Post-transcriptional regulation of myotube elongation and myogenesis by Hoi Polloi

Aaron N. Johnson1,2,*, Mayssa H. Mokalled3, Juliana M. Valera2, Kenneth D. Poss3 and Eric N. Olson1,∗

SUMMARY
Striated muscle development requires the coordinated expression of genes involved in sarcomere formation and contractility, as well as genes that determine muscle morphology. However, relatively little is known about the molecular mechanisms that control the early stages of muscle morphogenesis. To explore this facet of myogenesis, we performed a genetic screen for regulators of somatic muscle morphology in Drosophila, and identified the putative RNA-binding protein (RBP) Hoi Polloi (Hoip). Hoip is expressed in striated muscle precursors within the muscle lineage and controls two genetically separable events: myotube elongation and sarcomeric protein expression. Myotubes fail to elongate in hoip mutant embryos, even though the known regulators of somatic muscle elongation, target recognition and muscle attachment are expressed normally. In addition, a majority of sarcomeric proteins, including Myosin Heavy Chain (MHC) and Tropomyosin, require Hoip for their expression. A transgenic MHC construct that contains the endogenous MHC promoter and a spliced open reading frame rescues MHC protein expression in hoip embryos, demonstrating the involvement of Hoip in pre-mRNA splicing, but not in transcription, of muscle structural genes. In addition, the human Hoip ortholog NHPL2L1 rescues muscle defects in hoip embryos, and knockdown of endogenous nhp2l1 in zebrafish disrupts skeletal muscle development. We conclude that Hoip is a conserved, post-transcriptional regulator of muscle morphogenesis and structural gene expression.

KEY WORDS: Myotube elongation, Post-transcriptional regulation, Myogenesis, Drosophila, Zebrafish

INTRODUCTION
The transcriptional regulatory networks that direct muscle precursor cell specification and the expression of muscle structural genes have been well defined. However, the possible post-transcriptional contribution to mesoderm development is only beginning to come to light (Biedermann et al., 2010; Toledano-Katchalski et al., 2007; Yarnitzky et al., 1998). The unique properties of Drosophila, including external development and an extensive array of genetic tools, have allowed the discrete cellular processes directing muscle development to be dissected in detail (Guerin and Kramer, 2009a; Schejter and Baylies, 2010; Schnorrer and Dickson, 2004).

Embryonic somatic muscle development in Drosophila is a multistep process that initiates with the specification of founder cells from a field of myogenic competent cells in the mesoderm (Carmena et al., 1995; Jagla et al., 1998). Founder cells express a unique set of muscle identity genes, encoding transcription factors, that direct differentiation into one of 30 somatic muscles (de Joussineau et al., 2012). Once specified, muscle founders begin the process of migration and elongation that can be divided into three phases (Schnorrer and Dickson, 2004). During the first phase, founder cells migrate to their correct position within the segment. The second phase begins when the founder cells initiate myoblast fusion and form polarized myotubes that elongate along a single axis. The myotubes then form extensive filopodia in the direction of initial polarity, presumably in response to guidance cues from tendon cells in the overlying epidermis (Guerin and Kramer, 2009a; Schnorrer and Dickson, 2004). The center of the myotube remains localized while the ends of the myotube elongate towards their respective muscle-attachment sites (Schnorrer and Dickson, 2004). The final phase of elongation initiates when the myotube ends reach their muscle attachment sites and filopodia no longer form. The myotube then localizes integrin-mediated adhesion complexes with the overlying tendon cells to establish strong myotendinous junctions (Schejter and Baylies, 2010).

The mechanisms that control myotube elongation during somatic muscle morphogenesis are poorly understood. Slit is the single guidance molecule known to direct both myotube elongation and target site recognition, but loss of Slit modestly affects the elongation of only a subset of myotubes (Kramer et al., 2001). Nascent myotubes must undergo extensive cytoskeletal rearrangements during elongation, and recent work has focused on the role of microtubule dynamics in this process (Folker et al., 2012; Guerin and Kramer, 2009b). Tumbleweed (Tum) is a Rac family GTPase-activating protein that becomes localized to the nuclear periphery via its association with the microtubule-associated protein Pavarotti (Pav). Loss of pav or tum disrupts microtubule polarity and polarized growth, mislocalizes the minus-end microtubule nucleator γ-tubulin and causes modest myotube elongation defects (Guerin and Kramer, 2009b). A second regulator of microtubule dynamics, Dynemin heavy chain (Dhc64C), is also required for myotube elongation but its role is restricted to the final stages of elongation (Folker et al., 2012). Although microtubule dynamics plays a key role in the process, the mechanisms that initiate myotube elongation and the downstream targets of intracellular messenger proteins, such as Tum, remain largely unknown.

The cellular events that regulate myotube morphology are distinct from the molecular processes that direct terminal differentiation and structural gene expression. Embryos defective in myoblast fusion express Myosin Heavy Chain (MHC) in unfused mononucleate...
founder cells; this striking phenotype has been exploited in genetic screens to identify novel regulators of myoblast fusion (Chen and Olson, 2001). Components of the sarcomere, the basic unit of muscle contraction, are subject to extensive post-transcriptional regulation. For example, Drosophila MHC is encoded by a single genomic locus that can produce 480 unique protein isoforms (Zhang and Bernstein, 2001). These isoforms encode variant regions of the MHC globular head and provide diversity in contractile performance (Krontert et al., 1994). However, the RNA-binding proteins (RBPs) that regulate the production of different MHC isoforms have not been identified.

In a screen for genes that regulate somatic muscle morphology in Drosophila, we identified the putative RBP Hoip pollo (Hoip). hoip embryos show two dramatic phenotypes: myotube elongation does not initiate, even though founder cell specification and myoblast fusion initiate normally and striated muscles fail to express multiple sarcomeric proteins, including MHC and Tropomyosin (Tm). hoip expression is tissue specific and within the muscle lineage is restricted to striated muscle precursors. By RNA deep sequencing (RNA-seq), we found that known regulators of myotube elongation are expressed correctly in hoip mutant embryos, suggesting Hoip orchestrates a previously unrecognized post-transcriptional mechanism to initiate elongation. Functional rescue experiments demonstrate that Hoip directs pre-mRNA splicing during myogenesis. The human Hoip ortholog NHP2L1 can rescue the hoip phenotype in Drosophila, and morpholino (MO) knockdown experiments in zebrafish indicate that nhp2l1 is a conserved essential regulator of myogenesis. This is the first study to show a tissue-specific role for Hoip or its orthologs in vivo, to identify a robust genetic block in the second phase of myotube elongation, and to address post-transcriptional regulation of sarcomeric gene expression by a putative RBP during Drosophila embryogenesis.

MATERIALS AND METHODS

Drosophila genetics

All stocks were obtained from the Bloomington Stock Center unless otherwise noted. The stocks used in this study were: Df(2L)ED90, Df(2L)ED678, Df(2L)Exel0624, Df(2L)Exel7043, Df(2L)Exel7042, Df(2L)Exel8041, Df(2L)BSC216, Df(2L)BSC108, Df(3R)Exel6191, P[(lacW)hoip]5th, P[Mch.EMB] (Wells et al., 1996), Mhc(+) (Wells et al., 1996), P[Gal4-kirre]292P (Kirre292 lacZ) (Nose et al., 1998) and the Baylor P-element Mapping Kit (Zhai et al., 2003). The Cyo, P[Gal4-Twi], P[2x-UAS.eGFP]; Cyo, P[vg.lacZ]; and TM3, P[fit.lacZ] balancers were used to identify homozygous embryos.

EMS mutagenesis and gene mapping

Isogenic, starved w^{118} males were fed 25 mM EMS in 1% sucrose overnight and mated en masse as shown (supplementary material Fig. S1A). Mapping of hoip was performed as described previously (Zhai et al., 2003).

Immunohistochemistry, in situ hybridization and imaging

Antibodies used include anti-Mef2 (Lilly et al., 1995), anti-Nau (Wei et al., 2007), anti-Tin (Venkatesh et al., 2000), anti-MHC (Kiehart and Feghali, 1986), anti-Kr (Kosman et al., 1998), anti-MF20 (DSHB), anti-2007, anti-Tin (Venkatesh et al., 2000), anti-MHC (Kiehart and Feghali, 1994). However, the RNA-binding proteins (RBPs) that regulate the production of different MHC isoforms have not been identified.

RT-PCR

Appropriately aged embryos were dechorionated (Drosophila) or devitellinated (Danio), hand sorted to isolate homoyzogous mutants where needed, and homogenized in Trizol (Invitrogen). RNA was then extracted as per manufacturer’s specification. cDNA was generated using Superscript III (Invitrogen) and qPCR was performed with SYBR Green (Promega) using an ABI Prism 7000. Primers were designed to be intron spanning. qPCR reactions were run in triplicate and normalized to RpL32 (Drosophila) or GAPDH (zebrafish) expression. Primer sequences are provided in supplementary material Table S6.

Western blot

CO-1 cells were transfected with 1 μg of DNA according to manufacturer’s instructions (Fugene6, Roche), maintained for 48 hours, collected and lysed with NP40 lysis buffer. Western blots were performed as described previously (Mokalled et al., 2010).

Transgenes and site-directed mutagenesis

UAS constructs were generated by subcloning hoip (DGRC RE51843) and NHP2L1 (DF/HCC HSCD00326196) ORFs into pUASt. HA-tagged Hoip was made by PCR amplifying the hoip ORF and cloning the PCR fragment into pENTR (Invitrogen); after sequence verification, L/Re clone (Invitrogen) was used to recombine hoip into THW. For reporter genes, genomic DNA was PCR amplified, cloned into pCRII (Invitrogen), sequence verified, subcloned into pH.Stinger.eGFP (Barolo et al., 2004). Transgenic vectors were injected by standard methods to establish stable transgenic insertions. Multiple independent lines were characterized for each construct. Primer sequences are available upon request. Site-directed mutagenesis to generate Hoip-225GFP.AE was carried out as described previously (Johnson et al., 2011).

RNA-Seq

The SOLiD Total RNA-Seq Kit was used for RNA purification from 6-10 hour embryos and DNA library construction. Libraries were prepared in duplicate and sequenced on a 5500xl SOLiD Sequencer (Life Technologies) using a paired-end reading strategy. Sequencing reads were mapped to the UCSC reference genome using Lifescope (Life Technologies), then assembled and quantified using the Cufflinks algorithm (Trapnell et al., 2010). Assembled sequence reads were visualized using the Integrative Genomics Viewer (Robinson et al., 2011) and GO analysis was performed using the DAVID 6.7 bioinformatics resources (Dennis et al., 2003).

Zebrafish embryology

Fertilized, one-cell stage Tg(α-actin:GFP) zebrafish embryos were injected with 0.08 ng or 0.8 ng of nhp21b ATG-MO (GTTTCACTTCCGTTACGTCATCTT) or 0.8 ng of Cntrl-MO (Gene Tools). At 1 hpf, embryos were live imaged for GFP, fixed and analyzed for GFP and MF20 expression, or used for RNA isolation and qPCR analysis, by standard methods. Similar experiments were performed for nhp21b 5’UTR-MO (TTATTTTAATACACACGGTGCTCTTC). Annealed oligonucleotides encoding the ATG-MO target sequence were cloned into 5’ RV EGFPP T7TS (Small et al., 2005). Linearized transcript was transcribed with T7 mMessage machine (Ambion) to generate injectable RNA. Control and morphant embryos were prepared and imaged in parallel.

RESULTS

A forward genetic screen identified hoip as an essential regulator of myogenesis

We performed an EMS genetic screen to identify novel mutations that specifically affect somatic muscle morphology. Our screening strategy employed two reporter genes: MHC::GFP, which expresses a membrane-localized GFP in embryonic somatic muscles (Chen and Olson, 2001); and Hand.nGFP that expresses a nuclear-localized GFP in cardioblasts (CBs) by embryonic stage 12 (St12) (Johnson et al., 2011). The Hand.nGFP reporter served as a control to distinguish mutations that affected early mesoderm patterning events and cardiac cell fate specification, allowing us to focus on
mutations that specifically affected somatic muscle morphology. We screened ~10,000 genomes and identified 96 mutations on the second chromosome that caused somatic muscle defects (supplementary material Fig. S1A). We mapped two mutations to Kon-tiki (Kon), a known regulator of target site recognition (Schnorrer et al., 2007), and a second mutation to mind bomb 2, a known regulator of muscle attachment (Carrazco-Rando and Ruiz-Gómez, 2008).

One mutation that solely affected the pattern of MHC.τGFP expression mapped to a region of chromosome 2L that contains eight genes (Fig. 1A). Genetic analysis in this region showed that the mutation failed to complement P{lacW}hoip10704, so we named the EMS allele hoip1. The Hoip orthologs Snu13 in yeast and NHP2L1 (Non-Histone Protein 2 Like-1) in humans are RNA-binding proteins that bind noncoding RNAs associated with the spliceosome (Dobbyn and O’Keefe, 2004; Vidovic et al., 2000; Watkins et al., 2002). Compared with live wild-type (WT) embryos (Fig. 1B), somatic muscle morphology was aberrant in live hoip1 homozygous (Fig. 1C), hoip1/Df(2L)ED690 (Fig. 1D) and hoip1/P{lacW}hoip10704 transheterozygous embryos (Fig. 1E). In particular, the lateral transverse (LT1-4), lateral longitudinal (LL1), lateral oblique (LO1) and dorsal oblique (DO3-5) somatic muscles showed pronounced membrane extensions towards target sites, yet remained rounded in hoip1 mutant embryos (Fig. 1C-E; supplementary material Table S1). This failure in myogenesis prevented hoip1, hoip1/Df(2L)ED690 and hoip1/P{lacW}hoip10704 embryos from emerging from the chorion after embryogenesis.

The P{lacW}hoip10704 insertion was originally identified in a peripheral nervous system (PNS) screen (Kania et al., 1995; Prokopenko et al., 2000). P{lacW}hoip10704 embryos showed

---

**Fig. 1.** A mutation adversely affecting somatic muscle development maps to hoip. (A) The genomic region uncovered by Df(2L)ED690. Genes and direction of transcription are shown with blue arrows. Deficiencies that fail to complement hoip1 are shown in red; deficiencies that complement hoip1 are shown in dark blue. The minimal overlapping area among the deficiencies that fail to complement hoip1 contains eight genes. Of the four lethal transgene insertions (triangles) in the minimal overlapping area, only P{lacW}hoip10704 (red triangle) failed to complement hoip1. (B-E) MHC.τGFP, Hand.nGFP expression in St17 embryos. (B) Wild-type embryos express membrane-localized τGFP in each somatic muscle in all embryonic segments. Somatic muscles are severely rounded (arrowheads) in hoip1 (C), hoip1/Df(2L)ED690 (D) and hoip1/P{lacW}hoip10704 embryos (E). (B'-E') High-magnification views of embryos shown in B-E. (F) hoip1 is a G37E missense mutation (see supplementary material Fig. S1I). In this and subsequent figures, embryos are oriented with anterior towards the left and dorsal towards the top. Coordinates refer to base pair positions on chromosome 2L. Scale bars: 20 μm.
disorganized dorsal clusters of PNS neurons and axonal path finding defects (Kania et al., 1995). However, the P-element itself was not revertible (Kania et al., 1995) and P[Fr leth $3^{0/104}]$ embryos showed global patterning defects that we did not observe in other $hoip$ mutant combinations (supplementary material Fig. S1B,C). We assayed PNS morphology in $hoip^1$ and $hoip^1/P[Fr leth $3^{0/104}]$ embryos ($n\geq3$) but could not confirm the previously reported PNS defects (supplementary material Fig. S1D-F). These data suggest a lethal mutation on the $P[Fr leth $3^{0/104}]$ chromosome outside $hoip$ that disrupts a powerful regulator of embryonic patterning that could affect PNS development.

Sequencing the $hoip^1$ allele revealed a G37E missense mutation (Fig. 1F; supplementary material Fig. S1J) within the predicted Hoip RNA-binding domain (Schultz et al., 2006). Based on the predicted crystal structure of this domain (supplementary material Fig. S1G), the acidic amino acid substitution would be expected to eliminate Hoip RNA-binding activity (Vidovic et al., 2000). A tagged $hoip$ protein harboring the G37E mutation was detectable by western blot in transfected COS-1 cells, demonstrating that $hoip^1$ is not a protein null mutation (supplementary material Fig. S1H).

**Myotube elongation does not initiate in $hoip$ embryos**

To understand the stage of myogenesis regulated by Hoip, we performed time-lapse studies in embryos expressing τGFP under the control of the founder cell driver rp298.gal4. Our analysis focused on the LL, LT, LO and DO muscles, as these muscles were most often disrupted in $hoip$ mutant embryos (supplementary material Table S1). Our analysis began in late St12 embryos that expressed rp298>τGFP in nascent myotubes (Fig. 2A). In wild-type embryos, myotubes showed obvious polarization and had elongated to 50% of segment width after 30 minutes. By 60 minutes, the myotubes had largely completed their extension and created extensive filopodia for attachment site recognition (Fig. 2A; $n=2$). By contrast, nascent myotubes failed to elongate in $hoip^1$ embryos after 30 minutes even though the myotubes showed an initial polarity (Fig. 2B). By 60 minutes, myotubes failed to extend to 50% of segment width in $hoip^1$ embryos and had lost their polarity (Fig. 2B; $n=3$). These results demonstrate that myotubes failed to elongate and reach their attachment sites in $hoip^1$ embryos.

To understand whether Hoip controls target site recognition, we repeated time-lapse imaging at higher magnification to document filopodia in detail (supplementary material Fig. S2; $n=3$). $hoip^1$ embryos extended filopodia exclusively in the direction of myotube polarity. This phenotype contrasts with that of Kon mutant embryos, which initiate myotube elongation but, instead of orienting filopodia solely toward muscle attachment sites, extend ectopic filopodia in all directions (Schnorrer et al., 2007). Kon is a transmembrane receptor that regulates myotube target site recognition. As $hoip$ myotubes do not phenocopy Kon myotubes, we conclude that Hoip does not control attachment site recognition.

**Muscle attachment sites are specified in $hoip$ embryos**

To understand whether tendon cells, which mediate muscle attachment, were specified in $hoip^1$ embryos, we assayed expression of βPS (myospheroid), an effector of muscle attachment. βPS is expressed in tendon cells and localizes to myotendinous junctions after attachment-site recognition (Martin-Bermudo and Brown, 2000). βPS was clearly detectable in the epidermis of $hoip^1$ embryos, but showed diffuse localization along the dorsoventral axis compared with wild-type embryos (Fig. 2C,D). As βPS is also

---

**Fig. 2. $hoip$ embryos have myotube elongation defects.** (A,B) Time-lapse images of rp298>τGFP embryos initiated at late St12. (A) Wild-type embryos showed robust myotube elongation at 30 minutes (double arrows) and developed extensive filopodia for attachment site recognition at 60 minutes (white arrows). (B) Myotubes established polarity in $hoip^1$ embryos at 15 minutes but failed to elongate by 30 minutes. Polarized myofibers at 15 minutes compacted over time (double-headed arrows). (C,D) St16 rp298>τGFP embryos double labeled for GFP and βPS. (C) βPS localizes to myotendinous junctions in wild-type embryos. (D) Tendon cells express βPS in $hoip^1$ embryos but localization is diffuse (red arrowheads). (E,F) $hoip^1$ embryos. (E,F') Talin expression alone. mg, midgut. Scale bars: 20 μm.
expressed in somatic muscle, we next examined Talin expression. Talin is restricted to tendon cells and acts as a linker between βPS and the cytoskeleton. Similar to βPS, Talin is clearly expressed in the epidermis of hoip1 embryos (Fig. 2E,F). Taken together, these results show that the rounded muscle phenotype in hoip1 embryos is due to a block at, or prior to, myotube elongation and is not a result of tendon cell mis-specification.

**Hoip does not regulate founder cell specification or the first round of myoblast fusion**

To further characterize the myogenic phenotype in hoip1 embryos, we examined founder cell specification and myoblast fusion. After specification, founder cells undergo an initial round of fusion that is complete by the end of St12 (Bate, 1990). Subsequent fusion then determines final muscle size and each muscle undergoes a unique number of fusion events. MHC serves as a classic marker for identifying myoblast fusion defects; embryos with defects in the first round of myoblast fusion robustly express MHC in single unfused founder cells (Chen and Olson, 2001). To control against possible dominant mutations on the EMS chromosome, we compared MHC expression in hoip1/Cyo.lacZ embryos with heterozygous hoip1/Cyo.lacZ embryos. Strikingly, the somatic musculature of St16 hoip1 embryos showed almost no MHC protein expression, whereas hoip1/Cyo embryos showed normal MHC expression and somatic muscle morphology (Fig. 3A,B; Table 1).

Another method for identifying myoblast fusion defects is to quantify the temporal expression of muscle identity genes in the dorsal mesoderm (Chen and Olson, 2001). The identity gene nautilus (nau) is expressed in a subset of founder cells that give rise to the somatic muscles affected in hoip1 embryos, including DA3, DO3, DO4, DO5, VA1, LO1 (Wei et al., 2007). The number of Nau+ nuclei in hoip1 embryos was comparable with hoip1/Cyo.lacZ embryos at St12, but significantly less at St14 (supplementary material Fig. S3A,B). Thus, founder cell specification and the first round of myoblast fusion proceed normally in hoip1 embryos; however, the later rounds of fusion do not occur.

---

**Fig. 3. Hoip regulates somatic muscle and cardioblast maturation but not precursor specification.** (A,B) Mef2 and MHC protein expression in St16 embryos. Lateral views. Robust MHC and Mef2 expression is detectable in somatic muscles of hoip1/Cyo.lacZ embryos (A). Mef2 expression is unaffected in hoip1 embryos, whereas MHC is nearly absent from the somatic muscle (B). (C-F) St16 rp298.gal4>τ.GFP, rp298.nlacZ embryos double-labeled for GFP (red) and lacZ (green). (C-D) Dorsal muscles. The number of lacZ+ nuclei is reduced in hoip1 embryos (C) compared with hoip1/Cyo.lacZ embryos (D); however, binucleated dorsal muscles show complete elongation (arrowheads). (E-F) Lateral and ventral muscles. The number of lacZ+ nuclei is also reduced in lateral and ventral muscles in hoip1 embryos. Multinucleate lateral muscles show incomplete elongation (arrow). (G,H) Mef2 and MHC protein expression in St16 embryos. Dorsal views. (G) hoip1/Cyo.lacZ embryos express Mef2 and MHC in mature CBs. (H) hoip1 embryos express Mef2 but not MHC in a great majority of CBs. (I-L') Mef2 and Tin protein expression. (I,J) hoip1/Cyo.lacZ embryos express Mef2 in all myogenic precursors, including CBs. Tin is expressed in four Mef2+ CBs per hemisegment at St13 (I; lateral view) and St16 (K; dorsal view). Mef2 and Tin expression in hoip1 CBs is comparable with control embryos at St13 (I) and St16 (L). (K,L') Tin expression alone. (M,N) High magnification micrographs of visceral muscles in St16 embryos. MHC expression is comparable between hoip1/Cyo.lacZ embryos (M) and hoip1 embryos (N). Both genotypes develop LVMs and CVMs in the visceral mesoderm. (O) hoip1 rp298>Hoip embryos express MHC protein at near wild-type levels in the somatic mesoderm. SM, somatic muscle; VM, visceral muscle; LVM, longitudinal visceral muscle; CVM, circular visceral muscle; CBs, cardioblasts. Open arrowheads in UK show ectodermal cytoplasmic lacZ expression that distinguishes hoip1 heterozygotes from homozygotes. Scale bars: 20 μm.
To confirm this result, we used the founder cell transgene rp298.nlaCZ to assay founder cell specification and myoblast fusion. Similar to Nau, the number of lacZ-positive nuclei was comparable between hoip\(^1\) and hoip\(^1\)/CyO.nlaCZ embryos at St12 (supplementary material Fig. S3C,D), but significantly reduced at St14 and St16 (Fig. 3C-F; supplementary material Fig. S3E-H). However, the fusion defects in hoip\(^1\) embryos were not restricted to the muscles that showed elongation defects. For example, DO1 and DA1 muscles elongated normally in hoip\(^1\) embryos but showed dramatically fewer lacZ-positive nuclei than hoip\(^1\)/CyO.nlaCZ embryos (Fig. 3C,D). However, LL1 and LO5 were multineucleate in hoip\(^1\) embryos but failed to elongate (Fig. 3E,F).

**Somatic and cardiac muscle maturation is Hoip dependent**

One explanation for the somatic muscle defects in hoip\(^1\) embryos was that MHC itself is required for myotube elongation. However, embryos homozygous for the null mutation MHC\(^1\) (Wells et al., 1996) showed normal myotube elongation (supplementary material Fig. S3I,J). Sarcomere assembly occurs after myoblast elongation and myofiber attachment (Rui et al., 2010) and embryos defective for the first round of myoblast fusion do express MHC (Chen and Olson, 2001). Together, these observations demonstrate that muscle morphogenesis is genetically separable from muscle structural expression and prompted us to define a secondary role for Hoip during myogenesis.

In *Drosophila*, MHC is the single muscle myosin and is expressed in cardiac, somatic and visceral muscle (Bernstein et al., 1983). Mef2 is also expressed in all myogenic cells and is a direct transcriptional activator of MHC (Bour et al., 1995). Mef2 was expressed at comparable levels in the somatic and cardiac mesoderm of hoip\(^1\) and hoip\(^1\)/CyO.nlaCZ embryos (Fig. 3A,B,G-L), even though MHC was largely absent from both tissues in hoip\(^1\) embryos (Fig. 3B,H). The expression of a second transcriptional activator of MHC, Tinman (Tin), is restricted to and orchestrates the maturation of a subset of cardioblasts (CBs) (Reim et al., 2005). Cardiac Tin expression was also comparable between hoip\(^1\) and hoip\(^1\)/CyO.nlaCZ embryos (Fig. 3I-L). As *Drosophila* CBs are mononucleate, do not undergo elongation, yet fail to express MHC in hoip\(^1\) embryos, we conclude that Hoip regulates muscle maturation (i.e. muscle structural protein expression) independently of myotube elongation.

Even though MHC expression was largely absent from CBs and somatic muscles in hoip\(^1\) embryos, MHC expression in mature visceral muscle was comparable between hoip\(^1\) and hoip\(^1\)/CyO.nlaCZ embryos (Fig. 3M,N). These results suggest that Hoip performs tissue-specific functions during myogenesis to specifically regulate striated muscle maturation.

**Hoip regulates terminal muscle differentiation**

To confirm that Hoip regulates myogenesis after founder cell specification, we expressed Hoip with rp298.Ga4 in hoip\(^1\) embryos and assayed MHC expression. Founder cell-specific expression of Hoip was indeed sufficient to rescue myotube elongation and MHC protein expression in somatic muscles of hoip\(^1\) embryos (Fig. 3O; Table 1). Hoip therefore regulates myogenesis after founder cell specification in a mesoderm cell-autonomous manner.

**hoip is expressed in striated but not visceral muscle progenitors**

*In situ* hybridization using a probe antisense to the full-length hoip transcript (Fig. 4A) showed hoip expression initiates at low levels in the mesoderm and endoderm of St9 embryos (supplementary material Fig. S4A-C). Robust expression hoip mRNA could be detected in St11 embryos and, consistent with the tissue-restricted MHC phenotype in hoip\(^1\) embryos, is expressed in the Mef2-expressing cells of the somatic and cardiac mesoderm, but not in the Mef2-expressing cells of the visceral mesoderm (Fig. 4B). hoip mRNA is also expressed in the fat body and the endoderm at this stage, but is absent from the neuroectoderm ventral to the Mef2 expression domain. hoip mRNA continues to be expressed in the somatic musculature throughout embryogenesis (Fig. 4C; supplementary material Fig. S4D,E). We did not detect hoip in the PNS by *in situ* hybridization.

To confirm the *in situ* results, we generated a GFP reporter construct that contained 225 bp of genomic DNA upstream of the hoip-coding sequence (Hoip.-225.GFP). This reporter gene directed GFP expression in a pattern that recapitulated hoip mRNA expression in St11 and St13 embryos (Fig. 4D,E). Interestingly, the Hoip.-225 sequence contains a conserved E-box sequence (CANNTG, supplementary material Fig. S4F). Basic helix-loop-helix (bHLH) transcription factors bind E-box sequences and a hoip reporter gene with a mutated E-box (Hoip.-225AE.GFP) initiated GFP expression in a manner comparable with Hoip.-225.GFP (Fig. 4F) but did not maintain GFP expression in the mesoderm through St13 (Fig. 4G). These findings demonstrate that hoip is expressed in the striated muscle lineage after precursor cell specification, and that maintenance of hoip expression depends on a conserved E-box sequence that is likely a bHLH target.

We next assayed Hoip localization in the somatic musculature using an HA-tagged Hoip transgene. Surprisingly, Mef2\(\Delta\)Hoip-HA embryos showed both nuclear and cytoplasmic localization of Hoip-HA (supplementary material Fig. S5). Hoip may thus perform multiple molecular functions during myogenesis.

**Several sarcomeric genes are downregulated in hoip embryos**

The Hoip orthologs Snu13 in yeast and NHP2L1 in humans are spliceosomal RNA-binding proteins (Dobyn and O’Keefe, 2004; Vidovic et al., 2000; Watkins et al., 2002). In eukaryotes, spliceosomes that contain small nuclear (sn) RNAs are believed to remove intronic sequences from pre-mRNAs, whereas spliceosomes that contain small nucleolar (sno) RNAs orchestrate ordered cleavages along pre-rRNAs. Snu13/NHP2L1 proteins preferentially bind to GA-rich RNA sequences in the kink-turn motif of both snRNAs and snoRNAs (Cléry et al., 2007; Nottrott et al., 1999; Schultz et al., 2006; Vidovic et al., 2000).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>hoip(^1)/CyO</th>
<th>hoip(^1)</th>
<th>hoip(^1) Rp298 &gt;Hoip</th>
<th>hoip(^1) Rp298 &gt;NHP2L1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHC+ myofibers(\ast)</td>
<td>10.3±1.16</td>
<td>1.4±1.12</td>
<td>7.4±1.56</td>
<td>6.0±1.82</td>
</tr>
<tr>
<td>n</td>
<td>26</td>
<td>35</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td>#LT1-LT4</td>
<td>130/136</td>
<td>3/136</td>
<td>52/128</td>
<td>35/88</td>
</tr>
</tbody>
</table>

\(\ast\)Twelve lateral muscles assayed including DO3-5, DA3, DT1-4, LL1-LO1 and SBM.

**Table 1. MHC expression**
We took a non-biased approach to identify potential Hoip targets in the developing mesoderm. As robust hoip expression initiates at St11 and continues at high levels through St13, we performed RNA-seq in St11-13 (6-10 hour) embryos. Our analysis identified 353 transcripts that were differentially expressed and 60 transcripts that were expressed approximately at wild-type levels but inappropriately processed in hoip1 embryos (supplementary material Tables S2, S3). This RNA-seq analysis also identified the G37E missense mutation in hoip1 embryos, confirming the initial genomic sequencing data. In addition, 45S pre-rRNA was processed correctly in hoip1 St11-13 embryos, suggesting that Hoip does not regulate ribosome biogenesis during these stages of development (supplementary material Fig. S6A). These in vivo results demonstrate that Hoip is not required to process all pre-mRNA or pre-rRNA transcripts during embryogenesis.

We analyzed the misregulated transcripts in hoip1 embryos by Gene Ontology (GO) functional annotation clustering and found the most significant cluster associated with the GO term Contractile Fiber (Fig. 5A). Strikingly, transcripts within the Contractile Fiber cluster (Table 2) include Mhc and other sarcomere components, including inflated (if), Myosin light chain 2 (Mlc2), Tropomyosin 2 (Tm2), Troponin C at 47D (TpnC47D) and Troponin C at 73F (TpnC73F). We confirmed the RNA-seq data for these sarcomeric
genes by quantitative PCR (qPCR) and found that each transcript was dramatically downregulated in hoip\(^1\) embryos compared with controls (Fig. 5B). The RNA-seq data showed that the embryonic sarcomeric actins Act57B and Act87E, and Mef2, the only known robust transcriptional regulator of terminal muscle differentiation genes in Drosophila, were expressed at wild-type levels in hoip\(^1\) embryos (supplementary material Fig. S6B, Table S2). The developmental time point of our RNA-seq coincided with the onset of muscle structural gene expression at St12. However, hoip\(^1\) embryos fail to express MHC protein at all developmental stages (Fig. 3B). These observations suggest that Hoip is required to both initiate and maintain muscle structural gene expression during embryogenesis.

We examined 22 genes experimentally shown to regulate myotube elongation, attachment site recognition or myotendinous junction formation in our RNA-seq data (supplementary material Table S4). Surprisingly, pav expression was not changed in hoip\(^1\) embryos, whereas tum was upregulated (fold change=2.04). However, tum overexpression does not affect myotube elongation (Guerin and Kramer, 2009b). Of the remaining 19 genes, only MSP-300 showed significant downregulation in hoip\(^1\) embryos (fold change=0.27). Unlike hoip\(^1\) embryos, MSP-300 mutant embryos show only a modest somatic muscle phenotype that initiates late in embryogenesis (Rosenberg-Hasson et al., 1996); however, MSP-300 larvae do show defects in nuclear positioning and microtubule organization (Elhanany-Tamir et al., 2012). The 22 genes regulating somatic muscle morphology also showed normal splicing in hoip\(^1\) embryos, further arguing that Hoip regulates the expression of other transcripts to initiate myotube elongation.
Post-transcriptional regulation of MHC

The RNA analyses suggested that Hoip processes pre-mRNAs encoding sarcomeric proteins but does not regulate transcription or rRNA processing. To confirm these results, we performed a functional rescue experiment with the MHCemb transgene (Fig. 5C), which uses the endogenous MHC promoter to express a MHC cDNA specifically in somatic muscle (Hess et al., 2007; Wells et al., 1996). If Hoip regulated either ribosome biogenesis or processing of an mRNA whose protein product activates MHC transcription, then MHCemb would not be expected to rescue MHC protein expression in hoip1 embryos. However, if Hoip acts post-transcriptionally to splice the MHC pre-mRNA, then the MHC cDNA, which is expressed from the MHCemb transgene, would generate a functional mRNA that would be appropriately translated in hoip1 embryos.

Indeed, the MHCemb transgene restored MHC protein expression in somatic but not cardiac muscle of hoip1 embryos (Fig. 5D-F,I; supplementary material Fig. S7). This experiment corroborated our MHC:GFP expression studies in which the endogenous MHC promoter directed GFP protein expression throughout the somatic mesoderm of hoip1 embryos (Fig. 1C-E). In addition, the MHCemb transgene did not restore somatic muscle morphology in hoip1 embryos, further confirming that Hoip regulates myotube elongation independent of MHC expression. We conclude Hoip is required to perform at least one splice in the MHC pre-mRNA and functions independently of ribosome biogenesis to direct somatic muscle maturation.

MHC mRNA does not translocate out of the nucleus in hoip embryos

Two separate transgenes harboring MHC promoters (MHC:τGFP and MHCemb) were able to direct cDNA expression in hoip1 embryos (Fig. 1C, Fig. 5F). However, MHC mRNA was nearly undetectable in hoip1 embryos by RNA-seq and qPCR (Fig. 5B). We assayed MHC mRNA localization by in situ hybridization and found only punctate, nuclear MHC mRNA localization in hoip1 embryos, even though wild-type embryos showed MHC mRNA localization throughout the myofiber (Fig. 5G,H). These transgenic and in situ results clearly demonstrate that MHC is transcribed in hoip1 embryos but that the transcript fails to translocate out of the nucleus.

Hoip orthologs are essential regulators of myogenesis

The remarkable homology between Hoip and human NHP2L1 (supplementary material Fig. S11) suggested that the function of Hoip during myogenesis is conserved across species. To test this hypothesis, we expressed human NHP2L1 in founder cells of hoip1 embryos with rp298.gal4 and assayed MHC protein expression. Although human NHP2L1 did not rescue hoip1 embryos as effectively as Drosophila Hoip (Table 1), we observed a significant restoration of myotube elongation and MHC expression in the somatic musculature, demonstrating that Hoip and human NHP2L1 can perform similar functions during myogenesis (Fig. 6A,B).

Zebrafish nhp2l1b is expressed in the paraxial mesoderm at 10 hour post-fertilization (hpf) and throughout the myotome at 19 hpf.
(Thisse and Thisse, 2001). We asked whether nhp2l1b is essential for zebrafish muscle development using two independent MOs to knockdown endogenous nhp2l1b. One MO targets the nhp2l1b translational start site with 100% identity (ATG-MO), but does not target the highly divergent nhp2l1a (supplementary material Fig. S8A-C). The other MO targets the nhp2l1b 5’ UTR 57 bp upstream of the translational start site and shows no sequence similarity to nhp2l1a (supplementary material Fig. S8B). The α-actin:GFP transgenic (Tg) line harbors a skeletal muscle reporter (Higashijima et al., 1997) and Tg(α-actin:GFP) embryos injected with control-MO showed robust GFP expression in the somitic mesoderm by 14 hpf (Fig. 6C). However, both ATG-MO- and 5’UTR-MO-injected embryos showed little or no GFP expression (Fig. 6D,E). The MF20 antibody reacts with all muscle MyHC isoforms, and we observed robust somitic MF20 staining in control-MO but not ATG-MO-injected embryos at 14 hpf (Fig. 6F,G).

Tg(α-actin:GFP) expression showed a dose-dependent response to MO concentrations (Fig. 6H). Sixty-five percent (n=259) of embryos injected with 0.08 ng ATG-MO and 96% (n=233) of embryos injected with 0.8 ng ATG-MO failed to initiate α-actin:GFP expression; 43% (n=58) of embryos injected with 0.08 ng 5’UTR-MO and 72% (n=87) of embryos injected with 0.8 ng 5’UTR-MO failed to initiate α-actin:GFP expression; 11% (n=302) of embryos injected with 0.8 ng Ctrl-MO showed changes in α-actin:GFP expression. With the exception of skeletal muscle marker expression, 0.8 ng ATG-MO-treated embryos appeared normal through 16 hpf; however, MO treatment induced lethality after 16 hpf (supplementary material Fig. S8F,G). The ATG-MO also blocked eGFP translation when the nhp2l1b ATG-MO target site was placed upstream of the eGFP-coding sequence (supplementary material Fig. S8D,E). Thus, the ATG-MO targets nhp2l1b.

By qPCR, we found that several transcripts encoding sarcomeric proteins were present at reduced levels in ATG-MO 14 hpf embryos, including two slow MyHCs, two troponins and one tropomyosin. Importantly, the expression of other genes essential for mesoderm development, such as mef2 and notch1, was unaffected in ATG-MO embryos. Taken together, these results indicate that the function of Hoip is highly conserved and that nhp2l1b regulates myogenesis in vertebrates.

**DISCUSSION**

The results of this study reveal a specific and essential role for the putative RBP Hoip in the control of embryonic muscle development. Hoip is expressed in the striated muscle lineage and regulates two distinct processes: myotube elongation and sarcomeric protein expression. Using functional rescue experiments, we have established that Hoip regulates MHC pre-mRNA splicing but not MHC transcription. The human hoip ortholog human NHP2L1 can rescue myogenesis in hoip mutant embryos, and antisense nhp2l1b knockdown blocks muscle development in zebrafish. This study is the first to identify a tissue-specific function for hoip or its orthologs in vivo and highlights the essential role of post-transcriptional gene regulation during tissue morphogenesis.

Non-coding RNAs that function as ‘core’ spliceosome components can have tissue-specific functions in vivo. For example, the mouse genome encodes multiple U2 snRNA genes and the rnu2-8 U2 snRNA is differentially expressed in the mouse nervous system with peak expression levels in the cerebellum (Jia et al., 2012). Rnu2-8 knockout mice show normal splicing of constitutive exons, but incomplete splicing of alternative exons solely within the cerebellum (Jia et al., 2012), and rnu2-8 U2 snRNA is essential for neuron survival in the cerebellum.

The Hoip orthologs Snu13/NHP2L1 proteins have been characterized as ‘core’ spliceosome components that bind the kink turn motif of U4 snRNAs, U3 snoRNAs and U14 snoRNAs (Schultz et al., 2006). However, we found Hoip expression and function is restricted to the striated muscle lineage within the Drosophila mesoderm. RNA-seq and MHCm embryos rescue data clearly demonstrate that Hoip is not a global regulator of pre-mRNA splicing or ribosome biogenesis, but acts specifically on a set of RNAs that encode functionally related proteins.

The MHCm transgene contains a fully spliced MHC CDNA that encodes exons 2-19 (Wells et al., 1996). The 5’ end of the transgene contains genomic DNA that initiates 450 bp upstream of the first transcriptional start site and terminates in exon 2. This 5’ sequence contains the necessary enhancer and promoter elements for transgene expression in the somatic mesoderm, as well as three alternative transcriptional start sites. The 3’ end of the transgene contains a complete exon 19 with multiple polyadenylation (poly A) sites, but it does not contain an exogenous poly A signal (i.e. SV40). As MHCm rescues MHC protein expression in hoip embryos, Hoip must not regulate transcriptional start site selection or 5’UTR stability. In addition, the endogenous 3’UTR is sufficient to restore MHC protein expression, indicating that Hoip does not regulate poly A site choice, polyadenylation itself or 3’UTR stability. Hoip therefore acts post-transcriptionally to control at least one splicing event in exons 2-19.

The apparent specificity with which Hoip targets sarcomeric RNAs is striking. One explanation for this specificity is that the Hoip paralog Nhp2 fulfills the spliceosome functions that Hoip does not. A more intriguing hypothesis is that Hoip facilitates ribonucleotide modifications in a subset of transcripts. For example, NHP2L1/snoRNA processomes direct 2’-O methylation of ribosomal pre-RNAs (Watkins et al., 2002) and 2’-O methylation has been reported to enhance pre-mRNA splicing in some contexts (Ge et al., 2010). Perhaps Hoip confers a similar modification to sarcomeric RNAs that is permissive for pre-mRNA splicing; in this model, unmodified pre-mRNAs would not be spliced and would degrade in the nucleus. We envision Hoip-mediated modifications to be a rate-limiting step that ensures proper stoichiometry of sarcomeric proteins.

This study also identified myotube elongation defects in hoip embryos. To our knowledge, this is the first mutation reported that blocks the initiation of myotube elongation (supplementary material Table S4). Accordingly, we failed to identify robust misregulation of genes known to regulate myotube elongation, attachment site recognition or myotendinous junction formation in hoip embryos (supplementary material Table S4). The strength of the hoip phenotype suggests that a suite of proteins is required for myotube elongation. A second possibility is that some sarcomeric proteins could be required for myotube elongation. We have shown that MHC does not regulate elongation; however, tropomyosins regulate actin dynamics outside the sarcomere that influence cell polarity and cell outgrowth in Drosophila (Li and Gao, 2003; Zimyanin et al., 2008).

One known regulator of microtubule dynamics during elongation, Tum, is a Rac family GTPase-activating protein (RacGAP). We searched the RNA-seq data for other Rac/Rho family regulators and identified two RhoGAPs and one Rho guanine nucleotide exchange factor (RhoGEF) that were misregulated in hoip embryos (supplementary material Table S5). We also identified two Rab family GTPases, the function of which in synaptic endosome trafficking has been well established (Gurkan et al., 2005). Rab family members are essential for microtubule-dependent outgrowth.
of tracheal and bristle cells (Nagaraj and Adler, 2012; Schottenfeld-Roames and Ghribali, 2012) and the myotube guidance molecule Grip localizes to endosomes at the ends of extending myotubes (Swan et al., 2004). Finally, overexpressing the tendon cell regulator Stripe in the ectoderm upregulates a number of novel genes (Gilsolin and Volk, 2010a; Gilsolin and Volk, 2010b), including the transmembrane protein Tetraspanin 42Ea (Tsp42Ea). Tsp42Ea orthologs function in cell motility and signal transduction, and Tsp42Ea was downregulated in hoip embryos. It will be interesting to determine which of these genes are required for myotube elongation in vivo.

The zebrafish Hoip ortholog nhp21lb is expressed in the paraxial mesoderm during the 10- to 14-somite stage and in the myotome during the 20- to 25-somite stage (Thissle and Thissle, 2001). Thus, the expression and myogenic function of hoip/nhp21lb appears to be conserved in vertebrates. Despite being one of the most intensely studied developmental systems, skeletal muscle continues to reveal novel developmental mechanisms. In the future, it will be of particular interest to characterize the interactions between transcriptional and post-transcriptional mechanisms that coordinate final muscle morphology and function.

Acknowledgements

We appreciate Mike Buszczak for insights and discussions throughout this study. We thank Sandy Bernstein, Elizabeth Chen, Richard Cripps and Bruce Paterson for reagents.

Author contributions

Experiments were carried out by A.N.J., M.H.M. and J.M.V. Experimental design, analyses and preparation of the manuscript were carried out by A.N.J., M.H.M., K.D.P. and E.N.O.

Supplementary material

Supplementary material available online at http://dev.biologists.orglookup/suppl:doi:10.1242/dev.095596/DC1

References


