Myosin light chain proteins cooperatively promote sarcomere growth in fast-twitch muscle

Authors:
Tayo E Adekeye[1], Troy E Hupper[1], Teresa E Easterbrooks[1], Emily M Teets[2], Emily A Tomak[1], Sadie L Waterman[1], Kailee A Sprague[1], Angelina White[1], Maddison L Coffin[1], Daniel K Tanaka[1], Sabrina M Varga[1], Mason T Soares[1], Sarah J Shepherd[2], Jared D Austin[1], Dmitrii Krivorotko[1], Emma S Perry[3], Sharon L Amacher[4], Joshua B Kelley[5], and Jared C Talbot[1]*

[1] School of Biology and Ecology, The University of Maine, 04469, USA, [2] Molecular Genetics, The Ohio State University, 43210, USA, [3] Core Electron Microscopy Laboratory, The University of Maine, 04469, USA, [4] Departments of Molecular Genetics and Biological Chemistry and Pharmacology, The Ohio State University, 43210, USA, [5] Molecular and Biomedical Sciences, The University of Maine, 04469, USA.

* Corresponding author
Phone: 207-581-2835
Email: jared.talbot@maine.edu

Abstract

Muscle becomes stronger by expanding sarcomeres from the cell's edge to its center. Here we investigate how Mylpf (Myosin Light Chain Phosphorylatable Fast) abundance impacts sarcomere growth. The two zebrafish Mylpf genes (*mylpfa* and *mylpfb*) are exclusively expressed in fast-twitch muscle, with *mylpfa* expressed more abundantly than *mylpfb*. Mutations in both genes cause full mRNA and protein loss, and loss of sarcomere formation. The degree of sarcomere reduction in fast-twitch muscle is predicted by Mylpf protein dosage using a high-slope Hill equation. The sarcomere defect may be driven by thick filament localization, which is impaired in *mylpfa*$^{-/-}$ and obliterated in *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$. Slow-twitch muscle is spared, becoming larger and more active in the mutants. The *mylpfa*$^{-/-}$ defects can be rescued by Mylpfa-GFP, Mylpfb-GFP, and by human MYLPF-GFP with equivalent efficiency, but not by a MYLPF protein variant that causes Distal Arthrogryposis. These effects indicate that Mylpf dosage strengthens muscle by mediating myofilament localization, and thereby sarcomere and myofibril growth.

Introduction

Sarcomeres grow from the edge to the center of muscle cells (the 'width axis') and this growth is what strengthens the muscle. Despite the importance of the sarcomere's width to muscle strength, it remains unclear how its growth is regulated. One way to control a sarcomere's growth is to regulate the abundance of its structural proteins. Overall protein levels ('protein dose') are influenced by many factors including a gene's copy number ('gene dosage'), the amount of steady-state transcript produced by each gene ('mRNA abundance'), the measured amount of protein encoded by these genes ('protein abundance'), the activity of each encoded protein, and the protein's localization within a cell. In biochemical settings, the relationship between protein dose and its effect can be modeled using four-parameter logistic regressions, such as the Hill equation [1]. The same equation is often used to assess drug efficiency *in vivo*, though it is not commonly applied to investigate how native protein impacts muscle formation. In this study, we investigate how the abundance of one key sarcomeric component, Mylpf (Myosin Light Chain Phosphorylatable Fast), impacts sarcomere formation. We find that the protein is essential to an early step in sarcomere formation and its effects can be efficiently modeled as a Hill response.

Sarcomere growth is best understood in the context of its organelle, the myofibril, which contains chains of sarcomeres that stretch along the length axis of muscle. Myofibril formation is initially directed by actin-rich thin filaments aligned along the edge of the muscle cell [2–4]. These thin filaments contain F-actin strands bundled with nebulin and linked to actinin-rich Z-disks, producing I-Z-I bodies. The ends of I-Z-I bodies are temporarily connected to one another by non-muscle myosin to form pre-myofibrils, which act as a template for further growth [4–6]. Growth is also facilitated by integration of enormous contractile proteins titin and nebulin, which set the sarcomere's length [7]. Once pre-myofibrils have formed, the non-muscle myosin is quickly replaced by contractile myosin heavy chain (MyHC), bundled into double-headed thick filaments. These added thick filaments increase tension across the muscle cell, leading to myofibrillar growth [8–10]. Myofibrils grow to only around 2 μm wide, but multilayered arrays of myofibrils (myofibril bundles) continue this growth until they extend deep into

the central cytoplasm [6,10,11]. The myofibril bundles in vertebrate muscle are aligned and tightly interlinked along the width axis of a sarcomere, such that the sarcomere and myofibril bundle have matched width (Figure 1A). In this work, we refer to the sarcomere when examining localization of thick/thin filament components and the myofibril bundle when viewing the structure at a larger scale.

The class II MyHC in thick filaments is stabilized close to its force-generating head by a Regulatory Light Chain (RLC) and an Essential Light Chain (ELC) [12,13] (Figure 1B). These light chains regulate myosin movement and force generation but do not consume ATP nor produce force on their own [12]. Point mutations in the RLCs can reduce myosin step size [14], and outright removal of these light chains *in vitro* causes MyHC aggregation and partially reduces MyHC activity [15,16]. Light chains also have critical roles *in vivo,* as first shown in the *Drosophila* RLC mutant, which lacks skeletal muscle [17]. Similarly, the ELC gene *Myl1* is required for normal myogenesis in chicken, zebrafish, mouse, and human [18–20]. Because the primary effect of these light chains is to regulate and stabilize MyHC, we hypothesize that myosin light chain proteins influence the rate of cytoskeletal organization, leading to myofibril growth.

The only RLC with prominent expression in embryonic muscle fibers and postnatal fast-twitch skeletal muscle of mouse is Mylpf [21]. All skeletal muscle fibers are absent from the mouse *Mylpf* knockout at birth [22]. This defect is so dramatic that it is unclear what phase of myogenesis is disrupted by the mutation. *MYLPF* function is also critical for human development. Missense alleles in *MYLPF* cause Distal Arthrogryposis (DA), a congenital musculoskeletal disease characterized by inherited distal limb contractures [23]. DA is often caused by mutations in genes that encode contractile proteins [24,25]. However, neither muscle strength nor structure were examined in human patients with the *MYLPF* variants, and the impact of those variants has not been tested in animal models. Therefore, the impact of *MYLPF* missense alleles on muscle formation remains unresolved.

Muscle formation in vertebrates begins in segments of mesoderm called somites, which eventually produce most of the muscles in the body. Zebrafish somites produce two

93 fiber types, fast-twitch and slow-twitch, within the first day post fertilization (dpf) [26].
94 Whereas slow-twitch fibers contract and fatigue slowly, the zebrafish fast-twitch fibers
95 contract with such speed and power that they strain muscle cells nearly to the point of
96 snapping [27,28]. By 24 hours post fertilization (hpf), zebrafish fast-twitch muscle fibers are
97 positioned medial to a thin layer of slow-twitch fibers [26], enabling us to identify fiber
98 types by position in addition to molecular markers of fiber type. Embryonic muscle fibers
99 grow rapidly, more than doubling in size by the time of hatching (3 dpf) [29]. Hatched
100 larvae continue growing while subsisting on their yolk until 6 dpf, enabling study of early
101 larval growth independent of feeding. In zebrafish, Mylpf gene function is distributed
102 across two paralogs, *mylpfa* and *mylpfb* [30–32]. We previously showed that the zebrafish
103 *mylpfa*$^{-/-}$ mutant can form multinucleate muscle cells but have weakened muscle that
104 eventually deteriorates [33]. This mutant has severe physiological defects including
105 impaired escape response and 90% loss in muscle strength, suggesting that there may
106 be effects on the myofibril itself [33]. However, the impact of *mylpfa* on sarcomere
107 formation was untested and the requirements for *mylpfb* function in embryonic
108 development remained unexplored.

109 Here, we investigate the effect of Mylpf activity on sarcomere growth along the width
110 axis of muscle fibers. We generated a frameshifting *mylpfb*$^{-/-}$ mutant, which has no overt
111 muscle defect alone, but dramatically enhances the *mylpfa*$^{-/-}$ mutant phenotype. The
112 *mylpfa*$^{-/-}$ mutant has a severe reduction in myofilament overlap in fast-twitch muscle,
113 leading to impaired sarcomere growth and thereby less myofibril bundle growth within
114 that fiber type. Actin-myosin colocalization fails entirely in the fast-twitch muscle of the
115 *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant, causing a complete absence of fast-twitch contractile
116 structures. Both mutants are null for protein production due to nonsense mediated
117 decay. We find no evidence for compensation for Mylpf loss of function among *mylpfa*
118 and *mylpfb* genes, though the mutant animals do compensate in slow-twitch myofibril
119 bundle size and slow movement. Transgenic expression experiments in the *mylpfa*$^{-/-}$
120 mutant suggest that function and efficiency is conserved among the Mylpfa, Mylpfb, and
121 human MYLPF proteins. By contrast, a transgene that expresses a protein variant found
122 in Distal Arthrogryposis patients does not rescue muscle structure. Dose analysis in
123 knockout and transgenic animals suggests that Mylpf binding has a cooperative effect

124 on myofibril bundle growth, with initial binding enabling a more-than-additive
125 subsequent increase until saturation is approached. Together, these findings show a
126 critical impact of Mylpf dosage on sarcomere and myofibril production during early fast-
127 twitch myogenesis.

128

129 # Results

130 **Mylpfa is more abundant than Mylpfb during embryonic development**

131 Although *mylpfa* and *mylpfb* expression patterns have been described individually [34],
132 their overlap and relative abundance have not been clarified. Each gene produces one
133 detectable transcript exclusively in fast-twitch muscle fibers (Figures S1, S2).
134 Expression begins by 20 hpf in medial fast-twitch fibers (Figure S2). Neither gene is
135 expressed in slow-twitch muscle, which expresses a different regulatory light chain,
136 *myl10* (Figure 1C-C′′′) [35]. In both HCR and qPCR, *mylpfa* is expressed more abundantly
137 than *mylpfb*; the ratio between these genes is roughly 7:1 at 1 dpf and diminishes to 2:1
138 by 48 hpf, largely because of rising *mylpfb* mRNA abundance (Figure 1D-F). These
139 ratios can also be found in publicly available RNA-seq data [33]. These zebrafish
140 transcripts encode Mylpf proteins that are 94% identical to one another and each is
141 around 80% identical to human MYLPF. Using newly generated antibodies we can
142 distinguish the two proteins on western blot, which differ in size by 2 kilodaltons (kD).
143 The relative protein abundance correlates with transcript levels (Figure 1G-I). In
144 summary, the two zebrafish Mylpf genes are expressed in fast-twitch muscle, producing
145 Mylpfa protein more abundantly than Mylpfb during embryonic stages.

146

147 **The *mylpfa*^{oz30} and *mylpfb*^{oz39} alleles are effectively null**

148 We had previously made an allele of *mylpfa* (*oz30*) [33], and here provide evidence that it
149 causes severe sarcomere and myofibril bundle defects (Figure 2A-G). These myofibril
150 bundle defects are also apparent in confocal stack of phalloidin-labeled *mylpfa*$^{-/-}$ mutant
151 (Video 1) and in multiple quantification methods (Figures 2I-M S3, S4). Although the
152 *mylpfa*$^{-/-}$ mutant has severe myofibril reduction at all timepoints, it forms differentiated
153 somite muscle, which has wild-type size at 2 dpf and a modest growth defect by 6 dpf

154 (Figure 2H). These effects are confirmed by a second allele in the same gene (Figure
155 S5) [33]. Both *mylpfa* mutants can at least partially form myofibrils in fast-twitch muscle
156 (Figure S5) so we hypothesized that *mylpfb* may also contribute to function. We
157 mirrored the *mylpfa*^{oz30} mutant by generating a 5 bp deletion in *mylpfb* (allele *oz39*) that
158 frameshifts the *mylpfb* transcript at AA 78 out of 170 (Figure 3A-C). Using RT-qPCR we
159 find that mRNA levels for *mylpfa* and *mylpfb* are approximately halved in heterozygotes
160 and essentially undetectable in each mutant (Figure 3D-G), suggesting that both
161 mutations lose transcript through nonsense mediated decay. Any protein made by this
162 frameshifting allele should be nonfunctional because the encoded peptide has at most
163 only a short stretch capable of binding to MyHC (Figure 3H). However, we find no
164 Mylpfa protein produced in the *mylpfa*$^{-/-}$ mutant nor any Mylpfb in the *mylpfb*$^{-/-}$ mutant,
165 and neither protein in the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant at 48 hpf (Figure 3I-J).
166 Examination of all nine genotypes in western blot shows that overall Mylpf dosage
167 (Mylpfa & Mylpfb) corresponds closely to predictions made using the 2:1 dosage seen
168 at 48 hpf, suggesting that there is little intragenic compensation (Figure 3K, Table S1).
169 For instance, the *mylpfa*$^{+/-}$ and *mylpfb*$^{+/-}$ heterozygotes express around half the normal
170 protein dose (67% for *mylpfa*$^{+/-}$, 49% for *mylpfb*$^{+/-}$, Figure 3J). These findings indicate
171 that the mutants used in this study are essentially null for the single product of the
172 *mylpfa* and *mylpfb* genes, respectively.
173
174 **Together, *mylpfa* and *mylpfb* are essential to sarcomere formation**
175 Even though the Mylpfb protein sequence is very similar to Mylpfa, the two mutants
176 have strikingly different phenotypes. The *mylpfb*$^{-/-}$ mutant looks entirely wild-type on the
177 electron microscope (Figure 3L-M') and in every other assay we have tested thus far. By
178 contrast, the *mylpfa*$^{-/-}$ mutant has narrowed myofibrils, with poor myofilament alignment
179 on the sarcomere level (Figure 3N, N'). In contrast to the partial defects seen in the
180 *mylpfa*$^{-/-}$ mutant, the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant lacks sarcomeres entirely (Figure
181 3O, O'). This double mutant shows a catastrophic loss of striation in fast-twitch muscle,
182 more severe than either of the single mutants, suggesting that both genes contribute to
183 muscle structure (Figure 3O-Q). In this double mutant, the fast-twitch muscle shows no
184 true sarcomeres or myofibrils, only a scattering of thick filaments and I-Z-I bodies

185 (Figure 3O, O'). We observe equivalent effects on myofibril width (measured between
186 membrane triads) and myofibril bundle size (measured across interconnected
187 myofibrils) (Figure 3P, Q). Even in the double mutant, somite muscle is present and the
188 abundance of contractile MyHC is unchanged (Figure 3I, P>0.3 in ANOVA), suggesting
189 that muscle differentiation proceeds despite the loss of organized myofibrils. These
190 findings suggest that the two Mylpf genes are collectively essential to sarcomere
191 formation and therefore to myofibril bundle formation, with *mylpfa* playing a larger role.
192

193 **The *mylpf* genes are essential to myofilament colocalization**

194 Because Mylpf binds to MyHC, we hypothesized that the sarcomere defect may arise
195 through an initial failure in MyHC localization, independent of initial F-actin ordering. We
196 examined muscle structure in Mylpf mutants at an early stage of myofibril bundling (26
197 hpf). We find profound organizational defects in fast-twitch muscles, but not slow-twitch
198 fibers (Figure 4A-D). The *mylpfa*$^{-/-}$ mutant fails to properly localize MyHC to the fast-
199 twitch muscle cell's periphery where pre-myofibrils are forming, with stronger defects in
200 the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant (Figure 4C-D'). Consistent with this localization
201 defect, the correlation between F-actin and MyHC label is reduced in the *mylpfa*$^{-/-}$
202 mutant and worsened in the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant, to a degree that parallels
203 reductions in myofibril bundle width (Figure 4E-F). The confocal images suggested that
204 sarcomere and myofibril formation may fail in the mutant because F-actin and MyHC
205 are found in different portions of the cell. At this early stage, the pre-myofibrils form
206 rings around cells even in the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant (Figure 4A-D), consistent
207 with findings that actin localization precedes thick filament localization [36]. Using these
208 rings of F-actin label, we marked cell boundaries and then measured how far each
209 marker spreads from that boundary to the cell center in fast-twitch muscle cells (Figure
210 4G). By contrast, each genotype has a similar distribution of F-actin and myonuclei
211 along this cell edge-center axis (Figure 4H, I). MyHC is also localized at cell peripheries
212 in the wild-type embryo, but it spreads evenly through the cytoplasm of the *mylpfa*$^{-}$
213 $^{/-}$*;mylpfb*$^{-/-}$ double mutant (Figure 4J). These findings suggest that Mylpfa and Mylpfb
214 together are essential to localize MyHC to the fast muscle cell's periphery where the
215 thick filaments interdigitate with pre-myofibrils to initiate myofibril growth (Figure 4K).

216

217 **Mylpf dosage shows a cooperative effect on myofibril phenotypes**

218 The Mylpf phenotypes appear to arise in a dose sensitive fashion. To test how well
219 dosage predicts phenotype, we calculated dose predictions in mutants using the wild-
220 type ratio of Mylpfa and Mylpfb, ranging from zero in completely null *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ to
221 a dose of one in WT *mylpfa*$^{+/+}$*;mylpfb*$^{+/+}$ (Table S1). We then fit gene dosage and the
222 associated phenotypic outcomes to the Hill equation (Figure 5A), a standard approach
223 for assessing dose-response relationships, which allows us to find the half-maximal
224 dose and measure cooperativity of the response. The 26 hpf myofibril bundle widths
225 show a close correspondence (R^2=0.84) to dosage at that stage (7:1, 7 Mylpfa per 1
226 Mylpfb). We then measured muscle phenotype in animals fixed at 2, 3, and 6 dpf and
227 modeled the dose response using the 2:1 Mylpfa to Mylpfb ratio found at later
228 timepoints (Table S1, Figure 5B-D). At all timepoints we found a sigmoidal gradient of
229 fast-twitch myofibril bundle widths from the double mutant (narrowest/none) to wild-type
230 (widest) genotypes, including the 6 dpf timepoint where we analyzed all nine genotypes
231 of the *mylpfa;mylpfb* mutant series (Figure 5A-E). The degree to which myofibrils fill
232 muscle changes more dramatically than the somite muscle's size (Figure 5D-F). The
233 changes in myofibril bundle width can be explained by the underlying loss of localization
234 between F-actin and MyHC, which by 48 hpf both localize poorly with one another and
235 to the sarcomere in the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant (Figure 5G-J). Across all
236 genotypes, timepoints, and assays, mRNA and protein dose are better predictors of
237 phenotype than gene dose (Table S2). Early Mylpf dosage has long-lasting impact,
238 since the four genotypes with lowest Mylpf dose in embryos are inviable by adulthood,
239 whereas animals showing high embryonic Mylpf levels (such as the *mylpfb*$^{-/-}$ mutant)
240 can grow to a normal adult size (Figure S6). The response curves show a half-maximal
241 protein dose (EC50) of 0.38 on average, suggesting that the Mylpf protein is over
242 halfway to its maximal effect in the double heterozygote (*mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$, Mylpf
243 dose=0.5). Furthermore, by 48 hpf the response curves consistently have high Hill
244 slopes (>4), indicating that Mylpf levels have cooperative effects in the muscle (Figure
245 5K). This cooperativity suggests that the presence of Mylpf enhances the impact of

246 additional Mylpf levels on sarcomere growth and thereby on myofibril bundle size, with
247 lasting effects on animal health.
248
249 **Slow-twitch muscle compensates for impaired fast-twitch muscle**
250 Zebrafish movement is driven largely by force generated in somite muscle, which is
251 primarily fast-twitch [26,27]. Consistent with its fast-twitch impairment, our prior work
252 showed that *mylpfa* is essential to muscle strength and fish movement, leading to a
253 severely impaired escape response in the *mylpfa*$^{-/-}$ mutant [33]. In further support, high-
254 speed imaging shows that the sharp flexure driving initial escape response (the 'C-
255 bend') is consistently absent from this mutant (Figure 6A-C, Video 2). To learn how
256 behaviors change through time, we assessed voluntary movement by imaging larvae in
257 25-minute sessions daily from 2 to 6 dpf. Swimming became more frequent in both
258 genotypes through this developmental interval. Surprisingly, by 6 dpf the *mylpfa*$^{-/-}$
259 mutant moves further than their wild-type siblings (Figure 6D). As expected, the *mylpfa*$^{-/-}$
260 mutant cannot swim quickly at any of the timepoints, but by 5 dpf it increases the rate of
261 slow movement compared to its wild-type sibling (Figure 6E-F). To test whether the
262 remaining muscle function came from *mylpfb*, we also examined animals carrying all
263 nine combinations of *mylpfa* and *mylpfb* zygosity at 6 dpf (Figure 6G-J). Consistent with
264 the *mylpfa*$^{-/-}$ single mutant, the overall speed is not decreased in the *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$
265 double mutant, which travels at least as far as wild-type siblings during the imaging
266 period (Figure 6G). However, the highest movement speed is reduced in the *mylpfa*$^{-/-}$
267 animals and also in the *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$ double mutant, which do not swim faster than
268 70 mm/sec (Figure 6H), consistent with the escape response defect. The genotypes
269 with severe loss of Mylpf function increase the percent of time that they swim slowly
270 (20-40 mm/sec), enough to compensate for the loss of movement at high speed (>70
271 mm/sec) (Figures 6I, J, S7). This increased frequency of slow movement may be
272 explained in part by increased slow-twitch myofibril bundle size, which is unchanged
273 across the mutant series at 2 dpf but expands in the severe Mylpf mutants by 6 dpf
274 (Figures 6K-M, S8). We also find increased movement in the *mylpfa*$^{+/-}$;*mylpfb*$^{+/-}$ double
275 heterozygote, which moves faster on average (and therefore further) than wild-type
276 (Figure 6G, I, J Pink), by swimming more often at slow and rapid speeds. This double

277 heterozygote has a modest fast-twitch deficit and no slow-twitch hypertrophy at 6 dpf
278 (Figure 5D, 6M Pink). Together, these findings suggest that slow-twitch muscle
279 utilization and size compensate proportionately with fast-twitch muscle impairment
280 (Figure 6N).
281
282 **Mylpfa-GFP and Mylpfb-GFP restore *mylpfa*$^{-/-}$ muscle in proportion to dosage**
283 To further test parallels between Mylpf expression levels and myofibril width, we made a
284 *mylpfa-GFP* transgenic line, *Tg(mylpfa:mylpfa-GFP)mai102,* and found that it can
285 restore myofibril formation in the *mylpfa*$^{-/-}$ mutant (Figure 7A-H). The degree of rescue
286 (Figure 7G, H) and the brightness of Mylpfa-GFP (Figure 7B) expression varies among
287 individuals, and we reasoned that Mylpf dosage may impact the degree of rescue. This
288 transgene uses the *mylpfa* promoter and drives GFP-tagged Mylpfa protein at around
289 the same average abundance as the native wild-type Mylpfa (Figure 7C, D). The native
290 Mylpfa protein also varies in expression, though not as widely as Mylpfa-GFP (Figure
291 7D). The transgene's fluorescent signal corresponds with protein abundance as
292 measured in western blot, so we used fluorescence as a proxy for protein levels (Figure
293 7E). Both the myofibril bundle width and the sarcomeric fraction for F-actin correlate
294 closely with Mylpfa-GFP abundance within the *mylpfa*$^{-/-}$ genotype; the dimmer fish have
295 widths that are similar to their non-transgenic siblings and the brightest fish overlap with
296 the wild-type range (Figure 7I, J). However, in zebrafish that are wild-type for *mylpfa*,
297 the transgene causes no significant improvement in fast-twitch myofibril bundle size. We
298 found a similar restoration of myofibril bundle formation in the *mylpfa*$^{-/-}$ mutant animals
299 that express Mylpfb-GFP from transgene *Tg(mylpfa:mylpfb-GFP)mai103* (Figures 7F-J,
300 S9). The equivalent rescue by Mylpfb-GFP supports our proposal that gene expression
301 levels dictate the difference in *mylpfa* and *mylpfb* requirements, rather than protein
302 sequence. Likewise, we found comparable effects when we examined the effects of
303 Mylpfa-GFP at 6 dpf (Figure S10). Not only does the transgene increase the fast-twitch
304 myofibril bundle width in *mylpfa*$^{-/-}$ mutant, but it also suppresses the compensatory
305 hypertrophy of slow-twitch myofibrils in the same animals, restoring both fiber types to
306 their wild-type state (Figures 7, S10). Together, these findings support a model where

307 Mylpfa and Mylpfb are required in proportion to their wild-type dosage, but neither is
308 sufficient to promote excess myofibril bundle growth in a wild-type animal.
309
310 **Mylpf activity is conserved between zebrafish and human**
311 Since Mylpfa and Mylpfb have 80-81% amino acid identity with human MYLPF, we
312 reasoned that Mylpf protein function may be conserved in these two vertebrate species.
313 To test the compatibility of human MYLPF with zebrafish genetics, we generated a
314 transgene that expresses GFP-fused MYLPF under the *mylpfa* promoter,
315 *Tg(mylpfa:MYLPF-GFP)mai104* (Figure 8A). We find that this transgene is capable of
316 restoring myofibril bundle size in the *mylpfa*$^{-/-}$ mutant, to a degree that is linearly
317 proportionate with GFP fluorescence, up to the wild-type level (Figure 8B-E). To
318 compare rescue experiments, we scaled each to the non-transgenic wild-type mean for
319 that experiment (2 or 6 dpf). Across timepoints, we find no significant difference in slope
320 between animals expressing MYLPF-GFP, Mylpfa-GFP, and Mylpfb-GFP on scaled
321 myofibril bundle width (Figure 8F). This analysis shows that human MYLPF protein can
322 replace zebrafish Mylpfa and suggests that it acts with an efficiency that is similar to the
323 zebrafish Mylpfa and Mylpfb proteins.
324
325 **A MYLPF allele found in DA does not support myofibril formation**
326 We tested whether a DA-causing MYLPF allele can rescue muscle structure as
327 efficiently as its wild-type counterpart. This allele, encoding protein variant G163S, is
328 inherited heterozygously in two individuals with DA, suggesting that partial loss of
329 function may lead to the disease [33]. Analysis using F0 transient mosaic animals
330 suggests that the G163S MYLPF variant does not promote myofibril bundle formation in
331 the *mylpfa*$^{-/-}$ mutant (Figure 9A-E). Likewise, this did not rescue the F-actin sarcomeric
332 fraction, GFP sarcomeric fraction, nor the correlation between actin and MyHC labels
333 (Figures 9F-H, S11). We confirmed the failure to rescue by live imaging the *mylpfa*$^{-/-}$
334 mutant expressing a stable transgenic, *Tg(mylpfa:MYLPF*^{G163S}*-GFP)mai119* alongside
335 the actin transgenic label *TgBAC(six1b:Lifeact- mTurquoise2)mai121* (Figure 9I-I").
336 These findings suggest that a nonfunctional MYLPF allele can cause Distal
337 Arthrogryposis when heterozygous.

338

339 **Mylpf shows consistent effects on muscle structure from 2 to 6 dpf**

340 We hypothesized that Mylpf dosage may regulate sarcomere assembly with equal
341 potency throughout early development. To understand elements of Mylpf function that
342 may be constant through early development, we scaled our *mylpfa;mylpfb* mutant
343 datasets to the wild-type value at each timepoint and then overlaid measurements,
344 yielding a unified regression model with upper asymptote around 1, EC50 near 0.5, and
345 Hill slope >4 (Figure 10A). We then tested this model by comparing it to findings from
346 animals that express Mylpfa-GFP, Mylpfb-GFP, or MYLPF-GFP at 2 or 6 dpf. To model
347 dosage in individual animals, we converted GFP brightness in each animal to a
348 predicted Mylpf dosage, scaled using the 2:1 ratio of Mylpfa to Mylpfb protein (Figures
349 10B, S12). The transgenic dataset does not form a lower asymptote because it was
350 generated using the *mylpfa*$^{-/-}$ mutant line. We supplied a lower asymptote from the
351 *mylpfa;mylpfb* mutant series, to generate a sigmoidal curve suitable for testing the other
352 three Hill parameters (Figures 10B, S12). The transgenic regression shows an upper
353 asymptote of 1, supporting our claim that Mylpf is insufficient to expand myofibril widths
354 (Figure 10B). In support of our claim that 50% Mylpf dosage impacts phenotype, the
355 transgenic regression shows an EC50 that is close to 0.5. Myofibril bundle widths show
356 a higher EC50 and lower Hs value in transgenic analysis compared to knockout
357 analysis, suggesting that the GFP tag impedes Mylpf efficiency (Figure 10A, B), though
358 we do not detect this difference using the sarcomeric fraction measure (Figure S12).
359 Across all analyses, we find a Hill slope that is considerably larger than one, supporting
360 our claim that Mylpf proteins cooperatively promote the cytoskeletal arrangements that
361 strengthen muscle throughout the first week of development in zebrafish, including
362 myofilament localization, sarcomere formation, myofibril widening, and myofibril bundle
363 growth.

364

365 # Discussion

366 Our central finding is that Mylpf function is essential for the defining feature of striated
367 muscle, the overlap of sliding filaments [37,38]. We show that the protein is necessary but

368 not sufficient for myofibril formation in fast-twitch muscle; however, this binary language
369 does not fully capture the sensitivity of muscle development to Mylpf loss. Our work was
370 initiated by the intriguing finding that *mylpfa* is required for muscle structure, but *mylpfb*
371 seems not to be. This difference could be explained by the timing of gene expression,
372 since *mylpfa* mRNA abundance climbs faster than *mylpfb,* with a 7:1 ratio at 1 dpf and
373 2:1 ratio by 2 dpf. However, the large discrepancy at 1 dpf does not explain why the
374 phenotypic differences between *mylpfa*$^{-/-}$ and *mylpfb*$^{-/-}$ persist so long after *mylpfb* levels
375 increase. Another simple explanation is that Mylpfa protein has some function that
376 Mylpfb is lacking; however, that model would not explain our observation that both
377 Mylpfa-GFP and Mylpfb-GFP rescue the *mylpfa*$^{-/-}$ mutant with similar efficiency. Instead,
378 we propose that Mylpf protein dosage can explain the phenotypes we observed across
379 all Mylpf genotypes and timepoints, when modeled using a Hill regression (Figure 10A,
380 B). This regression model is a popular tool for biochemical analysis and has also been
381 used previously to investigate the dose sensitivity of gene promoters *in vivo* [1,39]. This
382 model suggests that the *mylpfb*$^{-/-}$ mutant expresses total Mylpf protein just slightly above
383 its functional saturation point, around two thirds the wild-type level. For Mylpf, we find an
384 overtly sigmoidal relationship between dosage and phenotype from the null to wild-type
385 genotypes; however, some Hill models can seem nearly binary or nearly linear in the
386 range between null and wild-type (Figure 10C). An overtly sigmoidal relationship could
387 arise because Mylpf is a structural protein that functions stoichiometrically in proportion
388 to other sarcomeric proteins, consistent with recent proteomic work in cardiac muscle [40].
389 These findings raise the intriguing possibility that many other muscle proteins have
390 stoichiometric dose-dependency, because they function together across entire myofibril
391 bundles during sarcomere formation and force production.
392
393 The relationship between Mylpf protein dosage and sarcomeric phenotype has a large
394 Hill slope, suggesting that Mylpf proteins function cooperatively during sarcomere
395 formation [41]. The Hill equation is traditionally used to describe cooperativity in ligand-
396 substrate binding, though later steps may also be synergistic (Figure 10D). For
397 instance, each Mylpf protein should help to stabilize MyHC during thick filament
398 production (step 2). We think that the core step regulated by Mylpf may be myofilament

399 localization (step 3), since we see poor colocalization of F-actin with MyHC across
400 multiple timepoints of development. Alternatively, the myofilaments may be returned to
401 the cytoplasm after briefly localizing. Both scenarios result in mislocalized thick
402 filaments, which cannot interdigitate with thin filaments (step 4) [4], leading to loss of the
403 tension-based feedback loops shown to regulate sarcomere growth (steps 5, 6) [6,42].
404 These feedback loops may be disrupted by impaired Mylpf stability, since point
405 mutations in RLCs are known to affect step size [14]. Prior work indicated that Mylpf also
406 controls overall muscle growth, since the mouse *Mylpf*$^{-/-}$ knockout is born without
407 skeletal muscle [22] and myosin heavy chains are needed to support myofibril
408 organization and muscle growth [43–45]. The *mylpfa*$^{-/-}$ mutant shows reduction in somite
409 growth by 6 dpf, when assessed in live animals. We find similar average measurements
410 using phalloidin on fixed samples, though here the difference in size between genotypes
411 does not rise to statistical significance. The *mylpfa*$^{-/-}$ muscle growth defect is consistent
412 with prior work showing that MyHC itself is required for myofiber growth and the known
413 feedback loops between strength and growth [43,44]. Furthermore, the effects on myofibril
414 bundle growth are consistently larger than those on somite growth. We think that the
415 primary role for Mylpf is in sarcomere formation, because this function occurs
416 immediately and has cascading effects.
417
418 Given its complete sarcomere loss in fast-twitch muscle (over 95% of total muscle), we
419 were surprised to find the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant moving as far as its wild-type
420 siblings. The *mylpfa*$^{-/-}$ mutant and *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant have classic signs of
421 fast-twitch deficiency including an absence of C-bends and impaired escape response,
422 consistent with previous work on zebrafish showing loss of function in fast-twitch muscle
423 [33,46,47]. We propose that the *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$ double mutant swims so far because they
424 increase the frequency of slow-twitch movements, which arise independently of fast-
425 twitch activity [48]. Some genotypes, such as the *mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$ double heterozygote
426 upregulate these slower movements and have functional fast-twitch muscle, allowing
427 them to outpace their wild-type siblings. Consistent with our model, zebrafish lacking
428 fast-twitch acetylcholine receptors upregulate slow-twitch synapses and behaviors by
429 adult stages [46]. Alternatively, the behavioral compensation in Mylpf-deficient animals

430 may be caused by excess exercise of the slow-twitch muscle fibers [28]. The importance
431 of muscle hypertrophy and neural activity could be resolved by investigating muscle
432 strength, calcium signaling, and synapse morphology through time in Mylpf mutants.
433 Investigating this compensatory mechanism may provide crucial insights for human
434 health, since many muscle diseases disproportionately affect the fast-twitch fiber type
435 [28].
436
437 The impact of Mylpf dosage on muscle formation may not be restricted to zebrafish. Our
438 G163S findings suggest that the allele has little or no activity, so a person heterozygous
439 for this variant may have dominantly inherited DA because they have around half the
440 normal MYLPF activity. Likewise, we show that zebrafish with 50% Mylpf dose (the
441 *mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$ double heterozygote) have impaired myofibril bundle assembly.
442 Similarly, a heterozygous effect is also found in *Drosophila,* wherein RLC$^{+/-}$ flight
443 muscles (fast-twitch) show impaired myofibril assembly [17]. No defects were reported in
444 a study of the mouse *Mylpf*$^{+/-}$ heterozygote, although embryonic myofibril bundle widths
445 were not measured [22]. However, dosage appears to be broadly relevant in mammals,
446 since studies of pigs and buffalo show that Mylpf mRNA levels can predict meat quality
447 [49,50]. Myosin light chain dosage also appears to impact heart muscles, since activation
448 of cardiac myosin light chain kinase is necessary and sufficient to initiate the formation
449 of sarcomeres in cultured heart cells [51]. We propose that the relationship between
450 sarcomeric protein dosage and myofibril expansion may have applications in many
451 contexts where muscle strength is desired including farming, exercise physiology, and
452 human health.

Acknowledgements

The fish facilities at the University of Maine and The Ohio State University (OSU) provided excellent animal care. Dr. Joy-El Talbot from Iris Data Solutions developed a FIJI macro for sarcomere fraction quantification. Dr. April DeLaurier provided feedback in the writing process. Dr. Clarissa Henry generously provided mentorship and resources to the Talbot lab. Dr. Thomas Gallagher and Collin Whitlock facilitated data sharing between OSU and the University of Maine. Gursimran Dhillon provided training on technique and data analysis for RT-qPCR. Paula Monsma and Dr. Sarah Cole supported imaging on the electron and confocal microscopes at OSU. Dr. Zhengxin Ma supported training and usage of the Zeiss LSM980 confocal in the Microscopy and Image Analysis Core facility at the University of Maine (RRID:SCR_025784). Antibodies mMac [52] and A4.1025 [53] were provided by DSHB at the University of Iowa. This project was supported by a University of Maine Institute of Medicine Seed Grant, and NIH grant R15AR081019 to JCT, NIH R15GM140409 to JBK, by NIH grants GM088041 and GM117964 to SLA, and COBRE 1P20GM144265 to JBK and JCT.

Materials and Methods

Zebrafish husbandry and established strains

Fish strains were maintained using standard methods [54]. All animal protocols used in this study were approved by the Institutional Animal Care and Use Committees at The Ohio State University (2012A00000113) and the University of Maine (A2022-09-05, A2025-10-01). Zebrafish staging followed established metrics [55]. We genotyped the *mylpfa*^{oz30} 20 bp deletion allele [33] using primers (F 5'-TCTCTACAGGCCAGCTGAATG'3', R 5'-ACCCTTCAACTTCTCTCCGAAC'3') that amplify a 116 bp product in wild-type fish and a 96 bp product in the *mylpfa*^{oz30} homozygote. We genotyped the *mylpfa*^{oz43} +1 bp insertion allele [33], (primers 5'-GCTTCATTGCTGTCAGGATAGAG-3' and 5'- ACCCTTCAACTTCTCTCCGAAC-3'), followed by digestion with BsmFI restriction enzyme (NEB, New England Biolabs) which cuts the wild-type but not mutant amplicon. Homozygotes for both *mylpfa*^{oz30} and *mylpfa*^{oz43} can be consistently sorted from wild-type siblings based on pectoral fin immobility, a slowed escape response, and reduced birefringence [33]. The *mylpfa*^{oz30} allele was used for all *mylpfa*$^{-/-}$ mutant analysis in this study, except for Figure S5E. Zebrafish lines in this study were maintained on the AB wild-type background.

Transgene construction

The four Mylpf-containing plasmids constructed for mosaic analysis in this study were *pMylpfa:mylpfa-GFP*, *pMylpfa:mylpfb-GFP*, *pMylpfa:WT-MYLPF-GFP*, and *pMylpfa:G163S-MYLPF-GFP.* Briefly, we obtained GeneArt-strings (Invitrogen) of the sequence encoding Mylpfa*,* Mylpfb*,* wild-type human MYLPF or MYLPF p.Gly163Ser (G163S)*,* proteins. These coding sequences were inserted into a plasmid containing the *mylpfa* promoter flanked by TOL2 sites [34,56,57]; GFP was linked to the 5' end of Mylpf genes, connected via a sequence encoding a short flexible peptide (GGGGSGAT). The LifeAct transgene was made by modifying a previously described six1b BAC [58] by integrating coding sequence for LifeAct-mTurquoise2 (LifeAct-mTq) protein after the *six1b* start codon. Plasmid and BAC injection was accompanied by co-injection of

497 zebrafish-optimized nuclear transposase ZT2TP [59,60]. For transient analysis, animals
498 with high levels of mosaicism were selected for imaging and all experiments were
499 repeated on at least two independent injection days. Stable transgenic lines were
500 generated for five transgenes, *Tg*(*mylpfa:mylpfa-GFP*)*mai102*, *Tg*(*mylpfa:mylpfb-*
501 *GFP*)*mai103, Tg(mylpfa:MYLPF-GFP)mai104, Tg(mylpfa:MYLPF^{G163S}-GFP)mai119,*
502 and *TgBAC(six1b:Lifeact- mTurquoise2)mai121*. For stable transgenic analysis,
503 progeny from one founder (*Lifeact- mTurquoise2*), two founders (*mylpfa-GFP, MYLPF-*
504 *GFP*) or four founders (*mylpfb-GFP, MYLPF^{G163S}-GFP*) were analyzed.

505

506 **Construction of *mylpfb^{oz39}***

507 The *mylpfb^{oz39}* mutation was generated using established CRISPR-Cas9 protocols [61].
508 One-cell embryos were co-injected with Cas9 mRNA and guide RNA targeting
509 sequences in exon 2 (5'- GGAAAACAGTAAAGTTGATG -3'), raised to adulthood, and
510 outcrossed to identify germline-transmitting founders. F1 progeny were screened using
511 High-Resolution Melting Analysis (HRMA) (primers 5'-
512 CCCTCTCTAAAACAAACAGGCTTTC-3' and 5'-GGTAAGTGAAGATTTGGACAACTC-
513 3') to identify founders carrying the frameshifting -5 bp deletion *mylpfb^{oz39}* allele. The
514 allele sequence was confirmed via Sanger sequencing of homozygotes and the founder
515 was outcrossed for at least two generations before conducting experiments. We identify
516 the *mylpfb^{oz39}* lesion via TaqMan real-time PCR genotyping assays. The TaqMan real-
517 time PCR method employs custom primers and TaqMan probes to distinguish between
518 wild-type and *mylpfb* mutant amplicons. The forward and reverse primers used are 5'-
519 AAGGATGATTTGAGGGATGTG-3' and 5'-TTTCACCAAACATGGTAAGGAA-3',
520 respectively; fluorogenic probes are (56-FAM/CC +CA+T +C+AA +CT+T TAC
521 /3IABkFQ/) for wild-type amplicons and (5SUN/CC C+CA +T+T+T A+CT G+TT
522 /3IABkFQ/) for mutant.

523

524 **RNA isolation and cDNA synthesis**

525 We determined absolute mRNA expression levels across developmental stages (24–72
526 hours post-fertilization, hpf) using total RNA isolated from AB wild-type embryos.
527 Embryo pools of 40 (24 hpf), 30 (36 hpf), 20 (48 hpf), and 10 (72 hpf) were
528 homogenized in Trizol using 0.5 mm zirconium beads (Laboratory Supply Network), and
529 RNA was subsequently purified using the Direct-zol RNA Miniprep Kit (Zymo Research).
530 We assessed the relative mRNA expression in the *mylpfa;mylpfb* mutant series using
531 decapitated 48 hpf embryos obtained from incross of adult *mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$ double
532 heterozygotes. The 'head' DNA was used for genotyping, while bodies were
533 immediately snap-frozen at −80°C in individual tubes. After genotyping, RNA was
534 extracted from pools of three embryos per genotype. Purified total RNA (1 µg for AB
535 wild-type or 100 ng of the *mylpfa;mylpfb* mutant series) was reverse transcribed using
536 the SuperScript IV Reverse Transcriptase Kit (Invitrogen), along with oligodT_{20} primers
537 (Invitrogen) and murine RNase inhibitor (NEB), according to the manufacturer's
538 protocol.

539

540 **RT-qPCR**

541 RT-qPCR was performed using SYBR Green master mix (Bio-Rad) on CFX384
542 thermocycler (Bio-Rad), with 60˚C annealing and 30 second extension. Each reaction
543 used 15 ng of cDNA in an 10 µl volume, performed in technical triplicate. The *mylpfa*
544 qPCR generates an 80 bp amplicon using the Forward primer, 5'-
545 TCAATGGGCCAGCTGAATGT -3′ and Reverse primer, 5′ -
546 ACGGTGAAGTTGATTGGGCC -3′. The *mylpfb* qPCR generates an 83 bp amplicon
547 using the Forward primer, 5'- AGGAATACAAGGAGGCTTTCACA -3′ and Reverse
548 primer, 5′ - TGCCAGCACATCCCTCAAAT - 3′. The *rpl13a* transcript was selected as
549 the reference gene due to its stable expression during zebrafish embryonic
550 development [62]. The *rpl13a* qPCR uses primers: Forward 5'-
551 TCTGGAGGACTGTAAGAGGTATGC-3′ and Reverse 5'-
552 AGACGCACAATCTTGAGAGCAG-3′. Relative mRNA quantification was performed
553 using the $2^{-\Delta\Delta Cq}$ method [63] after normalization to the mean of all wild-type groups for

554 each target gene. Each experiment was repeated on at least three biological replicates
555 before statistical analysis.

556

557 **Absolute quantification of *mylpfa* and *mylpfb* mRNA**

558 For absolute quantification of mRNA copy number, we generated *mylpfa* and *mylpfb*
559 DNA templates (called "standard DNA") from 72 hpf cDNA. For *mylpfa* standard DNA, a
560 563 bp was amplified from the zebrafish *mylpfa* gene using the Forward primer, 5'-
561 ACTACGCGGCTTCAGACTTC -3′ and Reverse primer, 5′ -
562 TCCTTCTCCTCTCCGTGTGT - 3′. For *mylpfb* standard DNA, a 555 bp was amplified
563 from the zebrafish *mylpfb* gene using the Forward primer, 5'-
564 TGTCAGTCACATCAACGTGGA -3′ and Reverse primer, 5′ -
565 CCTCTCCATGTGTGATGACGT - 3′. DNA standards were purified using a gel
566 extraction kit (Thermo Fisher) and sequence was confirmed. To calibrate for copy
567 number, we amplified a series of standard DNA concentrations spread across ten four-
568 fold dilutions on each qPCR plate, with predicted concentration ranging from 4.5
569 copies/reaction to 4.92×10^{6} copies/reaction. The graph of the curve was used to
570 calculate the number of cDNA molecules tested in the same reaction plate as the
571 standard, and a ratio between *mylpfa* and *mylpfb* copies is presented.

572

573 **Immunohistochemistry and HCR ISH**

574 Whole mount immunohistochemistry used established markers and techniques [64].
575 Briefly, the larvae were fixed in 4% paraformaldehyde and washed in PBST (1X PBS,
576 0.1% Tween-20) before immunolabel. Embryos older than 1 dpf were gently
577 permeabilized using a brief treatment with 0.001% Proteinase-K, washed repeatedly in
578 PBST, then blocked for at least two hours before overnight incubation at 4°C with
579 primary antibodies diluted in blocking solution. We used primary antibodies for the
580 myonuclei (1:500, Rbfox1l) [65]; α-Actinin (1:500, A7732, Sigma); Myomesin [1:30, mMac,
581 Developmental Studies Hybridoma Bank (DSHB)] [52], MyHC (1:1000, A4.1025, DSHB)
582 [53]. F-actin labeling uses Alexa Fluor-568 conjugated phalloidin, added during the

583 secondary antibody step (1:50, Thermo Fisher). HCR ISH procedures use established
584 protocols with probes and amplifiers supplied by Molecular Instruments [66]. Our HCR
585 mixtures included nine unique *mylpfa* probes, five *mylpfb* probes, and twenty *myl10*
586 probes. To compare HCR ISH signal intensity, separate fish were labeled with *mylpfa*
587 and *mylpfb* probes, both in the 488 nm excitation channel, and then imaged with
588 matched settings on a dissecting microscope before analysis in FIJI.

589

590 **Protein modeling**

591 Mylpf proteins were modeled in complex with the binding region of zebrafish Myhz1.3
592 (zebrafish) or MYH3 (human) using Robetta [67] and the models were visualized using
593 PyMol (Schrödinger).

594

595 **Antibody production, validation on western blot**

596 The zebrafish Mylpf antibody was generated by Boster Bio, using zebrafish Mylpfa (AA
597 14 to 167 of NP_571263.1) in complex with a Mylpf-binding region that is found in
598 several MyHC proteins (RRESIYTIQYNIRSFMNVKHWPWMKVYYKIKPL). Western blot
599 in the Talbot Lab showed only one band each for Mylpfa (16 kD) and Mylpfb (18 kD)
600 and no label in the MyHC size (252 kD). For western blotting, animals were raised to the
601 desired stage, then pools or single animal were homogenized in 100 µl (pools) or 20 µl
602 (single) 1X LDS sample buffer solution (Bolt; Invitrogen) using 0.5 mm zirconium beads
603 (Laboratory Supply Network). In the western blot showing Mylpfb and Mylpfa protein
604 abundance in wildtype (AB) animals, we used pools of 40, 30, 20 and 10 animals at 24,
605 36, 48, and 72 hpf respectively to normalize for zebrafish growth. Prior to western blot,
606 animals were decapitated and deyolked to enrich for somite muscle and provide
607 genotyping materials. For the *mylpfa;mylpfb* mutant series, we ran the protein from one
608 animal per genotype across the nine genotypes generated by incross of
609 *mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$ double heterozygotes. Homogenized samples (20 µl) were run on a 4-
610 20% gradient acrylamide gel (Bio-Rad). The protein samples were transferred to
611 Polyvinylidene difluoride (PVDF) membranes (Millipore Sigma) which were blocked (5%

612 nonfat dry milk), then replaced with primary antibody solution and then secondary
613 antibody solution. Primary antibodies included mouse polyclonal anti-MyHC (A4.1025,
614 DSHB, 1:1000) to detect MyHC protein while rabbit polyclonal anti-Mylpf (Boster,
615 1:1000) detected both Mylpfa and Mylpfb protein. The secondary antibodies used were
616 donkey-anti-mouse-800 (1:10000 dilution) and donkey-anti-rabbit-680 (1:10000
617 dilution). The resulting fluorescence was measured by scanning the blot with an
618 Odyssey Infrared Imager (LICOR). The Mylpfa to Mylpfb ratio was quantified using
619 GelBox software [68] and FIJI [69].

620

621 **Confocal imaging**

622 Confocal images in Figure S5 were collected using an inverted Nikon TiE microscope
623 equipped with an Andor Revolution WD spinning disk confocal system. Confocal images
624 in Figure 2A-D' and S8 were collected using a Zeiss LSM980 confocal microscope. All
625 other confocal images were acquired using a Leica TCS Sp8 confocal microscope and
626 processed with Lightning in LasX software. Myotomes were imaged over the mid yolk
627 tube region of embryos and larvae. Image settings including laser power, gain, and
628 confocal export protocols were standardized for each experiment. The channels tool is
629 used to generate colorblind friendly images, such as converting red to magenta.

630

631 **GFP brightness analysis**

632 We evaluated the GFP brightness using the mean grayscale value of the GFP channel
633 in FIJI. For stable transgenic lines, we drew ROI around the fast-twitch region of the
634 somite. In mosaic animals, we drew ROI around each muscle fiber because there was
635 often a high level of variation within a single image. Muscle was considered non-
636 transgenic if the GFP brightness is <10 (0-255) and considered transgenic when in the
637 10-255 range. MATLAB was used to calculate the average of each mean grayscale
638 value per image followed by statistical comparisons in JMP software.

639

640 **Sarcomeric fraction calculation**

641 We developed a metric to quantify the degree to which muscle proteins localize to
642 sarcomeres, termed the sarcomeric fraction (Figure S4). Grayscale intensities were
643 calculated across Region of Interest (ROI) lines along the length of a myofiber in FIJI,
644 with ten to thirty ROIs measured per image. The periodicity of these intensities was
645 determined using the 'distance between peaks' function in MATLAB. Peak-to-peak
646 distances per intensity were determined per image, and bootstrap confidence intervals
647 were calculated using variation between images, shown via histogram to represent the
648 Fraction of Peaks/micron (μm). Signal was considered 'sarcomeric' in different intervals
649 ('bin sets') for markers with full-sarcomere (F-actin, Myomesin, Actinin) or half-
650 sarcomere repeat (MyHC, Mylpf-GFP), or an even blend of the two patterns (LifeAct)
651 (Figure S3). Sarcomeric LifeAct shows peak signals in the 0.7-1.1 μm and 1.55-2.30 μm
652 range. Sarcomeric phalloidin (F-actin) also labels Z-disks and thin filaments, showing
653 peak signals in the 0.7-1.0 μm range and 1.55-2.15 μm range. For consistency, we
654 used the F-actin bin set for Myomesin and Actinin (Figure S4). Sarcomeric MyHC and
655 Mylpf-GFP show two peak signals in bins 0.6-1.2 μm and 1.75-2.05 μm (Figure S4).
656 These bins were determined by examining histogram plots of the localization of
657 sarcomeric signals during preliminary analysis (examples in Figure S4) and then applied
658 evenly to subsequent analyses. The sarcomeric bins occupy 0.9 μm of the 4 μm
659 sampled, so a uniformly distributed label would score 0.225 (shown with a brown line in
660 graphs) whereas a 1.0 indicates full localization of signal to sarcomeres. The
661 sarcomeric fraction per animal represents the ratio of image periodicity within the
662 sarcomeric bin to the overall periodicity of that marker, averaged across all ROIs (Figure
663 S4).

664

665 **Colocalization**

666 The extent of colocalization between confocal channels was determined using the
667 COLOC 2 tool in FIJI. Briefly, we identified large regions of interest (ROI) areas where
668 labeling for the markers compared was homogenously bright. This selection avoided
669 penetration issues, because the MyHC label is brighter at the surface of somites than it
670 is deeper. Typically, the area fills at least half of the fast-twitch muscle region of a

671 somite. We recorded the Spearman's rank correlation coefficient, Mander's tM1, and
672 Mander's tM2 for all images. Measurement was repeated if a large disconnect was
673 found between tM1 and tM2 suggesting that an ROI extended to a region of poor
674 labeling; slight ROI adjustments brought the Mander's measures into alignment and
675 produced a trustworthy Spearman's correlation. These Spearman's correlations were
676 tabulated and subjected to our standard statistical analysis.

677

678 **Cross-sectional analysis**

679 To analyze muscle size in cross-sectional view, confocal stacks were acquired with 0.5
680 µm step size, then orthogonal views were exported to FIJI. Somite area is the total area
681 enclosed within ROIs drawn around the somite. Based on cell position and markers we
682 can also identify fiber types, with the superficial layer being slow-twitch and all deeper
683 layers being fast-twitch [26].To measure the cross-sectional area of myofibrils within slow-
684 twitch fibers, we drew ROI encircling the desired fiber type and then performed auto
685 local thresholding with the Bernsend method, setting radius to 15 and the T parameter
686 to 7.5, before measuring area within threshold. To determine protein positions within
687 fast-twitch muscle fibers (Figure 4G-J), a mask was drawn manually, circling the
688 borders of each cell. A custom MATLAB script determined image brightness at each
689 pixel, moving from the periphery of the cells to the center. These edge-center distances
690 were scaled to micron lengths, averaged, and bootstrap confidence intervals were
691 calculated across samples of a given genotype. Graphs with confidence intervals were
692 then overlaid for each genotype.

693

694 **Myofibril bundle width measurements**

695 To measure myofibril width, we drew ROI lines across the narrow axis of the myofibril in
696 FIJI. We typically measured thirty myofibrils bundles per image. Each measurement
697 was made at a widened point of each myofibril, with measurements distributed evenly
698 across the fast-twitch muscle region. For mosaic animals we measured both GFP+ and
699 GFP-negative myofibril bundles within the same confocal slice. Images were included

700 only if they contained at least ten GFP+ myofibril bundles. For stable transgenic lines,
701 we compared measurements from non-transgenic and transgenic siblings. We
702 generated a custom FIJI macro to simplify measurement and compile all measures from
703 a given image. These measurements were averaged per image using MATLAB and
704 then imported to JMP software for statistical analysis.

705

706 **Transmission electron microscopy**

707 For TEM, embryos were raised to 72 hpf, fixed in phosphate-buffered glutaraldehyde,
708 resin embedded, thin sectioned, and contrasted using uranyl acetate and Reynold's
709 lead citrate before imaging an initial dataset on a Tecnai 30 G2 TWIN microscope at the
710 Ohio State University and later dataset on a Tecnai T12 BioTwin microscope at the
711 University of Maine. The two datasets were then merged. In total, we imaged 6 wild-
712 type animals, 6 *mylpfa*$^{-/-}$, 3 *mylpfb*$^{-/-}$, and 6 *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$ double mutant animals,
713 each with multiple sections. We set the mean of measurements within an animal as the
714 single N for statistical comparisons. Myofibril widths were measured between triads and
715 myofibril bundle widths were measured as in light microscopy; both were measured on
716 low-magnification TEM images to sample variation across several myofibers.

717

718 **Behavioral analysis**

719 For tail curvature assessment, individual larvae were placed in a petri dish and prodded
720 with a piece of fishing line, imaged using a GoPro Hero camera with a mounting ring
721 (H12Pro, Back-Bone) set to 240 frames/sec. The frame with highest curvature was
722 transferred to FIJI and curvature was measured along the length of the trunk using the
723 Kappa plugin [70]. For speed analysis, larvae were placed one per well into a 12-well
724 plate with 3 ml of facility water at 6 dpf. Then, they were imaged in a DanioVision
725 system (Noldus) equipped with a magnifier for improved resolution. The plates were
726 incubated at 28˚C in the preheated chamber. The larvae were imaged using the infrared
727 wavelength every 0.033 seconds for a total of 25 minutes, cycling lights off and on every
728 five minutes. Larvae were tracked using EthoVision XT Version 15.0 video tracking

729 software by Noldus. 'Distance traveled' is the sum of distances in this 25-minute time
730 interval. It is important to use software that includes low-speed movement, such as
731 Ethovision. Analysis was limited to speeds under 200 mm/sec because video
732 examination shows that higher-speed measurements are indications of tracking
733 glitches. The 'highest speed' was calculated by sorting data numerically and recording
734 the highest number beneath 200 mm/sec. Average speed was the average of all
735 momentary speeds. Speed is calculated per 0.033 second time interval. Movement is
736 summarized using the Microsoft Excel Histogram function, as the percentage of time
737 that fish swim in a wide range of speed 'bins'. The first bin is set to 1 mm/sec, then 1.09
738 mm/sec, then bins continue using the formula Y=X+Ln(X), where Y is the current bin
739 value and X is the previous bin value. The larvae swim exponentially less frequently at
740 high speeds than low speeds, so this natural-log based formula allows comparable
741 statistical sampling of the whole speed range (Figure S7). Shown plots integrate data
742 collected from the *mylpfb*$^{+/-}$ heterozygous incross, *mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$ double
743 heterozygous incross, and outcrosses of *mylpfa*$^{+/-}$*;mylpfb*$^{+/-}$ to AB (Figure 6D-G) or from
744 *mylpfa*$^{+/-}$ incrosses (Figure 6H-J). Each set of crosses was replicated at least three
745 times with consistent results across experiments.
746
747 **Statistical analysis**

748 We begin multiple comparisons using ANOVA and then use Tukey-Kramer's post-hoc
749 comparison of levels. Pairwise comparisons use a two-sided Student's T-test and
750 matching results are found with the non-parametric Kruskal-Wallis exact test. In all
751 figures, * P<0.05, ** P<0.01, *** P<0.001. Not significant (n.s.) is P>0.1 or P>0.05 as
752 defined in legend. For box plots, the central line shows the median, the upper bound
753 shows the upper quartile, the lower bound shows the lower quartile, and whiskers show
754 1.5x the interquartile range. Each N represents a different animal or in some cases a
755 pool of animals and are listed explicitly in the source data file. For mosaic animals, the
756 N represents the number of animals scored for GFP+ or GFP-negative cells,
757 respectively. Repeated measures of the same animal are averaged and included as a
758 single N. Statistical comparisons were made in Jmp Pro, GraphPad, or MATLAB
759 software. Logistic regression models used the four-parameter Hill equation in GraphPad

with formula y=L+((U-L)*X^Hs)/(X^Hc+EC50^Hs). Variables are defined as phenotypic response (y), dosage (x), upper asymptote (U), lower asymptote (L), EC50, and Hill slope (Hs). Dose modeling in transgenic animals uses the formula WT=1+0.667*(GFP/57), heterozygote=0.667+0.667*(GFP/57), and *mylpfa*$^{-/-}$ =0.333+0.667*(GFP/57), with 57 being the average fluorescence in our images.

Data availability

The data leading to conclusions in this paper are shown in the main text figures and supplementary figures, with source data provided. We are happy to accommodate requests for additional source files.

Code availability

Code used for data analysis in this manuscript is available upon request and from our GitHub account, https://github.com/MuscleZebrafish/MyofibrilQuant.

Competing Interest Statement

The authors declare that they have no competing interests.

Citations

1. Noviello, G. Transcriptional regulation as a dose-dependent process: insights from transcription factor tuning. *Open Biol.* **15**, 240328 (2025).
2. Luis, N. M. & Schnorrer, F. Mechanobiology of muscle and myofibril morphogenesis. *Cells & Development* **168**, 203760 (2021).
3. Berger, J., Berger, S. & Currie, P. D. Mob4-dependent STRIPAK involves the chaperonin TRiC to coordinate myofibril and microtubule network growth. *PLoS Genet* **18**, e1010287 (2022).
4. Sanger, J. W. *et al.* Assembly and Maintenance of Myofibrils in Striated Muscle. in *The Actin Cytoskeleton* (ed. Jockusch, B. M.) vol. 235 39–75 (Springer International Publishing, Cham, 2016).
5. Holtzer, H. *et al.* Independent assembly of 1.6 microns long bipolar MHC filaments and I-Z-I bodies. *Cell Struct Funct* **22**, 83–93 (1997).
6. Weitkunat, M., Brasse, M., Bausch, A. R. & Schnorrer, F. Mechanical tension and spontaneous muscle twitching precede the formation of cross-striated muscle *in vivo*. *Development* **144**, 1261–1272 (2017).
7. Tskhovrebova, L. & Trinick, J. Titin and Nebulin in Thick and Thin Filament Length Regulation. in *Fibrous Proteins: Structures and Mechanisms* (eds. Parry, D. A. D. & Squire, J. M.) vol. 82 285–318 (Springer International Publishing, Cham, 2017).
8. Loison, O. *et al.* Polarization-resolved microscopy reveals a muscle myosin motor-independent mechanism of molecular actin ordering during sarcomere maturation. *PLOS Biology* **16**, e2004718 (2018).

799 9. Lemke, S. B. & Schnorrer, F. Mechanical forces during muscle development. *Mechanisms of*
800 *Development* **144**, 92–101 (2017).
801 10. Hall, T. E. *et al.* In vivo cell biological screening identifies an endocytic capture mechanism
802 for T-tubule formation. *Nat Commun* **11**, 3711 (2020).
803 11. Mao, Q. *et al.* Tension-driven multi-scale self-organisation in human iPSC-derived muscle
804 fibers. *eLife* **11**, e76649 (2022).
805 12. Lowey, S. & Trybus, K. M. Role of Skeletal and Smooth Muscle Myosin Light Chains.
806 *Biophysical Journal* **68**, 7 (1995).
807 13. Heissler, S. M. & Sellers, J. R. Myosin light chains: Teaching old dogs new tricks.
808 *BioArchitecture* **4**, 169–188 (2014).
809 14. Sherwood, J. J., Waller, G. S., Warshaw, D. M. & Lowey, S. A point mutation in the
810 regulatory light chain reduces the step size of skeletal muscle myosin. *Proceedings of the*
811 *National Academy of Sciences* **101**, 10973–10978 (2004).
812 15. VanBuren, P. *et al.* The essential light chain is required for full force production by skeletal
813 muscle myosin. *Proceedings of the National Academy of Sciences* **91**, 12403–12407 (1994).
814 16. Pastra-Landis, S. C. & Lowey, S. Myosin subunit interactions. Properties of the 19,000-dalton
815 light chain-deficient myosin. *J. Biol. Chem.* **261**, 14811–14816 (1986).
816 17. Warmke, J., Yamakawa, M., Molloy, J., Falkenthal, S. & Maughan, D. Myosin light chain-2
817 mutation affects flight, wing beat frequency, and indirect flight muscle contraction kinetics
818 in Drosophila. *The Journal of cell biology* **119**, 1523–1539 (1992).

819 18. Ravenscroft, G. *et al.* Bi-allelic mutations in MYL1 cause a severe congenital myopathy.
820 *Human Molecular Genetics* https://doi.org/10.1093/hmg/ddy320 (2018)
821 doi:10.1093/hmg/ddy320.
822 19. Nawrotzki, R., Fischman, D. A. & Mikawa, T. Antisense suppression of skeletal muscle
823 myosin light chain-1 biosynthesis impairs myofibrillogenesis in cultured myotubes. *J Muscle*
824 *Res Cell Motil* **16**, 45–56 (1995).
825 20. Lossie, J. *et al.* Mutations of ventricular essential myosin light chain disturb myosin binding
826 and sarcomeric sorting. *Cardiovascular Research* **93**, 390–396 (2012).
827 21. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature*
828 **566**, 496–502 (2019).
829 22. Wang, Y. *et al.* Fast skeletal muscle regulatory light chain is required for fast and slow
830 skeletal muscle development. *The FASEB Journal* **21**, 2205–2214 (2007).
831 23. Hall, J. G. Arthrogryposis (multiple congenital contractures): Diagnostic approach to
832 etiology, classification, genetics, and general principles. *European Journal of Medical*
833 *Genetics* **57**, 464–472 (2014).
834 24. Bamshad, M., Van Heest, A. E. & Pleasure, D. Arthrogryposis: A Review and Update: *The*
835 *Journal of Bone and Joint Surgery-American Volume* **91**, 40–46 (2009).
836 25. Whittle, J., Johnson, A., Dobbs, M. B. & Gurnett, C. A. Models of Distal Arthrogryposis and
837 Lethal Congenital Contracture Syndrome. *Genes* **12**, 943 (2021).
838 26. Devoto, S. H., Melançon, E., Eisen, J. S. & Westerfield, M. Identification of separate slow and
839 fast muscle precursor cells in vivo, prior to somite formation. *Development* **122**, 3371–3380
840 (1996).

841 27. Mead, A. F., Kennedy, G. G., Palmer, B. M., Ebert, A. M. & Warshaw, D. M. Mechanical
842 Characteristics of Ultrafast Zebrafish Larval Swimming Muscles. *Biophysical Journal* **119**,
843 806–820 (2020).
844 28. Talbot, J. & Maves, L. Skeletal muscle fiber type: using insights from muscle developmental
845 biology to dissect targets for susceptibility and resistance to muscle disease. *Wiley*
846 *Interdiscip Rev Dev Biol* **5**, 518–534 (2016).
847 29. Roy, S. D. *et al.* Myotome adaptability confers developmental robustness to somitic
848 myogenesis in response to fibre number alteration. *Developmental Biology* **431**, 321–335
849 (2017).
850 30. Farnsworth, D. R., Saunders, L. M. & Miller, A. C. A single-cell transcriptome atlas for
851 zebrafish development. *Developmental Biology* **459**, 100–108 (2020).
852 31. Thisse, B. *et al.* Spatial and temporal expression of the zebrafish genome by large-scale in
853 situ hybridization screening. *Methods Cell Biol* **77**, 505–519 (2004).
854 32. Xu, Y., He, J., Wang, X., Lim, T. M. & Gong, Z. Asynchronous activation of 10 muscle-specific
855 protein (MSP) genes during zebrafish somitogenesis. *Developmental Dynamics* **219**, 201–
856 215 (2000).
857 33. Chong, J. X. *et al.* Mutations in MYLPF Cause a Novel Segmental Amyoplasia that Manifests
858 as Distal Arthrogryposis. *The American Journal of Human Genetics* **107**, 293–310 (2020).
859 34. Ju, B. *et al.* Recapitulation of fast skeletal muscle development in zebrafish by transgenic
860 expression of GFP under the mylz2 promoter. *Developmental Dynamics* **227**, 14–26 (2003).
861 35. Jackson, H. E. *et al.* The role of Sox6 in zebrafish muscle fiber type specification. *Skeletal*
862 *Muscle* **5**, 2 (2015).

863 36. Rhee, D., Sanger, J. M. & Sanger, J. W. The premyofibril: Evidence for its role in
864 myofibrillogenesis. *Cell Motility* **28**, 1–24 (1994).
865 37. Huxley, H. & Hanson, J. Changes in the Cross-Striations of Muscle during Contraction and
866 Stretch and their Structural Interpretation. *Nature* **173**, 973–976 (1954).
867 38. Huxley, A. F. & Niedergerke, R. Structural Changes in Muscle During Contraction:
868 Interference Microscopy of Living Muscle Fibres. *Nature* **173**, 971–973 (1954).
869 39. Noviello, G., Gjaltema, R. A. F. & Schulz, E. G. CasTuner is a degron and CRISPR/Cas-based
870 toolkit for analog tuning of endogenous gene expression. *Nat Commun* **14**, 3225 (2023).
871 40. Wood, N. B., Kelly, C. M., O'Leary, T. S., Martin, J. L. & Previs, M. J. Cardiac Myosin Filaments
872 are Maintained by Stochastic Protein Replacement. *Molecular & Cellular Proteomics* **21**,
873 100274 (2022).
874 41. Stefan, M. I. & Le Novère, N. Cooperative Binding. *PLoS Comput Biol* **9**, e1003106 (2013).
875 42. Kaya-Çopur, A. *et al.* The Hippo pathway controls myofibril assembly and muscle fiber
876 growth by regulating sarcomeric gene expression. *eLife* **10**, e63726 (2021).
877 43. Agarwal, M. *et al.* Myosin heavy chain-embryonic regulates skeletal muscle differentiation
878 during mammalian development. *Development* **147**, dev184507 (2020).
879 44. Hau, H.-T. A., Kelu, J. J., Ochala, J. & Hughes, S. M. Slow myosin heavy chain 1 is required for
880 slow myofibril and muscle fibre growth but not for myofibril initiation. *Developmental*
881 *Biology* **499**, 47–58 (2023).
882 45. Li, S., Wen, H. & Du, S. Defective sarcomere organization and reduced larval locomotion and
883 fish survival in slow muscle heavy chain 1 (*smyhc1*) mutants. *FASEB j.* **34**, 1378–1397
884 (2020).

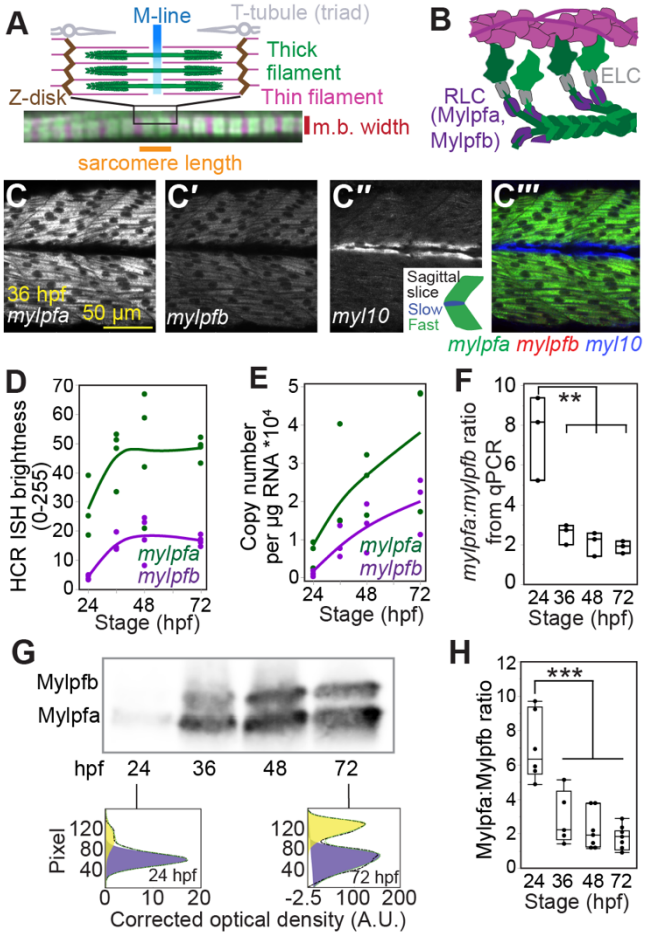
885 46. Zempo, B., Yamamoto, Y., Williams, T. & Ono, F. Synaptic silencing of fast muscle is
886 compensated by rewired innervation of slow muscle. *Sci. Adv.* **6**, eaax8382 (2020).
887 47. Naganawa, Y. & Hirata, H. Developmental transition of touch response from slow muscle-
888 mediated coilings to fast muscle-mediated burst swimming in zebrafish. *Developmental*
889 *Biology* **355**, 194–204 (2011).
890 48. Shimizu, S. *et al.* Spatially ordered recruitment of fast muscles in accordance with
891 movement strengths in larval zebrafish. *Zoological Lett* **11**, 1 (2025).
892 49. Lobjois, V. *et al.* A muscle transcriptome analysis identifies positional candidate genes for a
893 complex trait in pig. *Animal Genetics* **39**, 147–162 (2008).
894 50. Zhang, R. *et al.* The Expression Profiles of mRNAs and lncRNAs in Buffalo Muscle Stem Cells
895 Driving Myogenic Differentiation. *Frontiers in Genetics* **12**, (2021).
896 51. Aoki, H., Sadoshima, J. & Izumo, S. Myosin light chain kinase mediates sarcomere
897 organization during cardiac hypertrophy in vitro. *Nat Med* **6**, 183–188 (2000).
898 52. Grove, B. K. *et al.* A new 185,000-dalton skeletal muscle protein detected by monoclonal
899 antibodies. *J Cell Biol* **98**, 518–524 (1984).
900 53. Webster, C., Silberstein, L., Hays, A. P. & Blau, H. M. Fast muscle fibers are preferentially
901 affected in Duchenne muscular dystrophy. *Cell* **52**, 503–513 (1988).
902 54. Westerfield, M. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio*
903 *Rerio).* (University of Oregon Press, Eugene, 2007).
904 55. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of
905 embryonic development of the zebrafish. *Developmental dynamics* **203**, 253–310 (1995).

906 56. Kawakami, K. Transgenesis and Gene Trap Methods in Zebrafish by Using the Tol2
907 Transposable Element. in *Methods in Cell Biology* vol. 77 201–222 (Academic Press, 2004).
908 57. Ignatius, M. S. *et al.* In vivo imaging of tumor-propagating cells, regional tumor
909 heterogeneity, and dynamic cell movements in embryonal rhabdomyosarcoma. *Cancer Cell*
910 **21**, 680–693 (2012).
911 58. Talbot, J. C. *et al.* Muscle precursor cell movements in zebrafish are dynamic and require Six
912 family genes. *Development* **146**, dev171421 (2019).
913 59. Kwan, K. M. *et al.* The Tol2kit: a multisite gateway-based construction kit for Tol2
914 transposon transgenesis constructs. *Dev Dyn* **236**, 3088–3099 (2007).
915 60. Suster, M. L., Abe, G., Schouw, A. & Kawakami, K. Transposon-mediated BAC transgenesis in
916 zebrafish. *Nature Protocols* **6**, 1998–2021 (2011).
917 61. Talbot, J. C. & Amacher, S. L. A streamlined CRISPR pipeline to reliably generate zebrafish
918 frameshifting alleles. *Zebrafish* **11**, 583–585 (2014).
919 62. Tang, R., Dodd, A., Lai, D., McNabb, W. C. & Love, D. R. Validation of zebrafish (Danio rerio)
920 reference genes for quantitative real-time RT-PCR normalization. *Acta Biochim Biophys Sin*
921 *(Shanghai)* **39**, 384–390 (2007).
922 63. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time
923 Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408 (2001).
924 64. Bird, N. C., Windner, S. E. & Devoto, S. H. Immunocytochemistry to Study Myogenesis in
925 Zebrafish. in *Myogenesis* (ed. DiMario, J. X.) vol. 798 153–169 (Humana Press, Totowa, NJ,
926 2012).

927 65. Berberoglu, M. A. *et al.* Satellite-like cells contribute to pax7 -dependent skeletal muscle
928 repair in adult zebrafish. *Developmental Biology* **424**, 162–180 (2017).
929 66. Choi, H. M. T., Schwarzkopf, M. & Pierce, N. A. Multiplexed Quantitative In Situ
930 Hybridization with Subcellular or Single-Molecule Resolution Within Whole-Mount
931 Vertebrate Embryos: qHCR and dHCR Imaging (v3.0). in *In Situ Hybridization Protocols* (eds.
932 Nielsen, B. S. & Jones, J.) 159–178 (Springer US, New York, NY, 2020). doi:10.1007/978-1-
933 0716-0623-0_10.
934 67. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track
935 neural network. *Science* **373**, 871–876 (2021).
936 68. Gulbulak, U., Wellette-Hunsucker, A. G., Kampourakis, T. & Campbell, K. S. GelBox: Open-
937 source software to improve rigor and reproducibility when analyzing gels and immunoblots.
938 *American Journal of Physiology-Heart and Circulatory Physiology*
939 https://doi.org/10.1152/ajpheart.00144.2024 (2024) doi:10.1152/ajpheart.00144.2024.
940 69. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods*
941 **9**, 676–682 (2012).
942 70. Mary, H. & Brouhard, G. J. *Kappa (κ): Analysis of Curvature in Biological Image Data Using*
943 *B-Splines*. http://biorxiv.org/lookup/doi/10.1101/852772 (2019) doi:10.1101/852772.
944

Figures and legends

Figure 1: The *mylpfa* gene is expressed more abundantly than *mylpfb* in fast-twitch muscle. **(A)** Images illustrating myofibril and sarcomere structure. The m.b. width represents axis of measurements for myofibril bundle width. **(B)** Illustration of Mylpf localization within a thick filament. **(C-C''')** *In situ* hybridization chain reaction (HCR ISH) imaged in somites over the mid-yolk tube of a 36 hpf embryo. Shown as a single channel for *mylpfa* (C), *mylpfb* (C') or the slow muscle marker *myl10* (C''), and as a merged image (C'''). **(D)** Plot showing the brightness of *mylpfa* and *mylpfb* in the HCR ISH images. **(E)** Copy number of *mylpfa* and *mylpfb* per microgram of total RNA at 24, 36, 48, and 72 hpf as measured by qPCR and normalized using a standard curve–based normalization method. **(F)** The ratio of *mylpfa* to *mylpfb* transcript abundance calculated from the absolute copy numbers shown in (E). **(G)** Image of a western blot showing Mylpfa and Mylpfb protein abundance at these timepoints. Examples of Mylpfa and Mylpfb bands quantified in GelBox are shown beneath the image. **(H)** Ratio of Mylpfa to Mylpfb band intensity in western blot. **(I)** Tabular summary of means from D, F, H. The points represent statistical N, either pools of animals (E, F, H), or individual animals (D). Significance threshold determined by Tukey-Kramer comparison after one-way ANOVA; ** P<0.01, ***P<0.001.



I	24 hpf	36 hpf	48 hpf	72 hpf
mylpfa:mylpfb ratio				
HCR ISH	6.6	2.9	2.6	2.9
qPCR	7.5	2.5	2.1	1.9
Mylpfa:Mylpfb ratio				
Western Blot	7.0	2.9	2.2	1.8

Figure 2: Zebrafish *mylpfa* is necessary for fast-twitch myofibril formation. **(A-D)** Zoomed region of somite muscle, with myofibril bundles marked using *six1b:LifeAct-mTurquoise2* transgene (LifeAct-mTq protein) at 2 dpf and the same animal at 6 dpf in wild-type (A, B) or *mylpfa*$^{-/-}$ mutant siblings (C, D). **(A'-D')** Cross-sectional views of whole somites from the same animals. **(E)** Myofibril bundle widths were measured using the LifeAct images in the sagittal plane. Olive-colored lines connect the same individuals through time. **(F)** Plot of signal periodicity shows that the *mylpfa*$^{-/-}$ mutant has less LifeAct periodicity in sarcomeric intervals (gray bars). Lightly colored regions indicate bootstrap confidence intervals. **(G)** This readout is simplified as a 'sarcomeric fraction', which shows the fraction of signal in sarcomeric (1.85 µm) or half-sarcomeric (0.9 µm) lengths out of the total periodicity; further explanation is provided in Figure S4. The sarcomeric fraction remains constant through development but differs by genotype. **(H)** Somite area, measured from the cross-sectional views. The *mylpfa*$^{-/-}$ mutant somite is smaller than in wild-type or heterozygous siblings at 6 dpf, though not at 2 dpf. **(I-J''')** Fast muscle myofibers, labeled at 48 hpf, showing co-label for MyHC (A4.1025), the Z-disk protein Actinin (A7732), and F-actin (phalloidin), shown as single channel or overlays. The axes of width and length measurements are shown in I', J'. A representative confocal stack of this type of label is shown in Video 1. **(K)** The correlation between F-actin and MyHC label is reduced in the *mylpfa*$^{-/-}$ mutant, as assessed using Spearman's rank correlation coefficient (Spearman's correlation). **(L, M)** A similar pattern is seen in animals labeled for the M-line protein Myomesin (mMac), F-actin, and MyHC. Scalebar in A is for A-D, in A' is for A'-D', in I is for I-J''', in L is for L, M. Each point shown represents the average of measurements taken within one image of one animal. Significance thresholds for multiple comparisons determined by Tukey-Kramer HSD comparisons after one-way ANOVA (E, F, H); pairwise comparisons (K) use the Kruskal-Wallis exact test. Not significant (n.s.) is P>0.05, * P<0.05, ** P<0.01, *** P<0.001.

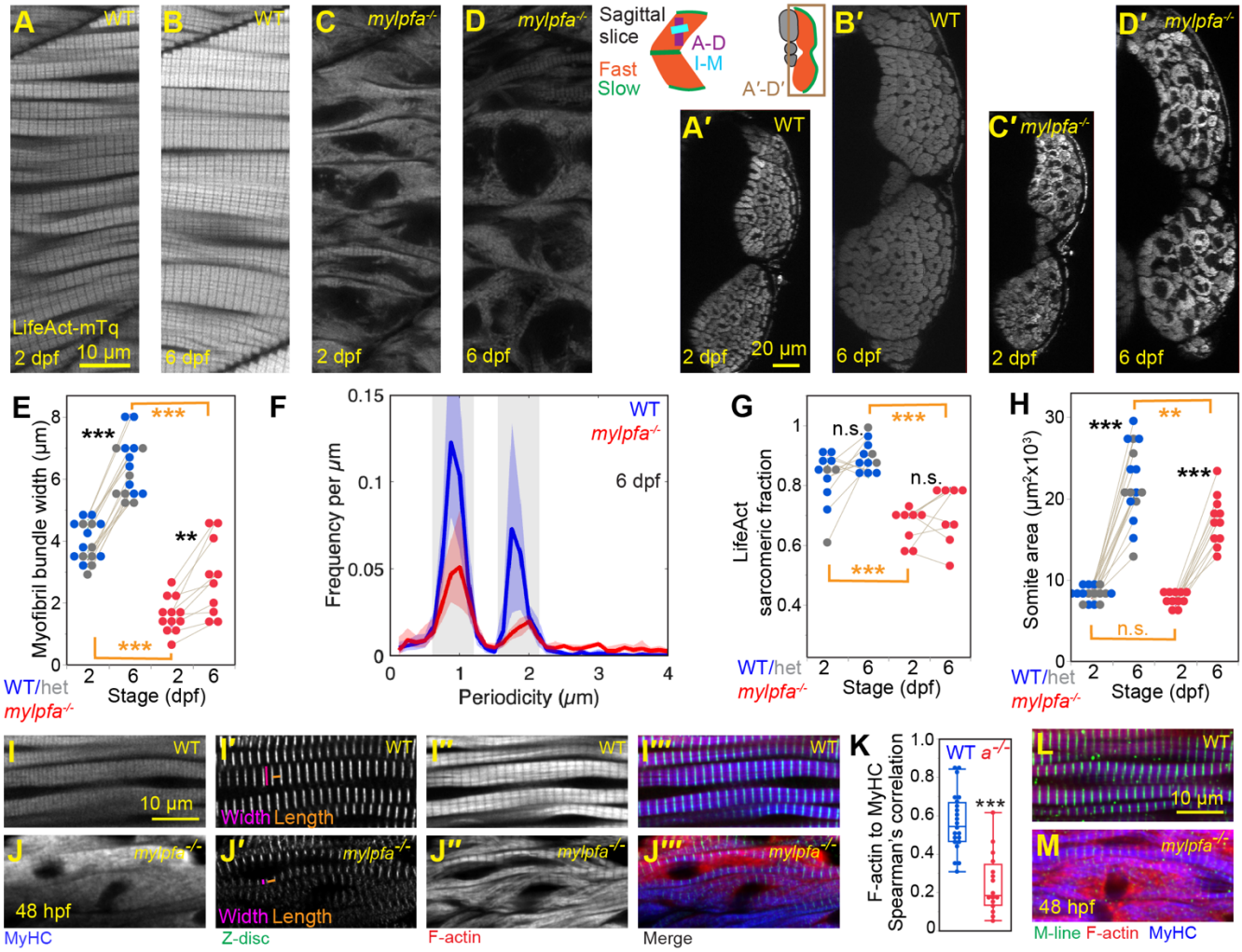


Figure 3: Mutations in *mylpfa* and *mylpfb* cause strong loss of function. **(A)** Illustration of *mylpfa* and *mylpfb* gene structure. Color code: 5' UTR (brown), coding sequence (purple), 3' UTR (gray), and frameshift locations (green arrowheads). **(B)** Chromatogram showing the gRNA target in wild-type sequence (top) and the 5 bp *mylpfb*^{oz39} lesion sequenced from a homozygous mutant (bottom). **(C)** Illustration of Mylpfa and Mylpfb proteins, with arrowheads marking frameshift locations. **(D-E)** Quantitative RT-PCR results from the *mylpfa;mylpfb* mutant series showing *mylpfa* expression per genotype (gen.) normalized to the average *mylpfa* abundance in *mylpfa*$^{+/+}$ genotypes. Result shown as individual genotypes (D) or as a merger of *mylpfa* genotypes (E). **(F-G)** Quantitative RT-PCR results for *mylpfb* expression, normalized to the average *mylpfb* abundance in *mylpfb*$^{+/+}$ genotypes, shown for individual genotypes (F) or as a merger of *mylpfb* genotypes (G). **(H)** Predicted protein structure generated using Robetta. Color code indicates the Mylpf peptides lost after *mylpfb*^{oz39} frameshift (black), those retained even in the mutant (green), and the MyHC (purple). **(I)** Western blot showing MyHC, Mylpfb, and Mylpfa bands from individual animals across the nine genotypes of *mylpfa*$^{+/-}$;*mylpfb*$^{+/-}$ incross. Three wild-type controls are loaded to increase sampling of the wild-type mean. **(J)** Plot showing the abundance of each protein in animals wild-type, heterozygous, or mutant for that gene. **(K)** Plot showing a strong correspondence between combined Mylpfa and Mylpfb band intensity and predicted dosage using a 2 to 1 ratio between Mylpfa and Mylpfb (Table S1). **(L-O)** TEM images of fast-twitch muscle showing myofibrils in wild-type and their defects in the single and double mutants. Lines illustrate measurements of muscle structure. **(L'-O')** Higher magnification views of different animals. **(P)** Measurements made between triads, showing myofibril width. **(Q)** Measurements of myofibril bundle widths. Each point (statistical N) in J, K, P, Q represent individual animals; in D-G, each dot represents a pool of three animals. Color code below D is for D-G, I-K, P-Q. Significance thresholds are determined by Tukey-Kramer comparison after one-way ANOVA; n.s. P>0.1, * P<0.05, ** P<0.01, *** P<0.001. Scalebar in L is for L-O, in L' is for L'-O'.

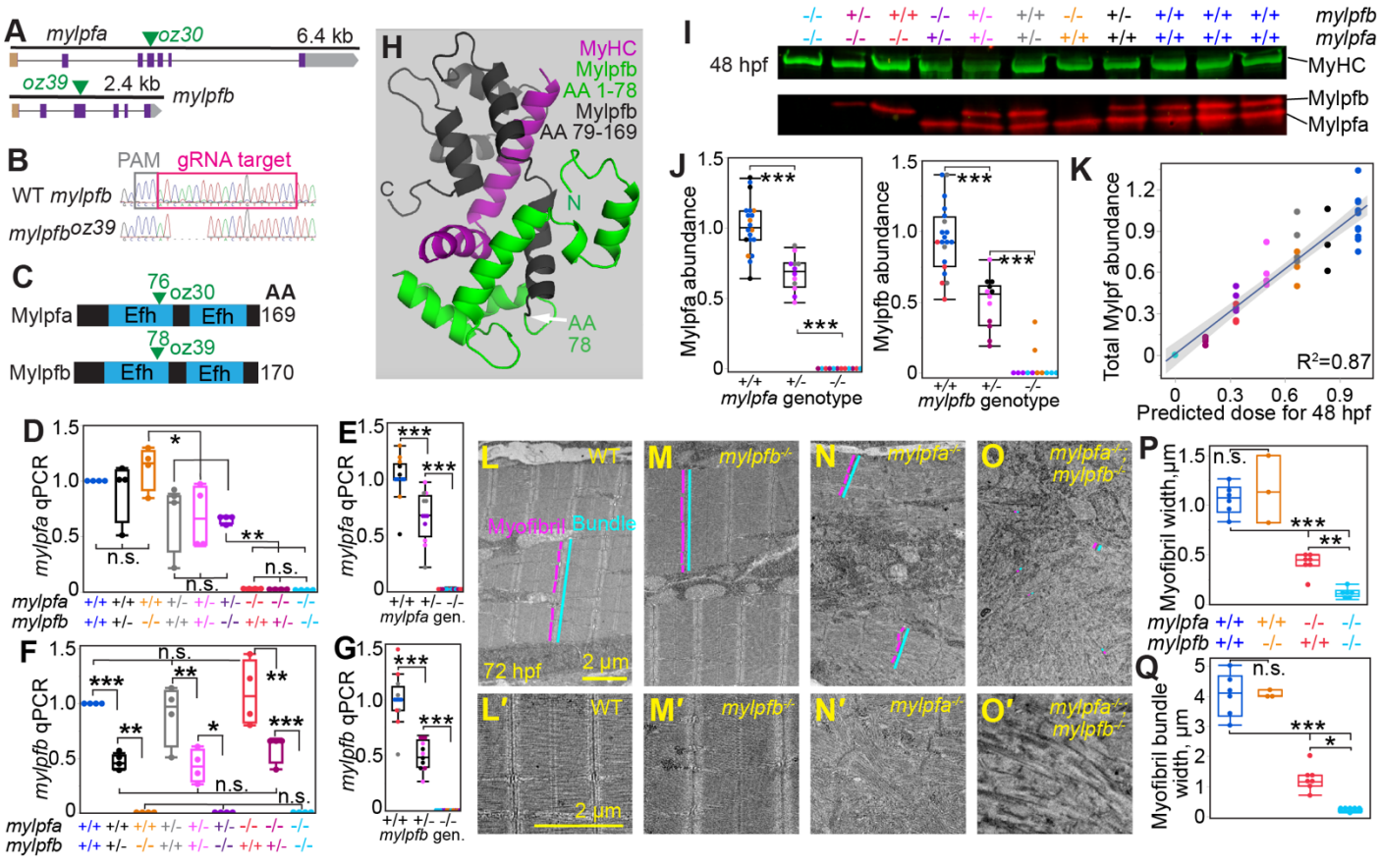


Figure 4: *mylpfa* and *mylpfb* are required for myosin heavy chain localization in fast-twitch muscle of young myofibers. **(A-D)** Immunolabel of fast muscle fibers at 26 hpf shown as orthogonal views of Z-stacks of wild-type (A), *mylpfb*$^{-/-}$ (B), *mylpfa*$^{-/-}$ (C), and *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$ double mutant zebrafish (D). **(A'-D')** Zoomed sagittal views of the same animals. **(E)** Box plots showing Spearman's correlation between the F-actin and MyHC label for the genotypes examined. **(F)** Plots showing the width of myofibril bundles in each genotype. **(G)** Overview of protocol to measure image brightness from edge-to-center. The fast-twitch muscle cells are (1) manually outlined, (2) segmented using these outlines, then (3) the outlines are used to determine edges when calculating the brightness of each channel from edge to center of the drawn cells. **(H-J)** Localization at 26 hpf of the F-actin marker phalloidin (H), the muscle nuclei label anti-Rbfox1l (I), and MyHC antibody A4.1025 (J). Data is plotted as percent of total image brightness from edge to center of segments, with bootstrapped 95% confidence intervals in shaded lines. Inset in J shows the stark difference between MyHC localization in the *mylpfa*$^{+/-}$;*mylpfb*$^{+/-}$ double heterozygote and the *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$ double mutant, unobscured by other genotypes. Color code in H is for H-J. **(K)** Illustration of the proposed localization defect. Points in E-F represent measurements from individual animals. Measurements for E-J were taken on 5 wild-type, 8 *mylpfb*$^{-/-}$, 9 *mylpfa*$^{+/-}$;*mylpfb*$^{+/-}$, 5 *mylpfa*$^{-/-}$, and 7 *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$ double mutants. Significance thresholds are determined by Bootstrap comparison (H-J) or by Tukey-Kramer comparison after one-way ANOVA (E, F); ** P<0.01, *** P<0.001. Scalebar in A is for A-D, in A' is for A'-D'.

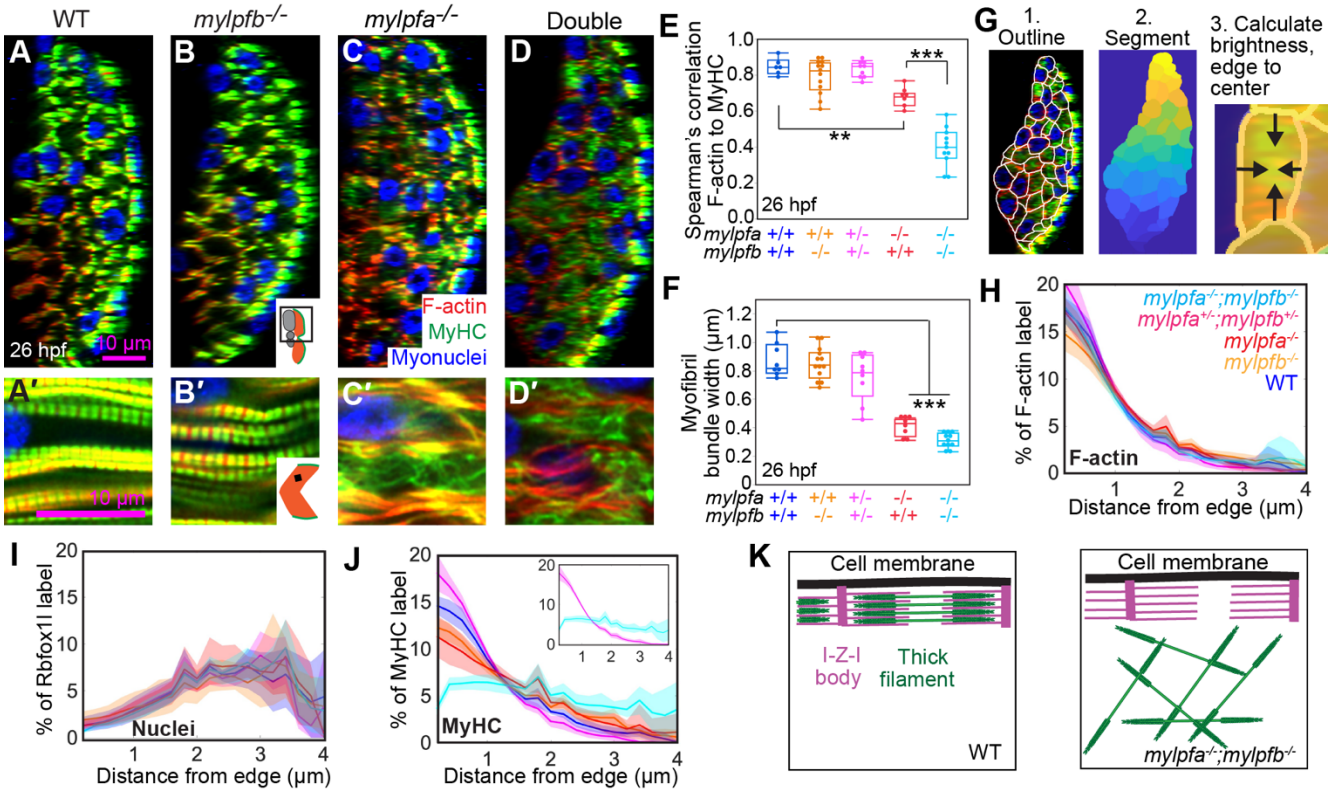


Figure 5: Levels of myofibril formation correspond to dosages predicted by *mylpfa* and *mylpfb* loss of function. **(A)** Scatterplot of myofibril bundle widths in all genotypes tested at 26 hpf, with the X-axis representing Mylpf dosage as predicted by a 7:1 ratio between Mylpfa and Mylpfb. This 26 hpf width data is also shown in Figure 4F, for Tukey-Kramer comparisons. **(B-D)** Scatterplots of myofibril bundle widths in all genotypes tested at 48 hpf, 72 hpf, and 6 dpf, with the X-axis representing Mylpf doses predicted by a 2:1 ratio between Mylpfa and Mylpfb. **(E)** Cross-sectional area measurements of the same somites at 6 dpf. The somite sizes do not differ significantly across genotypes. **(F)** Examples of fast-twitch muscle structure, labeled with phalloidin, in the nine genotypes arising from *mylpfa*$^{+/-}$;*mylpfb*$^{+/-}$ incross, viewed at high magnification. **(G)** Fast-twitch myofibril bundles at 48 hpf, labeled for the myonuclei (Rbfox1l, blue), F-actin (phalloidin, red), and MyHC (A4.1025, green). **(H)** The colocalization between F-actin and MyHC label corresponds well with modeled protein dosage, as measured using Spearman's correlation, and dose determined with a 2:1 Mylpfa to Mylpfb ratio. **(I-J)** Graph showing how the sarcomeric fraction of MyHC (I) or F-actin (J) corresponds with protein dosage in fast-twitch muscle, again calculated using a 2 to 1 skew. **(K)** Overlay of the 7 regression models with explanation of the Hill equation, scaled to have matching upper and lower limits. Overlay colors in K use regression color code from panels in A (brown), B-D (gray), H (gray), and I, J (blue). Formula abbreviations: Hill slope (Hs), dose of half-maximal effectiveness (EC50), upper asymptote (U), and lower asymptote (L). Each point represents a statistical N, the mean of measurements from one animal. R^2 and Hs plotted in each panel. The genotypic color code shown above F is used for panels A-E, H-J.

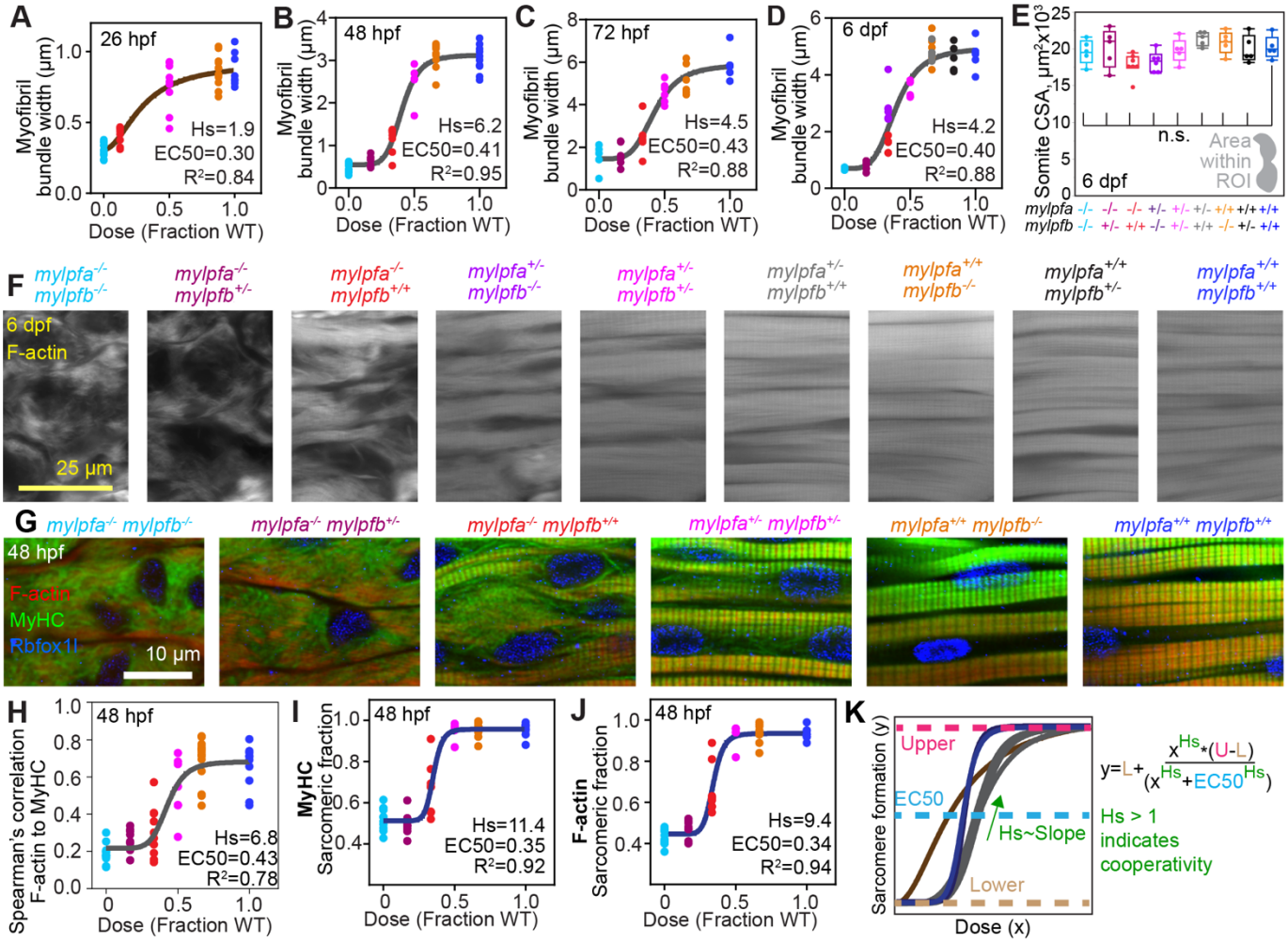


Figure 6: The *mylpfa*$^{-/-}$ and the *mylpfa*$^{-/-}$;*mylpfb*$^{-/-}$ double mutant animals do not swim at high speed but increase the frequency of slower movements. **(A-C)** Escape response in wild-type or in *mylpfa*$^{-/-}$ animals, showing images captured 1/60th of a second apart (A, B). Blue line shows maximal tail bend, quantified in (C). **(D)** Box plot of distances traveled when the same fish were imaged 25-minutes per day from 2 to 6 dpf. **(E-F)** A logarithmic plot of the proportion of time that select genotypes swim at different speeds, with bootstrap confidence intervals shown for each genotype (E). The inset region highlights differences in the *mylpfa*$^{-/-}$ mutant behaviors that become apparent by 5 dpf (E'). These differences are more striking when normalized to wild-type siblings (F). **(G, H)** Box plots show total distance traveled (G) or highest movement speed (H) of the *mylpfa;mylpfb* mutant series at 6 dpf. **(I)** A logarithmic plot of the proportion of time that select genotypes swim at different speeds, with bootstrap confidence intervals shown for each genotype. **(J)** The same data is shown after normalization to wild-type behaviors, plotted on a linear scale. **(K)** Quantification of slow-twitch myofibril size. **(L)** By 6 dpf the slow-twitch myofibrils bundles are wider in the non-transgenic mutant than their wild-type siblings, seen in cross sectional views, with inset showing the measured region. **(M)** Quantification of the myofibril bundle CSA within the selected slow-twitch regions. No significant difference is seen among high-dose Mylpf genotypes, but at low dose, there is a significant increase in myofibril bundle size, only within slow-twitch fibers. **(N)** Illustration summarizing the effects of Mylpf loss on slow and fast-twitch muscle. Data in (K-M) comes from datasets examined in Figure 5 for other measures. Scalebar in A is for A-B. Significance is determined using the Kruskal-Wallis exact test (C), comparison of bootstrap confidence intervals (E-F, I-J) or Tukey-Kramer comparisons after one-way ANOVA (D, G, H, K, M). Significance shown in comparison to wild-type siblings, with n.s. of P>0.1, * P<0.05, ** P<0.01 and *** P<0.001. Genotypic color code is matched across panels G-J, K, and M.

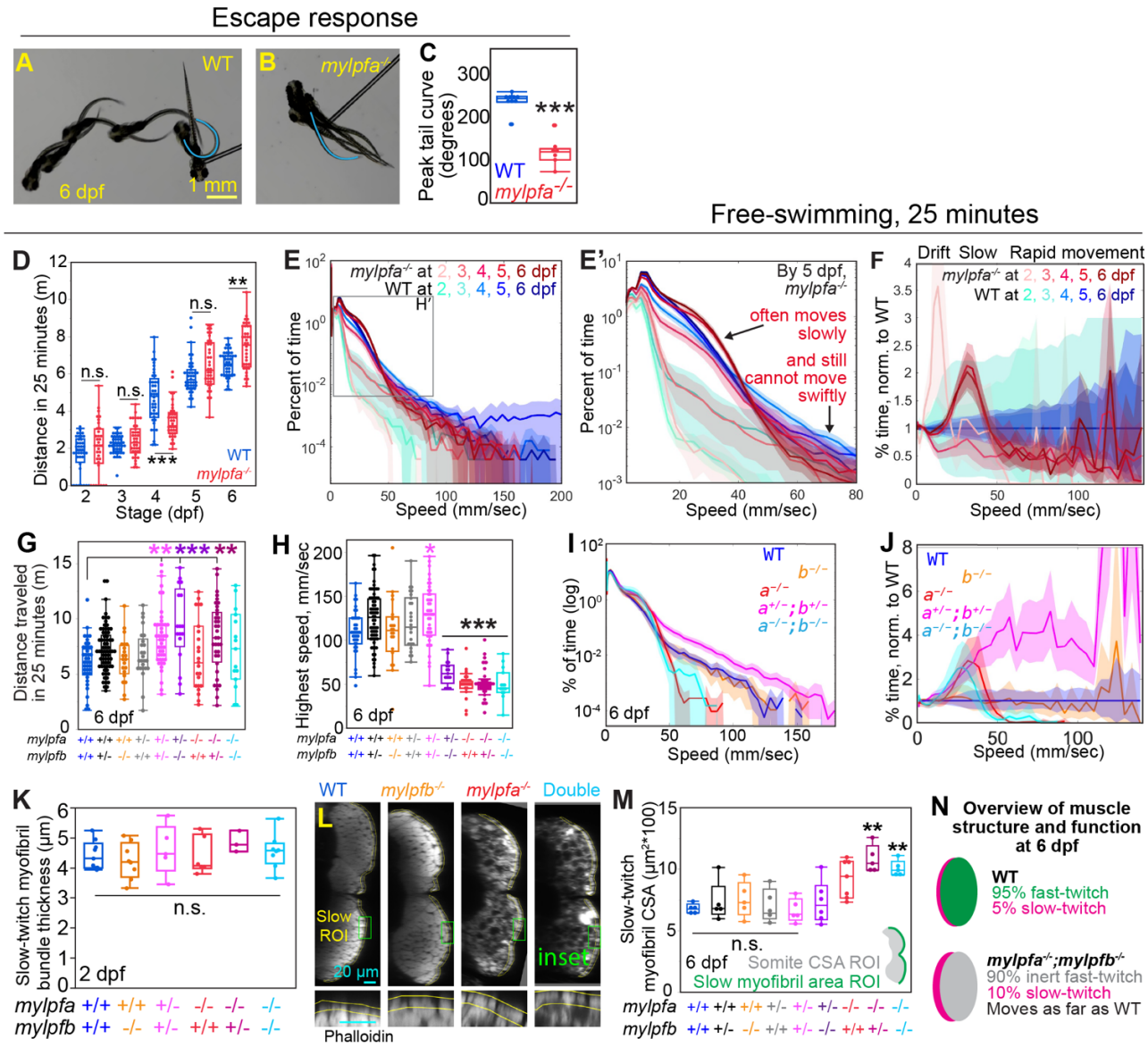


Figure 7: Expression of either *mylpfb-GFP* or *mylpfa-GFP* can rescue *mylpfa*$^{-/-}$ myofibrils to a degree proportionate with GFP abundance. **(A)** Schematic illustrating the *Tg(mylpfa:mylpfa-GFP)mai102* and *Tg(mylpfa:mylpfb-GFP)mai103* transgenes. **(B)** Box plot showing the variation of fluorescence brightness in the *mylpfa-GFP* transgenic animals and their non-transgenic siblings. The animals imaged for B were pooled into groups of 10 and run on acrylamide gels for western blotting. **(C)** Example of a western blot of Mylpf protein from the wild-type fish and their Mylpfa-GFP+ siblings at 3 dpf shows that the transgene produces one thin band that is shifted upwards by the addition of GFP. **(D)** Plot showing quantification of the Mylpfa, and Mylpfa-GFP bands, normalized to the Mylpfa. **(E)** Correlation plot showing that average fluorescence intensity from these pools corresponds precisely with band intensity in the western blots. **(F)** Confocal slices through fast-twitch muscle with green GFP and magenta phalloidin (F-actin). Zoomed inset shows the F-actin channel at 48 hpf. Panels show representative wild-type and *mylpfa*$^{-/-}$ animals that are non-transgenic (top), express the *mylpfa-GFP* transgene (middle), or the *mylpfb-GFP* transgene (bottom). **(G)** Statistical comparisons of myofibril bundle (m.b.) widths in 2 dpf animals. **(H)** Statistical comparisons of the F-actin sarcomeric fraction for these same samples. **(I)** Scatterplot showing correlations between GFP brightness and myofibril bundle width at 2 dpf. **(J)** Scatterplot shows correlations between GFP brightness and the F-actin sarcomeric fraction in the same samples. Points in these panels represent statistical N, which is the individual animal (B, G-J) or pool of animals (C-E). Significance thresholds are determined by Tukey-Kramer HSD comparisons after one-way ANOVA: n.s. P>0.1, * indicates P<0.05, *** P<0.001.

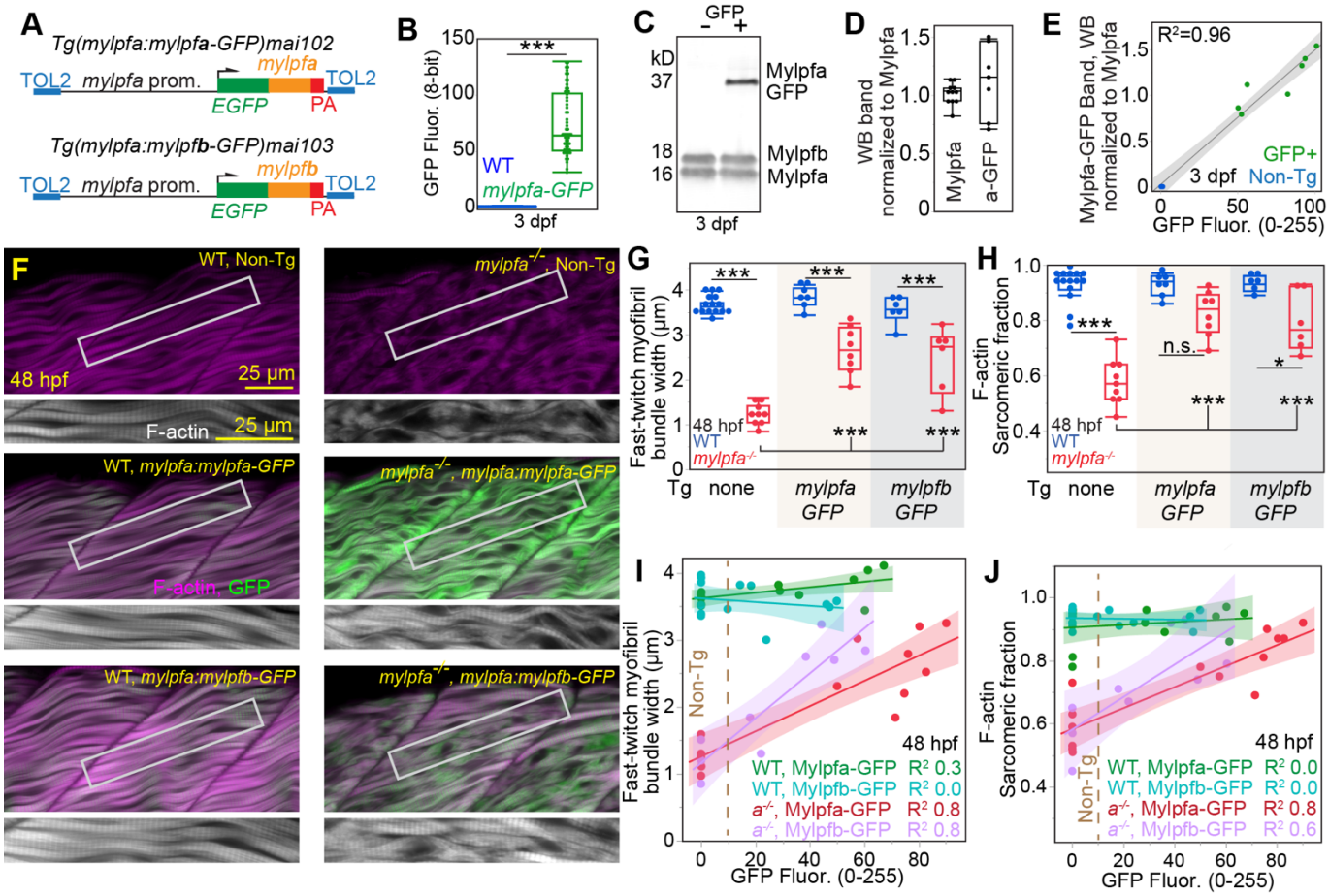


Figure 8: Expression of GFP-fused human MYLPF is sufficient to restore fast-twitch myofibril formation in the *mylpfa*$^{-/-}$ mutant. **(A)** Schematic illustrating the *Tg*(*mylpfa:MYLPF-GFP)mai104* transgene. **(B)** Scatterplot showing correlations between GFP brightness and the fast-twitch myofibril bundle width from an intercross of *mylpfa*$^{-/+}$ heterozygotes segregating this transgene. **(C-E)** Examples of confocal images of *mylpfa*$^{-/-}$ mutant animals that have no GFP expression (C), dim GFP expression (D), and bright GFP expression (E). Quantification of these individuals is shown in B. **(F)** Scatterplot of m.b. width data from all *mylpfa-GFP*, *mylpfb-GFP*, and *MYLPF-GFP* transgenic experiments in this study (Figure S12), normalized to the mean wild-type value per experiment. We find similar slopes in the *mylpfa*$^{-/-}$ mutant animals bearing these transgenes, suggesting that these three Mylpf genes act with equivalent efficiency. None of these three transgenic lines significantly influence the widths of myofibrils in wild-type (*mylpfa*$^{+/+}$) animals. Points in panels B and F represent individual animals.

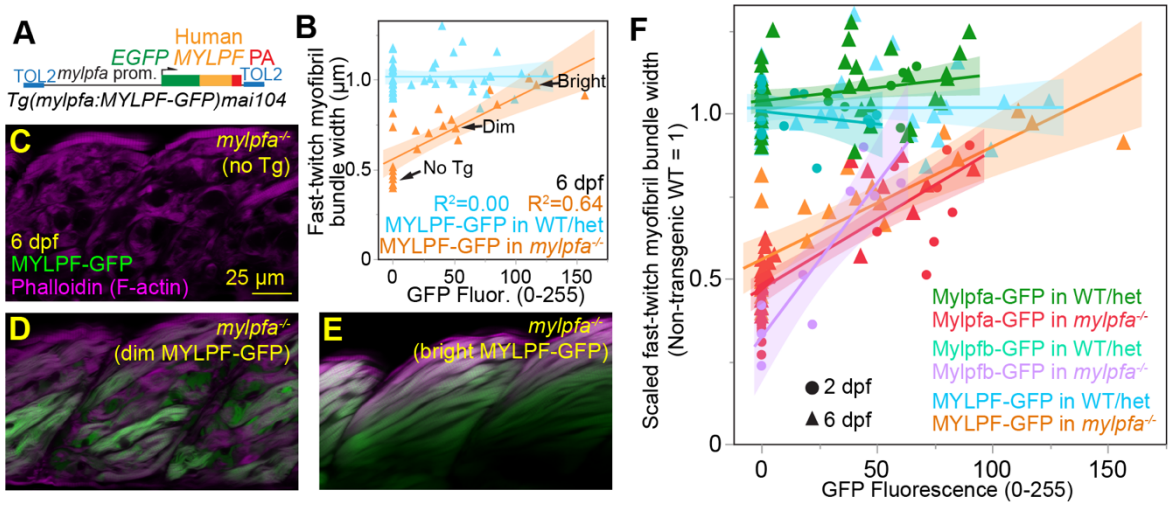


Figure 9: A MYLPF protein variant found in Distal Arthrogryposis fails to rescue the *mylpfa*$^{-/-}$ mutant. **(A)** Predicted MYLPF structure, with pink color and filled space highlighting the position of Glycine at residue 163. **(B)** Illustration of a construct that expresses GFP-fused human MYLPF or the G163S variant MYLPF. **(C-D)** Representative images of *mylpfa*$^{-/-}$ mutants that were injected with plasmid expressing MYLPF-GFP protein (C) or the G163S variant (D), resulting in mosaic GFP expression. Asterisks indicate balls of highly disordered GFP fluorescence signal, which are only seen in the G163S MYLPF protein variant. **(E)** Scatterplot showing correlates between myofibril width and GFP fluorescence for both injected constructs in the *mylpfa*$^{-/-}$ mutant and their WT/het siblings. **(F)** Scatterplot showing the correspondence between F-actin sarcomeric fraction and GFP fluorescence in the *mylpfa*$^{-/-}$ mutant and their WT/het siblings. **(G)** Box plots showing the degree to which wild-type MYLPF-GFP or the G163S MYLPF-GFP localizes to a sarcomeric pattern. **(H)** Plot of Spearman's correlation between GFP and F-actin in these GFP-positive animals. **(I-I")** Representative myofibers from a *mylpfa*$^{-/-}$ mutant expressing the G163S MYLPF variant from the stable transgenic allele *Tg(mylpfa:MYLPFG163S-GFP)mai119* alongside the live actin marker *TgBAC(six1b:LifeAct-mTq)mai121*. The GFP signal appears mildly sarcomeric but it is not ordered into myofibrils, despite expression at a high level (I'). The LifeAct-mTq signal also appears disordered (I"). Modest bleed-through from GFP into the mTq channel precludes colocalization analysis. Each point in E-H represents a statistical N, which is the mean of measurements from one animal. Graphs E and F show remeasurements of GFP-negative cells (GFP<10) and GFP-positive cells (GFP>10) within the same plot, divided by a dotted brown line. Significance thresholds in G and H are determined by Tukey-Kramer HSD comparisons after ANOVA: n.s. is P>0.1, ** indicates P<0.01, *** indicates P<0.001. Scalebar in C is for C-D, scalebar in I is for I-I".

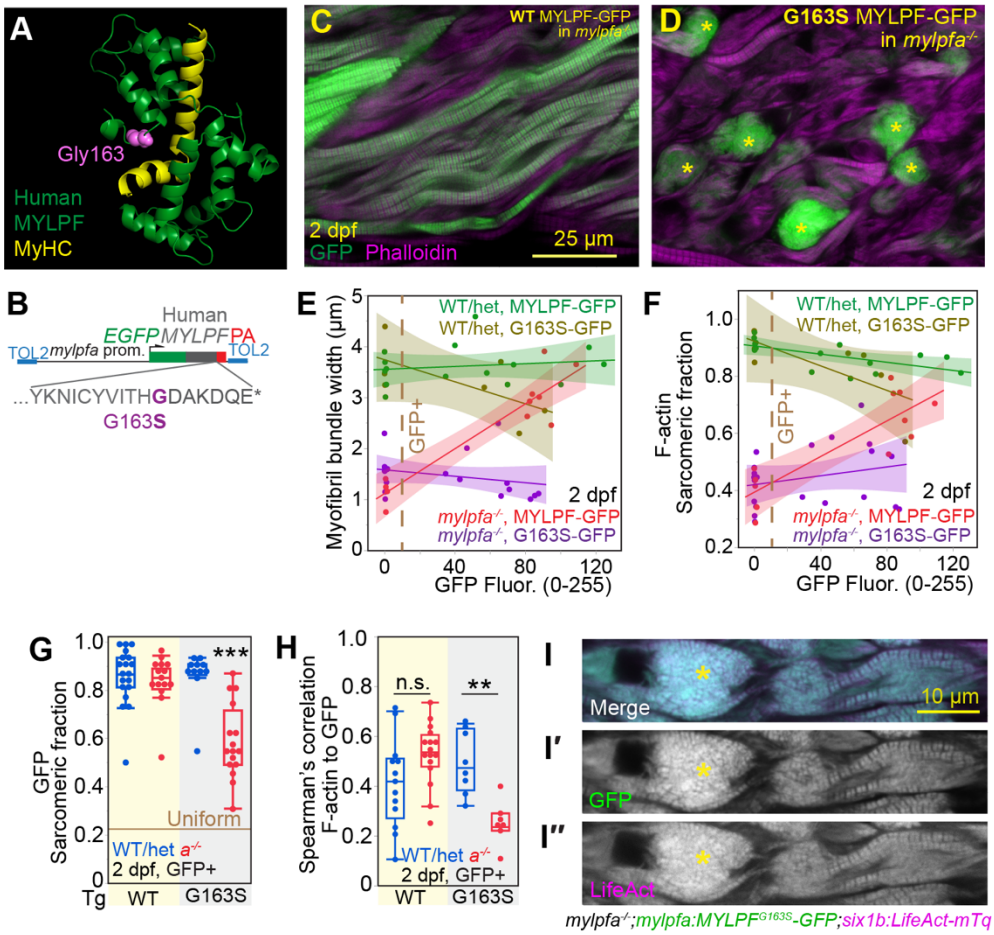
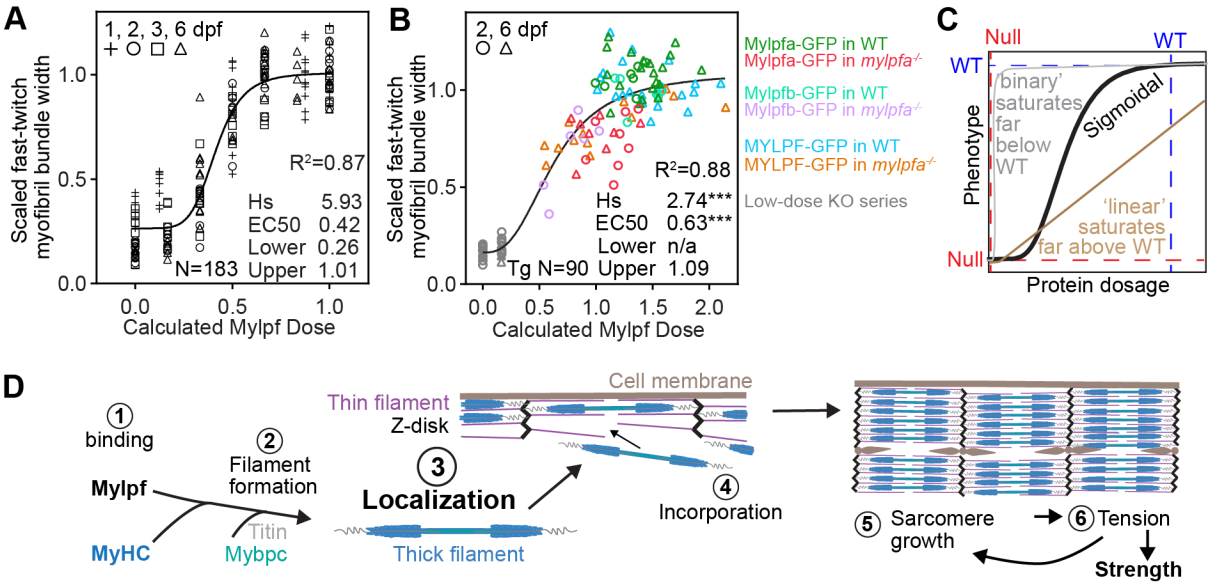



Figure 10: Modeling the impact of Mylpf on muscle structure. Scatterplot graphs, with Y-axis showing the myofibril bundle (m.b.) widths scaled to the non-transgenic wild-type level per experiment, and the X axis showing the Mylpf dosage as calculated using formulas shown in Figure S12. **(A)** A plot of data from the *mylpfa;mylpfb* mutant series shows strong overlap at each timepoint examined (1, 2, 3, and 6 dpf) using datapoints presented in Figure 5B-D. The Hill regression model shows a high correlate across timepoints. **(B)** A plot of the transgenic data, excluding GFP <10, alongside the most severe Mylpf loss of function alleles from the mutant lines (*mylpfa*$^{-/-}$*;mylpfb*$^{+/-}$ and *mylpfa*$^{-/-}$*;mylpfb*$^{-/-}$, gray dots). This dataset also produces a Hill regression with a high R^2, high Hill slope and moderate EC50. Asterisks on Hs and EC50 values reflect fit of the dataset in panel B for transgenic animals to the regression curve calculated in A for the *mylpfa;mylpfb* mutant series; significance threshold. Asterisks, *** indicates P<0.001. **(C)** Examples of three potential types of phenotypic responses to protein dose. Mylpf appears to have a sigmoidal response curve with a steep slope at inflection and an EC50 value somewhat below 50% protein dosage. **(D)** A model for how Mylpf may cooperatively promote sarcomere and myofibril bundle assembly. Our findings support a role for Mylpf in MyHC localization and Mylpf may cooperatively promote other steps in this process.