, et al. 2016, Storz, et al. 2016, Mason, et al. 2022). The lability of protein duplications and losses is therefore particularly striking in the case of MYH, since skeletal and cardiac muscles are core, highly interconnected physiological systems presumably subject to substantial endogenous selective constraints. While additional direct studies connecting MYH properties, fiber composition, and muscle performance are needed to understand how individual MYH genes contribute to changes in contractile

force, velocity, and metabolics (Seaborne & Ochala 2023, Martinez, et al. 2024), this investigation establishes the rampant evolutionary turnover of myosins in the evolution of vertebrates and their independent diversifications. This new perspective on both the present diversity and underlying evolution of MYH proteins we hope will stimulate new perspectives, experiments, and discussion of the evolution of vertebrate muscles and motivate renewed investigation of the diversification of other core animal gene families.

Acknowledgements

Thanks to Peri Bolton, Olivia Delgado, Lisle Gibbs, Caroline Kauh, Ellen Weinheimer, and Oli Wood for feedback on this manuscript.

Funding

This work was supported by National Science Foundation grant NSF-DEB#2217117 to JBP, NSF-OISE#1952542 and NSF-IOS#2423144 to MJF, and funding from the Howard Hughes Medical Institute to ERS.

Author Contributions

Conceptualization: CMH, JBP, MJF
Methodology: CMH, ERS, JBP, MJF,
Formal Analysis: CMH, JBP
Investigations: CMH, ERS, JBP
Resources: ERS, JBP, MJF, MSB
Data Curation: CMH, JBP
Writing - Original Draft: CMH, JBP, MJF
Writing - Review & Editing: CMH, JBP, MJF, MSB
Visualization: CMH, JBP
Supervision: JBP
Project Administration: JBP
Funding Acquisition: JBP, MJF, MSB

Competing Interests

The authors report no competing interests.

Data and Materials Availability

Lonchura striata syrinx raw RNA-seq sequences are available from NCBI SRA under BioProject #[accession pending]. All other data are available in the supplementary materials, or through NCBI Protein, Nucleotide, or SRA databases at the accession numbers specified therein.

Materials and Methods

Animals
Adult male Bengalese finches (<120 days post hatching) were from breeding colonies at UCSF. All experiments were conducted in accordance with UCSF policies governing animal use and welfare.

Tissue collection and RNA extraction
Birds were euthanized through an overdose of isoflurane, followed by rapid decapitation. Syrinxes were then removed and frozen on dry ice until further processing. At a later point, we homogenized samples using a rotor-stator homogenizer set to medium speed. As in Pease et al. (2022), we extracted total RNA using a Zymo® Direct-zol™ RNA miniprep kit (Zymo Research), in which we included an initial phenol-chloroform separation of RNA following the manufacturer's instructions. RNA quantity and integrity (RIN; all samples had RINs greater than 8.5) were measured using a Qubit and BioAnalyzer, respectively.

Library prep and sequencing
Libraries were prepared using the Zymo-Seq RiboFree total RNA library kit according to the manufacturer's instructions. The resulting libraries were diluted and pooled before sequencing on a NovaSeq X Plus (10B 151×12×24×151) at UCSF.

RNA-Seq Data
Lonchura striata syrinx raw RNA-seq sequences are available from NCBI SRA under BioProject #[accession pending]. All other expression data were obtained from public sets available on NCBI. The accession numbers to these sets are provided in Data S6 and paper DOIs are provided when available

Compiling Amino Acid Sequences
We compiled a sample of amino acid sequences by searching within the NCBI reference genomes of select species for myosins or their syntenic neighbors (Data S1 and Data S2). For species that did not have well annotated reference genomes, we sought out sequences by using NCBI's Basic Local Alignment Search Tool for proteins (BLASTp) or expanded searches to ENSEMBL. This resulted in a set of 1397 sequences prior to filtering and quality control.

Sequence Alignment and Phylogenetic Reconstruction
We filtered our dataset to remove sequences that were partial or of poor quality. We also made small manual edits as needed to remove regions prior to start codons, retained introns, and extended repetitive ends erroneously retained on the c-terminus (see Data S4 for a description of these edits). After filtering, our dataset included 1137 sequences from 112 vertebrate species (Data S1 and Data S2). We aligned these sequences using MAFFT's L-INS-i method and created phylogenetic trees using RAXML-ng with the QMaker PFAM model of protein evolution with gamma-distributed rates (Katoh, et al. 2005, Kozlov, et al. 2019). This was done with the following options:

```
raxml-ng --all --msa ALIGNMENT.msa --model Q.PFAM+G
```

RAXML-ng generates support values by default using the Felsenstein Bootstrapping protocol. Bootstrap replicates were generated until the majority rule extended (MRE) threshold was surpassed, up to a maximum of 1000 replicates.

Phylogenetic Relationships Between Myosins
We determined evolutionary relationships between myosins using both phylogenetic reconstruction and amino acid sequence analysis. First, we grouped phylogenetically monophyletic groups to identify clades of independently diverging myosins. We additionally compared sequence similarity by calculating the

pairwise difference between amino acids sequences (Fig. S1). To do this, we used a custom Python3 script to count the pairwise percent distances in amino acids. The numerator was the count of all amino acid sites without a gap in both sequences that were the same amino acid. The denominator was a count of all amino acid sites without a gap in both sequences. Additionally, we computed a multidimensional scaling plot (MDS) from the same pairwise distances using the Python3 *mds* function from the *scikit-sklearn* module (Fig. S2).

We used this combined data to make the best educated hypothesis of evolutionary relatedness between myosin sequences. To differentiate between non-orthologous myosins, we renamed myosins from several animal clades to better represent their evolutionary history. These names are described in detail in Supplementary Table S1, and further detailed in the Supplementary Text.

RNA Expression Quantification and Determination of Differentially Expressed Genes

We mapped RNA expression data to NCBI reference genomes using STAR and quantified expression with featureCounts (Dobin, et al. 2013, Liao, et al. 2014). We began by preparing genome directories using the following options:

```
STAR --runMode genomeGenerate --genomeDir ./GENOMEDIR --genomeFastaFiles
GENOMEFILE.fna --sjdbGTFfile GTFFILE.gtf
```

We then acquired SRA data using FasterQ from the Sequence Read Archive (SRA) Toolkit and trimmed the reads by using the following command:

```
fasterq-dump SRA_ID --p --W --split-3
```

We then used STAR to align expression data against the prepared genome using:

```
STAR --genomeDir ./GENOMEDIR --readFilesIn SAMPLES.fq
```

To quantify expression data, we used featureCounts using the following command:

```
featureCounts --p --a GTFFILE.gtf --o OUTPUT.txt STAR_OUTPUT.bam
```

We transformed the raw read counts into CPM (counts per million) by dividing the counts for each gene by the total number of counts per sample and multiplying by one million. Because myosins are often very similar in sequence and sometimes differ by only a few amino acids, we removed multi-mapped reads. This means that RNA expression depicted in our data is likely lower than is actually present within the tissue, but more accurately represents the diversity and relative abundance of transcripts.

For Sika, we quantified RNA expression using *kallisto* (Bray, et al. 2016):

```
kallisto quant -t 64 --index $INDEX -o $OUTPUT $FASTQ1 $FASTQ2
```

To identify patterns of tissue specificity and changes in transcript abundance, we quantified □ and log2 maximum expression (Yanai et al. 2004). We considered genes with $\tau > 0.75$ to exhibit high tissue-specific expression.

Calculating Conservation Score

To determine areas of high and low conservation, we calculated conservation scores based on an alignment of 624 sequences from the tetrapod core cluster (Data S7, Data S8). Conservation scores were obtained from JalViewJS using the Analysis of Multiple Sequences (AMAS) method (Livingston and Barton 1993, Waterhouse et al. 2009).

Determining Loop Lengths and Structure

Alignments of Loop 1 and Loop 2 were extracted from the primary alignment (Supplementary Data S1). We determined the start and stop of each loop based on previous work on humans and chicken myosins (Weiss et al. 1999). We calculated the ungapped length of these regions (see Data S5) and generated Logo-like plots from a sequence alignment of Loop 2 based on myosin type (Data S5, Fig 4).

Analysis of Amino Acid Diversity

The primary alignment (Supplementary Data S1) was filtered to four data subsets: (1) 227 sequences of MYH1–4, MYH6–8, and MYH13 from 35 mammal species, (2) 349 sequences of MYH3, MYH6, MYH7, MYH13, and MYH20–26 sequences from 36 bird species, (3) 130 sequences of MYH3, MYH6, MYH7, MYH13A, MYH13B, MYH13C, and MYH45–48 from 12 amphibian species, and (4) 106 sequences of MYH3, MYH52–55 and MYH60 from 6 ray-finned fish species. For each position in the amino acid alignment, the average Grantham's distance for each pair of amino acids at the site was calculated as:

$$\underline{d}_{ij} = \frac{1}{n(n-1)/2} \sum_{i<j} G_{x_i x_j}$$

where $\underline{d}_{ij}$ is the average Gratham's distance among all pairs of amino acids at a given position, $G_{x_i x_j}$ is Grantham's distance for a given pair of amino acids from sequences *i* and *j* at position *x,* and *n* is the total number of sequences at the position (Grantham 1974). Positions with more than one sequence having a gap were scored as zero. Eight sequences with partial missing sequences of >10 aa were excluded. This metric should be higher at aligned positions with both higher amino acid sequence diversity and greater amino acid side chain biochemical diversity. Calculations and plotting were performed with a custom Python3 script.

References

Bárány M, 1967. ATPase activity of myosin correlated with speed of muscle shortening. *The Journal of General Physiology* 50:197–218. 10.1085/jgp.50.6.197.

Berman H, Henrick K, Nakamura H, 2003. Announcing the worldwide Protein Data Bank. *Nature Structural and Molecular Biology* 10:980–980. 10.1038/nsb1203-980.

Bonine KR, Gleeson TT, Garland T, 2005. Muscle fiber-type variation in lizards (Squamata) and phylogenetic reconstruction of hypothesized ancestral states. *The Journal of Experimental Biology* 208:4529-4547. 10.1242/jeb.01903.

Bothe MS, Kohl T, Felmy F, Gallant J, Chagnaud BP, 2024. Timing and precision of rattlesnake spinal motoneurons are determined by the $KV7_{2/3}$ potassium channel. *Current Biology,* 34:286 - 297.e5 10.1016/j.cub.2023.11.062

Bray NL, Pimentel H, Melsted P, Pachter L, 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 34:525–527. 10.1038/nbt.3519.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR, 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21. 10.1093/bioinformatics/bts635.

Dos Santos M, Backer S, Auradé F, Wong MM, Wurmser M, Pierre R, Langa F, Do Cruzeiro M, Schmitt A, Concordet JP, Sotiropoulos A. A fast *Myosin* super enhancer dictates muscle fiber phenotype through competitive interactions with *Myosin* genes. 2022. *Nature Communications*. 13:1039. 10.1038/s41467-022-28666-1.

Dye DE, Azzarelli B, Goebel HH, Laing NG, 2006. Novel slow-skeletal myosin (MYH7) mutation in the original myosin storage myopathy kindred. *Neuromuscular Disorders* 16:357–360. 10.1016/j.nmd.2006.03.011.

Ennion S, Sant'Ana Pereira J, Sargeant AJ, Young A, Goldspink G, 1995. Characterization of human skeletal muscle fibres according to the myosin heavy chains they express. *Journal of Muscle Research and Cell Motility* 16:35–43. 10.1007/bf00125308.

Finno CJ, Gianino G, Perumbakkam S, Williams ZJ, Bordbari MH, Gardner KL, Burns E, Peng S, Durward-Akhurst SA, Valberg SJ, 2018. A missense mutation in MYH1 is associated with susceptibility to immune-mediated myositis in Quarter Horses. *Skeletal Muscle* 8. 10.1186/s13395-018-0155-0.

Foth BJ, Goedecke MC, Soldati D, 2006. New insights into myosin evolution and classification. *Proceedings of the National Academy of Sciences* 103:3681–3686. 10.1073/pnas.0506307103.

Friedman CE, Fayer S, Pendyala S, Chien WM, Loiben A, Tran L, Chao LS, McKinstry A, Ahmed D, Farris SD, *et al.*, 2024. Multiplexed functional assessments of MYH7 variants in human cardiomyocytes. *Circulation: Genomic and Precision Medicine* 17. 10.1161/circgen.123.004377.

Fuxjager MJ, Goller F, Dirkse A, Sanin GD, Garcia S, 2016. Select forelimb muscles have evolved superfast contractile speed to support acrobatic social displays. *eLife* 5. 10.7554/elife.13544.

Goodson HV, Warrick HM, Spudich JA, 1999. Specialized conservation of surface loops of myosin: evidence that loops are involved in determining functional characteristics. *Journal of Molecular Biology* 287:173–185. 10.1006/jmbi.1999.2565.

490 Grantham R. 1974. Amino acid difference formula to help explain protein evolution. *Science* 185:862--
491 864. 10.1126/science.185.4154.862.

492 Harvey CM, Fuxjager MJ, Pease JB, 2025. Deep-time gene expression shift reveals an ancient change in
493 avian muscle phenotypes. *PLoS Genetics* 21:e1011663. 10.1371/journal.pgen.1011663.

494 He ZH, Bottinelli R, Pellegrino MA, Ferenczi MA, Reggiani C, 2000. ATP consumption and efficiency
495 of human single muscle fibers with different myosin isoform composition. *Biophysical Journal* 79:945–
496 961. 10.1016/s0006-3495(00)76349-1.

497 Hedera P, Petty EM, Bui MR, Blaivas M, Fink JK, 2003. The second kindred with autosomal dominant
498 distal myopathy linked to chromosome 14q: genetic and clinical analysis. *Archives of Neurology* 60.
499 10.1001/archneur.60.9.1321.

500 Hoh JFY, 2023. Developmental, physiologic and phylogenetic perspectives on the expression and
501 regulation of myosin heavy chains in mammalian skeletal muscles. *Journal of Comparative Physiology B*
502 193:355-382 10.1007/s00360-023-01499-0.

503 Ikeda D, Ono Y, Snell P, Edwards YJK, Elgar G, Watabe S, 2007. Divergent evolution of the myosin
504 heavy chain gene family in fish and tetrapods: evidence from comparative genomic analysis.
505 *Physiological Genomics* 32:1–15. 10.1152/physiolgenomics.00278.2006.

506 Johnson CA, McGreig JE, Jeanfavre ST, Walklate J, Vera CD, Farré M, Mulvihill DP, Baines AJ, Ridout
507 M, Leinwand LA, *et al.*, 2021. Identification of sequence changes in myosin II that adjust muscle
508 contraction velocity. *PLoS Biology* 19:e3001248. 10.1371/journal.pbio.3001248.

509 Katoh K, Kuma K, Toh H, Miyata T, 2005. MAFFT version 5: improvement in accuracy of multiple
510 sequence alignment. *Nucleic Acids Research* 33:511–518. 10.1093/nar/gki198.

511 Korn ED, 2000. Coevolution of head, neck, and tail domains of myosin heavy chains. *Proceedings of the*
512 *National Academy of Sciences* 97:12559–12564. 10.1073/pnas.230441597.

513 Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A, 2019. RAxML-NG: a fast, scalable and user-
514 friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35:4453–4455.
515 10.1093/bioinformatics/btz305.

516 Laing NG, Laing BA, Meredith C, Wilton SD, Robbins P, Honeyman K, Dorosz S, Kozman H, Mastaglia
517 FL, Kakulas BA, 1995. Autosomal dominant distal myopathy: linkage to chromosome 14. *American*
518 *Journal of Human Genetics* 56:422–7.

519 Larsson L, Moss RL, 1993. Maximum velocity of shortening in relation to myosin isoform composition
520 in single fibres from human skeletal muscles. *The Journal of Physiology* 472:595–614.
521 10.1113/jphysiol.1993.sp019964.

522 Lee JH, Lewis KM, Moural TW, Kirilenko B, Borgonovo B, Prange G, Koessl M, Huggenberger S, Kang
523 C, Hiller M. 2018. Molecular parallelism in fast-twitch muscle proteins in echolocating mammals.
524 *Science Advances*.4:eaat9660. 10.1126/sciadv.aat9660.

525 Lee LA, Karabina A, Broadwell LJ, Leinwand LA, 2019. The ancient sarcomeric myosins found in
526 specialized muscles. *Skeletal Muscle* 9. 10.1186/s13395-019-0192-3.

527 Liao Y, Smyth GK, Shi W, 2014. featureCounts: an efficient general purpose program for assigning
528 sequence reads to genomic features. *Bioinformatics* 30:923–930. 10.1093/bioinformatics/btt656.

529 Lucas CA, Kang LHD, Hoh JFY, 2000. Monospecific antibodies against the three mammalian fast limb
530 myosin heavy chains. *Biochemical and Biophysical Research Communications*, 272:303-308.
531 10.1006/bbrc.2000.2768.

532 Luna VM, Daikoku E, Ono F, 2015. "Slow" skeletal muscles across vertebrate species. *Cell & Bioscience*
533 5:62. 10.1186/s13578-015-0054-6.

534 Martin JH, Bagby RM, 1973. Properties of rattlesnake shaker muscle. *Journal of Experimental Zoology*
535 185:293-300. 10.1002/jez.1401850303.

536 Martínez Mir C, Pisterzi P, De Poorter I, Rilou M, van Kranenburg M, Heijs B, Alemany A, Sage F,
537 Geijsen N. 2024. Spatial multi-omics in whole skeletal muscle reveals complex tissue architecture.
538 *Communications Biology*.7:1272. 10.1038/s42003-024-06949-1.

539 Mascarello F, Toniolo L, Cancellara P, Reggiani C, Maccatrozzo L, 2016. Expression and identification
540 of 10 sarcomeric MyHC isoforms in human skeletal muscles of different embryological origin. Diversity
541 and similarity in mammalian species. *Annals of Anatomy - Anatomischer Anzeiger* 207:9-20.
542 10.1016/j.aanat.2016.02.007.

543 Mason AJ, Holding ML, Rautsaw RM, Rokyta DR, Parkinson CL, Gibbs HL, 2022. Venom gene
544 sequence diversity and expression jointly shape diet adaptation in pitvipers. *Molecular Biology and*
545 *Evolution* 39. 10.1093/molbev/msac082.

546 McGuigan K, 2004. Evolution of sarcomeric myosin heavy chain genes: evidence from fish. *Molecular*
547 *Biology and Evolution* 21:1042–1056. 10.1093/molbev/msh103.

548 Mead AF, Osinalde N, Ørtenblad N, Nielsen J, Brewer J, Vellema M, Adam I, Scharff C, Song Y,
549 Frandsen U, *et al.*, 2017. Fundamental constraints in synchronous muscle limit superfast motor control in
550 vertebrates. *eLife* 6. 10.7554/elife.29425.

551 Meredith C, Herrmann R, Parry C, Liyanage K, Dye DE, Durling HJ, Duff RM, Beckman K, de Visser
552 M, van der Graaff MM, *et al.*, 2004. Mutations in the slow skeletal muscle fiber myosin heavy chain gene
553 (MYH7) cause Laing early-onset distal myopathy (MPD1). *The American Journal of Human Genetics*
554 75:703–708. 10.1086/424760.

555 Moore LA, Tidyman WE, Arrizubieta MJ, Bandman E. 1992. Gene conversions within the skeletal
556 myosin multigene family. *Journal of Molecular Biology* 223:383–387. 10.1016/0022-2836(92)90741-2

557 Moore LA, Tidyman WE, Arrizubieta MJ, Bandman E. 1993. The evolutionary relationship of avian and
558 mammalian myosin heavy-chain genes. *Journal of Molecular Evolution* 36:21–30. 10.1007/bf02407303.

559 Piazzesi G, Reconditi M, Linari M, Lucii L, Bianco P, Brunello E, Decostre V, Stewart A, Gore DB,
560 Irving TC, *et al.*, 2007. Skeletal muscle performance determined by modulation of number of myosin
561 motors rather than motor force or stroke size. *Cell* 131:784–795. 10.1016/j.cell.2007.09.045.

562 Piazzesi G, Reconditi M, Linari M, Lucii L, Sun YB, Narayanan T, Boesecke P, Lombardi V, Irving M,
563 2002. Mechanism of force generation by myosin heads in skeletal muscle. *Nature* 415:659–662.
564 10.1038/415659a.

565 Ramirez MD, Pairett AN, Pankey MS, Serb JM, Speiser DI, Swafford AJ, Oakley TH, 2016. The last
566 common ancestor of most bilaterian animals possessed at least 9 opsins. *Genome Biology and Evolution*
567 p. evw248. 10.1093/gbe/evw248.

568 Rayment I, Holden H, Whittaker M, Yohn C, Lorenz M, Holmes K, Milligan R, 1993. Structure of the

actin-myosin complex and its implications for muscle contraction. *Science* 261:58–65. 10.1126/science.8316858.

Resnicow DI, Deacon JC, Warrick HM, Spudich JA, Leinwand LA, 2009. Functional diversity among a family of human skeletal muscle myosin motors. *Proceedings of the National Academy of Sciences* 107:1053–1058. 10.1073/pnas.0913527107.

Robert-Paganin J, Pylypenko O, Kikuti C, Sweeney HL, Houdusse A, 2019. Force generation by myosin motors: a structural perspective. *Chemical Reviews* 120:5-35. 10.1021/acs.chemrev.9b00264.

Robinson WD, Rourke B, Stratford, JA, 2021. Put some muscle behind it: understanding movement capacity of tropical birds. *Ornithology* 138. 10.1093/ornithology/ukaa068.

Romiguier J, Ranwez V, Douzery EJP, Galtier N, 2010. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Research* 20:1001–1009. 10.1101/gr.104372.109.

Scala M, Accogli A, De Grandis E, Allegri A, Bagowski CP, Shoukier M, Maghnie M, Capra V, 2018. A novel pathogenic MYH3 mutation in a child with Sheldon–Hall syndrome and vertebral fusions. *American Journal of Medical Genetics Part A* 176:663–667. 10.1002/ajmg.a.38593.

Schiaffino S, Chemello F, Reggiani C. The diversity of skeletal muscle fiber types. 2024a. *Cold Spring Harbor Perspectives in Biology*. 12:a041477. 10.1101/cshperspect.a041477.

Schiaffino S, Hughes SM, Murgia M, Reggiani C, 2024b. MYH13, a superfast myosin expressed in extraocular, laryngeal and syringeal muscles. *The Journal of Physiology* 602:427–443. 10.1113/jp285714.

Schiaffino S, Reggiani C, 2011. Fiber types in mammalian skeletal muscles. *Physiological Reviews* 91:1447–1531. 10.1152/physrev.00031.2010.

Schiaffino S, Rossi, AC, Smerdu V, Leinward LA, Reggiani C, 2015. Developmental myosins: expression patterns and functional significance. *Skeletal Muscle* 5:22. 10.1186/s13395-015-0046-6.

Seaborne RA, Ochala J. The dawn of the functional genomics era in muscle physiology. 2023. *The Journal of Physiology*. 2023 601:1343-52. 10.1113/JP284206.

Sebé-Pedrós A, Grau-Bové X, Richards TA, Ruiz-Trillo I, 2014. Evolution and classification of myosins, a paneukaryotic whole-genome approach. *Genome Biology and Evolution* 6:290–305. 10.1093/gbe/evu013.

Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, Velankar S, Burley SK, Koča J, Rose AS, 2021. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research* 49:W431–W437. 10.1093/nar/gkab314.

Storz JF, 2016. Gene duplication and evolutionary innovations in hemoglobin-oxygen transport. *Physiology* 31:223–232. 10.1152/physiol.00060.2015.

Tajsharghi H, Darin N, Rekabdar E, Kyllerman M, Wahlström J, Martinsson T, Oldfors A, 2005. Mutations and sequence variation in the human myosin heavy chain IIa gene (MYH2). *European Journal of Human Genetics* 13:617–622. 10.1038/sj.ejhg.5201375.

Tajsharghi H, Hammans S, Lindberg C, Lossos A, Clarke NF, Mazanti I, Waddell LB, Fellig Y, Foulds N, Katifi H, *et al.*, 2013. Recessive myosin myopathy with external ophthalmoplegia associated with MYH2 mutations. *European Journal of Human Genetics* 22:801–808. 10.1038/ejhg.2013.250.

Uctepe E, Mancılar H, Esen FN, Unverengil GG, Vona B, Yesilyurt A, 2024. A homozygous MYH1 variant underlies autosomal recessive isolated recurrent rhabdomyolysis. *American Journal of Medical Genetics Part A* 197. 10.1002/ajmg.a.63952.

Usui K, Yamamoto T, Khannoon ER,Tokita M. 2024. Musculoskeletal morphogenesis supports the convergent evolution of bat laryngeal echolocation. *Proceedings of the Royal Society B*, *291*: 20232196. 10.1098/rspb.2023.2196.

van Rheede T, Bastiaans T, Boone DN, Hedges SB, de Jong WW, Madsen O, 2005. The platypus is in its place: nuclear genes and indels confirm the sister group relation of monotremes and therians. *Molecular Biology and Evolution* 23:587–597. 10.1093/molbev/msj064.

Veugelers M, Bressan M, McDermott DA, Weremowicz S, Morton CC, Mabry CC, Lefaivre JF, Zunamon A, Destree A, Chaudron JM, *et al.*, 2004. Mutation of perinatal myosin heavy chain associated with a Carney complex variant. *New England Journal of Medicine* 351:460–469. 10.1056/nejmoa040584.

Vilfan A, Šarlah A, 2023. Theoretical efficiency limits and speed-efficiency trade-off in myosin motors. *PLoS Computational Biology* 19:e1011310. 10.1371/journal.pcbi.1011310.

Wagoner JA, Dill KA, 2021. Evolution of mechanical cooperativity among myosin II motors. *Proceedings of the National Academy of Sciences* 118. 10.1073/pnas.2101871118.

Weiss A, McDonough D, Wertman B, Acakpo-Satchivi L, Montgomery K, Kucherlapati R, Leinwand L, Krauter K, 1999a. Organization of human and mouse skeletal myosin heavy chain gene clusters is highly conserved. *Proceedings of the National Academy of Sciences* 96:2958–2963. 10.1073/pnas.96.6.2958.

Weiss A, Schiaffino S, Leinwand LA, 1999b. Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity. *Journal of Molecular Biology* 290:61–75. 10.1006/jmbi.1999.2865.

Whalen RG, Sell SM, Butler-Browne GS, Schwartz K, Bouveret P, Pinset-Härström I, 1981. Three myosin heavy-chain isozymes appear sequentially in rat muscle development. *Nature* 292:805–809. 10.1038/292805a0.

Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E, *et al.*, 2004. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21:650–659. 10.1093/bioinformatics/bti042.