



Investigation of IFM specific splice factor Aret by two-steps CRISPR-RMCE based strategy in *Drosophila melanogaster*.

Masterarbeit aus dem Fachgebiet

Biochemie

Von cand. chem.

Olena Nikonova

Geboren am: 13.05.1988

In: Zaporozhye, Ukraine

Für die Masterprüfung in Biochemie an der
Ludwig-Maximilians-Universität München

Datum der mündlichen Prüfung:	14.11.2015
Beginn der Masterarbeit:	23.11.2015
Masterarbeit beim Prüfungsausschuß eingereicht am:	11.07.2016

Erklärung

Ich versichere, dass ich die vorgelegte Masterarbeit am Max-Planck-Institut für Biochemie unter der Anregung und Anleitung von Dr. Frank Schnorrer selbständig durchgeführt und keine anderen als die angegebenen Hilfsmittel und Quellen benützt habe.

München, den 11.07.2016

1. Berichterstatter: Prof. Klaus Förstemann

2. Berichterstatter: Prof. Roland Beckmann

Acknowledgments

I want to express my gratitude to Dr. Frank Schnorrer for a chance to perform my Master's Thesis in his lab and for his useful insights on the project. I am especially grateful to Maria Spletter for letting me to take part in her project, for all her help and support. Under her supervision I gained a valuable experience in fly genomics and research planning, as well as in scientific writing. I also want to thank all lab members for a perfect working climate. They were always willing to help with suggestions and advices. I especially want to thank Wouter Koolhaas for useful insights on the clonning strategy, Sandra Lemke for suggestions on implementation of Western Blot analysis and Bettina Stender for enormous help with construct injections.

Contents

Abstract	5
Introduction.....	6
Results	10
Part 1. Verification and characterization of five distinct CRISPR-mediated Aret deletions.....	10
1.1: Verification of Step I insertion location by PCR and sequencing	10
1.2. Viability assay for verification of Step I Aret deletion mutants.....	11
1.3: Fertility test to assay for known gametogenesis defects in new Step I Aret deletion alleles.....	12
1.4: Flight test to assay Aret function in indirect flight muscles	13
1.5: Recombination to remove the second site hit mutation causing lethality of line M3 3-1.....	15
1.6: Investigation of <i>aret</i> mutant phenotype in Step I fly lines by immunostaining.....	17
Part 2. Positioning of the epitope tag on Aret to build RMCE-ready Step II constructs.	29
Part 3. Investigation of Aret protein isoforms and mechanisms of function.....	36
3.1 Analysis of wild-type Aret isoforms present in IFM, ovary and testis by Western Blotting.....	36
3.2 The absence of Aret leads to structural rearrangement of essential sarcomere components. ...	40
Discussion	46
Materials & Methods	50
Bibliography.....	52
Supplement Materials	57

Abstract

Alternative splicing is a critical event in pre-mRNA processing, contributing to the diversity in isoform expression and influencing everything from cellular development to biophysical properties and metabolism. In muscle in particular, alternative splice isoforms are implicated in underlying differences in muscle function and facilitating muscle maturation, and defects in splicing likely underly disorders such as myotonic dystrophy and dilated cardiomyopathy. The flight muscles of *Drosophila melanogaster* are a useful model to both identify novel components of the splicing machinery in muscle and determine their influence on muscle function and development. Arrest, a homologue of the CELF family of splicing factors, was recently identified to control flight muscle specific alternative splicing, allowing proper sarcomere growth and function. Here I characterize 5 new CRISPR deletion alleles in the Aret locus, identifying a possible novel role for a subset of Aret isoforms in controlling myosin function independent of functions in sarcomere growth. I generate C-terminally tagged HA and V5 Aret alleles for future biochemistry experiments. I additionally confirm the Aret sarcomere growth and hypercontraction phenotypes in homozygous null mutants. Further, I perform Western Blot analysis and identify female and male specific Aret isoforms expressed tissue-specifically in ovary or testis. Excitingly, I also characterize sarcomere component localization and find that both bent (*Dm* titin homolog) and myosin (Mhc) are localized in a more tubular pattern in Aret knock-down flight muscle. This suggests that Aret may result in the formation of “mini-tubes” and imply a switch to a tubular-like pattern of sarcomere growth and maturation in Aret mutant muscle.

Introduction

After transcription, before translation takes place in the cytoplasm, every newly synthesized eukaryotic mRNA undergoes several steps of maturation. One critical event, called splicing, involves the removal of intronic sequences from the nascent pre-mRNA and ligation of exons. Splicing in eukaryotes involves a complex mechanism, comprising the joint function of spliceosome components (U snRNPs or Snurps) and spliceosome-associated proteins [1]. The spliceosome machinery not only needs to correctly identify the prominent splice sites, relying on internal splicing signals on a pre-mRNA sequence, it also needs to correctly position its components within an atomic distance range to enable the transesterification reaction. The mechanism becomes even more complex when it comes to alternative splicing, the process during which exons are either retained in the mRNA or targeted for removal in different combinations to create a diverse array of mRNAs from a single pre-mRNA. This gives rise to different protein isoforms in different tissues, developmental stages, or disease conditions. The choice of splicing sites on a pre-mRNA transcript is determined through interactions with RNA-binding proteins acting as splice factors (SFs; [2]). As has been shown for sex-specific splicing regulation in *Drosophila*, even one SF can act as a master regulator, controlling an entire chain of specific alternative splicing events [2], [3].

Regulation of tissue-specific mRNA and protein isoforms through alternative splicing has been documented in vertebrate muscle [4]. Mammalian skeletal muscle is composed of large, multi-nucleated cells called myofibers and can be classified as having slow (type 1, red muscle) or fast (type 2, white muscle) fibers. Fiber diversity is created on the level of both gene expression and alternative splicing, which create a unique fiber-type expression profile and allow for functional diversity. For example, Troponin T (TnT) in mammals encodes cardiac, slow and fast TnT isoforms. The fiber-type specific splicing of fast TnT enables the production of different isoforms with varied Ca^{2+} sensitivity and thus influences on actomyosin interactions [5]. Another example is muscle specific splicing of titin, a core component of sarcomere. The specification of heart muscle induces a very different splicing pattern compared to skeletal muscle, facilitating the different physiological properties between heart and skeletal striated muscle. Heart-specific splicing of titin by RBM20 promotes exon skipping of the flexible PEVK exons, resulting in shorter titin isoforms and increasing the stiffness of heart muscle [6]–[10]. The misregulation of alternative splicing may have severe effects, as for example mutations in RBM20 affect the length of titin's PEVK domain and are associated with dilated cardiomyopathy (DCM) [11]. Despite the contributions of alternative splicing to muscle development and disease, only a small handful of AS factors have been identified and their mechanisms of action are poorly understood.

Drosophila melanogaster is an ideal model organism to study both alternative splicing and muscle development. It has short generation times, is easily cultureable, has a broad array of genetic tools and shows remarkable conservation with many vertebrate systems. Notably, the muscles are highly conserved to the structure and function of individual proteins and many important mechanistic insights into myosin function, mechanisms of contractility, mechanisms of differentiation and insights into the cause of muscle diseases have been identified from *Drosophila*.

Drosophila possesses two functionally different types of body muscles. Tubular muscles in the head, legs and thorax act similarly to vertebrate skeletal muscles, contracting synchronously upon calcium release. In this type of muscles, myofibrils are laterally aligned, with the nuclei located in the center of the tube. On the other hand, *Drosophila's* indirect flight muscles (IFMs) are arranged differently to allow very fast oscillations and high power output, enabling flight [12]. They are composed of unfused, non-aligned myofibrils, with regularly spaced nuclei located between the fibers. Asynchronous contractions of fibrillar muscles are enabled by the combinatorial effect of calcium stimulation and stretch-dependent activation [13]–[15].

The differences between the muscle types arise developmentally based on transcriptome and proteome diversity arising from differential gene expression and alternative splicing. As has been shown by previous studies, Zn-finger transcriptional factor *spalt major (salm)* is necessary and sufficient for induction of the fibrillar muscle fate [16]. *Salm* activates a dedicated genetic network controlling flight muscle identity by inducing both expression of fibrillar specific genes and fibrillar specific splicing [17]. Prominent among those targets is the RNA-binding protein Arrest (Aret, Bruno), which has been shown to promote fibrillar exon inclusion and to inhibit the use of tubular exons during IFM specific splicing [18]. Arrest was first identified to have a cytoplasmic function as a translational repressor of *oskar* mRNA during early embryogenesis in *Drosophila* [19]–[22]. Thereafter, Aret has been suggested to have a similar function in translational control of *grk* mRNA in the oocyte [21]. Only recently has it been proposed as a prominent splice factor, controlling fibrillar muscle identity through regulation of IFM-specific splicing of a very significant subset of sarcomeric genes [17]. Among its targets are Stretchin (Strn-Mlck), Sls/Kettin, and WupA, which are incorporated into the growing sarcomeres during myofiber maturation [18]. Aret is a *Drosophila* orthologue of the CELF protein family, sharing the common structure composed of three evolutionary conserved RNA recognition motif (RRM) domains and a unique linker region called the “divergent domain” between RRM2 and RRM3 [23]. CELF proteins can be found both in the nucleus and cytoplasm where they regulate multiple aspects of gene expression, including pre-mRNA alternative splicing, RNA editing, deadenylation, mRNA stability and translation. Their role in vertebrate alternative splicing during cardiac development has been suggested by previous studies. It was shown that a developmental increase in CELF protein accumulation in the heart correlates with adult patterns of splicing prevailing over fetal patterns [24], [25].

Analysis of *aret* loss of function alleles in *Drosophila* (such as *aret*^{QB72} and *aret*^{PD41}) showed that loss of Aret does not affect viability, but results in female-sterility and a flightless phenotype [18]. These results coincide with the suggested dual function of the protein in the germline cytoplasm [19] and developing IFM nuclei [17], [18]. Previous studies also showed that Arrest is indeed required for myofiber maturation and sarcomere growth happening after 48 h APF [18]. However, little is known about mechanistic aspects of Aret function, including the functional importance of different *aret* isoforms, its potential interaction partners or the recognition specificity of RRM domains during IFM specific splicing. This Master’s thesis explores new genetic tools and foundational experiments to address these mechanistic questions.

To achieve a deeper insight into the function of Aret as a SF in fibrillar muscles, we decided to use the highly flexible two-step fly-genome engineering strategy originally introduced by Zhang *et.al.* 2014 [26]. This strategy is based on the combination of CRISPR with recombinase-mediated cassette exchange (RMCE) for targeted genome modification (Fig. 1, Fig. 2). In the first step, a cassette containing a splice acceptor, an SV40 terminator and a 3xP3-dsRed eye reporter flanked by attP sites is integrated at a precise genomic position using the CRISPR/Cas9 system. This enables both the efficient identification of the targeted event and the generation of a strong loss of function allele. During the second step, the flanking attP sites are used to replace the inserted cassette with any DNA of choice using RMCE technology, which is already well established in *Drosophila* [27].

When I started my project, five Step I CRISPR-mediated Arrest deletion fly lines had been generated (Spletter and Schnorrer, unpublished data), each targeting distinct regulatory regions of the Arrest locus (Fig. 2). Deletions 1 and 2 (further referred to as M1 and M2) are relatively large, removing both the RRM2 and RRM3 domains as well as the “divergent domain” that separates them. Deletion 3 (M3) takes out part of RRM3 as well as the 3'UTR. The M1-M3deletions are expected to affect all of the *aret* isoforms. Deletions 4 and 5 (M4 and M5) are much shorter and target the protein region prior to RRM1 that has been suggested to be required for Arrest dimerization [28], [29].

This Master’s thesis comprises three main aims:

- (1) The verification and characterization of Arrest mutants M1-M5 generated as Step I CRISPR alleles.
- (2) Cloning of Step II constructs containing a C-terminal epitope tag and generation of transgenic epitope-tagged flies using the Φ C31/RMCE strategy.
- (3) Investigation of the Aret phenotype and Aret’s mechanism of action on the protein and genetic level with Western blotting and immunohistochemistry.

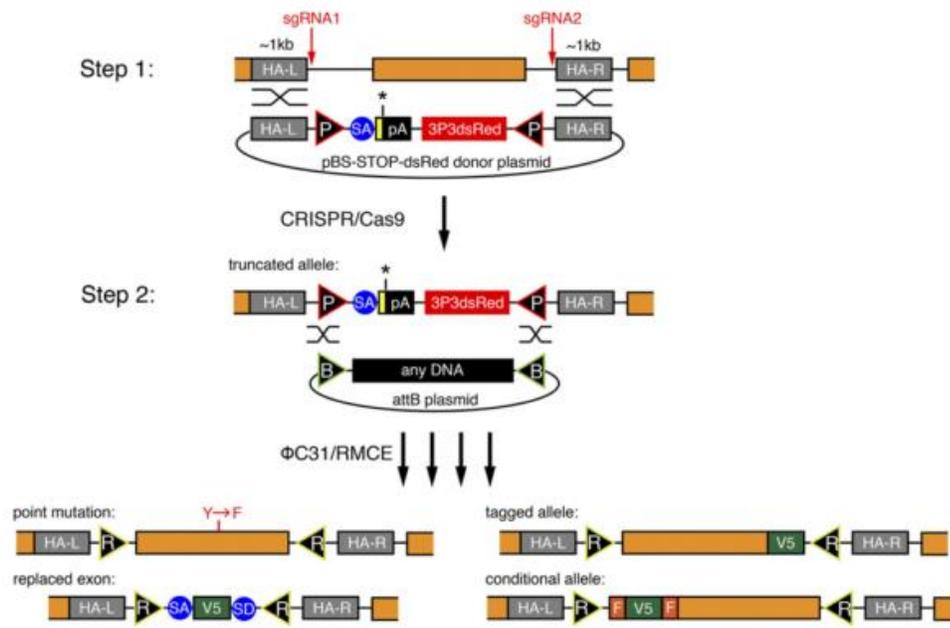


Figure 1. Overview of the two-step CRISPR/RMCE-based strategy for fly genome engineering with potential modifications. [© Zhang et.al 2014, p. 2412]

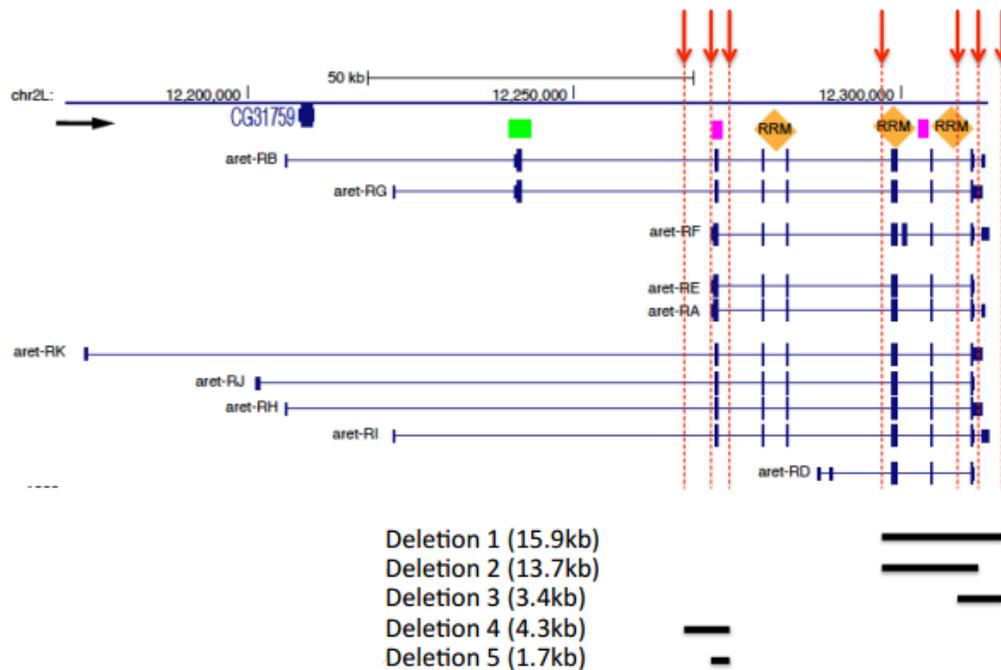


Figure 2. Representation of *aret* isoforms and regions deleted by Step I CRISPR-mediated strategy. [Spletter and Schnorrer, unpublished data]

Results

Part 1. Verification and characterization of five distinct CRISPR-mediated *Aret* deletions

The first aim of this thesis was to verify and phenotypically characterize five previously generated CRISPR-mediated *Aret* deletions, M1-M5. Several approaches were used to achieve this, including (1) verification of the correct Step I cassette insertion location by sequencing and characterization of founder lines through (2) a viability assay, (3) a fertility test, (4) a flight test and (5) immunostainings for *Aret* and actin at different developmental time points. Furthermore, by crossing *Aret* mutants to an *aret* deficiency allele (BL24911), we obtained transheterozygous flies that were also analyzed by flight tests and immunostainings to confirm the specificity of the mutant phenotype. Through the verification of the Step I flies, we aimed to both gain a deeper insight into the *Aret* phenotype as well as to form a basis for the insertion of Step II constructs.

1.1: Verification of Step I insertion location by PCR and sequencing

Stable stocks generated from at least one independent founder line for each Step I deletion were tested for correct localization of the DsRed cassette and absence of ends-in insertion by PCR amplification, followed by sequencing of the amplified band (Fig. 3). It was essential to verify the integrity of the Step I fly lines for any of the further experiments. During insertion of the Step I cassette, there is a risk of genomic abnormalities or off-target effects that could affect Step II cassette exchange or lead to misinterpretation of the phenotype. Full insertion of the donor vector (i.e. an ends-in insertion) may also lead to phenotypic variations and can considerably reduce the chances for successful Step II cassette exchange.

To identify the optimal Step I fly lines, two sets of specific primers were used. The first primer was designed to bind on the corresponding homology arm (HA) (Fig. 3A) and the second primer was designed to bind on the genomic region outside of the expected insertion site (Fig. 3B). Note that some of the deletions start or end on exactly the same base pair, consequently sharing the same homology arms (Fig. 3D). For the "ends-in test", we selected primers giving a positive-negative response (Fig. 3C). After several rounds of sequencing, we were able to verify 7 correct fly lines (Fig. 3E, full list of primers used for verification can be found in Supplement Table 2).

Although the presence of the cassette was confirmed for each line by primers on the homology arm, not all insertions proved to be precisely correct. Namely, M3 founder 1-3 contained a full backbone insertion (i.e. was ends-in), while amplification of the region downstream of the right homology arm of M1 line failed in all lines derived from two independent founders. A further PCR amplification within the deleted region of M1 gave a positive band, confirming that *aret* was not deleted in M1 mutants (data not shown). This suggests that potentially only one of the sgRNAs used to generate M1 cut efficiently, somehow resulting in an insertion of the Step I cassette into the *Aret* locus without deletion of the targeted native sequence. Furthermore, M5 line was confirmed by sequencing (Fig. 3E), but showed adult lethality, unexpected for *aret* mutant. Due to this, we decided to reinject Step I constructs 1 and 5 into Act5C-Cas9, lig⁴ flies. Each new injection gave of 1 founder line that remains to be confirmed and analyzed for consistency of mutant phenotype in follow-up studies.

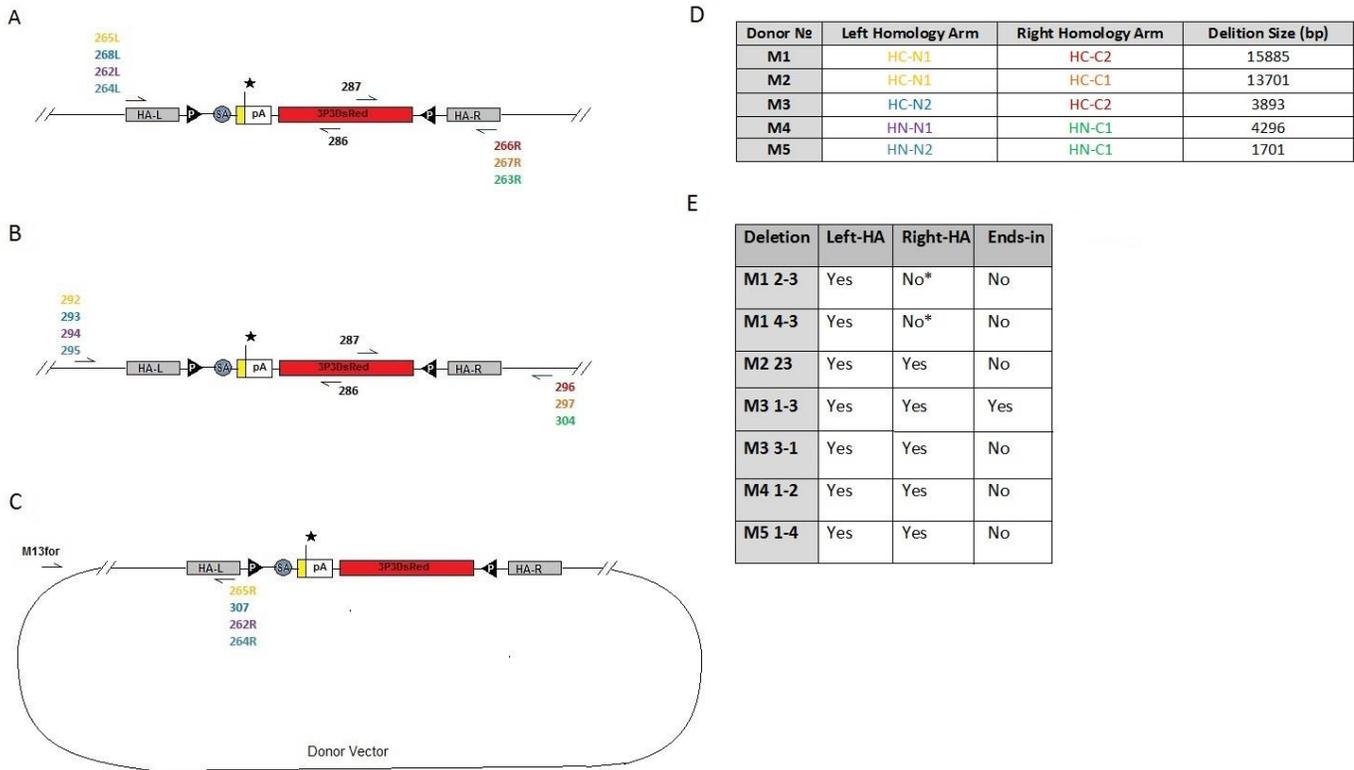


Figure 3. Verification of Step I construct insertion via sequencing

- A.** Schematic of the Step I cassette. Primer sets used for initial verification are indicated.
- B.** Schematic of the Step I cassette illustrating primer sets used for final verification and sequencing.
- C.** Schematic representation of the Step I insertion vector depicting primers used for the "ends-in" test.
- D.** List of Step I deletions with corresponding homology arms. Color coding corresponds to primer text color in other figure panels. The first two letters of the homology arm abbreviation indicate location on the *aret* locus (C- or N- terminal), while the last two symbols refer to the order within a region (i.e. the first (C1) or second (C2) C-terminal homology arm (HA)). Note that some of the deletions share the same HA.
- E.** List of confirmed founders for each Step I deletion. Results are based on sequencing data.

1.2. Viability assay for verification of Step I Aret deletion mutants

To assay whether the deletion mutants were truly viable and corresponded to predicted Mendelian ratios, a "viability test" was conducted. We aimed to determine how the absence of Arrest impacted survival rates, if at all, and to see how homozygous mutant animals performed in

a population. The test was performed as follows: heterozygous crosses were conducted for each deletion (M1-M5)line ($Aret^{Mx}/CyO \times Aret^{Mx}/CyO$) and the resulting F1 progeny were counted and analyzed by the chi-squared statistic and corresponding p-value for significance. Taking into account that CyO/CyO is lethal, the expected Mendelian ratios for the cross were 2 ($Aret^{Mx}/CyO$):1 ($Aret^{Mx}/Aret^{Mx}$). Accordingly, the expected value of completely viable homozygotes was estimated at 33.3% of the total fly count. From Table 1, we can conclude that homozygous mutants are semi-viable, as most of them, except lines M1 and M4, deviate significantly from predicted ratios. Thus, it can be suggested that the absence of Arrest or possibly the presence of the DsRed cassette may have a slightly negative effect on the survival of the flies. Nevertheless, it would be useful to perform the viability test on a higher number of F1 flies, to avoid potential statistical errors due to small sample size. Moreover, both the M3 3-1 and M5 1-4 lines appeared to be homozygous lethal, but as lethality is rescued when crossed to the deficiency, both lines appear to have a randomly occurring second site hit from off-target CRISPR sgRNA cuts in Step I, as discussed below.

Deletion	Total fly number	DsRed/CyO	DsRed/DsRed	Chi-square Test	P-value (p< 0.05)
M1 2-3	110	78	32	0.76025	0.383266
M1 4-3	131	101	30	6.01784	0.014162
M2 23	137	103	34	4.14167	0.041842
M3 1-3	590	445	145	18.9353	< 0.0001
M3 3-1*	130	130	0	64.02985	< 0.0001
M4 1-2	57	45	12	3.67004	0.0554
M5 1-4*	102	101	1	47.16719	< 0.0001

Table 1. Results of viability assay for Step I fly lines.

Table shows the number of observed flies of a specific genotype for each deletion line tested. Chi-square statistic and p-value are shown. P-values listed in blue deviate significantly from Mendelian ratios, while those depicted in red do not differ significantly from the predicted values. (P-value < 0.05)

1.3: Fertility test to assay for known gametogenesis defects in new Step I Aret deletion alleles

Previous studies showed that *aret* mutants have gametogenesis defects in males and females [30], [31], making the fertility test particularly useful for Step I fly characterization. Sterility would serve as a good indication of an absence of functional ovarian or testis Aret protein isoforms. A fertility test was performed by crossing male or female homozygous mutants ($Aret^{Mx}/Aret^{Mx}$)

with w- flies of opposite gender (Table 2). The M2, M3 and M4 lines were completely sterile, in agreement with previous data and confirming the loss of Aret function in those deletions [18]. Notably, all stocks derived from the M1 founders were fertile and their progeny were viable. Together with the sequencing results, this strongly indicates a problem with the M1 line and confirms a failure in the deletion of Arrest in those flies. The M3 3-1 and M5 1-4 lines were excluded from the analysis due to their lethality. These lines can be tested in the fertility test in the future after the second-site mutation has been recombined off of the targeted chromosome.

Deletion	Male cross ♂ ; dsRed:: × ♀ w ⁻	Female cross ♀ ; dsRed:: × ♂ w ⁻
M1 2-3	Fertile	Fertile
M1 4-3	Fertile	Fertile
M2 23	Sterile	Sterile
M3 1-3	Sterile	Sterile
M3 3-1	Non viable	Non viable
M4 1-2	Sterile	Sterile
M5 1-4	Non viable	Non viable

Table 2. Results of Fertility Test in Step I fly lines.

Table shows the summarized results for fertility for each line tested. Each cross was repeated twice. Lines highlighted in red could not be tested due to adult lethality.

1.4: Flight test to assay Aret function in indirect flight muscles

In addition to its function in early gametogenesis, Arrest is required for flight muscle–specific splicing in *Drosophila* [17]. Arrest mutant muscle shows defects both in maturation and growth of the sarcomere, as well as a hypercontraction phenotype resulting in loss of IFM fibers in adults, resulting in a complete loss of flight capability [18]. Therefore, we conducted a flight test to prove the loss of Aret function in IFMs of Step I fly lines. Adult homozygous male flies were collected and recovered overnight, then transferred to a flight test cylinder. The cylinder was divided into 5 regions and flies were scored based on which region they landed in. The bottom two regions were considered flightless, while the top regions indicated flight capability.

As expected based on the preceding sequencing, fertility and viability assays, M2, M3 and M4 homozygous mutants, for which the deletion was verified by sequencing, were completely flightless (Fig. 4). The M1 line was able to fly, although the performance was slightly reduced compared to wild-type. This is in agreement with results above suggesting Aret is not deleted in

M1. Once again, it was not possible to conduct a flight test for the M3 3-1 and M5 1-4 lines that displayed lethality.

In addition to testing homozygous mutant flies, we tested trans-heterozygous combinations of Step I mutants over the Df(2L)BSC407 allele (Bloomington number BL24911) to show our deletions and associated phenotypes are due to specific targeting of Arrest. Lines M2, M3 and M4 were flightless over the deficiency, confirming specific deletion of Aret in those alleles. Line M1 over the deficiency could fly, in agreement with above results that Arrest is not deleted in this line. Interestingly, the adult lethality of the M5 1-4 and M3 3-1 lines were rescued by the cross to Df allele, indicating that the lethality is due to a second site hit. The resulting viable adults were 80% flightless in the case of M5 1-4. These results may be especially interesting considering the region affected by the deletion. Nonetheless, it will be necessary to recombine off the second-site lethal mutations before reaching any phenotypic conclusions.

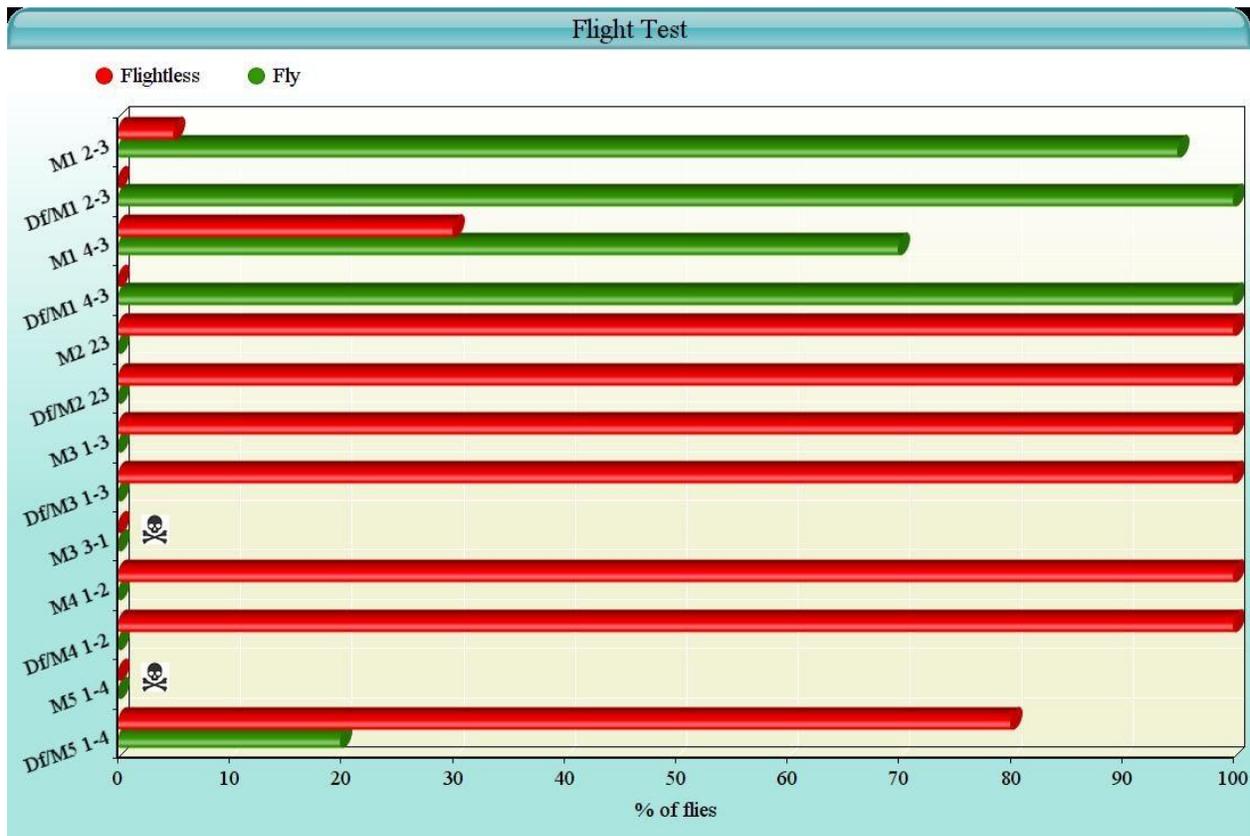


Figure 4. Results of Flight Test in Step I fly lines.

Flight ability by genotype, as listed. Red bars indicate flightlessness, while green bars indicate percent of flies that can fly. N= >50 for Aret^{Mx}/Aret^{Mx} and >15 for Aret^{Mx}/Df. M3 3-1 and M5 1-4 are lethal. Only homozygous males were assayed.

1.5: Recombination to remove the second site hit mutation causing lethality of line M3 3-1.

As mentioned above, we obtained two independent founders for the M3 deletion allele: M3 1-3 and M3 3-1. Both were verified for the correct DsRed cassette insertion. Unfortunately, the M3 1-3 line, although it showed the predicted *aret* mutant phenotype, contained an ends-in insertion, while the M3 3-1 line contains a homozygous lethal second-site mutation. This homozygous lethality makes it impossible to complete analysis of the fly line and could complicate interpretation of subsequent Step II alleles made with this line. Interestingly, from the cross of M3 3-1 to the deficiency line Df(2L)BSC407, we were able to get viable and flightless homozygotes (data not shown). This suggested that the lethality of the M3 3-1 line was coming from a second site hit on the same chromosome as the *aret* locus. Hence, we decided to clean-up the allele by recombining it back to w- and screening for viable fly lines, because the M3 3-1 line was not ends-in making it more convenient for subsequential Step II construct injections. Following the crossing scheme (Fig. 5B-C), we were able to get 5 lines with viable, DsRed positive homozygous flies. The genetic distance between *Aret* and the second site hit, based on the number of recombinants, was calculated at 5.2 centimorgans (cM), which agrees with predicted recombination rates on the 2nd chromosome arm (3.23 cM/Mbp - 7.62 cM/Mbp, [32]). Taking that into consideration, the second site hit was presumably located within a 1-2 Mbp region on the left arm of the second chromosome (Fig. 5A). The *aret* mutant phenotype was confirmed in two distinct lines: M3 3-1 11 and M3 3-1 32 (Fig. 5D-F), that were indeed viable, sterile and flightless

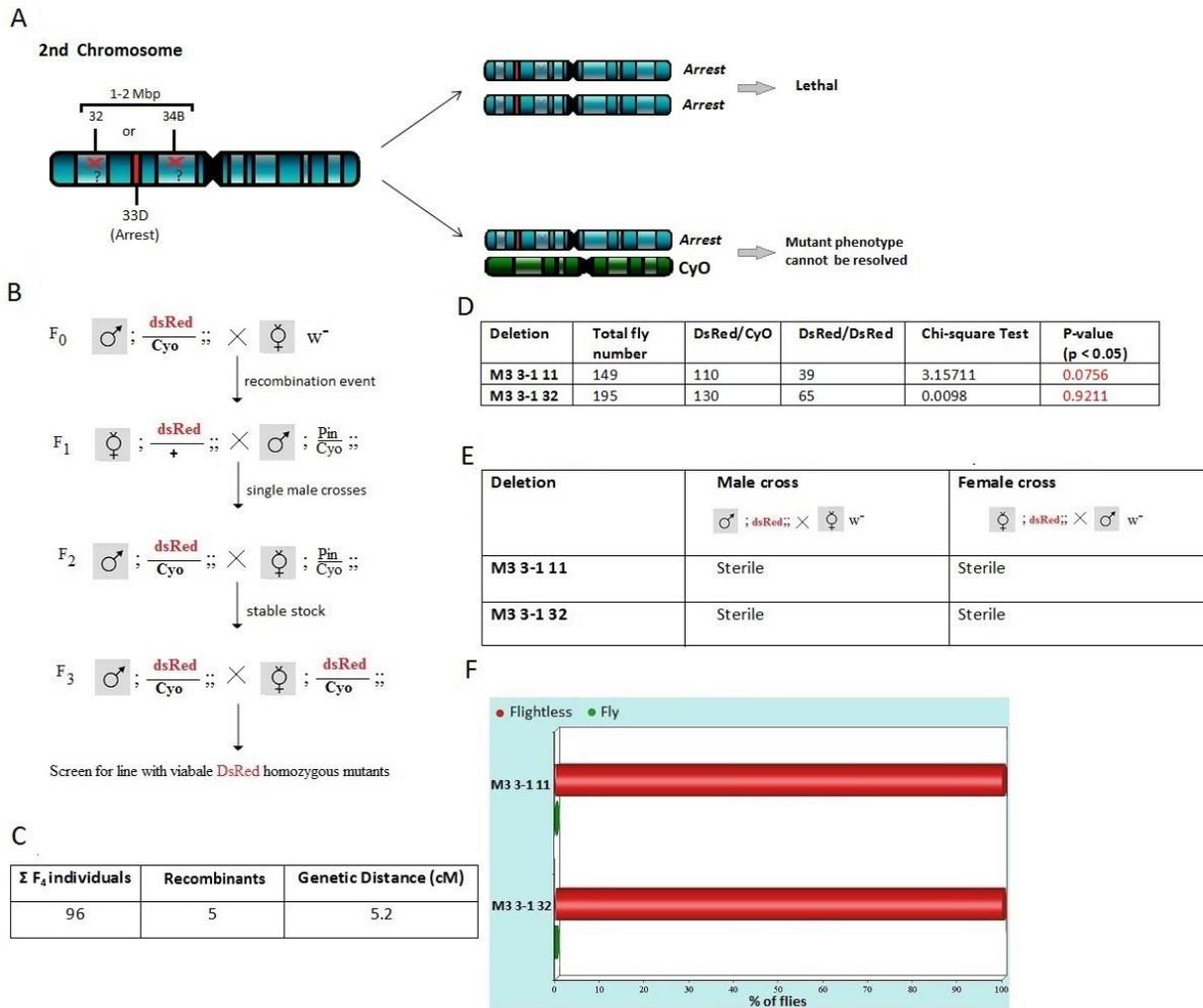


Figure 5. Lethality of M3 3-1 line comes from a second site hit on chromosome 2L.

- Schematic representation of the second site hit location and outcomes of its presence on the chromosome. Native chromosome carrying the construct depicted in blue, balancer chromosome depicted in green.
- Crossing scheme for recombining off the second site hit.
- Results of the recombinants cross.
- F Confirmation experiments on rescued M3 3-1 lines, showing the M3 3-1 allele to be viable (D), sterile (E) and flightless (F).

1.6: Investigation of *aret* mutant phenotype in Step I fly lines by immunostaining.

A notable characteristic of *aret* mutants is their flightlessness, implying severe alterations in muscle organization. IFM muscle fibers in *aret* mutant adult flies stretch and tend to rupture close to their thoracic attachment sites, subsequently leading to muscle degeneration [18]. We decided to perform an immunostaining analysis of Step I deletion fly lines to confirm the loss of Aret from the nuclei of mutant IFMs and to compare the phenotype of Step I deletion mutants to *aret* loss of function alleles and *aret* RNAi-induced knockdown. As Step I cassettes targeted different combinations of regulatory and coding regions on the *aret* locus, we might expect the existence of phenotypic variability between the mutant lines that may help to define the distinct splice-specific function of isoforms defined by each domain.

Immunostainings for Aret were performed at 90h after puparium formation (APF) and in 1-day old adults. Aret was previously shown to be almost entirely translocated to the nucleus by 72h APF, which coincides with the time when IFM sarcomere maturation occurs [18], [33]. By 90h APF, the fibrillar muscles are almost completely mature but the fly has not yet eclosed and regularly used the IFMs, making this particular time point ideal for phenotypic observation of myofibril growth and sarcomere organization. On the other hand, 1d adult stainings allow observation of phenotype progression in actively used IFMs, for example to observe hypercontraction. We used the w-fly line as a wildtype control and a transheterozygous combination of loss of function alleles *aret*^{QB72}/*aret*^{WH53} as a phenotype positive control. Comparison of Step I M1-M4 lines to control mutants was important to confirm if our new *aret* alleles phenocopy existing mutant phenotypes. Both existing *aret* alleles derived from EMS screens [31]. *aret*^{QB72} carries a nonsense codon upstream of the third RNA-binding domain, while the nature of *aret*^{WH53} has not been characterized. Additionally, we performed RNAi knockdown of Aret with Mef2-GAL4 driving the GD41568 hairpin from the Vienna collection.

The IFMs of wildtype flies are composed of six dorsal longitudinal muscles (DLMs) and six dorsal ventral muscles that are attached to the thoracic exoskeleton (Fig. 6A-A', Fig. 7A-A'). At 90h APF, a slight curvature of the IFM fibers is considered normal. By contrast, some fibers in trans-heterozygous *aret*^{QB72}/*aret*^{WH53} mutants already looked thinner and pulled at 90h APF (Fig. 6B-B') and were severely disrupted and even ruptured soon after eclosion (Fig. 7B-B'). Similar results were obtained for Aret-IR knockdown flies, with fibers starting to be visibly pulled from the anterior site already at 90h APF (Fig. 6C-C', tension sites indicated with white arrow heads) and all fibers showing severe hypercontraction and rupturing in 1d adults (Fig. 7C-C'). In both cases, the disruption in IFM organization was apparent and the width of the fibers varied considerably within and between the six fibers.

We also examined the fiber phenotype of M1-M5 mutant hemithoraces at 90h APF and in 1d adults. The phenotype of the M2 23 and M3 1-3 deletion mutants matched that observed in mutant controls. At 90h APF, IFM fibers looked strained but mostly remained attached to the thorax (Fig. 6E-F'). In 1 day adults, the fibers were pulled from the attachment sites and the majority of analyzed M2 and M3 mutant thoraxes displayed the formation of round bundles of muscle fiber due to fiber rupture (Fig. 7C-D'). We also confirmed the phenotype of the M3 3-1 mutant in 1d adults after recombination to remove the second site hit (Fig. 7E-E'). IFM fibers showed the canonical pulling and rupturing observed in other Aret mutants. We do observe some degree of variability in the severity of IFM fiber degeneration in 1d adults, which presumably correlates with the degree of hypercontraction or the time of active IFM usage. By examining 3 or 5-day old adults, we would be able to confirm that all flies show fiber rupture and

degeneration. Based on comparison with previous data, we conclude that deletions M2 and M3 are bona fide *aret* loss-of-function alleles resulting in hypercontraction and fiber loss in adults.

M4 1-2 mutant flies, which have a deletion affecting the promoter and first exon of only a subset of mid-length isoforms, had a different phenotype than other *Aret* mutants. At 90h APF, IFM fibers were considerably thinner than in the wild-type, but the fibers showed no sign of pulling or rupture (Fig. 6G-G'). In 1d adults, IFMs did detach from the posterior thorax (Fig. 7F-F'), but fibers were more organized as compared to M2 and M3 mutant lines and the IFM fibers did not appear to be overstretched.

We were not able to analyze the M5 fiber phenotype due to the second site lethal mutation. In M1 2-3 mutant flies, the fibers appeared completely wild-type at 90h APF (Fig. 6D-D') and in 1d adults (data not shown). Additionally, *Aret* signal was detected in IFM nuclei (Fig. 8D, indicated by red arrows), confirming *Aret* presence in the line. This data confirms the above data, requiring us to reinject the Step1 donor construct and potentially identify a new sgRNA to generate the desired M1 deletion allele.

Following our observation of the IFM fibers, we analyzed the myofibrillar structure of Step I mutant myofibrils at 100x to look for sarcomere and myofiber defects. We aimed to confirm the absence of *Aret* signal in the nuclei and determine if each deletion replicated the alterations in sarcomeric length that have been reported previously [18]. To investigate the real quantitative difference between mutant and wild-type sarcomeres, we applied a newly developed semi-automated script in Fiji (Image J) written by Giovanni Cardone for sarcomere length and width determination. The software uses a Fourier analysis to estimate the repeat length along the myofibril (i.e. sarcomere length) based on their periodic structure. The plug-in was run on 90h APF myofibrils, as progressive degeneration of in young adults makes it impossible to score the sarcomere parameters.

IFM myofibrils of wild type flies are fibrillar, meaning they are not laterally aligned and the nuclei are regularly spaced between the fibrils. As can be seen from Figure 8A, the nuclei of the wild-type control are regularly spaced between the fibrils and stain positive for *Aret*. By 90h APF, wild-type sarcomeres have reached the length of 2.93 μm (SD= 0.22). Sarcomere growth starts at 48h APF (2.06 μm (SD=0.08)) reaching 2.35 μm (SD= 0.10) at 56h APF. By 72h APF, sarcomeres are already 2.83 μm long (SD= 0.17) and continue to grow slowly until 90h reaching 2.93 μm (SD=0.22). After eclosion, sarcomeres elongate even further, reaching a length of 3.24 μm in 1d adults (SD=0.24). Similarly, the sarcomere width changes considerably between 48h and 60h APF (from 0.43 μm to 0.63 μm) and between 72h and 80h APF (from 0.78 μm to 0.92 μm). At 90h APF, myofibrils reach 0.90 μm (SD= 0.008) in width. (Time-course data generously provided by M. Spletter, unpublished data).

We next examined the detailed phenotype of *Aret* mutants M2 and M3 and *aret*^{QB72}/*aret*^{WH53}, as well as *aret*-IR control lines. We first confirmed that *Aret* staining was lost in the nuclei of M2 and M3 mutants IFMs as well as in *aret*-IR (Fig. 8C,E, F; Fig. 9C-E). Surprisingly, *Aret* signal was detected in transheterozygous *aret*^{QB72}/*aret*^{WH53} mutants. This signal is either from high background noise or if real, may suggest the potential existence of a truncated *aret* isoform in those flies (Fig. 8B, Fig. 9B). Next, we examined myofibril structure. In general, the length of sarcomeres in the mutants was indeed shorter than the wild-type control (Fig. 11, Table 3). At 90h APF, the sarcomere length of M2 was 2.29 μm , while the observed sarcomere length in M3 was 1.90 μm . Transheterozygous *aret*^{QB72}/*aret*^{WH53} mutants appeared to have somewhat longer sarcomeres (2.44 μm) than *aret*-IR (2.17 μm). Note that a length of ~2.4 μm is normally reached at around 60h APF. Although M3 1-3 sarcomeres seemed to be slightly shorter compared to other

mutants, the measurements for M2, M3 and *aret-IR* mutants coincide with the predicted wild-type sarcomeric length in between 48h and 56h APF. Interestingly, that is exactly the time when Aret translocates to the nucleus.

Additionally, we observed strong myofibril phenotypes. The fibrils had a tendency to fuse together and split apart along the myofibril length, resulting in a "wave-like" pattern already at 90h APF (Fig. 8C-F). By observation of young adult IFM fibrils, we conclude that active usage of IFM fibrils after eclosion led to amplification of the mutant phenotype, resulting in a gradual degradation of sarcomeric pattern, the formation of round bundles on the level of myofibrils and an increase in fibrillar fusion (Fig. 9B-E). These conclusions are supported by an increase in myofibril width observed in the mutants (Table 4). We measured a width of 1.30 μM at 90h APF for line M2 23 and a width of 1.32 μM at 90h APF for line M3 1-3, while the width of *aret-IR* flies was 1.29 μM at 90h APF. In some cases, fibrils were wide enough to contain two to four wild-type thickness fibrils. Additionally, sarcomeres of *aret*^{QB72}/*aret*^{WH53} appeared to be slightly thinner (1.08 μM) at 90h APF, compared to other mutants. The fibril fusion was especially obvious in M3 1-3 mutants (Fig. 8F, Fig. 11D), where the myofibrillar pattern was reminiscent of tubular muscle with its aligned sarcomeres. Combined with cross-section data showing hollow myofibrils and hook-like structures (M. Spletter and F. Schnorrer, unpublished data), we suggest that some kind of "mini-tubes" may be formed in *aret* mutants lacking the RRM3 domain. Additionally, we noted that the nuclei in Aret mutants appeared to be shifted closer together rather than regularly spaced. We would need to examine more hemithoraxes and quantify nuclear organization to confirm this phenotype in M2 and M3 mutants. Taken together, the sarcomere length and width of *aret-IR*, M2 and M3 mutants was comparable, indicating the reliability of the results.

The phenotype of M4 1-2 mutants was considerably different. Aret staining in IFM nuclei appeared to be positive (Fig. 8G, indicated by red arrows). This may indicate that short and long isoforms of Aret are still present, as the mutation should only affect mid-length isoforms. Additionally, myofibrils appeared much less affected than in M2 and M3 mutants, as the myofibrils were individually aligned and showed no sign of "mini-tubes" formation. The myofibrils of the M4 1-2 mutant were 0.79 μM in width, with a length of 2.83 μM at 90h APF that differed insignificantly from WT samples (Table 3, 4). Interestingly, these numbers correlated precisely with sarcomere parameters in wild-type flies at 72h APF, suggesting that M4 Step I deletion arrests sarcomeric growth around 72h APF. Moreover, after eclosion M4 sarcomeres progressively get shorter and thicker (Fig. 11G, Table 5), with 2.64 μM long and 1.17 μM thick sarcomeres at 1d and 2.56 μM long and 1.20 μM thick sarcomeres in 3d adults, indicating that M4 mutant sarcomeres still undergo hypercontraction. We also confirmed the specificity of the mutant phenotype for M2, M3 and M4 by immunostaining of IFMs from transheterozygous *Aret*^{MX}/*Df(2L)BSC407* flies in 1 day adults (Fig. 10). The examined transheterozygotes gave the same mutant phenotype both for IFM fibers and myofibrillar organization observed with homozygous mutant flies.

These results can be interpreted based on the location of the Step I M4 cassette insertion. The deleted N-terminal regulatory region located upstream of RRM1 does not affect all *aret* isoforms. It clearly should not affect the short isoform D (Fig. 2), and there is also a possibility that long *aret* isoforms are capable of splicing over the deleted region. This might mean that some isoforms may still be present in the IFMs of the M4 mutant line, and potentially could partially compensate for the loss of other isoforms. Our data suggest that the mid-length isoforms may contribute more minimally to sarcomere development, but are required for proper splicing necessary for correct myosin function. Nevertheless, the minimal knowledge about IFM-specific

aret isoforms and their potential splice targets prevents us from drawing any strong conclusions without further experiments.

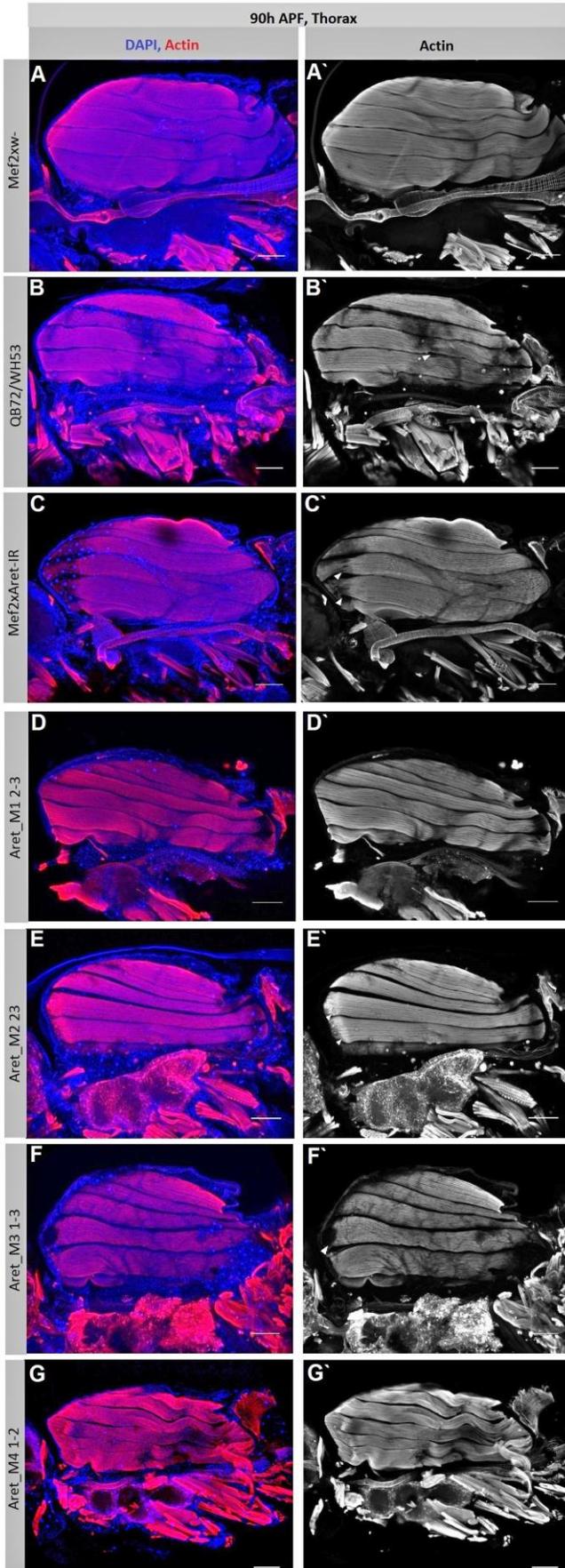


Figure 6. IFM fiber phenotypes at 90h APF

Hemithoraxes from wt control (A-A'), *aret* mutant controls (B-C') and Step I mutant lines (D-G'). IFM fibers in Mef2Gal4 x w- flies display a wild-type phenotype (A-A'). Fibers are pulled apart in transheterozygous *aret*^{QB72}/*aret*^{WH53} flies (B-B') and the M2 23 Step I line (E-E'), or even ruptures close to attachment sites as in Mef2Gal4 x *aret*-IR flies (C-C') and M3 1-3 Step I line (F-F'). IFM fibers of M1 2-3 Step I line do not display a mutant phenotype and look wild type (D-D'). Thinner but unruptured fibers are detected in M4 1-2 Step I line hemithoraxes (G-G'). White arrows indicate stretched fibers, presumably due to hypercontraction. Scale bars are 100 μ m.

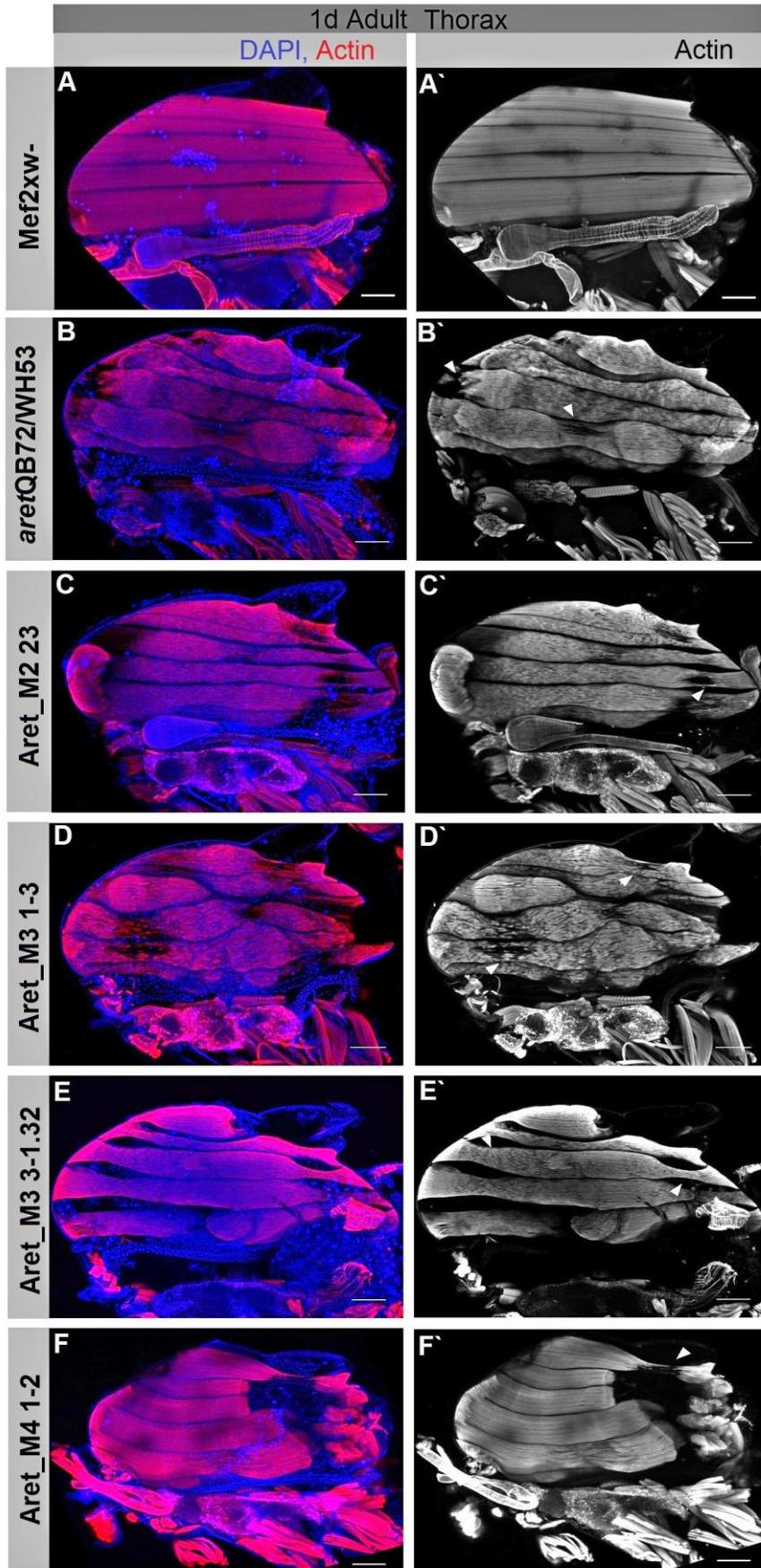


Figure 7. IFM fiber phenotypes in 1 day adult hemithoraxes

Hemithoraxes from wild-type (A-A'), *aret* mutant controls (B-B') and Step I Deletion flies (C-F, C'-F'). Note the pulling and rupture of IFM fibers upon *aret* removal (B'-F', white arrow heads). Scale bars are 100 μ m.

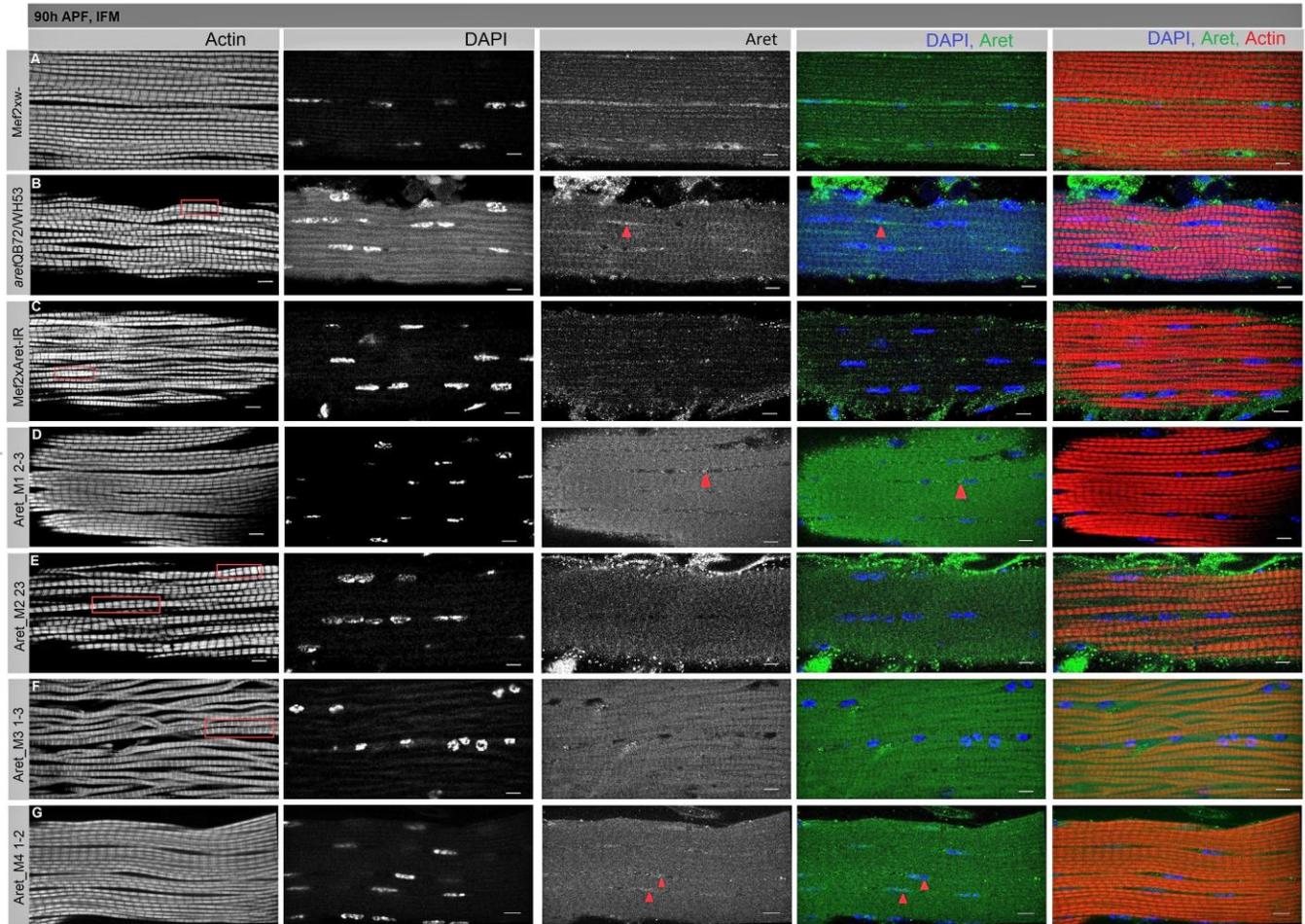


Figure 8. Detailed IFM myofibril phenotypes at 90h APF.

Wild-type (A), *aret* mutant controls (B,C) and Step I Deletion fly lines (D-G) are illustrated. Myofibrillar fusion is detected in flies missing Aret expression (B-C, E-G, red boxes). "Mini-tubes" formation appears more severe in the M3 1-3 Step I line (F). Aret staining could be detected in M1 2-3 and M4 1-2 Step I flies (D-G, red arrow heads). Scale bars are 5 μ m.

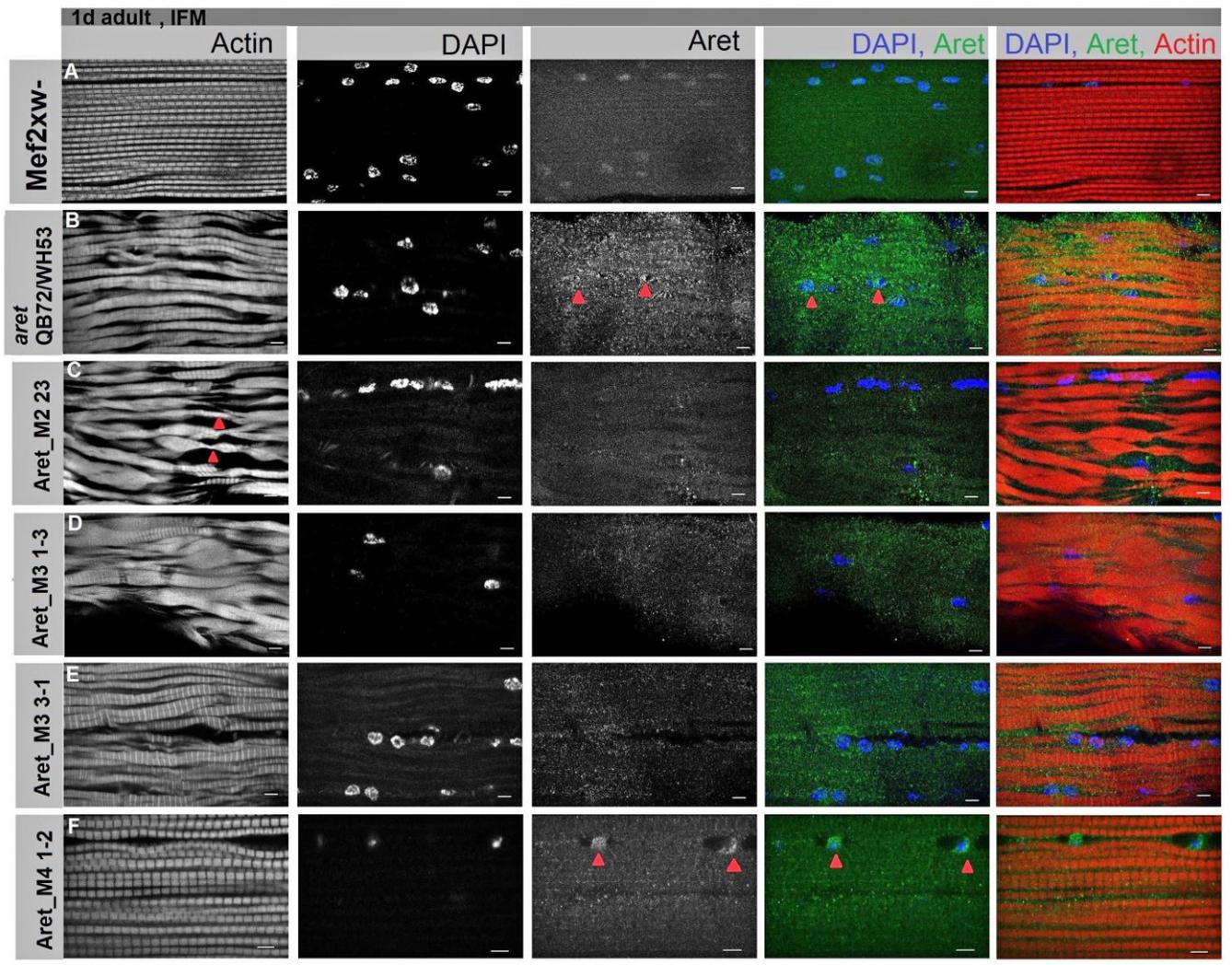


Figure 9. IFM myofibril detailed phenotypes in 1-day adults.

Wild-type (A), mutant control (B) and Step I Deletion fly lines (C-F). Complete degeneration of IFM myofibrils is observed in *aret^{QB72}/aret^{WH53}* mutants (B), M2 23 (C), M3 1-3 (D) and M3 3-1 (E) Step I lines. Positive Aret signal is detected in the nuclei of *aret^{QB72}/aret^{WH53}* mutants and the M4 1-2 Step I line (B, F, red arrow heads). Note that sarcomeres of the M4 1-2 appear considerably shorter than the other mutants (F). Scale bars are 5 μ m.

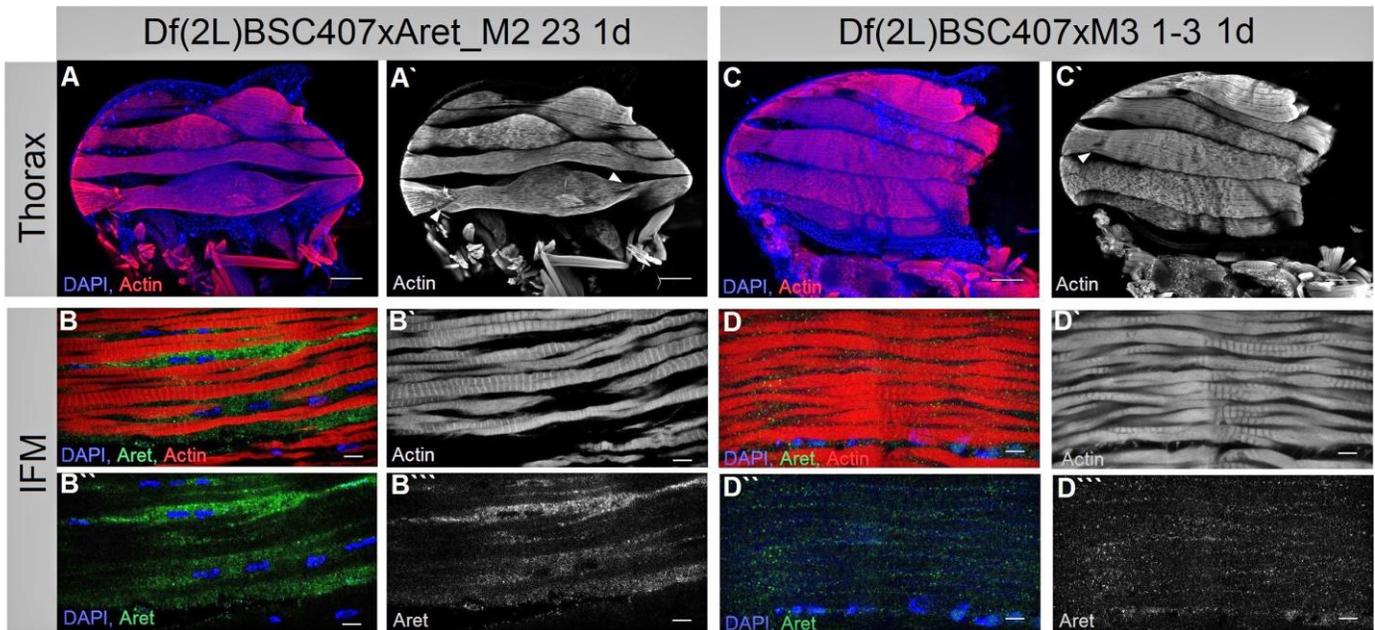


Figure 10. Fiber and myofibril phenotypes of transheterozygous Aret mutant flies in 1d adults.

Hemithoraces of 1-day adult Aret_M2 23/Df(2L)BSC407 (A-A') and Aret_M3 1-3/Df(2L)BSC407 (C-C') flies and their IFM myofibrils (B-B'', D-D''). IFM fibers rupture close to attachment sites (A', C', white arrow heads). Myofibrils show fusion and loss of sarcomeres (B-B', D-D'). No Aret staining is detected (B''-B''', D''-D'''). Scale bars are 5 μ m.

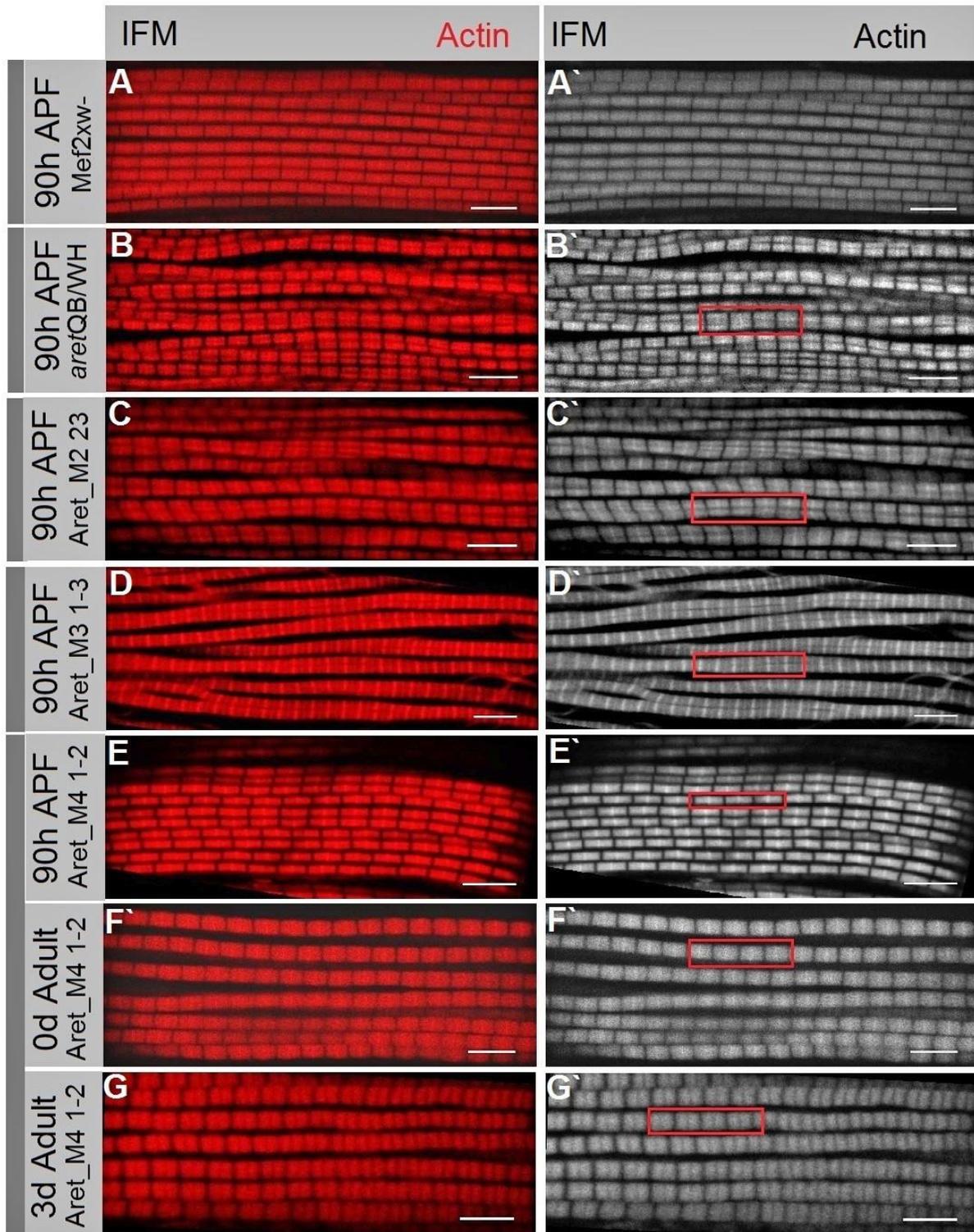
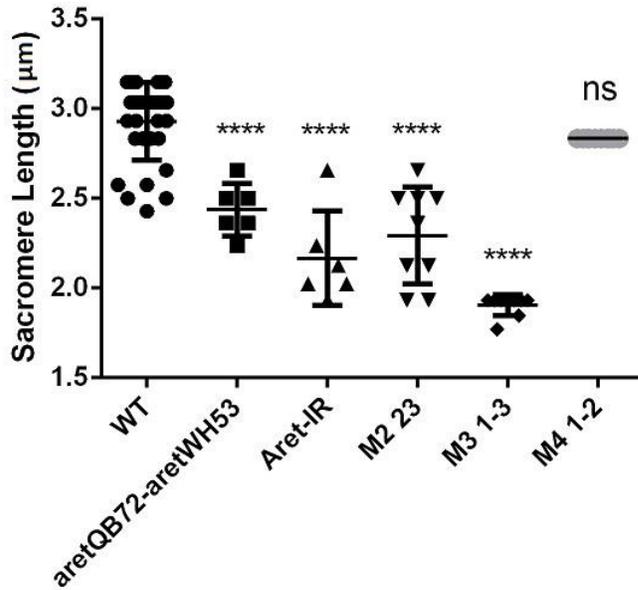


Figure 11. Comparison of sarcomeric structure in Step I deletion flies.

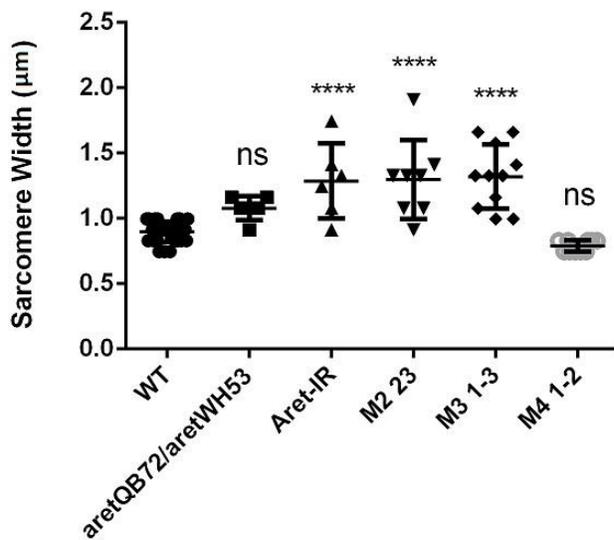
Wild-type control (A-A'), *aret*-IR control (B-B') and M2-M4 Step I lines (C-E') IFMs at 90h APF. M4 1-2 line represented at three developmental points: at 90h APF, in newly-eclosed adults (F-F') and in 3 day old adults (G-G'). Scale bars are 5 μ m.

Sarcomeres length at 90h APF



	Control (w-)	AretQB/AretWH	Aret-IR	Aret M2	Aret M3	Aret M4
# crops	32	6	6	9	9	10
N (flies)	10	2	2	2	2	2
Mean Sarc Length	2.93	2.44	2.17	2.29	1.90	2.83
Std. Deviation	0.22	0.15	0.26	0.27	0.06	0.00

Sarcomeres width at 90h APF



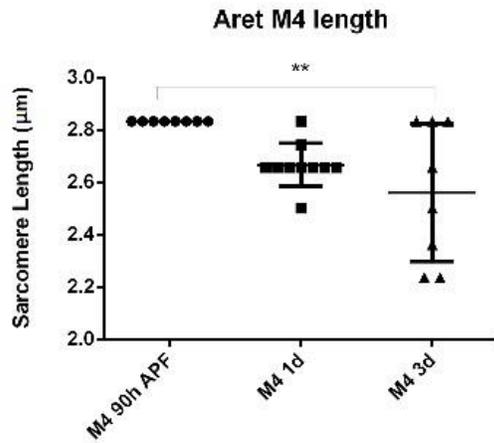
	Control (w-)	AretQB/AretWH	Aret-IR	Aret M2	Aret M3	Aret M4
# crops	30	6	6	8	11	10
N (flies)	10	2	2	2	2	2
Mean Fibril Width	0.90	1.08	1.29	1.30	1.32	0.79
Std. Deviation	0.08	0.09	0.29	0.3	0.25	0.04

Table 4. Estimation of sarcomere width in Step I mutants and control lines at 90h APF.

Step I deletion mutants and aret mutant controls were compared to WT sample for identification of difference among the means by ANOVA test. The difference in sarcomeric width between WT and Aret-IR, M2 23 and M3 1-3 is statistically significant (as indicated by asterisks). Difference among the means of WT, *aret*^{QB72}/*aret*^{WH53} and M4 1-2 is not significant.

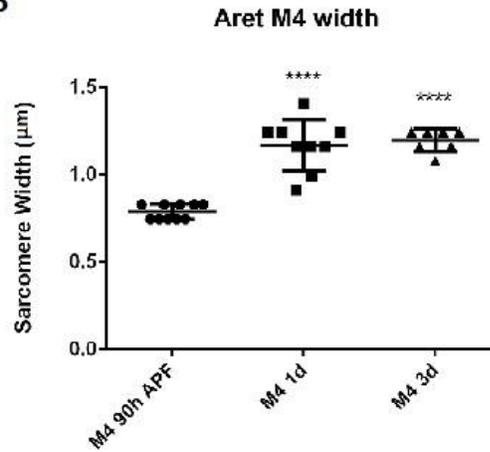
Table below indicates the number of flies and crops measured by Giovanni stript in Fiji.

A



	Aret M4 90h APF	Aret M4 1d	Aret M4 3d
# crops	10	10	8
N (flies)	2	2	2
Mean Fibril Length	2.83	2.67	2.56
Std. Deviation	0	0.08	0.26

B



	Aret M4 90h APF	Aret M4 1d	Aret M4 3d
# crops	10	9	7
N (flies)	2	2	2
Mean Fibril Width	0.79	1.17	1.20
Std. Deviation	0.04	0.15	0.07

Table 5. Estimation of sarcomere length and width in Step I M4 mutants at different developmental stages.

- A.** Sarcomere length was estimated at 90h APF, 1d and 3d adult flies. The graphic indicates that the difference between 90h APF and 3d APF is significant based on ANOVA predictions.
- B.** Sacromere width was estimated at 90h APF, 1d and 3d adult flies. The difference between the 90h APF and 1d adults and between 90h APF and 3d adults is highly significant.

Part 2. Positioning of the epitope tag on Aret to build RMCE-ready Step II constructs.

The second aim of the study was to clone an epitope tag on the protein in a donor vector to allow RMCE-tagging of the protein at the endogenous locus. This strategy, if successful, can open many new possibilities for the investigation of Aret function and localization. It remains to be determined whether Aret controls all fiber-specific splicing events directly or influences some events through alternatively spliced SF proxies, and the direct binding sequences of Aret for splicing are unknown. Biochemical approaches, such as IP-mass spec, can potentially give insight into the interaction partners of Aret while RIP-Seq can give insight into which sequences are bound by Aret to regulate splicing. These approaches would be greatly facilitated by an endogenously epitope-tagged protein.

The *Aret* locus codes for many protein isoforms that can be separated into 3 larger groups: the long (G/B/K/J/H/I), the mid-length (A/ E/ F) and the short (D) isoforms (Fig. 18). We carefully considered where to place the epitope tag to minimize any affect on protein functionality *in vivo*, as well as to universally tag *aret* isoforms. We decided on a C-terminal tag, directly after the 3rd RRM followed by the stop codon and the native 3'UTR. Step II constructs with both HA and V5 tags were designed, as well as a wild-type control construct without a tag. To generate an endogenously tagged fly line, we chose the Step I M3 line as it was not only universal for all *aret* isoforms (Fig. 13A), but also deleted a considerably shorter region as compared to the M1 and M2 deletions (Fig. 2), making the construct design easier. Considering that Aret, like most of the CELF protein members, forms a complex structural fold to interact with targeted RNA sequences [34], we decided to fuse the tag via a linker domain that may provide enough flexibility to not disrupt the folding of RRM3. A schematic representation of Step I and Step II constructs integrated into the fly genome of the M3 mutant is presented in Figure 13.

The region deleted by insertion of the Step I M3 construct contains part of RRM3 followed by the stop codon and the 3'UTR. Interestingly, based on mRNA-Seq data, the 3'UTR also contains spliced introns (M. Spletter and F. Schnorrer, unpublished data). To retain this complex, native structure, we decided to use a standard 3-piece ligation approach for the Step II construct design. We split the Aret region into two pieces (a “Left-Exons” and “Right-Exons” region) with the tag inserted between and ligated this into the Step II RMCE vector.

The first Aret piece contains the “Left-Exons” region (including exons 18, 19, 20 and 21 which cover part of RRM3) connected to the tag via the linker domain (Fig. 14A). The genome region containing the exons was PCR amplified from the Act5C>>Cas9, lig⁴ flies used to generate the Step I insertion, to minimize SNPs and genomic variation. The linker-tag sequence was ordered as an oligo from Eurofins Genomics. The two parts were connected via overlap PCR. Using specially designed primers, overhangs containing BsaI restriction enzyme (RE) sites were added to the 5` and 3` ends using extension PCR. This strategy allowed us to generate a backbone compatible end on the 5`-end and a 4 bp overhang on 3`-end that matched the “Right-Exons” region.

The second piece, the “Right-Exons” region, contains exons 22-24 and is relatively long (3160 bp). Its assembly required several steps, as shown in Figure 15. We tested several strategies to

clone the region as it was not possible to PCR amplify the genomic region in its entirety. Here I summarize the final strategy that generated the desired construct. We were able to amplify the “Right-Exons” region without matching overhangs from w- flies with primers 314 & 315 (Fig. 14 B). Next, we created 4 bp TAGA overhang using extension PCR (primers: 299 & 320), allowing complementarity with the 3'-overhang on the “Left-Exons” region. The amplified region was subcloned into pJet, taking the advantage of available NotI and BstBI restriction sites (Fig. 15). Finally, a second overhang containing BsmBI RE site was added on the 3'-end allowing ligation to the backbone through a HindIII site (Fig. 14B, primers 299 & 317).

Overall, three Step II constructs containing an HA or V5 tag or the wild-type (untagged) Aret sequence were generated. It was of particular interest to recombine the WT region back into Step I flies as a control to verify that the attR “scar” sites remaining in the genome sequence in the modified fly did not affect transcription or splicing *in vivo*. In the cloning strategy, both Type II and Type IIS restriction enzymes were used, providing the necessary directionality. Prior to injection into the Step I M3 line, the completed constructs were confirmed by restriction enzyme digest (Fig. 16A) and sequencing (Fig. 16B).

Due to the second-site lethal mutation in M3 3-1, injections for Step II constructs were initially performed in the M3 1-3 line. Thus, we need to take into consideration that M3 1-3 is ends-in positive, meaning that in addition to the DsRed cassette, the whole vector backbone was inserted into the fly genome (Fig. 17B). In this case, recombination with a Step II construct could lead to three possible exchange variants, including the optimal outcome: generation of a DsRed negative, ends-in free fly line. As can be seen from the injection table (Fig. 17A), we obtained only 2 DsRed negative founder lines for the V5-tag construct. Unluckily, as was confirmed by verification PCR, the combination of Exchange 1 and 2 took place in those lines (Fig. 17C), leading to retention of the backbone sequence.

To increase the probability of successful recombination, we wanted to inject into a line that was not ends-in. Therefore, we recombined the lethal mutation away from line M3 3-1. The Step II constructs were subsequently injected into this clean M3 3-1 line. As this thesis was being submitted, we obtained 6 founder lines for the HA-tagged construct and 2 founder lines for the V5-tagged construct. Unfortunately, we did not obtain founders for the WT control construct. This construct will be reinjected, and all lines remain to be verified and phenotypically characterized.

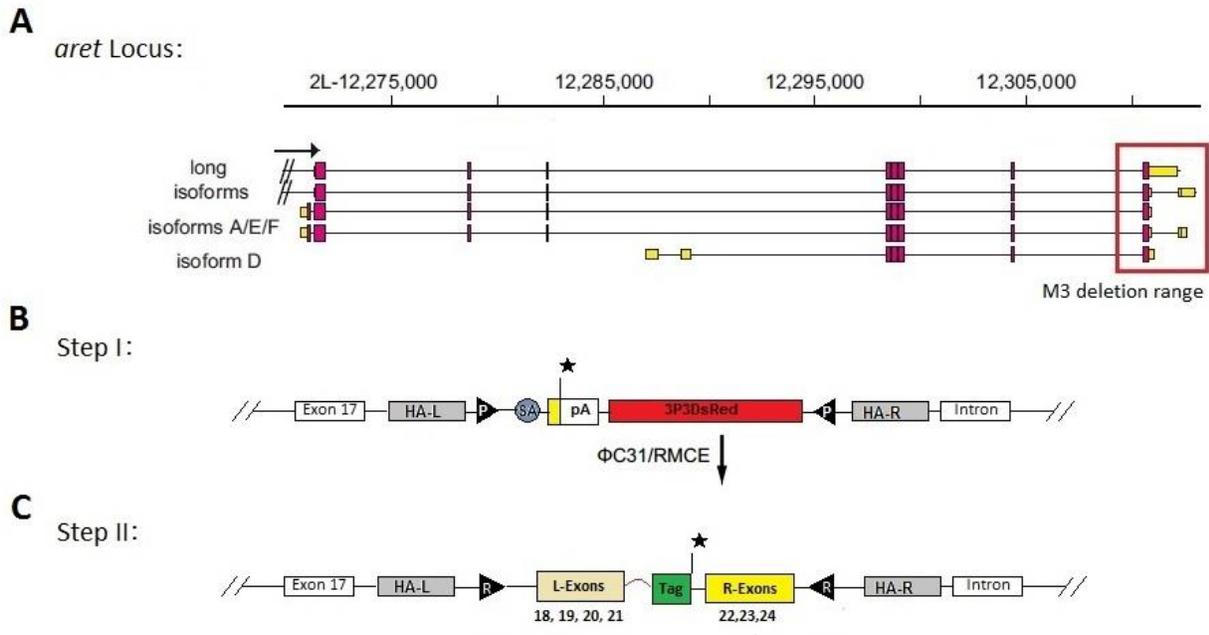


Figure 13. Localization of the epitope tag on the C-terminal end of the Arrest protein using the Step I Deletion 3 mutant.

A. Schematic representation of *aret* isoforms. The region deleted by M3 construct insertion is boxed in red. [modified from Spletter *et.al.*, 2014.]

B. Localization of the Step I DsRed cassette in the fly genome of M3 mutants.

C. Localization of the Step II cassette containing the tag in the fly genome after recombination.

Note: the black asterisk on a sequence indicates the stop codon.

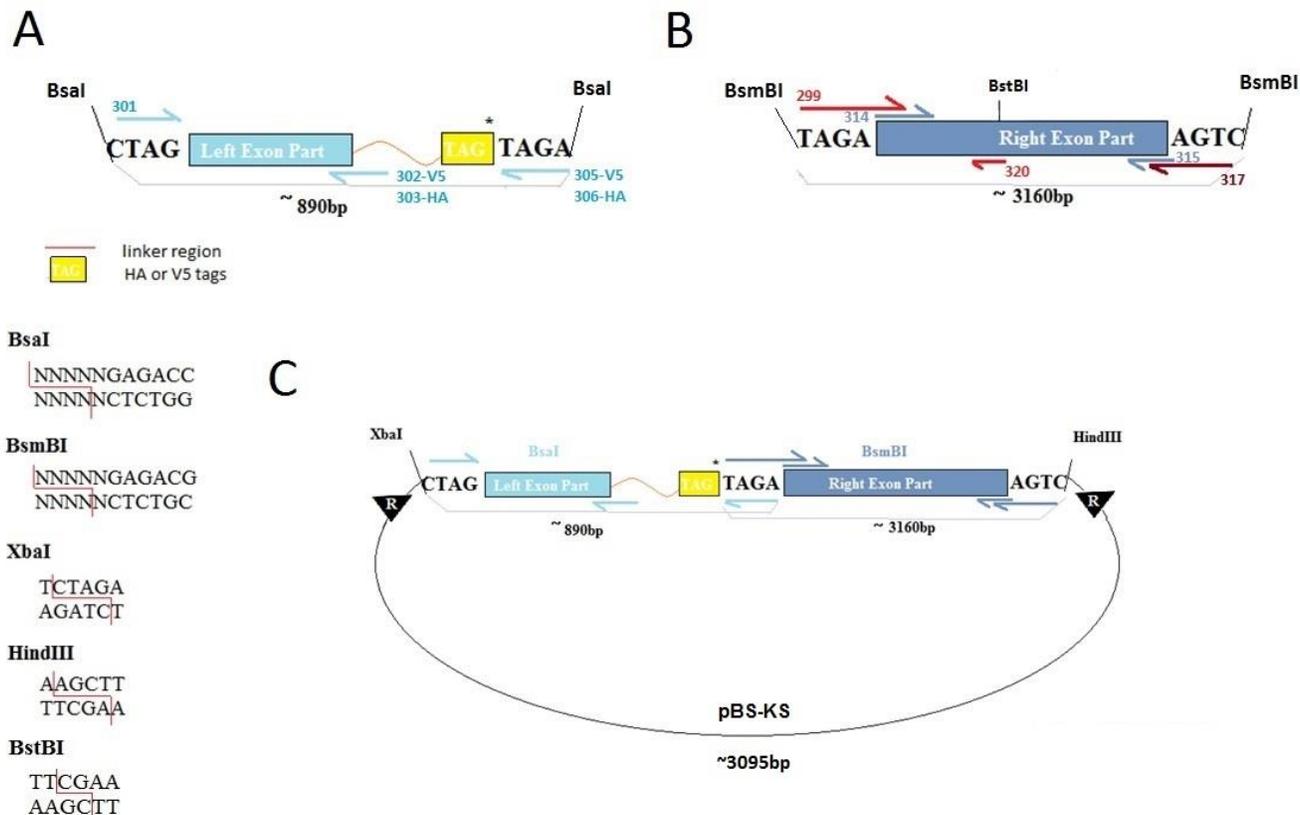


Figure 14. A canonical 3-pieceligation cloning strategy for epitope tag positioning.

- A.** “Left Exons” piece connected via a flexible linker to the tag constitutes the first insert. BsaI RE sites were generated on both ends to provide overhangs compatible with backbone on the 5'-end and with the second insert on the 3'-end.
- B.** “Right exons” piece was generated by several steps. Primers 299 and 317 created overhangs containing BsmBI RE sites, allowing ligation to the “Left Exons” insert on the 5'-end and the backbone on the 3'-end.
- C.** The standard pBS-KS vector was used as a backbone. The ligation reaction was performed at a 1:3:3 ratio.

The cutting sites for all restriction enzymes used are indicated. Primers used in this approach are listed in Supplement’s Table 2.

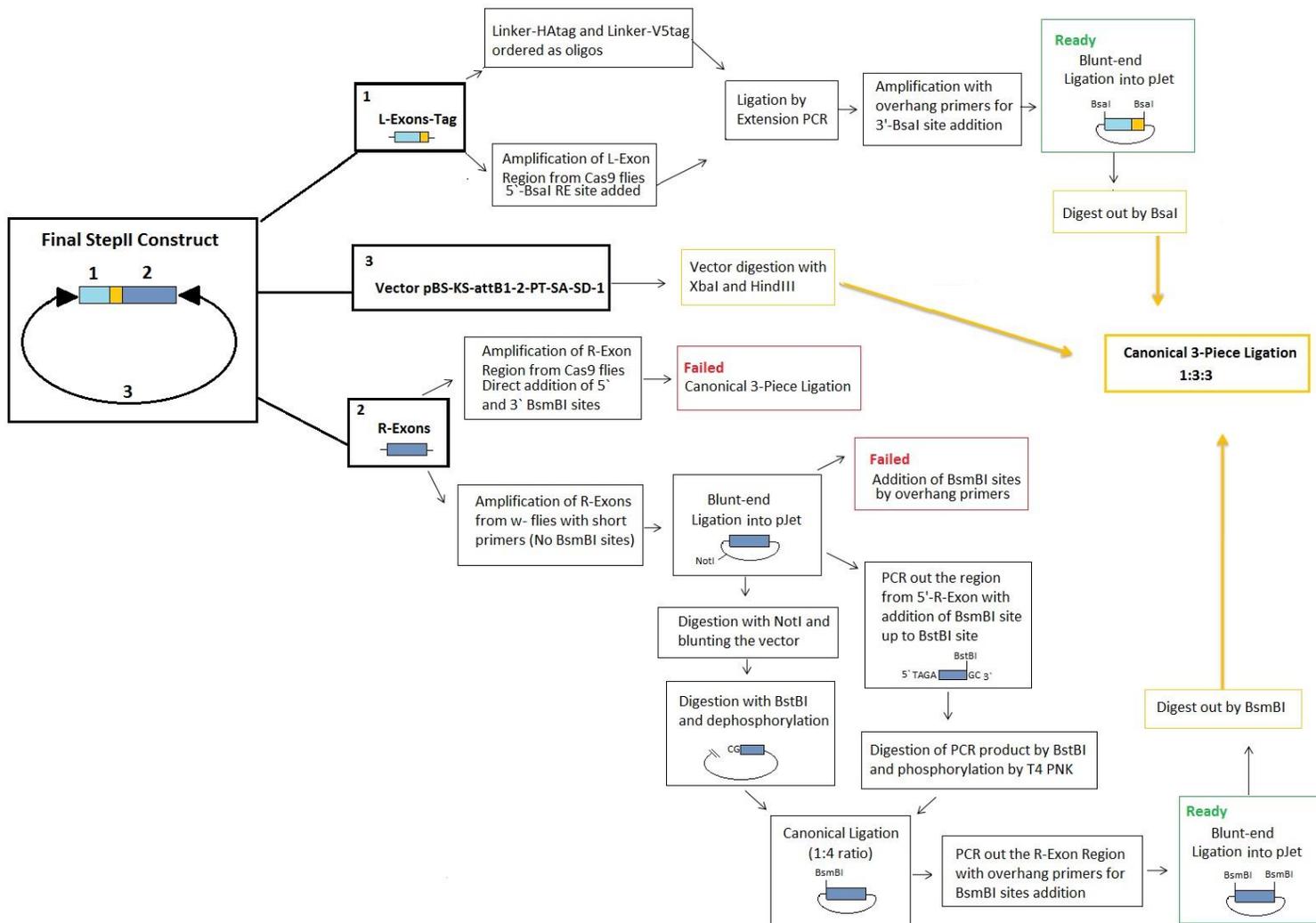
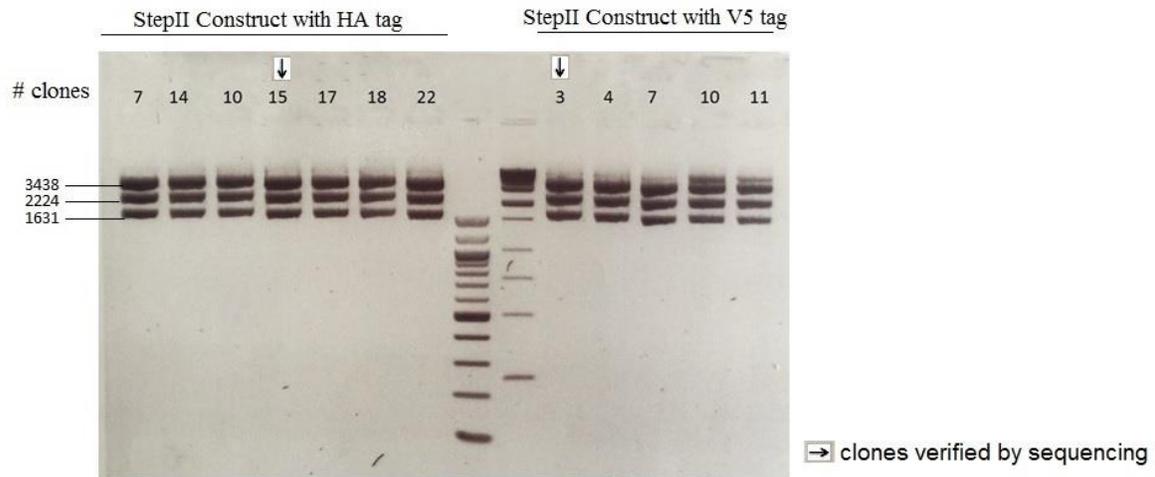
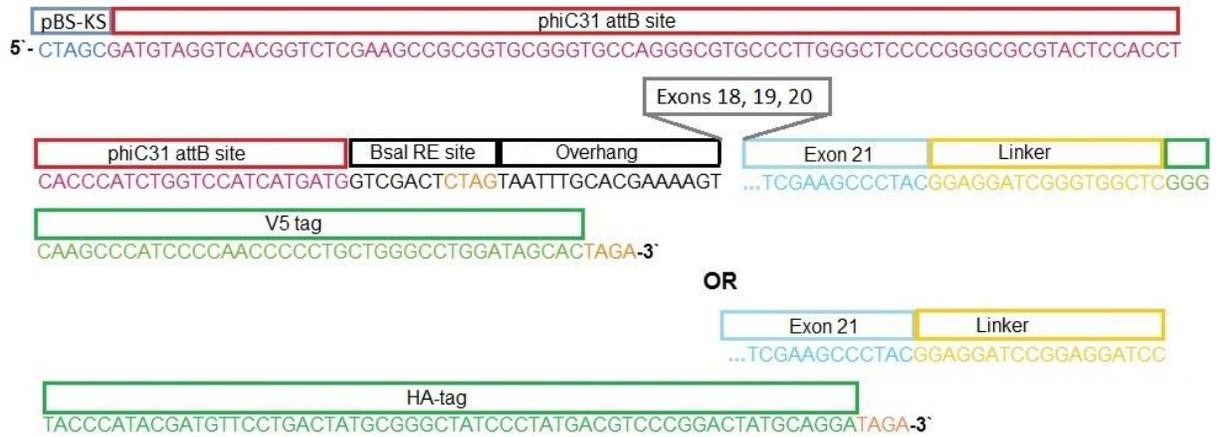


Figure 15. Flow-diagram of troubleshooting during Step II construct design.

The final Step II constructs were obtained by several steps. Intermediate pieces were cloned into pJet. Several approaches did not result in desired products, requiring alternative strategies to obtain the desired product.

A**B****Figure 16. Confirmation of Step II construct integrity.**

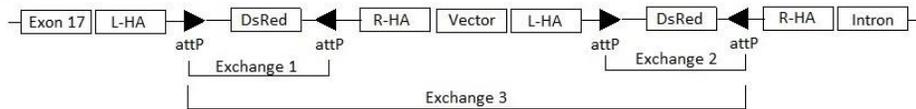
- A. Restriction enzyme digest of HA and V5 final Step II constructs shows the expected 3-band pattern. Arrows indicate clones verified by sequencing.
- B. Schematic representation of Left-Exon region with V5 or HA tags. Sequence integrity was confirmed by PCR and sequencing.

A

Injected Fly Line	Construct Injected	# G0 Larvae	#G0 adults	# Founder Lines (transgenic)
yw;lig4 ^{-/-} ;act5C-cas9	Step I M1	229	114	1
	Step I M5	201	75	1
M3 1-3	Step II HA-tag	137	81	0
	Step II V5 tag	328	77	2
	Step II wt	235	98	0
M3 3-1 (11&32)	Step II HA-tag	148	64	6
	Step II V5 tag	216	103	2
	Step II wt	104	47	0

B

Step I



C

Step II

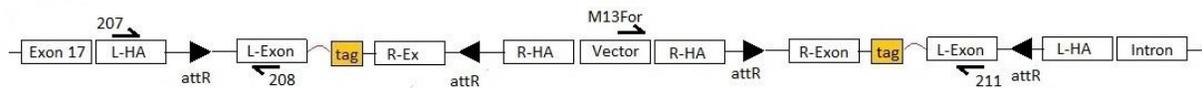


Figure 17. Injection list and schematic representation of ends-in line recombination with Step II construct.

- A.** Step I constructs with appropriate sgRNAs were co-injected into Act5C>>Cas9, lig⁴ flies. Step II constructs and Vasa plasmids, carrying ϕ C31 integrase, were coinjected to the M3 1-3 and M3 3-1 fly lines at the concentration of 200ng/ μ l.
- B.** The integration of the Step 1 Deletion 3 cassette into the M3 1-3 line was ends-in and duplicated the attP sites. There are three possible alternative recombination events occurring after the introduction of an attB carrying cassette and integrase.
- C.** Schematic representation of the M3 1-3 genome after Step II induction. Marked primers were used for region verification. The most likely possibility, a combination of exchange 1 and 2 with inversion of the duplicated cassette took place.

Part 3. Investigation of Aret protein isoforms and mechanisms of function

Aret is only detectable in two distinct fly organs: the gonad and the IFMs [17], [19]. Aret has a defined function in the cytoplasm of ovaries [23] as a translational repressor and in the nuclei of 24h APF fibrillar muscles [18] as a splice factor. There are several Arrest splice isoforms of different molecular weights (Fig. 18) that differ in their enhancer/promoter sequences and amino-terminal region, but share a common C-terminal region. mRNA-Seq data indicates that in contrast to ovary and testis, which are reported to express specifically an isoform in the mid-length range at 68 kDa [17], [19], IFM most highly expresses a short isoform with only two RRM3 predicted to be 37.7 kDa, as well as long and mid-length isoforms at lower levels (M. Spletter and F. Schnorrer, unpublished data). The N-terminal region has been reported to mediate Aret dimerization and interactions with binding partners [28], [29], [35], so different isoforms may bind different partners or have different regulatory properties. Aret does not have an obvious NLS sequence, and it is attractive to hypothesize that different Aret isoforms may have different functions and be expressed in different tissues.

Here we performed Western Blotting to verify which Aret protein isoforms are expressed in IFM at the protein level. Additionally, we wanted to further understand how loss of Aret leads to defects in sarcomere growth and hypercontraction. We specifically wanted to examine if the localization of sarcomere proteins is disrupted in Aret mutants. Informative phenotypic experiments can be repeated in the future with isoform specific Aret alleles.

3.1 Analysis of wild-type Aret isoforms present in IFM, ovary and testis by Western Blotting.

Aiming to identify the IFM specific Aret isoform(s), we conducted a Western Blot analysis using available rat and rabbit anti-Bruno antibodies (provided by Anne Ephrussi). To determine the specificity of the antibodies, we compared wild-type whole fly samples from male or female to dissected ovary and testis or whole fly samples from the M2 23, M3 1-3 or *aret*^{QB72}/*aret*^{WH53} mutant backgrounds. Unfortunately, we found that the existing polyclonal antibodies are quite dirty. The two antibodies are generated against slightly different regions of Aret, the rat against RRM3 and the rabbit against the entire ovarian isoform, and the specific epitopes recognized are unknown. The most informative comparison is to the M2 deletion mutant as a negative control, because it lacks almost 14 kb of the *aret* sequence, covering a large region of the C-terminus. We were able to identify multiple non-specific bands, identified by the black boxes in Figure 19. Despite the background, we were also able to identify several bands matching predicted Aret isoforms.

We next wanted to detect which Aret isoforms are present in specific tissues. For that purpose, we used whole fly extracts (male or female) as well as dissected IFM, leg, ovary and testis. We were able to distinguish the testis isoform running approximately in the 88 kDa range and two ovary isoforms, one running slightly higher than the testis isoform at ~ 90 kDa and the other previously reported ovary isoform at ~ 68 kDa [19] (Fig. 19 A).

Our IFM samples ran poorly, and thus we cannot draw any conclusions about which Aret isoforms are IFM specific. Most of the Westerns were negative (see also Supplement Figure 1)

with only one blot (Fig 19B, indicated in red) revealing potential bands that were not observed in the mutant samples. However, the band sizes do not coincide with those predicted for Aret isoforms. We can concretely say that the ovary or testis isoforms were not detected in the IFM sample, indicating that germline isoforms are tissue-specific (Fig. 19, S. Fig. 1). This experiment needs to be repeated to detect and identify Aret isoforms in IFM.

In addition to using mutant samples as negative controls, we also wanted to determine if various mutations result in the complete loss of Aret protein or the generation of stable truncated Aret isoforms. We were able to identify two potential Aret bands for *aret*^{QB72}/*aret*^{WH53} whole fly samples, indicating the possible presence of truncated protein. This is also supported by some signal in the nuclei of Aret mutants observed in immunostaining of hemithoraxes (Fig. 9B). Overall, because of the lack of specificity and the preliminary status of our Western Blot results, we cannot yet reach any solid conclusions about Aret protein in the mutant alleles. In the future, Western Blotting should be repeated with dissected IFM, ovary and testis from each genotype of interest. Additionally, endogenous Aret protein fused to an epitope tag should significantly improve the specificity and signal-to-noise detection of Aret isoforms.

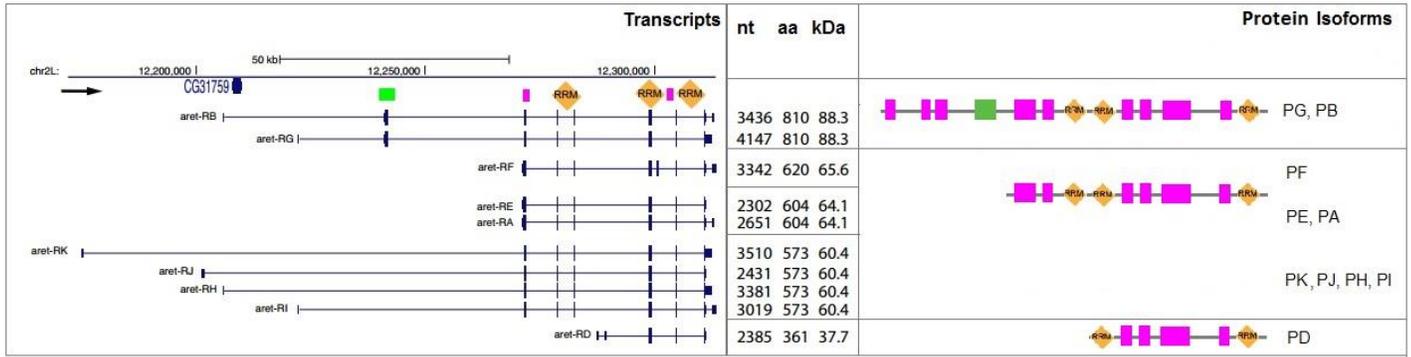


Figure 18. Schematic of Aret transcripts and corresponding protein isoforms in *Drosophila*.

Diagram of the Aret locus. Associated transcript isoforms generated by alternative promoters and alternative splicing are shown. Localization of protein domains to specific exons indicated by the green, pink and orange boxes. We refer to the isoforms by length as long (B/G/K/J/H/I), mid-length (F/E/A) and short (D) isoforms. Transcript length and predicted protein molecular weight as indicated.

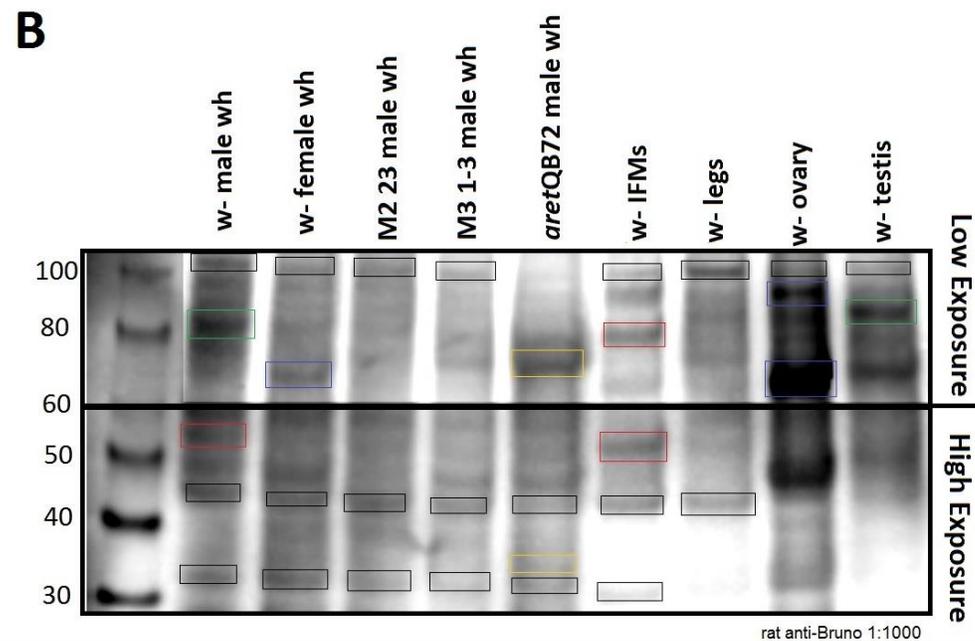
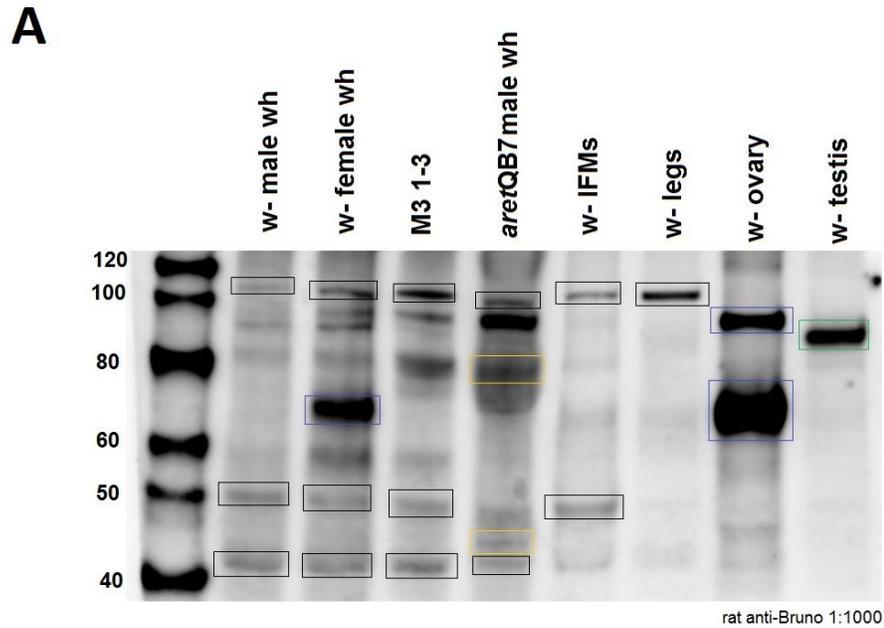


Figure 19. Investigation of Aret isoforms by Western Blotting.

A –B. Stainings of Aret isoforms in wild-type and mutant samples with rat anti-Bruno antibody (1:1000). Sample labels indicated above the blots. Both blots compare dissected tissue-specific samples to whole-fly (indicated as wh) and homozygous mutant controls. Ovary and testis specific isoforms are clearly visible and absent in mutant samples. The blot in (B) is shown in two separate panels, with the higher molecular weight region at a low exposure and the lower molecular weight region at a high exposure, as indicated. Background bands are marked by black boxes, prominent female isoforms by blue, male isoforms by green, potential truncated isoforms by yellow and potential IFM-specific isoforms by red.

3.2 The absence of Aret leads to structural rearrangement of essential sarcomere components.

Fibrillar and tubular muscle have distinct patterns of alternatively spliced proteins, likely contributing to the differences in their physiological properties. Among the structural genes that are reported to be alternatively spliced are Myofilin (Mf), Troponin-T (*upheld*), Tropomyosin 1 (Tm1), Myosin alkali light chain (Mlc1), Myosin heavy chain (Mhc) and Projectin (*bent*, *bt*) [36], [37]. Projectin is one of the largest *Drosophila* proteins. It possesses a proline-glutamic acid-valine-lysine (PEVK)-rich domain suggested to form a flexible spring-like structure. The IFM splice variant contains a short PEVK domain only a few amino acids long that provides the high passive resting stiffness required for stretch-activation of the IFMs [7], [38]–[40]. Myosin heavy chain (Mhc), the motor protein that produces contractile force in muscle, is a second example of a differentially spliced transcript that produces proteins shown to have distinct physiological force profiles [41], [42].

Based on our hypothesis that myofibrils in *aret* mutants form mini-tubes, we wanted to test if any sarcomeric proteins are mislocalized. In particular, we wanted to see how the distributions of Mhc tagged by weeP24-GFP and bt-HA, which are differentially localized in tubular and fibrillar muscle, might change in the *Aret* mutant background. As RNAi gives the same null phenotype as M2 and M3 and to easily test protein localization before building complicated mutant stocks, we crossed flies carrying HA-bt and Mhc-weeP-GFP with *aret-IR* and *salm-IR* flies and induced knockdown at two different developmental time points using the Mef2-GAL4 and Act88F-GAL4 drivers. Mef2-Gal4 is expressed from the myoblast stage, resulting in an early knockdown, while Act88F-Gal4 is first expressed around 20h APF, resulting in later knockdown after IFM fate determination.

In wild-type IFM, projectin normally localizes to the I-Z-I region, associated with the forming Z-band (Fig. 20A-A''), while weeP-GFP localizes in two dots on either side of the M-line (Fig. 21 A-A''). In the case of weeP-GFP, this localization pattern is due to an isoform switch in Mhc expression specifically in IFM where the early GFP-tagged Mhc isoform is replaced by an untagged Mhc isoform during sarcomere maturation [42]. According to previous studies, the absence of *salm* during early IFM myogenesis leads to complete tubular conversion of the muscle [16]; thus, we expected both projectin-HA and weeP-GFP to change their localization pattern from fibrillar to tubular in Mef2-Gal4 driven *salm-IR* flies (Fig. 20B-B'', Fig. 21B-B''). This is indeed what we observed.

We then compared the localization of bt-HA and weeP-GFP in Mef2-Gal4 *aret-IR* flies to the wild-type and the *salm-IR* patterns. We observed the same “mini-tube” phenotype we saw with previous stainings (Fig 21 vs Fig. 8), confirming that our *Aret* RNAi worked. Strikingly, we observed that the pattern of bt-HA and weeP-GFP changed in the “mini-tubes” (Fig. 20C-C'', Fig. 21C-C''). Although bt still colocalizes with Z-disc and Mhc on the A-band, we see that bt now forms two clearly distinguishable bands and weeP-GFP is no longer restricted to two dots. This pattern is more similar to the pattern observed in *salm-IR* or in WT leg (data not shown), supporting the mini-tubes hypothesis. Although *aret-IR* disrupts IFM specific splicing, mRNA-Seq data suggest that the Mhc isoform switch still occurs, suggesting the weeP-GFP pattern is at

least partially due to a change in how Mhc is incorporated into the sarcomere. The bt-HA allele tags all bt isoforms, regardless of splicing alterations.

We next analyzed bt-HA and weeP-GFP localization in Act88F-Gal4 driven RNAi flies. As the Act88F-Gal4 *salm*-IR phenotype has not been reported, we first characterized this phenotype. The phenotype of *salm*-IR driven by Act88F is considerably less severe than the one induced by the Mef2-GAL4 driver. Namely, the IFMs of 88F>>*salm*-IR flies did not convert into tubular muscle. Thus, it seems likely that initial induction of the Salm transcription factor from 0-24h APF appears sufficient to initiate the IFM fate. It should be noted that there may be perdurance of the Salm protein in 88F>>*salm*-IR flies, resulting in a hypomorphic phenotype. Regardless, it is clear there is a late function for *salm* during sarcomere maturation based on the phenotype (or hypomorphic phenotype) observed in this experiment. Sarcomeres were disorganized and thinner than wild-type, and blobs of actin accumulated at the Z-line (i.e. so-called Zebra-bodies) (Fig. 22B-B'', Fig. 23 B-B''). This suggests that Salm regulates gene expression contributing to sarcomere growth and maturation in addition to its early function regulating the IFM fate. Regarding projectin and Mhc distribution, we could not resolve the bt localization due to poor staining. However, the localization of weeP-GFP in Act88F>>*salm*-IR flies is identical to wt patterns seen in the w- controls, suggesting that late *salm* knockdown is insufficient to change the localization pattern of sarcomere components.

Finally, we examined localization of bt-HA and weeP-GFP in Act88F driven *aret*-IR. Similarly to Salm, initial Aret induction occurs in Act88F>>*aret*-IR flies, presumably allowing the sarcomeres to initiate normal growth. However, by 90h APF the phenotype of Act88F>>*aret*-IR flies is indistinguishable from Mef2-Gal4 driving RNAi or mutant lines. Sarcomeres are too short, the myofibrils appear to fuse forming “mini-tubes” and hypercontraction is observed in 1d adults. This suggests that the *aret* hairpin is very efficient, and that the main function of Aret occurs during later stages of sarcomere growth and maturation. It may also imply that Aret protein is turned-over reasonably quickly. In *aret*-IR flies, bt-HA appears in the form of a double band on the both sides of a Z-disc (Fig. 22C-C''), rather than a single band axis observed in the wild-type control (Fig. 20A-A''). Interestingly, the bt-HA pattern in wild-type leg muscles is also two stripes located on the border of the A band (Fig. 22D-D''). Similarly, weeP-GFP localization also changes in *aret*-IR flies (Fig. 23C-C'') showing a tubular-like pattern spanning the entire sarcomere instead of two localized dots flanking the M-line. Also, note that weeP-GFP is aligned in “mini-tube” regions, instead of displaying a non-laterally aligned wild-type IFM fibril pattern. These results need to be repeated in the homozygous mutant M2, M3 and M4 lines, but they may serve as an indication that the localization patterns of essential sarcomeric proteins partially switches from fibrillar to tubular in the absence of Aret. This is an exciting and novel finding, suggesting there may be a fundamental difference in the construction of fibrillar and tubular sarcomeres, in particular regarding the processes of lateral alignment and myofibril fusion.

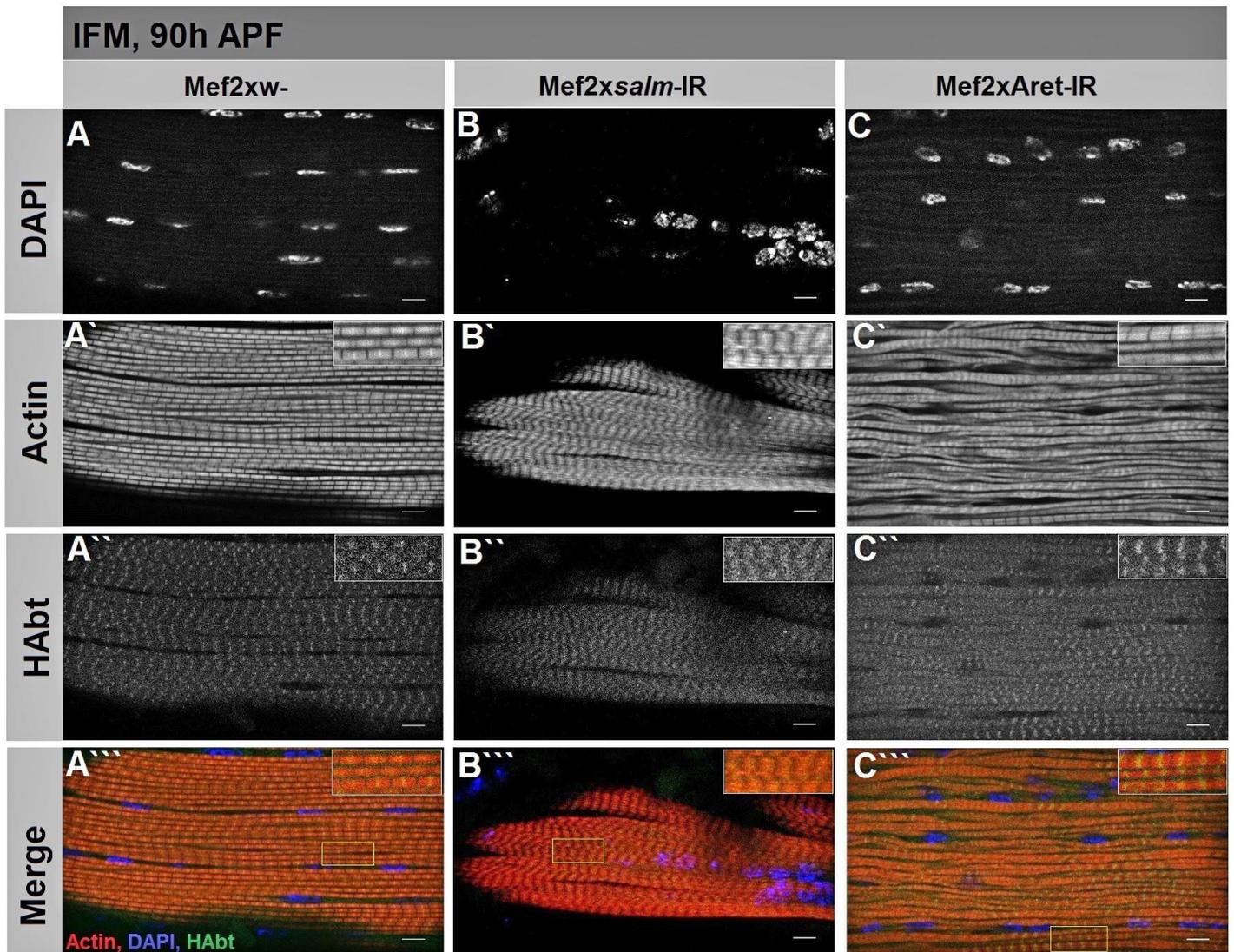


Figure 20. Projectin/bt-HA localization in Mef2-GAL4 driven RNAi lines at 90h APF.

Wild type (A-A'''), *salm*-IR (B-B''') and Aret-IR (C-C'''). In WT, projectin-HA is observed as a single band at the Z. In *salm*-IR, two bands localized adjacent to each Z-line are visible. Aret-IR shows two clear projectin-HA bands at each Z-line, suggesting a more-tubular localization. Scale bars are 5 μ m.

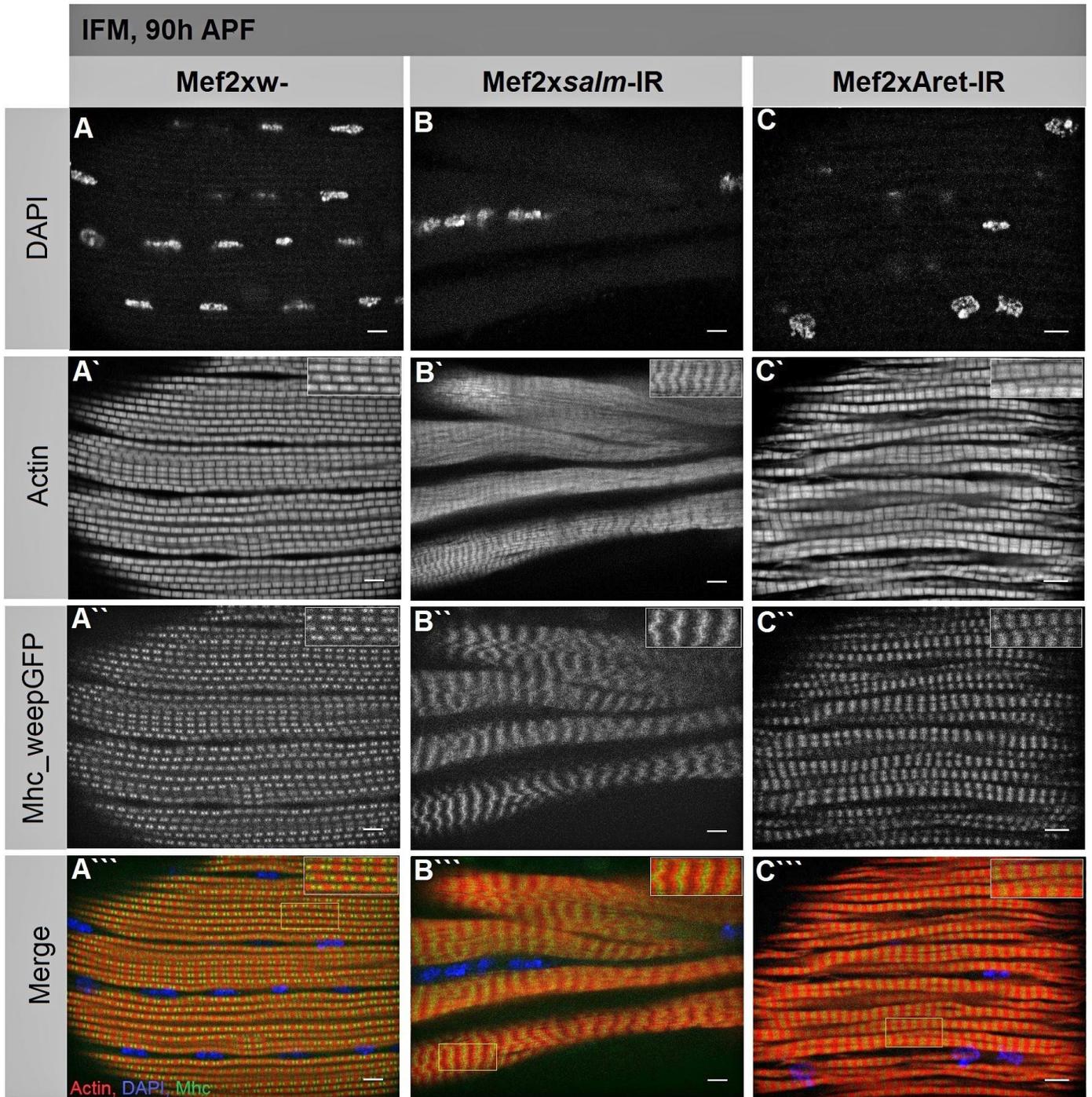


Figure 21. Mhc (weep-GFP) localization in Mef2-GAL4 driven RNAi lines at 90h APF.

Wild type (A-A'''), *salm*-IR (B-B''') and Aret-IR (C-C''') are present. In WT, weep-GFP is observed as two dots on either side of M-line. In *salm*-IR, Mhc is no longer restricted to two dots and is similar to localization in WT legs (data not shown). Aret-IR shows the “mini-tubes” distribution, similar to one observed in *salm*-IR. Scale bars are 5 μ m.

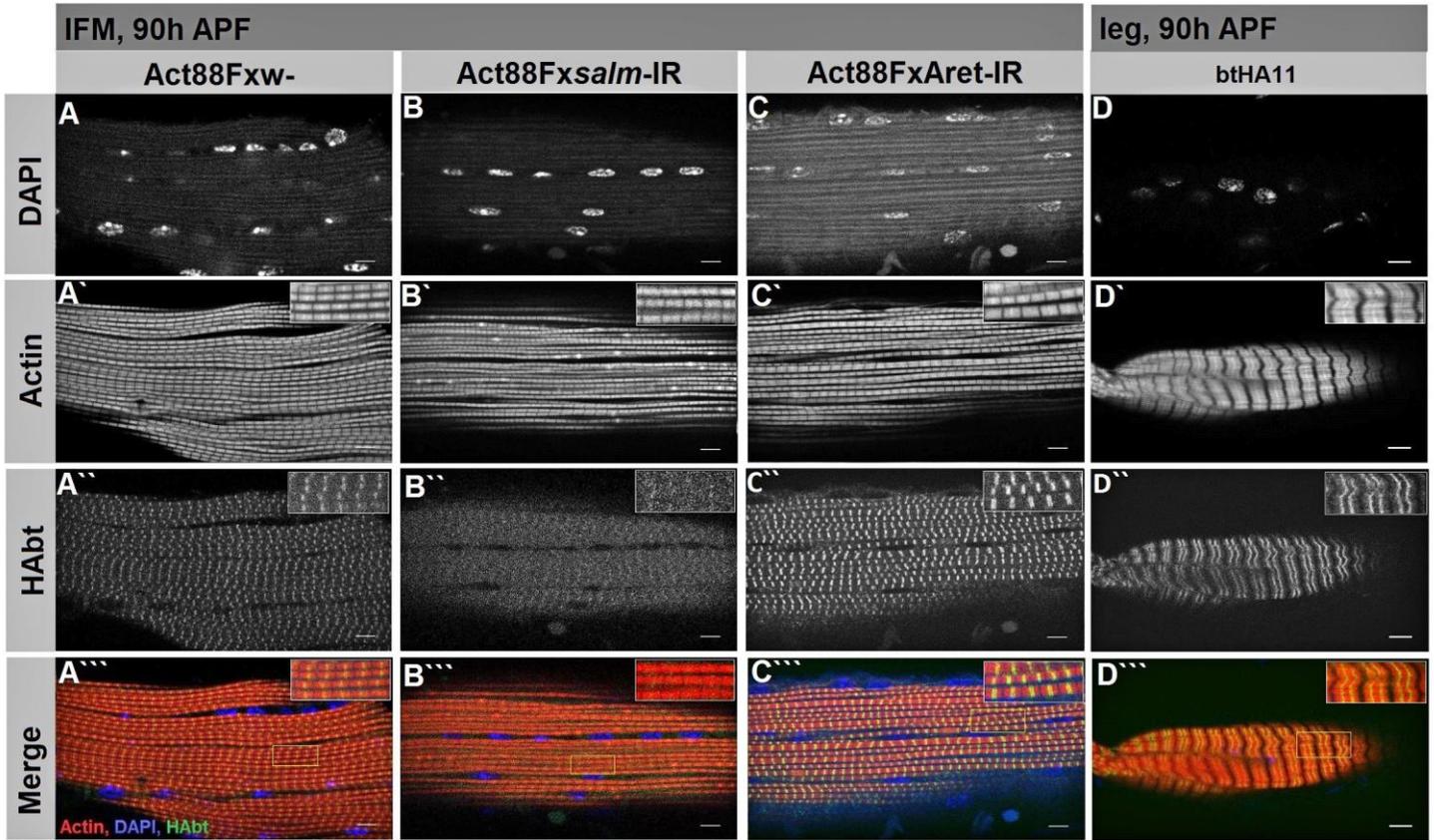


Figure 22. Projectin/bt-HA localization in Act88F-GAL4 driven RNAi lines at 90h APF.

Wild type (A-A'''), *salm*-IR (B-B'''), *Aret*-IR (C-C''') and tubular leg muscle (D-D'''). The btHA11 fly line expresses the HA::bt endogenously (D-D''', photo credits to Wouter Koolhaas and Chris Barz). Act88F>>*salm*-IR leads to a possible hypermorphic phenotype characterized by thin myofibrils and actin accumulations (Zebra bodies) at the Z-line, while localization of bt has not been changed. In *aret*-IR, bt appears in form of double band on both sides of Z-disc, that is similar to bt destitution in WT leg (D-D'''). Scale bars are 5 μ m.

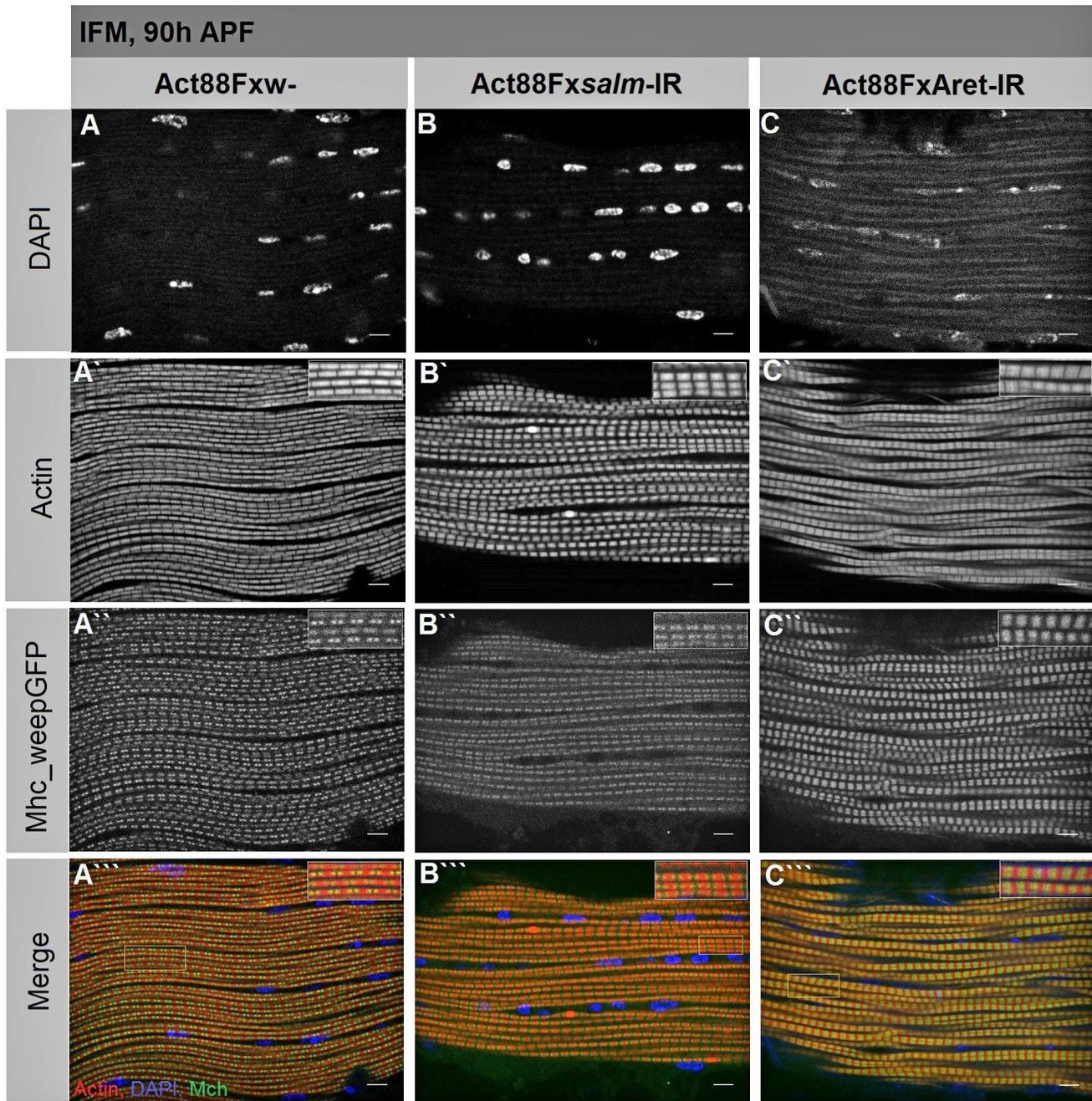


Figure 23. Mhc (weeP-GFP) localization in Act88F-GAL4 driven RNAi lines at 90 h APF.

Wild type (A-A'''), *salm*-IR (B-B''') and *Aret*-IR (C-C'''). Summary of localization patterns. Note that Act88F-Gal4>>*aret*-IR phenocopies Mef2-Gal4>>*aret*-IR and mutants, while Act88F>>*salm*-IR leads to a possible hypermorphic phenotype characterized by thin myofibrils and actin accumulations (Zebra bodies) at the Z-line. Scale bars are 5 μ m.

Discussion

Alternative splicing (AS) is an important process that likely contributes to the specification of many tissues, in particular muscle. AS leads to unique tissue-specific patterns of protein isoform expression, presumably with distinct biological functions. As such, it serves as an important mechanism to increase genetic complexity by generating multiple transcripts from a single gene. Muscle was one of the first tissues in which alternative splicing was observed and characterized. AS has been shown to play an important developmental role in vertebrate heart muscle, with MBNL family members driving embryonic splicing patterns and a shift to CELF family members driving mature splicing patterns in adult muscle [43], [44]. This transition in splicing is also important, as loss of function of either MBNL or CELF leads to heart defects [24]. AS also influences the specification of contractile protein genes in both flies and mammals, instructing the physiological and biophysical differences between fiber-types. In addition to the CELF (CUGBP and ETR3 like) and MBNL (muscleblind-like) families of regulatory RNA-binding proteins (RBP), members of the polypyrimidine tract binding proteins (PTB or PTBp1) [45], hnRNP H [46] and FOX [47] families have been shown to be involved in the regulation of muscle-specific AS events [48].

In the present study, we focused on the *Drosophila* CELF homolog Aret, which is essential for indirect flight muscles specific alternative splicing. Aret initially localizes in both the cytoplasm and nucleus of growing myofibers and is required for sarcomere growth happening after 48 h APF [18]. By 72h APF, Aret has translocated almost completely to the myofiber nuclei, where it contributes to myofibril maturation and the proper regulation of myosin function [17], [18]. In this thesis, I first confirmed several new deletion alleles in Aret, including lines M2, M3 and M4. I show that these alleles phenocopy existing Aret mutants and Mef2-Gal4 driven RNAi. I then introduce an RMCE-based strategy to epitope-tag the endogenous Aret protein. I constructed the RMCE-based flies, introducing both an HA and a V5 epitope tag C-terminally on the Aret protein. Finally, I perform a more detailed analysis of Aret function and phenotype, including Western Blotting and sarcomere structural marker characterization. In addition to contributing new tools to study the mechanism of Aret function, my data suggests several interesting directions for further study of Aret isoform function, Aret protein domain function and Aret function during muscle development.

The main goal of this master's thesis was to design and implement an RMCE-based strategy to epitope-tag the endogenous Aret protein. This will facilitate both studies of endogenous protein localization by immunofluorescence as well as multiple biochemical approaches. The identification of Aret IFM-specific isoforms, interacting partners and binding sites on target RNAs are all feasible with epitope-tagged protein. In addition to the highly specific identification of Aret on Western Blot, the tag will enable the use of advanced biochemical techniques such as IP-mass spectrometry, RIP-Seq, etc.

Like many RNA binding proteins, the Aret protein binding sites in mRNA lack a tight consensus sequence. Oskar mRNA contains several BRE (Bruno response elements) in the 5' UTR that are recognized by the Aret protein, leading to translation repression. However, *in vitro* assays showed that Arrest binds to the 3' UTRs of several additional regulatory targets, including the

gurken (*grk*), cyclin A (*cycA*), Sex lethal (*Sxl*) and germ cell-less (*gcl*) mRNAs, for which no specific Bruno binding sites were specified [34]. These targets were all identified in the context of translation repression, and it is completely unknown what sequences Aret binds to regulate AS. Thus it will be essential to perform RIP-Seq or CLiP with the epitope-tagged alleles to identify the Aret target binding sequences in mRNA *in vivo* in IFM.

It does seem clear that Aret interacts with a number of sequences depending on temporal and tissue context. This suggests that Aret may interact with binding partners that confer temporal or tissue-specific binding preferences. This is also suggested by the lack of overlap in the AS exons that are regulated by Aret at 30h, 72h and in 1d adults [18]. Aret interaction partners in IFM remain to be identified. This will be greatly facilitated by an epitope-tagged Aret allele that will allow for IP-mass spec from endogenous, *in vivo* tissues. It may also be possible to identify interaction partners via a candidate approach. For example, muscleblind (*Mbl*) is one candidate that has been suggested as a potential partner by co-localization analysis and observation of genetic interaction [17], [49].

As a first RMCE approach, we chose to integrate the epitope tag on the C-terminus of the Aret protein. This has the advantage of tagging all possible Aret isoforms, as they share the same stop codon. However, there is a concern that the tag location on the C-terminal end of RRM3 may prevent the protein from folding correctly, despite the addition of a linker sequence. Thus, we examined other possible locations for tag insertion based on available sequencing data. One possibility would be an N-terminal location upstream of RRM1, but this location does not target all *aret* isoforms. Two other options are upstream or downstream of RRM2. Taking into consideration previous studies showing the importance of the divergent domain in protein-protein interactions [50], it would be strategic to avoid adding the tag downstream of RRM2 near the divergent domain. Our back-up tag location is thus in a non-conserved region upstream of RRM2. The decision about the epitope-tag location is also complex because the influence of intronic sequence on *aret* transcription is not known. Replacement of native intron-exon sequence with cDNA may provide essential information on this matter, for example replacing the sequence deleted in M1 or M2 with a cDNA by RMCE.

Most of the Aret isoforms, with an exception of isoform D, contain three RRM domains, organized similarly to the CUG-BP1 protein and other RRM protein families [19], [23], [34]. RRM domains typically bind single stranded RNA, although there are examples of members that can bind single stranded DNA or proteins [51]. The RRM domain itself consists of four anti-parallel beta strands and two alpha helices, arranged in an alpha/beta sandwich [52]. Interactions between the core RRMs and RNA typically span up to four nucleotides, providing only limited specificity in binding [53]. Many RRMs rely on additional structural features to expand the RNA binding surface and increase specificity and affinity. Another notable feature on the common portion of the protein is the divergent domain that separates RRM2 and RRM3, potentially serving as a platform for binding of other factors resulting in functional divergence of Aret [54].

The presence of multiple RRMs in Aret raises questions about the contribution of individual RRMs to RNA binding and the differences in the requirements for RRMs in translation repression versus AS. Webster *et. al.* showed that RRM3 along with a short stretch of upstream

sequence is sufficient for the specific binding of the BREs (Bruno response elements) on Oskar mRNA *in vitro* [19]. The amino-terminal domain also proved to be essential for Aret dimerization and for facilitating several specific functions including Cup binding and transcript repression [28], [34]. However, there is no information on the function of individual domains or on domain specificity during alternative splicing in IFMs. It remains to be tested whether each RRM sequence *per se* is required for AS target identification or whether it is the complete protein fold provided by the presence of all three RRM domains. The Step 1 alleles characterized are essential to answering these questions, as they allow generation of point mutant or deletion alleles affecting specific RRMs or other Aret domains, such as the divergent domain. One particularly interesting experiment that would give insight into complementarity and specificity of RRM domains is the substitution of RRM2 and RRM3 domains by RRM1 by a RMCE-based strategy applied to the Step I M1 line.

The *aret* gene is located on the left arm of the second chromosome and codes for several protein isoforms (Fig. 18). Two transcripts have been found in ovaries and one in testis by Northern Blot analysis [19]. Moreover, Oas *et. al.* have suggested that the RA isoform is more likely a flight muscle-specific one. Nevertheless, there are still no clear evidence of what isoforms are required for annotated cytoplasmic and nuclear functions and if those isoforms can substitute for each other's activity. Specific isoforms are known to be expressed exclusively in the germline [19], however limited information is available on the isoform-specific expression in IFM and on the function of specific isoforms. By Western Blot analysis, we identified two female-specific ovary isoforms of ~ 90 kDa (long) and of ~60 kDa (mid-length) that are not found in other tissues or in male samples. Analogously, we observed one male-specific testis isoform of ~88 kDa (long). We were unable to confirm an IFM-specific isoform by Western Blot due to the preliminary nature of these experiments in this thesis. However, mRNA-Seq data suggests strong expression of the short isoform in IFMs, as well as lower expression of mid-length and long isoforms (M. Spletter and F. Schnorrer, unpublished data). The identification of the IFM-specific isoforms on the protein level remains a priority for future experiments.

Our phenotypic characterization of Step I Aret mutants also suggests that different Aret isoforms may play different functional roles in IFM. The Step I mutants we were able to characterize either affect all isoforms through the terminal RRM2/RRM3 domains (M2 and M3) or only a subset of mid-length isoforms by targeting a specific N-terminal regulatory region (M4). In M2 and M3, we showed that Aret protein is completely absent from IFM nuclei by immunostaining. We find that M2 and M3 phenocopy older Aret alleles and Mef2-Gal4 driven Aret-RNAi [18]. The mutant sarcomeres were ~2.09 μm in length at 90 h APF (Fig. 11, Table 3), a size corresponding to the length of wild-type sarcomeres at 48 h APF. We also observed a tendency of IFM myofibrils to fuse together and form a "mini-tube" pattern in M2 and M3 (Fig. 8D-E, 9C-E), which was also observed in RNAi experiments and with other Aret alleles. Interestingly, we also observe a change in the localization of projectin and Mhc in M2 and M3 (Fig. 22, 23), suggesting that the sarcomere organization in these lines may be more similar to that normally observed in tubular muscle. This is an interesting observation, as IFMs develop relatively normally until 48h APF, and the absence of Aret does not lead to a complete IFM conversion into tubular muscle. There may be some IFM-specific maturation processes controlled by Aret targets that are disrupted in

M2 and M3, leading to a more tubular muscle pattern of sarcomere protein incorporation. Further experiments are required to test this hypothesis.

In the case of the M4 mutant which targets presumably only mid-length Aret isoforms, and in contrast to the M2 and M3 alleles, *aret* signal was detectable in the nuclei of the IFMs (Fig. 8G) by immunostaining. Additionally, the fiber and myofibril phenotype were less severe at 90h APF than those observed in M2 and M3, as myofibrils were separately aligned, hosting longer sarcomers and showed no sign of “mini-tube” formation. Based on sarcomere measurement data, it seems that sarcomeres grow decently until 72h APF (Fig. 11; Table 5), implying the presence of a functional Aret isoform or isoforms sufficient to induce initial fibrillar-specific splice patterns, but failing to support the final stages of sarcomere growth and maturation. We hypothesize that either the short isoform D, which is downstream of the domain targeted by Step I cassette integration, or longer isoforms that skip the deleted region, may still be present in M4. These isoforms either compensate for the missing isoforms or normally perform a subset of the AS regulation seen in wild-type. It will be interesting to further characterize the M4 allele and perform mRNA-Seq to see how the AS landscape differs from WT and from other Aret alleles.

To understand the role of individual Aret isoforms in IFM development, we need to identify which Aret isoforms are present in Step I M4 mutants, either by Western Blot or by RT-PCR. Potentially interesting results could be obtained from characterizing M5 mutants that possess a deletion similar to M4. If phenotypes of the M5 line match those observed in M4, it would allow us to confirm our observation and match function to a specific Aret isoform. Altogether, the repeat of Western Blot analysis on wild-type, M4 and M5 mutants is essential to identify the native IFM-specific *aret* isoforms and the isoforms present in the mutants. Eventually, this may allow us to identify the different *aret* isoforms present in fibrillar muscle and their specific functions during sarcomerogenesis.

An important challenge to understand how different muscle types develop is the generation of a complete network of both transcriptional and splicing regulators involved in the process. Research in model organisms, such as flies, provides an important complement to vertebrate studies. It has been suggested that isoform switching is a basic principle of muscle fiber-type specification in both vertebrates and *Drosophila* [4]. Similarly to IFMs and leg muscles in the fly, alternative splicing in the vertebrate heart is very distinct to that observed in skeletal muscles. Notably, heart muscle shares common features with insect flight muscle. This indicates that the function of CELF family members in muscle development may be evolutionarily conserved and implies that the data obtained from *Drosophila* studies may be correlated to processes during mammalian muscle morphogenesis.

Materials & Methods

Fly work and flight tests

All fly work was performed at 25 °C on standard medium. All fly stocks used in this study are listed in Supplementary Information (Table 1). For the crosses, only virgin females were selected. Step I flies were maintained in Aret^{Mx}/CyO stable stocks.

Flight tests were performed in a flight chamber as previously described [55]. In brief, a 1m cylinder was divided into 5 regions. Flies released from the top of the cylinder were scored for which zone they landed in. The bottom 2 zones indicate flightlessness, while the upper zones indicate flight capability. Flight tests were performed with a minimum of 10 and maximum of 100 flies per test. Tests were performed with males of the desired genotype, selected on CO₂ anesthesia and recovered for at least 24 hours at 25 °C.

Cloning and verification experiments

For verification of Step I Deletion flies, genomic DNA was extracted from homozygous mutants by a simple fly genomic DNA preparation with the use of Buffer A (0.1M Tris, 0.1 M EDTA, 1% SDS) and 5M KoAc (pH 4.8) [56]. All primers used for verification and cloning approaches, as well as linker-HA and linker-V5 tag oligos, are listed in Supplementary Information and were ordered from Eurofins Genomics. Primers were designed in the IDT OligoAnalyzer Tool (<http://eu.idtdna.com/calc/analyzer>). Gel purification was done with the MinElute Gel Extraction kit from Qiagen. Purified DNA was sent for sequencing using the premixed Mix2Seq Kit (Eurofins Genomics). Sequence prediction and verification was done with SeqMan and SeqBuilderSoftwares (version: 8.1.2) by Lasergene.

The following Restriction Enzymes were used in the cloning approach: HindIII, XbaI, BsmBI, BsaI, BstBI, NotI (all from NEB). Additionally, T4 polynucleotide kinase, DNA polymerase I (Klenow) Fragment and Antarctic Phosphatase (all from NEB) were used to enable the 3-piece ligation. Full-length inserts were cloned into pBS-KS-attb1-2-PT-SA-SD-1 [26] receiving vector, while intermediate pieces were cloned into pJet1.2 (CloneJet PCR Cloning Kit by Thermo Scientific). The integrity of the constructs was verified by RE digest by XhoI and ScaI enzymes (NEB). Plasmids were transformed into DH5alpha competent cells (provided by Sandra Lemke and WouterKoolhas) using a standard transformation protocol [57]. To obtain the necessary high concentration and pureness for injections, Step II constructs were purified using the Invitrogen PureLink Kit.

Immunostainings

Immunostainings were performed using standard protocols [58]. Briefly, to stain IFMs, fly thoraxes were dissected from 90 h APF pupae and 1-day adults. Thoraxes were fixed in 0.5% PBS-Triton containing 4% PFA for 20-30 min. at RT, washed in 0.5% PBS-Triton and cut sagittally with a razor blade. Blocking was performed in 5% NGS in 0.5% PBS Triton for at least 1 hour at RT. Primary antibodies were used at the following dilutions: rabbit anti-bruno (1:200) provided by Anna Ephrussi, rat anti-kettin (1:50) from Babraham Institute, mouse anti-GFP

(1:1000) by Roche, mouse anti-HA (1:1000) by HISS Diagnostics. Secondary antibodies (Life Technologies), all used at 1:500, included Alexa Fluor 488 goat anti-rabbit, Alexa Fluor 488 goat anti-mouse, Alexa Fluor 633 goat anti-rat, Rhodamine-Phalloidin. Samples were mounted in Vectashield (Vector Laboratories) containing DAPI for DNA staining. Slides were imaged on a Zeiss 780 laser scanning confocal microscope using 10x and 100x objectives. Images were processed using Fiji (Image J). Sarcomere length and width was quantified based on phalloidin staining in Fiji using software written by Giovanni Cardone. Sarcomere length was estimated using a Fourier transform to ascertain the length of the repeat along the myofibril, while width was estimated using an auto-correlation function across multiple myofibrils. Significance was evaluated with ANOVA using GraphPad Prism software.

Western Blotting

Drosophila tissues were dissected manually as described [58], [59] in cold PBS containing 0.1% Triton. The samples were lysed in the presence of 6xSDS buffer at 95^oC for 5 minutes and were subsequently diluted to 2xSDS for loading on the gel. Each well contained the amount of sample estimated to correspond to 0.5 fly. We used ready-made 4-12% Bis-Tris Gel with MOPS Running Buffer (Sigma-Aldrich), run for 45 minutes at 180 Volts/mA. The nitrocellulose membrane was activated by a 1 min. incubation in MeOH prior to blotting. Blotting was performed in precooled 1x Blotting Buffer (1.92 M Glycine, 250 mM Tris in 1L d₂H₂O) for 1 hour at 4^oC. Subsequent membrane washes were performed in 1xPBS containing 0.1% Tween. As primary antibodies, we have used rabbit anti-Bruno (1:1000) and rat anti-Bruno (1:1000), both generously provided by Anne Ephrussi. Secondary antibodies against rat (by Sigma-Aldrich) and rabbit (by Jackson) were applied in 1:10,000 ratio. Membranes were developed with the chemilluminant kit from Millipore. Chemiluminescence was detected on a LAS 4000 machine, with 10 sec exposure interval time.

Bibliography

- [1] C. L. Will and R. Lührmann, “Spliceosome structure and function.,” *Cold Spring Harb. Perspect. Biol.*, vol. 3, no. 7, Jul. 2011.
- [2] D. L. Black, “Mechanisms of alternative pre-messenger RNA splicing.,” *Annu. Rev. Biochem.*, vol. 72, pp. 291–336, 2003.
- [3] P. Förch and J. Valcárcel, “Splicing regulation in *Drosophila* sex determination.,” *Prog. Mol. Subcell. Biol.*, vol. 31, pp. 127–51, 2003.
- [4] M. L. Spletter and F. Shnorrrer, “Transcriptional regulation and alternative splicing cooperate in muscle fiber-type specification in flies and mammals,” *Exp Cell Res.*, vol. 321, no. 1, pp. 90–98, 2014.
- [5] B. Wei and J. Jin, “Troponin T isoforms and posttranscriptional modifications: Evolution, regulation and function.,” *Arch Biochem Biophys*, vol. Jan 15, no. 505(2), 2011.
- [6] S. Li, W. Guo, C. N. Dewey, and M. L. Greaser, “Rbm20 regulates titin alternative splicing as a splicing repressor.,” *Nucleic Acids Res.*, vol. 41, no. 4, pp. 2659–72, Feb. 2013.
- [7] M. Gautel and D. Goulding, “A molecular map of titin/connectin elasticity reveals two different mechanisms acting in series,” *FEBS Lett.*, vol. 385, no. 1–2, pp. 11–14, Apr. 1996.
- [8] W. Guo, S. J. Bharmal, K. Esbona, and M. L. Greaser, “Titin diversity--alternative splicing gone wild.,” *J. Biomed. Biotechnol.*, vol. 2010, p. 753675, 2010.
- [9] W. A. Linke, D. E. Rudy, T. Centner, M. Gautel, C. Witt, S. Labeit, and C. C. Gregorio, “I-band titin in cardiac muscle is a three-element molecular spring and is critical for maintaining thin filament structure.,” *J. Cell Biol.*, vol. 146, no. 3, pp. 631–44, Aug. 1999.
- [10] W. Linke, M. Kulke, H. Li, S. Fujita-Becker, C. Neagoe, D. Manstein, M. Gautel, and J. Fernandez, “PEVK domain of titin: an entropic spring with actin-binding properties.,” *J Struct Biol*, vol. Jan-Feb, no. 137(1–2), pp. 194–205, 2002.
- [11] D. S. Herman, L. Lam, M. R. G. Taylor, L. *et.al.*. “Truncations of titin causing dilated cardiomyopathy.,” *N. Engl. J. Med.*, vol. 366, no. 7, pp. 619–28, Feb. 2012.
- [12] F. O. Lehmann and M. H. Dickinson, “The changes in power requirements and muscle efficiency during elevated force production in the fruit fly *Drosophila melanogaster*.,” *J. Exp. Biol.*, vol. 200, no. Pt 7, pp. 1133–43, Apr. 1997.
- [13] B. Bullard and A. Pastore, “Regulating the contraction of insect flight muscle.,” *J. Muscle Res. Cell Motil.*, vol. 32, no. 4–5, pp. 303–13, Dec. 2011.
- [14] J. W. S. PRINGLE, “The Bidder Lecture, 1980 The Evolution of Fibrillar Muscle in Insects,” *J. Exp. Biol.*, vol. 94, no. 1, pp. 1–14, 1981.
- [15] O. W. Tiegs, “The Flight Muscles of Insects-Their Anatomy and Histology; with Some Observations on the Structure of Striated Muscle in General,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 238, no. 656, pp. 221–348, Feb. 1955.

- [16] C. Schönbauer, J. Distler, N. Jährling, M. Radolf, H.-U. Dodt, M. Frasch, and F. Schnorrer, “Spalt mediates an evolutionarily conserved switch to fibrillar muscle fate in insects,” *Nature*, vol. 479, no. 7373, pp. 406–409, 2011.
- [17] S. T. Oas, A. L. Bryantsev, and R. M. Cripps, “Arrest is a regulator of fiber-specific alternative splicing in the indirect flight muscles of *Drosophila*,” *J. Cell Biol.*, vol. 206, no. 7, pp. 895–908, 2014.
- [18] M. L. Spletter, C. Barz, A. Yeroslaviz, C. Schönbauer, I. R. S. Ferreira, M. Sarov, D. Gerlach, A. Stark, B. H. Habermann, and F. Schnorrer, “The RNA-binding protein Arrest (Bruno) regulates alternative splicing to enable myofibril maturation in *Drosophila* flight muscle,” *EMBO Rep.*, vol. 16, no. 2, p. e201439791, 2014.
- [19] P. J. Webster, L. Liang, C. A. Berg, P. Lasko, and P. M. Macdonald, “Translational repressor bruno plays multiple roles in development and is widely conserved,” *Genes Dev.*, vol. 11, no. 19, pp. 2510–2521, 1997.
- [20] J. Kim-Ha, K. Kerr, and P. M. Macdonald, “Translational regulation of oskar mRNA by Bruno, an ovarian RNA-binding protein, is essential,” *Cell*, vol. 81, no. 3, pp. 403–412, 1995.
- [21] P. Filardo and A. Ephrussi, “Bruno regulates gurken during *Drosophila* oogenesis,” *Mech. Dev.*, vol. 120, no. 3, pp. 289–297, 2003.
- [22] J. Moore, H. Han, and P. Lasko, “Bruno negatively regulates germ cell-less expression in a BRE-independent manner,” *Mech. Dev.*, vol. 126, no. 7, pp. 503–516, 2009.
- [23] M. Snee, D. Benz, J. Jen, P. M. Macdonald, “Two distinct domains of Bruno bind specifically to the oskar mRNA,” *RNA Biol.*, vol. 5, no. 1, pp. 1–9.
- [24] A. N. Ladd, G. Taffet, C. Hartley, D. L. Kearney, and T. A. Cooper, “Cardiac tissue-specific repression of CELF activity disrupts alternative splicing and causes cardiomyopathy,” *Mol. Cell. Biol.*, vol. 25, no. 14, pp. 6267–78, Jul. 2005.
- [25] A. Kalsotra, X. Xiao, A. J. Ward, J. C. Castle, J. M. Johnson, C. B. Burge, and T. A. Cooper, “A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 51, pp. 20333–8, Dec. 2008.
- [26] X. Zhang, W. H. Koolhaas, and F. Schnorrer, “A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in *Drosophila*,” *G3 (Bethesda)*, vol. 4, no. 12, pp. 2409–18, 2014.
- [27] K. Venken, “MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes,” *Nat Methods*, vol. 8, no. 9, pp. 737–743, 2011.
- [28] G. Kim, “Region-Specific Activation of oskar mRNA Translation by Inhibition of Bruno-Mediated Repression,” *PLOS Genetics*, 2015. [Online]. Available: <http://www.plosgenetics.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pgen.1004992&representation=PDF>. [Accessed: 17-Feb-2016].
- [29] M. Chekulaeva, M. W. Hentze, A. Ephrussi, “Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles,” *Cell*,

vol. 124, no. 3, pp. 521–33, Feb. 2006.

- [30] D. H. Castrillon, P. Gönczy, S. Alexander, R. Rawson, C. G. Eberhart, S. Viswanathan, S. DiNardo, and S. A. Wasserman, “Toward a molecular genetic analysis of spermatogenesis in *Drosophila melanogaster*: characterization of male-sterile mutants generated by single P element mutagenesis.,” *Genetics*, vol. 135, no. 2, pp. 489–505, Oct. 1993.
- [31] T. Schüpbach and E. Wieschaus, “Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology.,” *Genetics*, vol. 129, no. 4, pp. 1119–36, Dec. 1991.
- [32] A.-S. Fiston-Lavier, N. D. Singh, M. Lipatov, and D. A. Petrov, “*Drosophila melanogaster* recombination rate calculator,” *Gene*, vol. 463, no. 1, pp. 18–20, 2010.
- [33] M. C. Reedy and C. Beall, “Ultrastructure of Developing Flight Muscle in *Drosophila*. I. Assembly of Myofibrils,” *Dev. Biol.*, vol. 160, no. 2, pp. 443–465, Dec. 1993.
- [34] B. Reveal, C. Garcia, A. Ellington, and P. M. Macdonald, “Multiple RNA binding domains of Bruno confer recognition of diverse binding sites for translational repression.,” *RNA Biol.*, vol. 8, no. 6, pp. 1047–60, 2011.
- [35] R. J. Bandziulis, M. S. Swanson, and G. Dreyfuss, “RNA-binding proteins as developmental regulators.,” *Genes Dev.*, vol. 3, no. 4, pp. 431–7, Apr. 1989.
- [36] J. P. Venables, J. Tazi, and F. Juge, “Regulated functional alternative splicing in *Drosophila*.,” *Nucleic Acids Res.*, vol. 40, no. 1, pp. 1–10, Jan. 2012.
- [37] A. Ayme-Southgate, R. A. Philipp, and R. J. Southgate, “Projectin PEVK domain, splicing variants and domain structure in basal and derived insects.,” *Insect Mol. Biol.*, vol. 20, no. 3, pp. 347–56, Jun. 2011.
- [38] M. C. Leake, D. Wilson, M. Gautel, and R. M. Simmons, “The elasticity of single titin molecules using a two-bead optical tweezers assay.,” *Biophys. J.*, vol. 87, no. 2, pp. 1112–35, Aug. 2004.
- [39] W. A. Linke, M. Ivemeyer, P. Mundel, M. R. Stockmeier, and B. Kolmerer, “Nature of PEVK-titin elasticity in skeletal muscle.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 14, pp. 8052–7, Jul. 1998.
- [40] K. Trombitás, M. Greaser, S. Labeit, J. P. Jin, M. Kellermayer, M. Helmes, and H. Granzier, “Titin extensibility in situ: entropic elasticity of permanently folded and permanently unfolded molecular segments.,” *J. Cell Biol.*, vol. 140, no. 4, pp. 853–9, Feb. 1998.
- [41] L. Wells, K. A. Edwards, and S. I. Bernstein, “Myosin heavy chain isoforms regulate muscle function but not myofibril assembly.,” *EMBO J.*, vol. 15, no. 17, pp. 4454–9, Sep. 1996.
- [42] Z. Orfanos and J. C. Sparrow, “Myosin isoform switching during assembly of the *Drosophila* flight muscle thick filament lattice.,” *J. Cell Sci.*, vol. 126, no. Pt 1, pp. 139–48, Jan. 2013.

- [43] A. N. Ladd, N. Charlet, and T. A. Cooper, "The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing.," *Mol. Cell. Biol.*, vol. 21, no. 4, pp. 1285–96, Feb. 2001.
- [44] Y.-H. Goo and T. A. Cooper, "CUGBP2 directly interacts with U2 17S snRNP components and promotes U2 snRNA binding to cardiac troponin T pre-mRNA.," *Nucleic Acids Res.*, vol. 37, no. 13, pp. 4275–86, Jul. 2009.
- [45] R. Spellman, C. W. J. Smith, S. M. Berget and et al., "Novel modes of splicing repression by PTB.," *Trends Biochem. Sci.*, vol. 31, no. 2, pp. 73–6.
- [46] C. D. Chen, R. Kobayashi, and D. M. Helfman, "Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene.," *Genes Dev.*, vol. 13, no. 5, pp. 593–606, Mar. 1999.
- [47] H. Kuroyanagi, "Fox-1 family of RNA-binding proteins.," *Cell. Mol. Life Sci.*, vol. 66, no. 24, pp. 3895–907, Dec. 2009.
- [48] M. Llorian and C. W. Smith, "Decoding muscle alternative splicing," *Curr. Opin. Genet. Dev.*, vol. 21, no. 4, pp. 380–387, 2011.
- [49] M. Vicente-Crespo, "Drosophila Muscleblind Is Involved in troponin T Alternative Splicing and Apoptosis," *PLoS ONE*, 2008. [Online]. Available: <http://www.plosone.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pone.0001613&representation=PDF>. [Accessed: 10-Feb-2016].
- [50] T. Dasgupta and A. N. Ladd, "CELF control," *Wiley Interdiscip Rev RNA*, vol. 3, no. 1, pp. 104–121, 2012.
- [51] S. D. Johnston, J. E. Lew, and J. Berman, "Gbp1p, a protein with RNA recognition motifs, binds single-stranded telomeric DNA and changes its binding specificity upon dimerization.," *Mol. Cell. Biol.*, vol. 19, no. 1, pp. 923–33, Jan. 1999.
- [52] A. Cléry and F. H.-T. Allain, "FROM STRUCTURE TO FUNCTION OF RNA BINDING DOMAINS," 2013.
- [53] J. L. Smith, J. E. Wilson, P. M. Macdonald, "Overexpression of oskar directs ectopic activation of nanos and presumptive pole cell formation in Drosophila embryos," *Cell*, vol. 70, no. 5, pp. 849–859, Sep. 1992.
- [54] A. Nakamura, K. Sato, K. Hanyu-Nakamura, "Drosophila Cup Is an eIF4E Binding Protein that Associates with Bruno and Regulates oskar mRNA Translation in Oogenesis," *Dev. Cell*, vol. 6, no. 1, pp. 69–78, Jan. 2004.
- [55] D. R. Drummond, E. S. Hennessey, and J. C. Sparrow, "Characterisation of missense mutations in the Act88F gene of Drosophila melanogaster," *MGG Mol. Gen. Genet.*, vol. 226–226, no. 1–2, pp. 70–80, Apr. 1991.
- [56] A. M. Aransay, E. Scoulica, and Y. Tselentis, "Detection and Identification of Leishmania DNA within Naturally Infected Sand Flies by Seminested PCR on Minicircle Kinetoplasmic DNA," *Appl. Environ. Microbiol.*, vol. 66, no. 5, pp. 1933–1938, May 2000.

- [57] H. Inoue, H. Nojima, and H. Okayama, “High efficiency transformation of *Escherichia coli* with plasmids,” *Gene*, vol. 96, no. 1, pp. 23–28, 1990.
- [58] M. Weitkunat and F. Schnorrer, “A guide to study *Drosophila* muscle biology,” *Methods*, vol. 68, no. 1, pp. 2–14, 2014.
- [59] L. C. Wong and P. Schedl, “Dissection of *Drosophila* ovaries.” *J. Vis. Exp.*, no. 1, p. 52, Dec. 2006.

Supplement Materials

Characteristics	Name of the line	Source
Common stocks	w-	Bloomington Collection
	yw; Pin/CyO	
	yw; Pin/CyO, Tb	
	Act5C-Cas9, lig ⁴	Zhang <i>et. al.</i> , 2014
Step I lines	M1 2-3; AretTm/CyO	M. Spletter, F. Schnorrer (unpublished)
	M 1 4-3 ; AretTm/CyO	
	M2 23; AretTm/CyO	
	M3 1-3; AretTm/CyO	
	M3 3-1; AretTm/CyO	
	M3 3-1.11; AretTm/CyO	
	M3 3-1.32; AretTm/CyO	
M4 1-2; AretTm/CyO		
M5 1-2; AretTm/CyO		
RNAi lines	UAS-salm-IR-TF3029	Vienna Collection (VDRC)
	aret-IR-GD41568	
Defficiency line	Df(2L) BSC407; BL24911	Bloomington Collection
Driver lines	Mef2-Gal4;HA-bt	W. Koolhaas, F. Schnorrer (unpublished)
	Mef2-Gal4;Mhc-weeP26	Sparrow <i>et.al.</i> , 2013
	Act88F-Gal4;HA-bt	Oas <i>et. al.</i> , 2014 (Act88F-Gal4 line), W. Koolhaas (HA-bt line, unpublished)
	Act88F-Gal4;Mhc-weeP26	Oas <i>et. al.</i> , 2014 (Act88F-Gal4 line), Sparrow <i>et.al.</i> , 2013 (Mhc-weeP26)
Aret loss of function alleles	<i>aret</i> ^{QB72}	Anne Ephrussi
	<i>aret</i> ^{WH53}	

Table 1. Fly stocks used in the study.

The source of flies and their genotypes are indicated.

	Primer name	Primer Sequence	Comments	
Primers for Step II Constructs verification	MS 262L	atcaaaacatcgctccttgg	On HN-N1 HA	
	MS 262R	cttcattgtcgaaggagcc	On HN-N1 HA	
	MS 264L	cctagtattgccacgatt	On HN-N2 HA	
	MS 264R	aatcgacttctacgagtcca	On HN-N2 HA	
	MS 263L	ggaatggaaatctctgcgagt	On HN-C1 HA	
	MS 263R	ctgaccccctaacctcaca	On HN-C1 HA	
	MS 265L	gtgtgtgtgtgggttttgc	On HC-N1 HA	
	MS 265R	tggtcgcctttatagagaga	On HC-N1 HA	
	MS 266L	tgttttatgattggcggaa	On HC-C2 HA	
	MS 266R	acgccgctaaattgtttgtc	On HC-C2 HA	
	MS 267L	tgcgattgtaaaggaaacg	On HC-C1 HA	
	MS 267R	gcaaatcgcaagtttagcaa	On HC-C1 HA	
	MS 268L	tttgcggattccaagaagt	On HC-N2 HA	
	MS 268R	ctgtcaacgtcactggcaat	On HC-N2 HA	
	MS 286	gactagttgatcataatcagcca	On DsRed cassette	
	MS 287	ccacaaggccctgaagctga	On DsRed cassette	
	MS 288	atgctgatggcgggtatta	On HC-N1 HA	
	MS 292	gcactgacagtttgcgataa	Step I Constructs 1,2 L-HA	
	MS 293	cgacactgaaaaattccaa	Step I Constructs 3 L-HA	
	MS 294	gtggaaggacgaaagcagag	Step I Constructs 4 L-HA	
	MS 295	tgctgaaacctgaagaatg	Step I Constructs 5 L-HA	
	MS 309	tggctgtttcgttgctctc	Step I Constructs 1,3 R-HA	
	MS 297	gaatctcgccagaaccaaga	Step I Constructs 2 R-HA	
	MS 304	ccgaccccgagaaaataacgcc	Step I Constructs 4,5 R-HA	
	MS 307	cacaaaacaggagaaaggg	On HC-N2 HA	
	MS 308	tattcggaaggatgctgc	On HC-C2 HA	
	MS 312	ttctcgctgttactgcc	On HC-C1 HA	
	MS 313	actcgataactaattggcagg	On HC-N1	
	MS 322	cgactcccgtgacaggt	On HN-C1 HA	
	MS 323	tgtaatacgcctgctaacc	On HN-N2 HA	
	Primers for Step II constructs design	MS 299	attcgtctcatagaggccaccatcggcggcaaggt	Step II. ExRight-5'. BsmBI RE site
		MS 317	actcgtctcgagctgcattcgattgatgatcaatgagccgcttatcgatg	Step II. ExRight-3'. BsmBI RE site.
MS 301		atggctcactagtaattgcacgaaaagtgaaccgc	Step II. ExLeft-5' BsaI RE site	
MS 302		gcccagaccaccgatcctccgtagggttcgagtccttg	Step II. ExLeft-V5 tag overhang	
MS 303		gtaggatcctccgatcctccgtagggttcgagtccttg	Step II. ExLeft-HA tag overhang	
MS 305		atggctcttctacgtgcta	Step II. 3`V5 tag. BsaI RE site	
		atggctcttctatcctgca	Step II. 3`HA tag. BsaI RE site	
MS 314		cggcaaggttaaacggac	Step II. ExRight-5'. No overhangs	
MS 315		aatgagccgcttatcgatgc	Step II. ExRight-3'. No overhangs	
MS 320		agtcgttaaagattggg	Step II. ExRight. Cover BstBI RE site	
MS 324		ggtctcatctagtagggcttcgagtccttg	Step II. ExLeft-3` BsaI RE site	

Table 2. Summary of the primers used for verification of Step I mutants and for the cloning of Step II constructs.

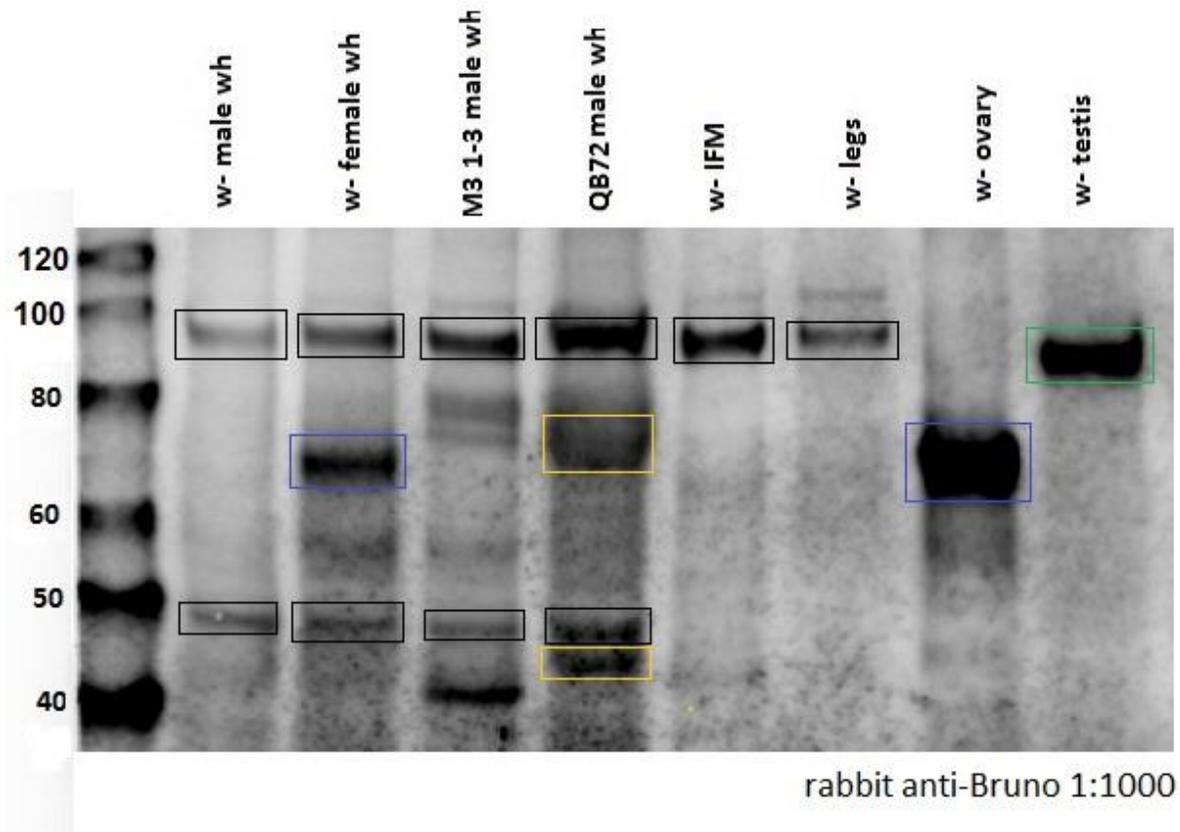


Figure 1. . Investigation of Aret isoforms by Western Blotting.

Staining of Aret isoforms in wild-type and mutant samples with rabbit anti-Bruno antibody (1:1000). Sample labels indicated above the blot. Blot compares dissected tissue-specific samples to whole-fly (indicated as wh) and homozygous mutant controls. Ovary and testis specific isoforms are clearly visible and absent in mutant samples. Background bands are marked by black boxes, prominent female isoforms by blue, male isoforms by green, potential truncated isoforms by yellow.