A candidate RNAi screen reveals diverse RNA-binding protein phenotypes in *Drosophila* flight muscle

Shao-Yen Kao1†, Elena Nikonova1†, Sabrina Chaabane1, Albiona Sabani2, Alexandra Martitz3, Anja Wittner1,4, Jakob Heemken1, Tobias Straub5, Maria L. Spletter1,*

1 Biomedical Center, Department of Physiological Chemistry, Ludwig-Maximilians-Universität München, Großhaderner Str. 9, 82152 Martinsried-Planegg Germany
2 Department of Biology, University of Wisconsin at Madison, 1117 W. Johnson St., Madison, WI, 53706, USA
3 Molecular Nutrition Medicine, Else Kröner-Fresenius Center, Technical University of Munich, 85354 Freising, Germany
4 Current: Gene Center Munich, Department of Biochemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany
5 Biomedical Center, Bioinformatics Core Facility, Ludwig-Maximilians-Universität München, Großhaderner Str. 9, 82152 Martinsried-Planegg Germany

† These authors contributed equally.

* Correspondence: maria.spletter@bmc.med.lmu.de; Tel.: +49 89 2180 77086

Abstract: The proper regulation of RNA processing is critical for muscle development and the fine-tuning of contractile ability between muscle fiber-types. RNA binding proteins (RBPs) regulate the diverse steps in RNA processing including alternative splicing, which generates fiber-type specific isoforms of structural proteins that confer contractile sarcomeres with distinct biomechanical properties. Alternative splicing is disrupted in muscle diseases such as myotonic dystrophy and dilated cardiomyopathy, and is altered after intense exercise as well as with aging. It is therefore important to understand splicing and RBP function, but currently only a small fraction of the hundreds of annotated RBPs expressed in muscle have been characterized. Here we demonstrate the utility of *Drosophila* as a genetic model system to investigate basic developmental mechanisms of RBP function in myogenesis. We find that RBPs exhibit dynamic temporal and fiber-type specific expression patterns in mRNA-Seq data and display muscle-specific phenotypes. We performed knockdown with 105 RNAi hairpins targeting 35 RBPs and report associated lethality, flight, myofiber and sarcomere defects, including flight muscle phenotypes for *Doa*, *Rm62*, *mub*, *mbf*, *sbr*, and *clu*. Interestingly, knockdown phenotypes of spliceosome components, as highlighted by phenotypes for A-complex components SF1 and Hrb87F (hnRNPA1), revealed level- and temporal-dependent myofibril defects. We further show that splicing mediated by SF1 and Hrb87F is necessary for Z-disc stability and proper myofibril development, and strong knockdown of either gene results in impaired localization of Kettin to the Z-disc. Our results expand the number of RBPs with a described phenotype in muscle and underscore the diversity in myofibril and transcriptomic phenotypes associated with splicing defects. *Drosophila* is thus a powerful model to gain disease-relevant insight into cellular and molecular phenotypes observed when expression levels of splicing factors, spliceosome components and splicing dynamics are altered.

Keywords: RNA binding proteins; SF1; Hrb87F; Bru1; *Drosophila*; flight muscle; RNAi; splicing

1. Introduction

Animals from flies to humans possess hundreds of distinct muscles that are characterized by differences in both morphology and contractile properties [1]. An outstanding question in developmental biology is how this diversity is generated, and in particular which cytoskeletal adaptations support such a diverse range of contractile abilities.
Sarcomeres, the simplest contractile subunits of muscle, are mini-motors driven by the interaction between antiparallel filaments of actin, anchored at the z-disc, and myosin, anchored at the M-line [2]. Sarcomeres are built of more than a hundred proteins and assemble end-to-end into long myofibrils that span the length of a muscle [3]. Small changes in the structure and ratios of these structural proteins through altered gene expression, alternative splicing, or mRNA regulatory dynamics can alter the biomechanics of muscle contraction, and thereby serve as a mechanism to fine-tune contractile properties [4–7]. For example, expression of short, stiff Titin isoforms in the heart and longer, more flexible Titin isoforms in skeletal muscle contributes to physiological differences between these muscle types [8,9]. Misregulation of RNA processing therefore leads to a vast collection of muscle diseases from myotonic dystrophies to cardiomyopathies [10–13]. Moreover, exercise, adaptation, aging and recovery after injury also involve changes in RNA regulation [14–16]. Thus, understanding normal RNA regulatory dynamics in developing muscle is foundational to understanding normal muscle physiology as well as disease.

1.1. RNA-binding proteins with diverse structures and functions regulate RNA dynamics

RNA binding proteins (RBPs), which recognize and bind defined RNA sequences, control all steps in the mRNA life-cycle, from exon definition and alternative splicing to mRNA trafficking, nuclear export, localization, translation dynamics and turn-over [17,18]. RBPs therefore play a key role in determining which structural gene isoforms are expressed in muscle, in which ratios and with which temporal and fiber-type specific dynamics [7,11]. Canonical RBPs have one or more RNA-binding domains, for example an RNA-recognition motif (RRM), hnRNP homology domain (KH), double-stranded RNA-binding domain (dsRBD) or DEAD/DEAH box [4]. Other RBPs bind RNA through intrinsically disordered, or unstructured regions. RNA interactome capture (RIC) studies, which unbiasedly identify RNA-bound proteins, suggest that hundreds of such proteins lacking canonical RNA-binding domains are capable of binding RNA in a developmentally or spatially regulated manner [19,20]. RBPs often contain additional protein domains, from Zinc-fingers and PDZ domains that facilitate protein-protein interactions and the assembly of multimeric RBP regulatory complexes such as LASR (large assembly of splicing factors) [21] to G-patch and serine-rich (RS) domains frequently associated with splicing factors [22] to prion-like domains (PrLDs) that together with RNA-binding facilitate the formation of phase-separated liquid compartments [23]. This illustrates that a major challenge in deciphering RNA regulatory dynamics in healthy and diseased muscle is the tissue-type specific identification of RBPs and their functions.

1.2. Regulation of RBPs modulates RNA processing in muscle

Functionally, RBPs often have manifold RNA targets and can regulate multiple steps in RNA processing. A survey of 56 RBPs in Drosophila S2 cells found that individual RBPs can regulate from tens to hundreds of splice events [24], while studies in vertebrates suggest that CELF, RBFOX and MBNL may regulate hundreds to thousands of events in muscle [25,26]. TDP-43, which is associated with neurodegenerative disorders as well as inclusion body myositis and rimmed vacuole myopathies, regulates not only alternative splicing but also mRNA trafficking and translation and mRNP granule formation [27,28]. CELF and MBNL proteins regulate mRNA stability, translation and localization in addition to alternative splicing [26,29,30]. RBPs function as part of regulatory networks subject to cooperativity, feedback and cross-regulation, as exemplified in muscle by interactions between Rbfox, CELF and MBNL family members [11,25,31,32]. RBPs themselves are also subject to multiple levels of regulation. For example, RBP activity or expression level can be temporally regulated: higher levels of CELF1 expression during early myogenesis promote embryonic splicing, while increased levels of MBNL1 at later stages of differentiation promote post-natal splicing patterns in skeletal muscle [33]. RBPs can be alternatively spliced to produce different isoforms that localize to distinct sub-cellular compartments: nuclear-localized MBNL isoforms regulate alternative splicing, while cytoplasmic-localized MBNL isoforms regulate mRNA trafficking and stability. RBPs can be post-
translationally modified to regulate their activity and stability: serine/threonine kinase 38 (Stk38) binds and modifies Rbm24 in cardiomyocytes, stabilizing the protein and promoting Rbm24 mediated splicing of structural proteins necessary for sarcomere assembly [34]. RBPs are thus central to defining RNA regulatory dynamics in muscle, both through their regulation of target mRNAs as well as through regulation of their own expression and activity.

1.3. Drosophila as a model to identify and study muscle-specific RBP function

Despite the hundreds of proteins identified to bind RNA or contain canonical RNA-binding domains [20,35,36], our understanding of RNA regulatory dynamics is still vastly incomplete as relatively few RBPs have been studied extensively in muscle. Most well-studied RBPs in muscle are associated with or causal for muscle diseases, notably CELF1 and MBNL1 (myotonic dystrophy) [37], SMN (spinal muscular atrophy) [38], TDP-43 [28], and Rbm20 and Rbm24 (dilated cardiomyopathy) [39]. We recently found that only 82 RBPs have been studied in muscle in any model organism, which is far less than the hundreds of annotated RBPs expressed in transcriptomics data from muscle [11]. Even well-characterized RBPs are often incompletely understood because of tissue-specific regulation, a lack of attention to RBP isoform-specific function and pleotropic function at multiple-levels in the RNA regulatory hierarchy [40].

To address this need to both identify new RBPs and characterize mechanisms of RBP function and regulation in muscle, we present the Drosophila adult musculature as an attractive genetic model system to investigate RBP function. The adult musculature in Drosophila displays functional diversity and developmental plasticity similar to vertebrate muscles [15]. Structural components of the sarcomere and many aspects of muscle physiology are highly conserved [2,41]. Disease models in fly have proven highly informative for many genes implicated in human muscle disease, for example CELF1, MBNL (Muscleblind), SMN (Survival motor neuron) and DMD (Dystrophin) [42–44]. Flies have a particularly diverse and powerful genetic toolbox that can be employed in vivo, including well-established tools to regulate spatial and temporal expression and to fine-tune levels of both RNAi-mediated knockdown and transgene expression [45–47]. These characteristics make flies an attractive model to both screen for RBPs with muscle phenotypes as well as investigate conserved molecular mechanisms of RBP function.

Here we demonstrate the utility of the Drosophila model through a candidate RNAi screen for muscle-specific RBP function. Our bioinformatic characterization of RBPs expressed in adult muscles identifies hundreds of diverse RBPs with temporal and spatially regulated expression dynamics, many of which additionally have muscle-specific RNAi phenotypes. We verify phenotypes for 35 of these RBPs in a candidate RNAi screen, uncovering a wide diversity in associated myofiber and myofibril defects. We find that phenotypes in many cases depend on RNAi knockdown efficiency, suggesting that muscle is sensitive to RBP expression levels. We also see phenotypic differences with temporally-restricted RNAi knockdown. Interestingly, we uncover phenotypes for spliceosome components, indicating they may have additional regulatory functions beyond basal splicing activity during muscle development. We further characterize the knockdown phenotypes for SF1 and Hrb87F in greater detail and show that both genes, which are known components of the spliceosomal A complex, regulate muscle-specific alternative splicing necessary for proper myofibril assembly, z-disc integrity and Kettin localization in flight muscle. Taken together, our data identify new RBPs with muscle-specific functions, provide insight into the regulatory function of individual spliceosome components and demonstrate the efficacy of Drosophila as a genetic model to study muscle-specific RNA regulatory dynamics.

2. Materials and Methods

A key resources table is provided in Table S1, listing the sources of all genetic and chemical reagents and bioinformatic tools used in this manuscript.
Fly stocks and crosses. Fly stocks were grown on normal cornmeal medium and maintained under standard conditions. w^{1118} was used as the wildtype control. RNAi lines used in the screen were obtained from the Vienna Drosophila Resource Center (VDRC), Bloomington Drosophila Stock Center (BDBSC) or the National Institute of Genetics (NIG-FLY) and are listed in Table S1. Gal4 drivers used include Mef2-Gal4 [48] which drives in all muscle; UAS-Dcr2, Mef2-Gal4 which enhances RNAi efficiency; and Act88F-Gal4 [49] which drives largely in IFM from 24h APF. Mef2-Gal4 crosses were grown at 27°C, while Act88F-Gal4 crosses were grown at 25°C.

Lethality and behavior assays. Flight behavior was tested as described previously [52,53]. Briefly, flies are introduced into the top of a 1-meter-long cylinder divided into 5 zones. Flies landing in the top two zones 4 and 5 are ‘normal fliers’, in the middle two zones 2 and 3 are ‘weak fliers’ and at the bottom (zone 1) are ‘flightless’. The flight index used in Figure 2 and Figure 4 is defined as \( F_i = \frac{[n_{zone1} + n_{zone2} + n_{zone3}]}{n_{total}} - \frac{[n_{zone1} + n_{zone2} + n_{zone3}]}{n_{total}} \), where \( n = \) number of flies. We considered \( F_i < 0.75 \) as flightless, \( -0.75 < F_i < 0.75 \) as weak fliers and \( F_i > 0.75 \) as normal flight ability. Lethality stage was monitored and recorded as embryonic (no hatching larvae), early (no pupa formed), early pupal (lethal before P10), late pupal (pupae reach P13-14, but do not eclose) or pharate/adult lethal. Pupal eclosion (Figure 6) was determined by counting the number of flies that eclose from 80-100 pupae of the appropriate genotype.

Immunofluorescence and microscopy. IFMs were dissected and stained as previously described [47]. After removing head and abdomen, thoraxes were fixed for 15-30 minutes in 4% PFA in 0.5% PBS-T (1x PBS + Triton-X100). Thoraces were then cut longitudinally with a microtome blade, blocked in 5% normal goat serum in PBS-T for 2 hours and stained overnight at 4 °C. Primary antibodies used include rabbit anti-GFP (1:500) and rat anti-Kettin 1:100 (Babraham). Samples were washed three times in 0.5% PBS-T for 10 min and incubated overnight at 4 °C with secondary conjugated antibodies (1:500) from Invitrogen (Molecular Probes), including Alexa488 goat anti-rabbit IgG, Alexa647 goat anti-mouse IgG, and rhodamine-phalloidin. Samples were washed three times in 0.5% PBS-T and mounted in Vectashield containing DAPI. Confocal images were acquired on a Leica SP8X WLL and analyzed with Image J/Fiji [50]. Sarcomere length and width were measured in Fiji based on rhodamine-phalloidin staining using the measure function because of the variability in myofibril structure after SF1 and Hrb87F knockdown. Plotting and statistical analysis were performed in GraphPad Prism 9.

mRNA-Seq. IFMs were dissected from 72 h APF pupal thoraxes as described previously [51], producing samples highly enriched for fibrillar muscle with low levels of contamination from tendon, motor neurons and trachea. Tissue was dissected for a maximum of 20-30 minutes and frozen in Trizol at -80 °C to minimize RNA degradation and changes to the transcriptome. IFMs from ~100 flies were combined into a single sample prior to total RNA isolation. Library preparation and sequencing was performed by LC Sciences (Houston, Texas, USA). Poly-A selected and stranded libraries were sequenced as Illumina paired-end 150 base pair reads with a depth > 60 million reads per library.

Bioinformatics. mRNA-Seq data from SF1-IR and Hrb87F-IR IFM was generated for this manuscript and is available from GEO under accession number GSEXXXXX (Note: data will be uploaded and made publicly available prior to acceptance). All other mRNA-Seq data used in this manuscript have been published previously [52,53] and are available from GEO under accession numbers GSE63707, GSE107247 and GSE143430. For all datasets, reads were mapped (or remapped) with STAR to ENSEMBL genome assembly BDGP6.22 (annotation dmel_r6.32 (FB2020_01)), indexed with SAMtools and features counted with featureCounts. Analysis and visualization were performed in R using the packages listed in Table S1. Soft clustering to identify temporal expression patterns was performed with Mfuzz [54]. Differential expression was analyzed on the gene level with...
DESeq2 [55] and on the level of exon use with DESeq [56]. GO term enrichments were performed with GOrilla [57] and term complexity reduced in R with rvrgo (https://ssayols.github.io/rvrgo).

For most analyses, the significance threshold included all exons with a p-value < 0.01 or all genes with an adjusted p-value < 0.01, to capture trends in the data independent of the magnitude of fold change. Plots with an additional fold change threshold are noted in the text and figure legends. Fiber-type specific genes were defined as all genes with a DESeq2 adjusted p-value < 0.01 and abs(log2FC) > 1.5 in salm-IR IFM and leg or salm-IR IFM and jump muscle (TDT) as compared to wildtype (WT) IFM. Fiber-type specific exons were defined as all exons with a DESeq p-value < 0.01 in salm-IR IFM and leg or salm-IR IFM and jump muscle (TDT) as compared to wildtype (WT) IFM.

Gene sets of annotated RBPs were obtained from AmiGO [36,58], GLAD [35] or a paper that compiled RBPs identified by RNA-interactome capture (RIC) experiments [20]. A list of spliceosome components, including many RBPs, was obtained from Spliceosome database [59]. Genome-wide phenotypic data for RBPs was obtained from a published RNAi screen [60]. A list of Drosophila sarcomeric proteins (SPs) has been compiled previously [53]. Exon classification as CDS, 5′-UTR or 3′-UTR was downloaded from Flybase.

3. Results

Muscle-specific functions are known for only a small portion of the hundreds of annotated RBPs in the eukaryotic genome [11]. To identify additional RBPs and splicing associated proteins in Drosophila that are expressed and thus potentially have a function in adult muscle, we looked for annotated RBPs in published mRNA-Seq datasets from fibrillar indirect flight muscle (IFM), tubular tergal depressor of the trochanter muscle (TDT, jump) and whole leg [52,53]. We considered four different lists of RBPs, including genes with the terms “RNA binding” or “mRNA binding” in AmiGO [36,58], genes identified as “RNA binding” in GLAD [35], a curated list of proteins identified by Drosophila RNA-interactome capture (RIC) experiments to bind RNA [20], and spliceosomal proteins listed in Spliceosome database [59] (Figure S1A, Table S2). We found that 64% or more of RBPs contained in any one list are expressed in adult muscle (Figure S1B). These curated lists vary in length and content and partially overlap (Figure S1A), reflecting the inherent challenge in identifying and classifying proteins as RBPs. We therefore conservatively define 568 “muscle RBPs” that are found in at least two annotations and are expressed in muscle (Figure 1A). Most muscle RBPs are expressed in all fiber types, as 95% (537 out of 568) can be found in all three (IFM, leg and TDT) mRNA-Seq datasets (Figure S1C). We conclude that there are indeed hundreds of RBPs expressed in muscle that may contribute to RNA regulation.

3.1. Muscle RBPs are expressed in distinct spatial and temporal patterns

IFMs are known to undergo dynamic shifts in gene expression that are correlated with the different phases in myofibril formation and sarcomere assembly, with a marked switch in expression as the myofibrils undergo maturation and establish stretch-activation properties [53]. To test if RBPs are regulated in a similar manner, we evaluated the temporal expression dynamics of muscle RBPs. RBPs display a dynamic range of expression level across adult muscle fiber types and throughout IFM development (Figure S1D, E). Using a published dataset collected from eight timepoints across IFM development [53], we used Mfuzz [54] to define 9 RBP gene clusters with distinct temporal expression profiles (Figure 1B, Figure S1E, Table S2). These clusters are described by three main patterns: 1) RBPs with preferential expression in myoblasts or during early pupal development (16-48 h after puparium formation (APF)); 2) RBPs with peak expression during pre-myofibril formation and initial stages of myofibril assembly (30-48 h APF); 3) RBPs upregulated during later stages of myofibril maturation and in adult flies (48 h APF – 1 d adult). This data illustrates that RBPs undergo temporal switches in expression, and suggests that different sets of RBPs may be necessary for RNA processing at different stages of muscle development.
Figure 1. RNA binding proteins (RBPs) have fiber-type and temporal dependent expression patterns in muscle. (A) Comparison of RBPs from AmiGO (purple), GLAD (yellow) or RNA-interactome capture (RIC, cyan) annotations that are expressed in mRNA-Seq data from 1 d adult Drosophila indirect flight muscle (IFM), tergal depressor of the trochanter (TDT) or leg. A core set of 568 RBPs (red) are expressed in muscle and present in at least 2 annotations. (B) Heatmap of RBP temporal expression profiles identified from Mfuzz clustering of IFM mRNA-Seq timecourse data. Core expression profiles, the average of standard normalized count values for each cluster, cluster into three distinct temporal expression patterns. (C) Heatmap and hierarchical clustering of log2 fold change (log2FC) values from DESeq2 differential expression analysis of RBP genes between wildtype IFM and leg, TDT, salm−/− IFM and bru1-IR IFM. (D) Heatmap and hierarchical clustering of log2FC values from DEXSeq differential exon usage analysis of RBP gene exons between wildtype IFM and leg, TDT, salm−/− IFM and bru1-IR IFM. (E) RBPs with exons that are differentially expressed between tubular and fibrillar fiber types. Exons with positive log2FC values are fibrillar preferential while negative log2FC values are tubular preferential, as detected in DEXSeq comparison of wildtype IFM and leg (purple), TDT (yellow) or salm−/− IFM (cyan). (F) Sunburst plot of RBP phenotypes in muscle from a published genome-wide RNAi screen [60]. The inner ring summarizes...
lethality and flight phenotypes. The second ring depicts the stage of lethality. The third and fourth rings summarize myofiber (magenta) and sarcomere (yellow) phenotypes, respectively.

*Drosophila* have two recognized muscle fiber-types with distinct morphologies and functional properties, namely the fibrillar IFMs and the tubular body muscles [6,61]. Spalt major (Salm), the master transcriptional regulator of the fibrillar fate [62], induces expression of the RBP Bru1 in IFM to promote many fibrillar-specific alternative splice events [52,63]. To determine if other RBPs have fiber-type specific expression profiles, we evaluated RBP expression at the gene level using DESeq2 [55] (Figure 1C) and at the exon-usagen level using DEXSeq [56] (Figure 1D, Table S2). We found that 56% of muscle RBP genes (318 of 568) are significantly differentially expressed (DE) between fibrillar IFM and tubular TDT or leg muscle, or have significantly altered expression levels in tubular-converted salm− IFM or bru1 knockdown IFM (Figure 1C). 36% (206 of 568) of muscle RBP genes have at least one exon that is significantly differentially expressed between muscle fiber types or in salm− or bru1-IR IFM (Figure 1D). Interestingly, we observe that RBP expression levels and exon-usagen patterns also differ between the TDT and leg samples, likely reflecting sub-specialization of tubular muscle types. 138 genes show both gene and exon-usagen level regulation (Figure S1F), so in total 68% (386 of 568) of muscle RBPs are differentially expressed or differentially spliced between muscle fiber-types or upon loss of Salm or Bru1. Notably, 45 genes had exons significantly differentially regulated between IFM and tubular muscle in at least two comparisons (Figure 1E), displaying a clear pattern of muscle-type specific exon use, and thus likely muscle-type specific splice isoform use. These RBPs are diverse and include splicing factors (i.e., WDR79, qkr54B, Pep, Pcf11, mbl), spliceosome components (i.e., U2af38, Prp8), nuclear export factors (sbr), ribosome components (i.e., Rps4, Rps9, Rpl30), and proteins associated with mRNA stability, localization or translation (i.e., qkr54B, how, gw, elf1A, egl), suggesting that many RNA regulatory processes might be impacted by muscle-type specific gene or isoform expression. Considering that we could detect mRNA from most RBPs in all muscle types (Figure S1C), these findings suggest that the regulation of RBP gene and splice isoform expression levels plays an important role in the regulation of RNA processing in muscle.

3.2. RNAi-mediated knockdown of hundreds of RBPs produces muscle phenotypes

Having shown that hundreds of RBPs are expressed in muscle in temporal and spatially regulated patterns, we next evaluated how many RBPs might have muscle-specific phenotypes. As a first look, we analyzed published data from a genome-wide RNAi screen using the muscle-specific Me2-Gal4 driver [60]. A large portion of muscle RBPs (246 of 568, 43%) as well as of all annotated RBPs (665 of 1924, 35%) displayed a lethal phenotype, and another 5% (31 of 568, 98 of 1924) had impaired flight ability, when knocked down specifically in muscle (Figure 1F, Figure S2A, Table S2). A smaller portion of RBPs with flight or lethality phenotypes had detectable myofiber defects (irregular or missing myofibers) or defects in myofibril or sarcomere structure (actin blobs, fuzzy Z-discs, degenerate or missing sarcomeres) (Figure 1F, Figure S2A). IFMs were examined for RBPs with impaired flight or late pupal lethal phenotypes, while larval muscle was examined for lines with larval lethality phenotypes [60]. This suggests that 40-50% of RBPs, or somewhere between 280 to 773 different RBP genes, are required for proper muscle development. A large portion of these genes are also required for viability, either to enable larval hatching, locomotion or adult eclosion.
In both vertebrates and invertebrates, alternative splicing (AS) serves as a mechanism to fine-tune contractile properties through the production of fiber-type distinct isoforms of many sarcomere proteins [4,6,7], and thus plays an important role in both muscle development and maintenance. We were particularly interested in identifying RBPs that might be involved in fiber-type specific alternative splicing, so we selected a set of 35 candidate genes (Figure 2A, Figure S2B, Table S3), of which 17 were spliceosome associated components and 18 were regulated by Salm or Brul or differentially regulated between fiber-types in the mRNA-Seq data. We then performed a muscle-specific candidate RNAi screen using Mef2-Gal4. In addition to the hairpins used in the Schnorrer et al. screen [60], we included the second-generation KK hairpins from VDRC as well as the TRiP hairpins available from Bloomington, and screened in total 105 RNAi lines (Figure S2C). We assayed both lethality and flight ability, characterizing lethality phenotypes for 46 hairpins and impaired flight ability for an additional 40 hairpins (Figure S2C). If we consider the phenotype of the strongest hairpin for each gene, 23 genes have lethal phenotypes and 9
genes have impaired flight ability (Figure 2B, C). A cross of the driver to ap1715 served as a wildtype control, and we used salm-IR [62] as a flightless control. We also observed a notable level of variation in phenotype between hairpins targeting the same gene (Figure 2C). Based on 76 hairpins where we tested more than one hairpin per gene, we estimate that around 50% of hairpins produce efficient knockdown leading to strong phenotypes, 30% of hairpins produce weak or moderate knockdown and weaker phenotypes and around 20% of hairpins are ineffective (Figure S2D). More generally, our candidate selection approach successfully enriched for RBPs with muscle function.

3.4. Muscle-specific RBP knockdown leads to diverse myofibril and sarcomere defects

To evaluate myofiber and myofibril defects caused by RBP knockdown, we performed confocal microscopy with 35 hairpins targeting 17 RBP genes. There are six IFM myofibers that span and fill the volume of the thorax in wild-type controls (Figure 3B, F), and we analyzed IFM morphology in 1 d old adults for crosses with impaired flight ability, and at 90 h APF for crosses leading to late pupal or pharate lethality. We observe a broad diversity in the severity and types of IFM phenotypes after RBP knockdown (Figure 3A). In severe cases, for example with knockdown of Sβa2 with line 55650 (Sβa2-55650) (Figure 3C), IFM myofibers were completely missing. A missing myofiber phenotype at 90h APF can result from defective myoblast migration or fusion or impaired tendon attachment [64,65], as well as defective regulation of cellular growth and tension [66–68], which suggests that RBP function is necessary for one or more of these cellular processes.

We frequently observed detached or torn myofibers, for example mbl-29585 (Figure 3D), CG9346-27013 (U2SURP) (Figure 3G), bru1-41586 (Figure 4E), and Rbfox1-27586 (Figure 3H). A less severe myofiber phenotype in several flightless lines was characterized by stretched and atrophic fibers, for example as observed with Doe-50902 (Figure 3E) and mub-34870 (Figure 3I). Both stretched and torn myofibers are commonly associated with hypercontraction, where misregulated myosin activity leads to aberrant contractility [69,70]. The Bru1 and Rbfox1 phenotypes served as positive controls, as we have previously characterized that loss of either protein leads to missplicing and hypercontraction phenotypes [52,71]. We conclude that on the fiber level, RBP knockdown can alter the number of myofibers and lead to myofiber hypercontraction, tearing or loss.

We next analyzed myofibril and sarcomere structure after RBP knockdown. Myofibrils in wild-type controls have a fibrillar structure (myofibrils are organized in parallel and not laterally aligned) with regular, crisply-defined rectangular sarcomeres (Figure 3, P). Myofibrils and sarcomeres after RBP knockdown display a variety of defects, from myofibril splitting and fraying to short sarcomeres and actin inclusions. Hairpins with the strongest myofiber defects also typically displayed severe myofibril and sarcomere defects. For example, IFM myofibers are frequently missing after knockdown with Sβa2-55650, and remaining myofibrils, when not severely degraded, are split and disorganized with short sarcomeres (Figure 3K). Knockdown of mbl, bru1 and Rbfox1 (Figure 3O2, T, U), as we have previously reported [11,52,71], generates short sarcomeres and thick myofibrils, characteristic of hypercontraction. We see similar defects for knockdown with Rm62-46908 (Figure 3N1), Rm62-110102 (Figure 3N2), mub-34870 (Figure 3Q), snf-330329 (Figure 3R), and sbr-103715 (Figure 3S). This may indicate that these RBPs also regulate splicing of cytoskeletal components necessary for the proper regulation of myosin activity.
Figure 3. Mef2-Gal4 knockdown of RBPs produces diverse myofibril and sarcomere phenotypes. (A) Graphical summary of confocal phenotypes from 35 hairpins targeting 17 RBP genes. Dots represent hairpins and are colored by flight (flightless, red; weak flier, orange) and lethality (blue shades) phenotype. Hairpins may appear in more than one morphological phenotype category. (B-I) Confocal z-projections of IFM hemi-thoraxes showing myofiber phenotypes at 90 h APF (B-E) or in 1 d adult (F-I). Torn, detached or missing fibers are observed in severe examples (yellow arrow heads in D, E, G, H) and classified as “detached myofiber” in (A). Other myofibers are stretched (yellow arrows in E, I), likely reflecting hypercontraction. White dotted lines mark boundary of the thorax in C, D, G and H. Scale bars = 100 μm. (J-AA) Confocal single-plane images showing myofibril and sarcomere structure at 90 h APF (J-O2) or in 1 d adult (P-AA). As compared to w1118 control images (J, P), knockdown of RBPs can result in split or frayed myofibrils (K, L, N2, Q, R, V), actin blobs which are commonly known as zebra bodies (M, N, O, R, S, U-AA), thick or thin myofibrils (K, L, O2, Q, R, S, U, V, X), as well as sarcomeres that are disorganized, short (N1, N2, O2, Q, S, T, U) or degenerate (K, N Q, V). Scale bars = 5 μm; phalloidin stained F-actin, white.

The most common defect we observe with Mef2-Gal4 mediated RBP knockdown, even in some lines with largely normal flight ability, are abnormal actin inclusions that
form at the z-disc, or so-called “zebra bodies.” The frequency and size of zebra bodies increases as flight ability decreases. We observe these actin inclusions in lines where sarcomere structure appears otherwise normal or merely thin, for example Dros-46449 (Figure 3M), clu-42318 (Figure 3X), Prp6 (CG6841)-34254 (Figure 3Y), Rbn17 (Spf5)-32949 (Figure 3Z) and Hrb87F-52937 (Figure 3AA), as well as in combination with other myofibril and sarcomere defects as with sbr-103715 (Figure 3S) or Rfox1-27586 (Figure 3U). In some cases, it appears the formation of actin inclusions may potentially drive or at least precede myofibril splitting and fraying, for example snf-330329 (Figure 3R) or CG9346 (U2SURP)-27013 (Figure 3V). Taken together, we interpret this morphological data to reflect the diversity in RBP targets and RBP functions. Different RBPs regulate disparate aspects of muscle development, and therefore knockdown would be predicted to result in the observed phenotypic variability between RBP genes.

We also noted that phenotypic severity often differs between hairpins targeting the same gene. For example, although both Rm62-46908 and Rm62-110102 are pupal lethal, sarcomeres from Rm62-46908 knockdown are short with sporadic irregular actin densities (Figure 3N1), while sarcomeres from Rm62-110102 knockdown are short and thick and myofibrils are frayed with actin blobs (Figure 3N2). A similar phenotypic variation is observed with mbl knockdown: mbl-29585 is pupal lethal with zebra bodies (Figure 3O1); mbl-28752 is pupal lethal with short, thick sarcomeres and partially degraded myofibrils (Figure 3O2); and mbl-105486 is flightless with short, thick sarcomeres (Figure 3T). We propose this variability arises from target sensitivity to knockdown efficiency. Many RBPs are able to bind multiple or degenerate binding motifs with different levels of affinity [20], thus RBP function would be more or less impaired depending on knockdown efficiency.

3.5. Temporally restricted knockdown of RBPs in IFM using Act88F-Gal4

Mef2-Gal4 is expressed throughout muscle development, including during embryonic and larval stages, which can lead to early lethality if RBPs are required in both embryonic and adult muscle. As nearly 50% of the hairpins in our screen were lethal when crossed to Mef2-Gal4, we performed a follow-up screen with Act88F-Gal4, which is reported to drive in IFM starting around 24h APF [49]. In theory, this should allow us to bypass lethality and assay muscle-specific phenotypes in IFMs which are not required for viability. We screened 65 hairpins targeting 27 RBP genes with Act88F-Gal4 and assayed both lethality and flight ability (Figure 4A, Table S3). We found that knockdown with 25 hairpins (38%) resulted in impaired flight ability, 13 (20%) produced no phenotype and 27 (42%) still resulted in lethality, although mostly late pupal or pharate lethality (Figure 3A). Taking the strongest hairpin phenotype for each gene, knockdown of 8 RBPs (30%) impaired flight ability and knockdown of another 16 RBPs (59%) resulted in pupal lethality (Figure 4B), largely due to failed eclosion. This lethality phenotype occurs later than lethality observed with Mef2-Gal4, suggesting that knockdown is indeed restricted to the pupal stage. However, as IFMs are not required for eclosion, we checked if we could detect Act88F-Gal4 driver expression in tubular muscle. When crossed to a nuclear GFP, we could detect strong Gal4 expression in IFM as well as weak signal in a subset of fibers in adult jump, leg and gut muscle, but not in abdominal muscle (Figure S3B-G). This implies that lethality phenotypes observed with Act88F-Gal4 likely reflect RBP requirement in these tubular muscles.

We next analyzed the morphology of IFM myofibers and myofibrils by confocal microscopy. We classified the phenotypes into different morphological categories, as defined above with Mef2-Gal4 knockdown (Figure 4C). Myofibers, as compared to the wild-type control (Figure 4D), were detached or torn, for example after knockdown with bru1-41568 or sbr-46117 (Figure 4E, G). Knockdown of Atx2-108843 led to stretched myofibers (Figure 4F). We observed myofibril and sarcomere phenotypes including frayed or thin myofibrils with short, disorganized or at times degenerate sarcomeres. For example, knockdown of bru1-41568, sbr-46117 and hov-37342 all resulted in thick myofibrils with short sarcomeres (Figure 4I-K), likely indicating hypercontraction. In addition, sbr-46117 myofibrils were frayed (Figure 4J) and bru1-41568 myofibrils were degenerate (Figure 4I). As compared to
Mef2-Gal4, Act88F-Gal4 knockdown infrequently resulted in actin inclusions at the z-disc, but more frequently resulted in thin myofibrils with otherwise normal sarcomeres, for example Prp19-32865, snf-51459, fant-104186 and Hrb87F-100732 (Figure 4L-O). We emphasize that Act88F-Gal4 expression, at least in IFM, can first be detected around 24h APF, so knockdown effects will only be evident later in pupal development after the target protein has turned-over. Thus, these phenotypes indicate that RBPs are required for developmental processes related to mid- or late-steps in myofibril and sarcomere growth, organization and maturation.

Figure 4. Temporally restricted knockdown of RBPs in IFM with Act88F-Gal4 generates assorted myofibril phenotypes. (A) Plot summarizing Act88F-Gal4 phenotype for 65 hairpins targeting 27 RBP genes. Dots represent individual hairpins organized into columns by gene, as labeled. Squares indicate spliceosome membership and are colored as in Figure 2 A. Plot background color marks normal flight (yellow), weak/impaired flight (orange), flightless (red) and lethal (blue) phenotypes. (B) Pie chart summary of RBP gene-level (strongest hairpin) phenotype when crossed to Act88F-Gal4, which drives in IFM from 24 h APF. (C) Graphical summary of confocal phenotypes from 27 hairpins in (B). Dots represent hairpins and are colored by flight (flightless, red; weak flier, orange) and lethality (blue shades) phenotype. Hairpins may appear in more than one morphological phenotype category. (D-G) Confocal z-projection images showing myofiber phenotypes in 1 d adult flies, including detached or torn myofibers (yellow arrow heads in E, G) and stretched (yellow
arrows in F) myofibers. Scale bars = 100 μm. (H-O) Single-plane confocal images of 1 d adult IFM myofibril and sarcomere structure. RBP knockdown results in diverse phenotypes, such as frayed (j) or thin (L-O) myofibrils as well as degenerate (l) or short (I-K) sarcomeres. Scale bars = 5 μm; phalloidin stained F-actin, white.

3.6. Diverse phenotypes associated with altered levels of spliceosome components

Although the loss of core spliceosome components is lethal, the efficiency of the splicing reaction as well as splice site recognition are sensitive to spliceosome component levels [72]. To ascertain how spliceosome component expression levels might affect myogenesis, we compiled the data from 44 hairpins in our RNAi screen targeting 17 spliceosome components (Figure 2A, C, Figure S4A, Table S3), including confocal data from 27 hairpins. We tested components from different spliceosome complexes, including: Hrb87F and SF1 (A complex); snf (U1); Sf3a2, Sf3b1, Sf3b2, Spx and noi (U2); Sfp45 (Rbm17) and (U2SURP) (U2 associated); Cdc5, famd, Prp19 (Prp19 complex); CG6686 (Sart1/SNu66), CG6841 (PRP6), Prp8 (tri-snRNP) (Figure S4A, B). Components of the E, U1 and U2 complexes (A complex components) play important roles in splice site recognition and spliceosome assembly and positioning, while the U2, Prp19 and tri-snRNP complexes form the catalytic core after B-complex assembly and perform the splicing reaction (Figure S4B).

As expected based on the essential requirement for splicing in all cells, we observed that muscle-specific knockdown of many spliceosome components is lethal. 50% (22 of 44) of hairpins crossed to Mef2-Gal4 were lethal in early pupal or pre-pupal stages and 62% (18 of 29) of hairpins crossed to Act88F-Gal4 were lethal at late pupal stages (Figure 5A, Figure S4C). We observed that only 15% of hairpins (2 of 13) targeting A-complex components as compared to 65% of hairpins (20 of 31) targeting B-complex components had lethal phenotypes when crossed to Mef2-Gal4 (Figure S4C). We observed the opposite with impaired flight ability, where 38% of hairpins (5 of 13) targeting A-complex components as compared to 6% of hairpins (2 of 31) targeting B-complex components were flight impaired (Figure S4C). One possible explanation for this result that remains to be tested in future experiments is that loss of specific E or A-complex components might lead to alternative splicing defects that impair muscle development and function, while loss of B-complex components results in a general block of splicing.

We also observed diverse myofibril and sarcomere defects after knockdown of spliceosome components. Phenotypes ranged from split and frayed myofibrils with degenerate sarcomeres (Figure 5D, K), to actin blobs at the Z-disc (Figure 5 E, F, H, I), to thin myofibrils (Figure 5G, J, L, M, N-P). Notably, all myofibrils from Act88F-Gal4 knockdown of spliceosome components were thin (Figure 5A). By contrast, Act88F-Gal4 mediated knockdown of non-spliceosome associated RBPs such as bru1-41568, sbr-46117 or how-37342 (Figure 4I-K) resulted in thick myofibrils with short sarcomeres, in addition to myofibril tearing and abnormal actin structures. This suggests that decreased levels of core spliceosome components during IFM maturation may generally reduce mRNA levels and thus translation of most sarcomere components, while disrupting splicing factors alters isoform levels of specific sarcomere proteins and biomechanical properties of sarcomeres.

For four spliceosome components including SF1, Hrb87F, snf and Prp19, we were able to evaluate myofibril phenotypes after knockdown with three or more independent hairpins (Figure 5E-P). We observed that in several cases, different hairpins targeting the same gene give different apparent phenotypes. For example, Mef2-Gal4 mediated knockdown of SF1-13425 and SF1-13426 results in actin accumulations at the z-disc, torn myofibrils and aberrant actin distribution along the myofibril, while Act88F-Gal4 mediated knockdown of SF1-52938 results in thick myofibrils (Figure 5E-G). Similar defects are observed with both Hrb87F (Figure 5H-J) and snf (Figure 5K-M). We propose two possible explanations. First, the level of knockdown efficiency might vary between the hairpins, resulting in phenotypes of different severity. Second, knockdown with Mef2-Gal4 at early stages of muscle development may impair pre-myofibril formation leading to myofibril splitting, fraying, tearing and abnormal actin accumulations, while knockdown with Act88F-Gal4 rather impairs myofibril growth. Future studies will be required to clarify the detailed developmental mechanism.
3.7. SF1 and Hrb87F are necessary for myofibril development and z-disc integrity

Based on the results of our screens above, we selected SF1 and Hrb87F for further characterization. Neither gene has been previously characterized in Drosophila muscle, and both genes are reported to be A-complex components that contribute to splice site...
recognition (Figure S4B). More specifically, SF1 is responsible for branch-point sequence recognition [73,74], while Hrb87F as a homolog of the hnRNPA1 family is involved in 3'-SS recognition [75,76]. For each gene, we selected the hairpin with the strongest flight phenotype that was still viable when crossed with Mef2-Gal4, SF1-13426 and Hrb87F-31244 (Figure 2C, Figure S4C, Table S3), and examined the IFM phenotype in 1 d old adults. All 6 IFM myofibers were present and attached after knockdown of SF1-13426 and Hrb87F-31244, although Hrb87F-31244 myofibers had an uneven actin density and appeared stretched (Figure 6A, C, E). At a higher magnification, myofibers from both SF1-13426 and Hrb87F-31244 were thick with aberrant sarcomere structure. In SF1-13426 flies, the myofibrils were torn and frayed, and z-discs as marked with Kettin staining were wavy or bulged (Figure 6D1-D4). In Hrb87F-31244 flies, the myofibrils were thick and frequently split with prominent zebra-bodies often spanning 2-3 myofibrils (Figure 6F1-F2). Kettin staining revealed that z-discs were spotty and in the vicinity of zebra-bodies discontinuous (Figure 6F3-F4). IFM sarcomeres in Hrb87F-31244 flies are significantly shorter than the wildtype control (Figure 6G), while myofibrils from both SF1-13426 and Hrb87F-31244 are significantly thicker than the wildtype control (Figure 6H). These phenotypes are similar but not identical, suggesting that both SF1 and Hrb87F mediate splicing necessary for proper myofibril development.

One or more of the hairpins we tested against SF1 and Hrb87F resulted in lethality (Figure 2C), indicating that SF1-13426 and Hrb87F-31244 mediate only a partial knockdown of the target gene. UAS-Dcr2, Mef2-Gal4 was previously reported to enhance RNAi knockdown efficiency [77], so we next tested how a stronger knockdown with the same hairpins influenced the muscle phenotypes. While >90% of SF1-13426 and Hrb87F-31244 Mef2-Gal4 knockdown pupae eclosed as adults, both hairpins are pupal lethal when crossed to UAS-Dcr2, Mef2-Gal4 (Figure 6I). This confirms that Dcr2 indeed enhances knockdown efficiency and phenotypic severity.

We then examined IFM phenotypes at 72 h APF, a timepoint where sarcomere structure is well defined but before knockdown induced pupal death. We found that IFM myofiber structure is severely disrupted, with the majority of fibers torn or missing with both SF1-13426 and Hrb87F-31244 (Figure 6J, M, P). Myofibril and sarcomere structure in remaining fibers are severely affected. Myofibrils in IFM from SF1-13426 flies are thin and frequently split or frayed (Figure 6N1-N3). Sarcomeres have notable z-disc defects, including misorientation, abnormal actin accumulation and poorly refined structure (Figure 6N2). Strikingly, Kettin localization was severely impaired with the majority of Kettin staining localizing to cytoplasmic puncta (Figure 6N3). Myofibrils in Hrb87F-31244 IFM were split, frayed and torn with poorly defined sarcomeric structures (Figure 6Q1-Q3). Kettin was mostly localized to cytoplasmic puncta (Figure 6Q3). Myofibril width as well as sarcomere length was significantly shorter with SF1-13426 and Hrb87F-31244 knockdowns at both 72 h APF and 88 h APF (Figure 6G, H). As these flies are unable to eclose, we also examined tubular leg muscle structure. Myofibrils in tubular leg muscle are disorganized with poorly-defined sarcomeres in SF1-13426 and Hrb87F-31244 flies (Figure 6L, O, R). We conclude that SF1 and Hrb87F are required for myofibril development and likely for z-disc integrity, and a stronger knockdown enhances the severity of myofibril and z-disc phenotypes. We also infer that fibrillar IFM is more sensitive to alterations in expression levels of both proteins, as functional and morphological defects in tubular leg muscle are first observed with Dcr2-enhanced knockdown.
Figure 6. SF1 and Hrb87F are required for proper myofibril development and z-disc stability. (A-F) Confocal z-projection (A, C, E) and single-plane images (B, D, F) showing IFM phenotypes in 1 d adults after Mef2-Gal4 driven knockdown of SF1 (C-D) and Hrb87F (E-F). Although gross myofiber structure is intact (A, C, E), compared to w1118 controls (A-B), knockdown myofibrils are thick and split with short sarcomeres and actin-inclusions at the z-discs (B2 compared to D2 and F2). Kettin staining at the z-disc is discontinuous and less refined, and is aberrant in the vicinity of zebra-bodies (compare B3-4 to D3-4 and F3-4). Scale bars = 100 µm (A, C, E), 5 µm (B1, D1, F1) and 2 µm (B4, D4, F4). (G-H) Quantification of sarcomere length (G) and myofibril width (H) after SF1 (orange) and Hrb87F (purple) knockdown at 72 h APF, 88 h APF or in 1 d adults. Significance is from post-hoc Tukey test after ANOVA; *, p < 0.05; **, p < 0.01; ***, p < .001. (I) Plot of the percent of pupae with SF1 or Hrb87F knockdown that are able to eclose (white) or remain trapped in the pupal case (red). Genotypes as labeled. Error bars show standard deviation. (J-R) Confocal z-projection (J, L, M, O, P, R) and single-plane images (K, N, Q) showing muscle phenotypes at 72 h APF with UAS-Dcr2, Mef2-Gal4 driven knockdown of SF1 (M-O) and Hrb87F (P-R). Enhanced knockdown efficiency results in degraded (Q2) and split (N2) myofibrils with short sarcomeres. Kettin localization at the z-disc is impaired and z-discs are poorly defined (N1-3, Q1-3). Sarcomere structure and organization
in tubular leg muscles is also disrupted (L compared to O and R). Scale bars = 100 µm (J, M, P), 5 µm (K1, L, N1, O, Q1, R).

3.8. Transcriptomic analysis reveals molecular defects in SF1 and Hrb87F knockdown IFM

To characterize the transcriptomic phenotype and identify possible developmental mechanisms that disrupt myofibril structure after knockdown with SF1-13426 and Hrb87F-31244, we performed mRNA-Seq. We dissected IFM at 72h APF from Me2-Gal4 driven knockdown flies, as well as from Me2-Gal4 crossed to w1118 as a wildtype (WT) control. We analyzed differential gene expression between knockdown and WT IFM using DESeq2 [55] and differential exon use with DEXSeq [56]. Differential exon use reflects differential mRNA isoform use due to either alternative splicing or alternative promoter use. We also specifically examined a set of previously annotated genes encoding Drosophila sarcomere proteins [53] in this data. We were able to detect changes in both gene expression (Figure 7A, B) and exon use (Figure 7D, E) in IFM after knockdown with SF1-13426 and Hrb87F-31244.

We started by characterizing transcriptomic defects in SF1-13426 (SF1-IR) and Hrb87F-31244 (Hrb87F-IR) IFM at the gene level. Although knockdown of SF1 leads to both up- and downregulation of differentially expressed (DE) genes, sarcomere genes tend towards weak downregulation (Figure 7A, Table S4). Sarcomere genes as well as DE genes in general tend to be weakly downregulated in Hrb87F knockdown IFM (Figure 7B). The small fold changes we observe may be the result of a hypermorphic phenotype due to partial RNAi knockdown, or may alternatively reflect an overall decrease in all mature mRNA. We then performed a Gene Ontology (GO) analysis of enriched terms in up- or downregulated genes. Upregulated genes in SF1-IR are enriched in Biological Process terms such as “response to stress,” “mature ribosome assembly” and “telomere organization,” while downregulated genes are enriched for terms including “adult somatic muscle development,” “myofibril assembly,” “translation,” “endoplasmic reticulum unfolded protein response,” and “regulation of mitochondrial organization” (Figure 7C). Upregulated genes in Hrb87F-IR are enriched for “developmental process” and “cytoskeleton organization” terms, while downregulated genes are enriched for “myofibril assembly,” “cytoplasmic translation,” “ATP metabolic process” and “mitochondrion organization,” among others (Figure 7C). This suggests that both SF1 and Hrb87F regulate muscle structural genes among many other targets, and a decrease in SF1 and Hrb87F levels more generally leads to a decrease in translation and mitochondrial function and activation of the stress response, which remains to be verified in future experiments.

We then characterized changes in exon use in SF1-13426 (SF1-IR) and Hrb87F-31244 (Hrb87F-IR) IFM. Knockdown of both SF1 and Hrb87F results in expression changes in hundreds of exons (Figure 7D, E, Table S4). Notably, we observe up- and downregulation of different sets of sarcomere gene exons (Figure 7D, E, Figure 8 A) that are comparatively larger than overall changes in sarcomere gene expression (Figure S5 A). This is further reflected in the enrichment for GO Biological Process terms such as “muscle system process,” “myofibril assembly” and “flight” in upregulated SF1-IR exons and “cytoskeleton organization” and “muscle structure development” in Hrb87F-IR downregulated exons (Figure 7C). Interestingly, we see exons from the same sarcomere gene regulated in opposite directions. For example, one set of exons from br, sls, Mhc, Ryr, in, Mj and Zasp66 are upregulated while a different set of exons is concurrently downregulated (Figure 8 A), indicating that ratios of splice isoform expression in IFM for multiple sarcomere components are altered when levels of SF1 and Hrb87F are reduced. As fiber-types have distinct isoform expression profiles [6], splicing alterations likely contribute to the observed defects in IFM structure and function.
Figure 7. Knockdown of SF1 and Hrb87F alters exon usage of sarcomeric genes. (A) Volcano plot of all DESeq2 gene expression data for SF1-IR versus WT IFM. Genes with an adjusted p-value < 0.01 are marked in blue, and genes encoding sarcomere components are labeled red. The gray shaded region marks all genes with a fold change (FC) > 1 or < -1. (B) Volcano plot colored as in (A) of all DESeq2 gene expression data for Hrb87F-IR versus WT IFM. (C) Heatmap of select GO Biological Process term enrichments for up and downregulated genes and exons for SF1-IR and Hrb87F-IR versus WT IFM. A full list of terms can be found in Table S4. Terms were manually grouped into categories related to a developmental process (red), muscle and cytoskeletal organization (orange), translation (yellow), stress and protein folding (green), a metabolic process (magenta) or mitochondria (brown). (D) Volcano plot of DEXSeq changes in exon use for SF1-IR versus WT IFM. Exons with a p-value < 0.01 are marked in blue, and exons from sarcomere genes are labeled red. The gray shaded region marks all exons with a FC > 1 or < -1. (E) Volcano plot colored as in (D) of DEXSeq exon use data for Hrb87F-IR versus WT IFM.
As our immunostaining of both SF1-IR and Hrb87F-IR IFM revealed mislocalization of Kettin (Figure 6D3, F3, N3, Q3), we evaluated expression changes in z-disc components that directly bind to or influence Kettin localization including bt (Projectin), sls, Zasp52, Zasp66, Zasp67, cher (filament) and zormin [78–80] (Figure S5 B). The sls gene, which encodes the Kettin protein, is expressed at similar levels in SF1-IR and WT IFM (FC=0.17, adjusted p-value .6627) and at a slightly higher level in Hrb87F-IR IFM (FC=0.63, adjusted p-value .0017). Although Zasp66 expression is weakly but significantly decreased in both SF1-IR and Hrb87F-IR and Zasp67 is slightly decreased in SF1-IR, changes in gene-level expression of bt, Zasp52, zormin, and cher are non-significant (Figure S5 B, Table S4). While we do not detect strong changes in gene expression, we do see changes in exon use within Kettin itself as well as in bt, Zasp52 and Zasp66 in both SF1-IR and Hrb87F-IR (Figure S5 B). These differences would likely alter the ratio of z-disc protein isoforms in the IFM sarcomere, but it remains to be tested if any exon-use changes detected here are sufficient for Kettin mislocalization, or if a different mechanism is responsible. Taken together, our mRNA-Seq analysis identified altered gene expression and exon use in SF1-IR and Hrb87F-IR IFM that is predicted to affect muscle structure, sarcomere contractility, mitochondrial function and metabolism.

3.9. SF1, Hrb87F and Bru1 regulate distinct but overlapping splice events

SF1 and Hrb87F are both A-complex components required for splice site recognition, but one recognizes the branch point and the other the 3’-SS. To test if their splicing signatures are distinct as well as how they differ from another muscle-specific splicing factor, we compared gene expression and exon use patterns in SF1-IR, Hrb87F-IR and bru1-IR. Bruno1 (Bru1) is a CELF-family homologue known to regulate alternative splicing important for sarcomere assembly and contractility in IFM [52,63]. We started by calculating the Pearson coefficient for pairwise comparison of the significant fold change values and found a weak positive correlation of 0.3666 between SF1-IR and Hrb87F-IR at the gene level which increased to 0.6956 at the exon level (Figure 8 B). By comparison, bru1-IR has little or no correlation to SF1-IR or Hrb87F-IR. Interestingly, genes that are downregulated in SF1-IR are also downregulated in Hrb87F-IR, but genes that are upregulated are largely unchanged in Hrb87F-IR (Figure S6 A). As Bru1 contributes to the regulation of fiber-type identity, we also looked at genes that are significantly differentially expressed between tubular and fibrillar muscle. Although bru1-IR leads to upregulation of tubular genes and downregulation of fibrillar genes, SF1-IR and Hrb87F-IR do not (Figure S6 B), indicating they are not involved in regulation of fiber-type identity.

We next evaluated exon usage patterns in the DEXSeq data. The majority of exons that are significantly changed in SF1-IR and Hrb87F-IR are regulated in the same direction (Figure 8 C, Figure S6 B). This can also be clearly observed for sarcomeric gene exons, where SF1-IR and Hrb87F-IR have a similar regulatory pattern which is distinct from that of bru1-IR (Figure 8 A). We see regulation of both coding exons as well as 5’- and 3’-UTR exons, indicating these changes reflect bona-fide splicing changes as well as use of alternative promoters (Figure 8 A). If we look at all significantly DE exons, we find only 5 events regulated by all 3 proteins, and a majority of events selectively regulated by one of the proteins (Figure 8 C). Events regulated by Bru1 are largely distinct from those sensitive to decreased expression of SF1 and Hrb87F. For the small number of events significantly regulated in both bru1-IR and either SF1-IR or Hrb87F-IR, we see much larger fold-changes in the bru1-IR IFM (Figure 8 C). This is also observed when we look selectively at exons differentially regulated between tubular and fibrillar muscle (Figure S6 D), indicating that SF1 and Hrb87F do not play a large role in fiber-type specific splicing. Exons regulated in both SF1-IR and bru1-IR change in the same direction, while exons regulated in both Hrb87F-IR and bru1-IR tend to change in opposite directions (Figure 8 C, Figure S6 D). These patterns indicate that RBPs that regulate splicing at the same step will have more similar splicing profiles than those that regulate splicing at a different step in the splicing process (Figure S4 B). However, individual exons and genes can be more or less sensitive to decreased expression of a specific RBP. Strikingly, a large number of fiber-type specific
exons (148 of 216 in this dataset) are not significantly altered in *bru1-IR, SF1-IR*, or *Hrb87F-IR* (Figure S6 D). This may reflect use of a stringent p-value threshold or development differences in fiber-type specific splicing between 72h APF and 1 d adult, which we used to identify fiber-type specific exons. More interestingly, this likely indicates that additional, uncharacterized RBPs contribute to fiber-type specific splicing and motivates future research into RBP function in muscle.

**A. log_{2}FC of DE sarcomeric gene exons**

\[ p < 0.01, n = 144 \text{ exons} / 28 \text{ genes} \]

**B. correlation of log_{2}FC values**

- **bru1-IR vs WT**
- **Hrb87F-IR vs WT**
- **SF1-IR vs WT**
- **bru1-IR vs SF1-IR**
- **bru1-IR vs Hrb87F-IR**
- **SF1-IR vs Hrb87F-IR**

**C. comparison of all significant DE exons**

\[ n = 1112 \text{ exons} / 743 \text{ genes} \]

**Figure 8.** Changes in exon use profiles in *SF1-IR* and *Hrb87F-IR* are distinct from those in *bru1-IR*. (A) Heatmap and hierarchical clustering of log_{2}FC expression values for all 144 sarcomeric gene exons significantly differentially expressed (DEXSeq p-value < 0.01) in either *bru1-IR, SF1-IR* or *Hrb87F-IR*. Note the similarity in expression between *SF1-IR* and *Hrb87F-IR* and difference to *bru1-IR*, even though all three knockdowns affect myofibril and sarcomere development. Exons are classified as coding (CDS, green), 5’-UTR (pink) or 3’-UTR (magenta) based on Flybase annotation. Black squares denote individual exons from the same gene. For multiple sarcomere genes, different sets of exons are up- and downregulated indicating alternative isoform use. (B) Plot of Pearson’s correlation coefficients of log_{2}FC values from gene expression and exon use analyses between *bru1-IR,*
SF1-IR and Hrb87F-IR. Coefficient values for comparison of all significant genes (DESeq2 adjusted p-value < 0.01) and exons (DESeq p-value < 0.01) are shown in black, and for comparison of all genes or exons irrespective of significance is shown in blue. (C) Plot of log2FC values for all 1112 significantly DE exons (DESeq p-value < 0.01) genome wide (bru-IR exons, green; Hrb87F-IR exons, purple; SF1-IR exons, orange). The number of exons is noted in the top left of each plot region. Gray shaded regions contain exons that are downregulated in the knockdown condition. Note the different regulatory dynamics for individual exons between bru1-IR and SF1-IR or Hrb87F-IR, the similar dynamics in regulation between SF1-IR or Hrb87F-IR and the large number of exons regulated in only one genotype.

4. Discussion

RNA binding proteins (RBPs) play important roles in tissue development through their regulation of mRNA alternative splicing, trafficking, localization and stability. Their coordinated regulation fine-tunes ratios of splice-isoform expression and helps define muscle-type specific morphology and contractility [6,81]. Here we provide insight into the breadth of this RNA regulatory network by identifying 568 RBPs expressed in transcriptomics data from *Drosophila* adult muscle and revealing their patterns of temporal and fiber-type specific expression. We screened 105 RNAi hairpins targeting 35 RBPs for muscle-specific phenotypes, including their effects on viability, flight ability, and the morphology of IFM myofibers, myofibrils and sarcomeres. Our data illustrates the diversity of muscle phenotypes that can be attributed to loss or altered levels of RBP expression, and identifies previously uncharacterized flight muscle-related functions for Clueless (Clu), Darkener of apricot (Doa), Heterogenous nuclear ribonucleoprotein at 87F (Hrb87F), Matal expression at 31B (Me31B), Mushroom-body expressed (Mub), Rm62, Small bristles (Sbr), Splicing factor 1 (SF1), and Splicing factor 3a subunit 2 (Sf3a2). We examined developmental phenotypes for SF1 and Hrb87F in greater detail, showing that exon use profiles of sarcomeric genes are misregulated upon knockdown of both SF1 and Hrb87F resulting in severe developmental defects in myofibril and sarcomere structure, including impaired localization of Kettin to the z-disc. Gene expression and exon use profiles in SF1-IR and Hrb87F-IR are largely distinct from those in bru1-IR, illustrating that splicing regulators have selective targets and control discrete portions of the alternative splicing network in muscle.

4.1. Muscle fiber-type specific expression and function of RBPs

Differential splicing of structural proteins between muscle fiber-types in both flies and vertebrates has been previously described, and is thought to underly their distinctive morphologies and contractile properties [6,7,82]. One mechanism to achieve fiber-type specific splicing patterns is through expression of fiber-type specific splicing factors. An example from *Drosophila* is Bruno1 (Bru1), a CELF-family homolog which is predominately expressed in IFM and regulates fibrilar-specific splicing [52,63]. Interestingly, however, our data suggest that the majority of RBPs are expressed in all types of muscle (Figure S1 C). This is similar to RBP expression in vertebrates, where most splicing factors are expressed in all skeletal muscle fiber-types and achieve fiber-type specific function through a combination of isoform-specific expression, regulated subcellular localization or post-transcriptional modification [4,83]. For example, MBNL1 isoforms containing exon 7 are preferentially localized to the nucleus and inclusion of exon 5 further enhances splicing regulatory activity [84]. Such regulation is also disease relevant, as for example CUG repeat expansions found in patients with myotonic dystrophy sequester MBNL1 resulting in the abnormal stabilization of CELF1, disrupting normal RBP function [85–87]. We also observe differential use in RBP exons between muscle fiber-types in fly (Figure 1D, E), suggesting basic mechanisms regulating RBP localization and function may be conserved and experiments in the fly model have the potential to expand our understanding of fiber-type specific RBP regulation.

4.2. Developmental regulation of RBP expression and function during myogenesis
Our data further indicate that there are developmental shifts in RBP expression and function during IFM myogenesis. This is evidenced by the clustering of distinct RBP temporal expression profiles (Figure 1B, Figure S1 E) as well as by differences in Mef2-Gal4 and Act88F-Gal4 driven knockdown phenotypes (Figures 2-5). Temporal shifts in RBP expression levels and activity have been characterized in vertebrate muscle, notably with CELF-family proteins promoting embryonic splicing events and MBNL-family members promoting mature, post-natal splicing patterns [31,33]. Alternative splicing moreover refines contractile properties as well as transcriptional programs during myogenesis [88]. For example, a developmentally regulated switch in alternative splicing of the inhibitory MEF2Dα1 and MEF2Cα1 isoforms to the MEF2Dα2 and MEF2Cα2 isoforms which activate muscle-gene expression is necessary during vertebrate skeletal muscle differentiation, and this switch is disrupted in muscle diseases including myotonic dystrophy [89-91]. A transcriptional switch in Drosophila mediated by Spalt major (Salm) [62] as well as E2F [92] has been characterized to promote maturation of IFM, including regulation of the splicing network that favors fibrillar-specific events [52]. Additionally, alternative splicing of “early” and “mature” forms of sarcomere proteins such as Myosin heavy chain (Mhc) [93] and Zasp52 [94], among others, are reported in IFM, suggesting that transcriptional and alternative splicing switches are likely a conserved feature of muscle differentiation. It will be interesting to determine through future experiments how temporally regulated RBPs identified in our analysis contribute to the differentiation process.

4.3. Diversity in muscle phenotypes likely reflects RBP target specificity and expression level

A salient observation from our screen was the wide diversity in RBP-associated myofibril defects and the phenotypic dependence on knockdown strength. We showed that RNAi knockdown of RBPs can result in actin inclusions, torn, frayed, split, thick or thin myofibrils, short sarcomeres and disorganized or discontinuous z-discs (Figures 3-5). This phenotypic diversity most likely reflects the functional diversity of RBPs as well as their target specificity. As we show by comparing transcriptomic signatures of bru1-IR, SF1-IR and Hrb87F-IR (Figure 8 and S6), different RBPs regulate distinct, but often partially overlapping, subsets of exons. This has been systematically demonstrated by a screen of 56 Drosophila RNA binding proteins in S2 cells that found extensive combinatorial regulation of target exons and notably wide-spread cross-regulatory interactions between RBPs [24]. Extensive combinatorial and cross-regulatory interactions, in addition to factor-specific targets, are also observed in vertebrate muscle, for example between CELF, MBNL and RBFOX family proteins [25,32]. This implies that muscle RBPs participate in carefully balanced regulatory networks and provides a possible mechanism for the RBP level-dependent regulation we observe: target exons could be bound by different RBPs with different affinities, and competitive or cooperative interactions with other RBPs are likely to depend on expression levels. RBPs that function at overlapping nodes in this network would be predicted to have similar phenotypes and potentially cross-regulate, as we have recently shown in a pre-print for Bru1 and Rbfox1 [71]. Here we find that similar to Bru1 and Rbfox1, knockdown of muscleblind (mbl) and held out wings (how), previously identified to play a role in muscle development [11,95,96], as well as knockdown of Rm62, mushroom-body expressed (mub), sans fille (snf) and small bristles (sbr) produce similar hypercontractility-like phenotypes characterized by short sarcomeres and thick myofibrils (Figure 3 and 4). Further experiments, including cross-linking immunoprecipitation (CLIP) to identify RBP direct targets and binding-site preferences, will be necessary to define the alternative splicing network in IFM and determine if and how these additional RBPs cooperate with Bru1 and Rbfox1 to regulate structural gene expression in muscle.

4.4. Knockdown of spliceosome components results in alternative splicing defects in muscle

In addition to splicing regulatory factors, we evaluated 17 spliceosome components in our screen. Our data show that muscle specific knockdown with Mef2-Gal4 led to early lethality as well as severe defects in muscle structure, in particular for B and C complex members (Figure 5, Figure S4). By contrast, knockdown restricted to later stages of muscle
development using Act88F-Gal4 resulted predominately in thin myofibrils. Knockdown of A complex components caused less lethality but disrupted both myofiber and myofibril structures. Our results suggest both temporal and knockdown-level dependent phenotypes of spliceosome components in IFM. There is precedence for this observation, as several core spliceosome components have been reported to regulate alternative splicing in addition to serving as components of the canonical spliceosome, and spliceosome components including SF3B1, U2AF1, SRSF2 and hnRNP A1, among others, are often targeted by somatic mutation in cancer [97–99]. Altered levels of snRNAs themselves produce defects in alternative splicing [100]. Levels of snRNPs and spliceosome complexes are decreased in spinal muscular atrophy (SMA), resulting in alternative splicing defects and contributing to disease pathology in both muscle and neurons [38,99]. Together with our data, these observations blur the line between “basal” spliceosome components and “regulatory” splicing factors, implying a far more dynamic regulation of splicing in muscle.

Interestingly, knockdown of A complex members SF1 and Hrb87F results in myofibril defects including splitting, fraying, abnormal z-disc structures and impaired localization of Kettin, and these phenotypes can be enhanced using stronger knockdown conditions (Figure 6). Related sarcomere phenotypes have been previously characterized for proteins whose loss specifically compromises z-disc integrity, for example Zasp52 [79,94,101], Zasp66 [78], filamin [102] or zornmin [80,103]. Indeed, we find that knockdown of both SF1 and Hrb87F resulted in specific defects in sarcomere gene expression and exon use, including in Zasp66 and Zasp52 (Figure 7 and 8). Strikingly, however, the regulated exons were largely distinct from events altered in bru1-IR and did not contribute significantly to fiber-type specific exon use or gene expression (Figure 7 and 8). This shows that both factors mediate alternative splicing and have broad effects on muscle-specific gene expression, but SF1 and Hrb87F function in parallel to Bru1 and other fiber-identity factors in the splicing network.

The defects we observe in Hrb87F-IR, a Drosophila homolog of the hnRNP-A family, are reminiscent of phenotypes in hnRNP-A1 knockout mice. These mice are embryonic lethal due to defects in heart and skeletal muscle, most prominently aberrant splicing of both muscle structural genes and key regulators of myofiber differentiation including mef2c, notch4, Myo5a, Myo6, Pax6, and Foxo3, among others [104]. Conditional knockout of hnRNP-A1 in mouse muscle promotes insulin resistance and more generally disrupts glucose metabolism [105], a parallel to GO enrichments for metabolic terms we see in DE genes in Hrb87F-IR. hnRNP-A1 is also abnormally expressed in myotonic dystrophy, where like CELF1 it promotes fetal splicing patterns and antagonizes MBNL activity [106]. Our data on Drosophila Hrb87F are thus consistent with a conserved role for hnRNP-A family members in regulating splicing during muscle development. Although SF1 has been characterized to bind the branch point site and regulate alternative 3’-5S selection [107,108], to our knowledge, this is the first report of a striated muscle-specific splicing phenotype for SF1. In summary, our results illustrate the power and versatility of Drosophila as a well-conserved genetic model to identify factors with RNA-regulatory functions in muscle and to explore basic, disease relevant mechanisms of RBP function in myogenesis.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Dynamics of RBP expression in muscle., Figure S2: RBP phenotypes with Me2-Gal4 mediated knockdown in Drosophila muscle., Figure S3: Act88F-Gal4 is expressed in a subset of tubular muscles., Figure S4: Knockdown of spliceosome components produces muscle-specific phenotypes., Figure S5: SF1-IR and Hrb87F-IR cause changes in gene expression and exon use in z-disc components., Figure S6: SF1 and Hrb87F do not regulate fiber-type specific gene expression or exon use., Table S1: Key Resources Table, Table S2: Expression data of muscle RBPs, Table S3: Data from candidate RBP screen., Table S4: mRNA-Seq data from SF1-IR and Hrb87F-IR.

Author Contributions: Conceptualization, Maria Spletter; Data curation, Tobias Straub and Maria Spletter; Formal analysis, Shao-Yen Kao, Elena Nikonova, Tobias Straub and Maria Spletter; Funding acquisition, Maria Spletter; Investigation, Shao-Yen Kao, Elena Nikonova, Sabrina Chaabane, Albiona Sabani, Alexandra Martitz, Anja Wittner and Jakob Heemken; Methodology,
Maria Spletter; Project administration, Maria Spletter; Supervision, Shao-Yen Kao, Elena Nikonova and Maria Spletter; Validation, Sabrina Chaabane, Albiona Sabani, Alexandra Martitz, Anja Wittner and Jakob Heemken; Visualization, Shao-Yen Kao, Elena Nikonova, Anja Wittner and Maria Spletter; Writing – original draft, Maria Spletter; Writing – review & editing, Shao-Yen Kao, Elena Nikonova and Maria Spletter.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (MLS, 417912216), the Center for Integrated Protein Science Munich (CIPSM) at the Ludwig-Maximilians-University München (MLS), the Deutsche Gesellschaft für Muskelerkrankungen e.V. (MLS), and the International Max Planck Research School (EN).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Raw and processed values used to generate plots are available in Table S2-4. mRNA-Seq data are publicly available from GEO with accession numbers GSExxxxx, GSE63707, GSE107247 and GSE143430.

Acknowledgments: We sincerely thank Frank Schnorrer, the Bloomington Drosophila Stock Centre (BDSC), the Vienna Drosophila Resource Center (VDRC), and the Fly Stocks of the National Institute of Genetics (NIG-FLY) for providing flies. MLS is grateful to Andreas Ladurner for generous support. We thank Sandra Esser for excellent technical assistance as well as Sophie Göttke and Kelley Manbeck for preliminary work with the RBP lines. Vanessa Todorow generously shared scripts for mapping and splicing analysis. We acknowledge the Core Facility Bioimaging and the Core Facility Bioinformatics at the LMU Biomedical Center (Martinsried, DE) for confocal imaging and bioinformatics support, respectively.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References


Taylor, M.V. Comparison of Muscle Development in Drosophila and Vertebrates; Landes Bioscience, 2013;


