The DEAD box RNA helicase Ddx39ab is essential for myocyte and lens development in zebrafish

Linlin Zhang[1, #], Yuxi Yang[1, #], Beibei Li[1], Ian C. Scott[2, 3] and Xin Lou[1]

[1]Model Animal Research Center, Nanjing University, China. [2]Program in Developmental and Stem Cell Biology, The Hospital for Sick Children; [3]Department of Molecular Genetics, University of Toronto, Canada.

[#] These authors contributed equally to this work.

[*]Correspondence should be addressed to

Xin Lou, Model Animal Research Center, Nanjing University

Address: Rm 403B, 12 Xuefu Road, Nanjing, Jiangsu, China

Phone: +86-2558641547.

e-mail: xin.lou@nju.edu.cn

Keywords: RNA helicase, myocyte, lens, RNA splicing, zebrafish

Abstract

RNA helicases from the DEAD-box family are found in almost all organisms and have important roles in RNA metabolism including RNA synthesis, processing and degradation. The function and mechanism of action of most of these helicases in animal development and human disease are largely unexplored. In a zebrafish mutagenesis screen to identify genes essential for heart development we identified a mutant which disrupts the gene encoding the RNA helicase DEAD-box 39ab (*ddx39ab*). Homozygous *ddx39ab* mutant embryos exhibit profound cardiac and trunk muscle dystrophy, along with lens abnormalities, caused by abrupt terminal differentiation of cardiomyocyte, myoblast and lens fiber cells. Further investigation indicated that loss of *ddx39ab* hindered mRNA splicing of members of the *kmt2* gene family, leading to mis-regulation of structural gene expression in cardiomyocyte, myoblast and lens fiber cells. Taken together, these results show that Ddx39ab plays an essential role in establishment of proper epigenetic status during differentiation of multiple cell lineages.

Introduction

The DEAD-box RNA helicase family comprises a large protein group characterized by the presence of an Asp-Glu-Ala-Asp (DEAD) motif that is highly conserved from bacteria to man (Bleichert and Baserga, 2007; Rocak and Linder, 2004). Using the energy derived from ATP hydrolysis, these proteins modulate RNA topology and association/dissociation of RNA-protein complexes. DEAD-box RNA helicases play important roles in all aspects of RNA metabolism, including transcription, pre-mRNA splicing, rRNA biogenesis, RNA transport and translation (Calo et al., 2015; Jarmoskaite and Russell, 2011, 2014; Linder and Jankowsky, 2011). Recently, a fuller appreciation of the functions of these RNA helicases in variant physiological or developmental scenarios has started to emerged. Numerous reports have described deregulation of expression or function of DEAD-box RNA helicases in cancer development or progression, indicating that DEAD box proteins may be involved in processes that are key to cellular proliferation (Fuller-Pace, 2013; Sarkar and Ghosh, 2016). Recent reports have further shown that DEAD-box RNA helicases play diverse roles in developmental events ranging from body axis establishment (Meignin and Davis, 2008) to germ cell, blood, digestive organ and brain development (Hirabayashi et al., 2013; Hozumi et al., 2012; Payne et al., 2011; Zhang et al., 2012). These studies indicated that DEAD-box RNA helicase family members could regulate specific developmental processes by affecting pre-mRNA splicing, ribosomal biogenesis or RNA transport.

As a member of DEAD box family of ATP-dependent RNA helicases, Ddx39a was originally identified in a screen for proteins that interact with the essential splicing factor U2AF65 (Fleckner et al., 1997). Studies in different organisms indicated that besides playing an important role in pre-mRNA splicing (Fleckner et al., 1997; Shen et al., 2008; Shen et al., 2007), DDX39a also acts in mRNA nuclear export (Luo et al.,

2001), cytoplasmic mRNA localization (Meignin and Davis, 2008) and maintenance of genome integrity (Yoo and Chung, 2011). However, how Ddx39a functions during embryogenesis, and which RNAs are processed by Ddx39a in different developmental scenarios remains open to investigation.

In the current study, we examined the function of *ddx39a* in vertebrate development by using a newly identified zebrafish *ddx39ab* gene trap line. Transcriptome profiling and phenotypic analysis of the *ddx39ab* mutant showed that *ddx39a* is indispensable for the development of heart, trunk muscle and eyes. Further experiments revealed that Ddx39ab can bind to mRNAs encoding a set of epigenetic regulatory factors, including members of KMT2 family. Loss of *ddx39ab* leads to aberrant pre-mRNA splicing of these epigenetic regulator transcripts, leading to a failure in establishment of proper epigenetic status of multiple structural genes, and eventually hampers the terminal differentiation of cardiomyocyte, myoblast and lens fiber cell lineages.

Materials and methods

Ethics approval

The zebrafish used in this study were regularly maintained and handled in accordance with approved guidelines of the Institutional Animal Care and Use Committee of the Nanjing University and per Canadian Council on Animal Care and Hospital for Sick Children Laboratory Animal Services guidelines.

Zebrafish lines

Zebrafish embryos were maintained and staged using standard techniques (Westerfield, 1993). The RP-T gene-trap vector was modified from RP2 (Clark et al., 2011) by switching the monomeric RFP to monomeric GFP. RP-T plasmid (25 ng/μl)

and Tol2 transposase mRNA (50 ng/μl) were injected (1 nl each) into 1-cell stage embryos as previous described (Hou et al., 2017). To identify the affected gene in the gene trap lines, inverse PCR and 5'RACE were applied as previously described (Clark et al., 2011). The gene trapping insertion position in the *ddx39ab* locus was determined by sequencing. For PCR genotyping single embryos, primers were designed to differentially amplify the wild type allele or gene trap allele. After immunostaining or *in situ* hybridization, genotyping PCR was applied to as previously described (Kawakami et al., 2016). The primer pairs and detailed PCR conditions for PCR genotyping are listed in Supplementary Table 1.

Reverse transcription PCR

Total RNA was prepared using TRIzol (Invitrogen, Life Technologies Corp.), with DNase-treated RNA reverse transcribed using random 16-mer priming and SuperSript II reverse transcriptase (Life Technologies Corp.). PCR was applied with primers specifically amplifying cDNA from *ddx39ab* mRNA or fusion transcript. The primer pairs and detailed PCR conditions are listed in Supplementary Table 1.

Reverse Transcription Quantitative PCR

Embryos from 3 different clutches were collected as biological replicates and total RNA was prepared using TRIzol (Invitrogen, Life Technologies Corp.). RNA was reverse transcribed with mix of oligo(dT) and random 16-mer priming and SuperSript II reverse transcriptase (Life Technologies Corp.). Quantitative PCR (qPCR) assays were performed in triplicate with the SYBR Green Master Mix (Takara) according to the manufacturer's instructions. Melt curves were examined to ensure primer specificity. Results were analyzed using the standard ΔΔCT method (Schmittgen and Livak, 2008). 18S rRNA served as the reference gene in all analyses, and changes in mRNA levels relative to 18S rRNA were confirmed using *actb2* as alternate reference genes in

independent experiments. Primers used for qPCR analysis are listed in Supplementary Table 1.

Statistics

Two-sided, paired Student's t-tests were applied for RT-qPCR and CHIP-qPCR results; two-sided, unpaired Student's t-tests were applied for other quantification experiments. Significant differences were calculated with GraphPad Prism6. The significance level was set to $P < 0.05$.

mRNA injections

pCS2+ vectors carrying a cDNA fragment encoding membrane RFP, Ddx39ab, Flag-Ddx39ab and Tol2-transposase were used in this study. Capped mRNA was synthesized using an SP6 mMESSAGE mMACHINE kit (Ambion, Life Technologies Corp.). For phenotypic rescue experiments, *ddx39ab* mRNA (100 pg) was injected at the one-cell stage.

Immunocytochemistry

Embryos were fixed in 4% paraformaldehyde (PFA) at 4 °C overnight. Embryos were blocked with blocking solution (1x PBS, 1% BSA, 1% Triton-X100, 0.1% DMSO) for 2 hours then incubated with primary antibodies diluted in block solution overnight at 4 °C. Embryos were washed 3 times in 1x PBS with 1% Triton-X100 for 15 min. Embryos were incubated with secondary antibodies diluted in blocking solution for 2 hours. Primary antibodies specific to MYH1A, MYH1E, Cardiac troponin T (cTnT), and Acetylated Tubulin were used. Fluorescent immunocytochemistry was performed using anti-mouse antibody conjugated with Actin filaments were visualized with Rhodamine-conjugated phalloidin. Unless otherwise stated, manipulations were performed at room temperature. Detailed information of antibodies and dilution was listed in Table S2.

Imaging

Embryo whole mount imaging was performed using a Leica DFC320 camera on a Leica M205FA stereomicroscope. All confocal images were taken using a Zeiss LSM880 confocal microscope.

RNA *in situ* hybridization

Transcription of DIG-labeled antisense RNA probes was performed using standard methods. RNA *in situ* hybridization (ISH) was carried out as previously described (Thisse and Thisse, 2008).

Characterization of lens defects

To quantify lens equatorial width, bright field images from live embryos were taken with Nikon ECLIPSE Ni microscope and measured with the Nikon Elements BR measurement tool. At least 20 embryos of each genotype were measured under identical conditions. To quantify number and convexity of lens fiber cell, mRFP mRNA was injected into embryo at 1-cell stage and at 28 hours post-fertilization (hpf), confocal images for equatorial sections from control and *ddx39ab* mutant embryos were taken. Cell number on optical section was counted manually. Cell convexity was calculated with 3D Convex Hull plug-in on ImageJ (Sheets et al., 2011). The convexity was defined as ration between convex surface area and surface area.

RNA-seq for Expression and Splicing Analysis.

Total RNA from 100 control and *ddx39ab* mutant 24 hpf embryos was isolated by TRIzol reagent (Invitrogen). Two biological replicates for each group (control and *ddx39ab* mutant) were processed and sequenced. Sequencing libraries were prepared using the Nextera sample preparation kit (Illumina) and subjected to HiSEQ paired-end 100 bp plus sequencing. Resulting reads were aligned to the zebrafish reference genome (GRCz10) and gene expression quantified using TopHat V2.2.1 and Bowtie2 v2.2.3 (Kim

et al., 2013; Langmead and Salzberg, 2012). Gene differential expression was analyzed using HTSeq v 0.6.1p1 (Anders et al., 2015). Genes showing altered expression with adjusted $P < 0.05$ were considered differentially expressed. For the set of differentially expressed genes a functional analysis was performed using Ingenuity Pathway Analysis Software and DAVID (Huang da et al., 2009), and some of the enriched processes were selected according to relevant criteria related to the biological process studied. Using a R visualization package called GOPlot (Walter et al., 2015), a chord plot was generated to better visualize the relationships between genes and the selected enriched processes. OLego (Wu et al., 2013) and Quantas pipelines were used for alternative splicing analysis. Transcript structure was inferred between paired-end reads. Alternative splicing (AS) was quantified by separating genomic and junction reads and scoring the output from transcript inference. Finally, statistical tests were run to filter the significant AS events [Fisher's exact test and Benjamini false discovery rate (FDR)]. RNA-seq data have been uploaded to the Gene Expression Omnibus (GEO: GSE97067).

RNA immunoprecipitation and RIP-seq

RNA immunoprecipitation (RIP) was performed as previously described (Jain et al., 2011) with the specific modifications below. Ddx39ab-Flag mRNA was injected into embryos from *ddx39ab* heterozygous fish in-crosses, with mutant embryos sorted based on GFP brightness. De-yolked embryos were homogenized in RIP buffer and briefly sonicated using a probe-tip Branson sonicator to solubilize chromatin. Each sample was normalized for total protein amount then Flag-Ddx39ab and associated RNA was isolated via incubation with anti-Flag agarose beads (Sigma) for 6 hours at 4 °C with gentle rotation. Samples were washed sequentially in high stringency buffer, high salt buffer and RIP buffer. Ddx39ab-associated RNA was extracted with Trizol and then processed for sequencing. Sequencing libraries were prepared using the Nextera

sample preparation kit (Illumina) and subjected to HiSEQ paired-end 100 bp plus sequencing. Data analysis was performed as previously described.

Reverse transcription-PCR analysis of splicing

Total RNA was prepared from control and *ddx39ab* mutant larvae at 36 hpf. Reverse transcription (RT)-PCR was performed using total RNA prepared as described above to monitor the splicing of pre-mRNAs. The primer pairs and detailed PCR conditions used to amplify each of the genes are listed in Supplementary Tables 1 and 2.

ChIP-qPCR

Chromatin immunoprecipitation (ChIP) assays were performed as previously described (Lindeman et al., 2009). In brief, at 24 hpf, *ddx39ab* mutant and control embryos were collected, de-yolked and cross-linked with 1% formaldehyde for 10 min at room temperature and subsequently quenched with glycine to a final concentration of 0.125 M for another 10 min. Chromatin was sonicated with a Bioruptor (Diagenode), cleared by centrifugation, and incubated overnight at 4 °C with 5 mg anti-H3K4me1 antibody (Abcam). Immunocomplexes were immobilized with 100 ul of protein-G Dynal magnetic beads (Abcam) for 4 hours at 4 °C, followed by stringent washes and elution. Eluates were reverse cross-linked overnight at 65 °C and deproteinated. DNA was extracted with phenol chloroform, followed by ethanol precipitation. H3K4me1 occupied regions at 24 hpf were retrieved from a previously reported data set (Bogdanovic et al., 2012). ChIP-qPCR analyses were performed using a Light Cycler 480II machine (Roche). ChIP-qPCR signals were calculated as percentage of input. All primers used in qPCR analyses are shown in Table S1.

Results

Loss of ddx39ab leads to an embryonic lethal phenotype in zebrafish

In a *Tol2* transposon-mediated gene-trapping screen to identify novel genes involved in cardiovascular system development (Hou et al., 2017), we identified a zebrafish line, RT-011, in which embryos demonstrated a dynamic GFP expression pattern, with strong signal evident in somites and eyes from the 8 somite stage (13 hours post-fertilization, hpf). As development proceeded, GFP signal also emerged in heart (Fig. 1A). Multiple incrosses of RT-011 heterozygotes yielded wild type (WT), heterozygous, and homozygous mutant embryos at expected Mendelian ratios, with all homozygous mutant embryos dying by 4 days post-fertilization (dpf). This indicated that the RT-011 trap line carried a recessive lethal allele. Homozygous mutant embryos from incrosses of heterozygous RT-011 fish showed no obvious morphological defects until after 24 hpf (Fig. 1B), at which point contraction of the definitive heart tube was extremely weak and irregular in mutants (Movie S1). Homozygous mutant embryos were completely paralyzed at 24 hpf, and did not exhibit spontaneous tail movements or response to tactile stimulation (Movie S2). At later stages of development further defects became apparent, including a failure to establish blood circulation, a curved body axis, a prominent cardiac edema, disorganized myotome and extensive cell death (Fig. 1B and Fig.S1). 5' RACE was used to identify the gene trapped in the RT-011 line. Sequencing indicated that the gene-trapping element was integrated within the 2nd intron of *ddx39ab* locus (Fig. 1C). The zebrafish Ddx39ab protein contains DEXDc (for ATP binding and hydrolysis) and HELICc (helicase superfamily c-terminal) domains; both of which are highly conserved (over 90% amino acid identity) between zebrafish and

human (Fig. S2). The gene trap insertion resulted in a transcript that encoded a fusion protein containing the first N-terminal 69 amino acids of Ddx39ab (Fig. 1D). Since this fusion protein lacked both the DEXDc and HELICc domains, the allele we identified from the RT-011 line should act as a true null allele. RT-PCR result clearly showed the presence of the *ddx39ab-GFP* fusion transcript in heterozygous RT-011 embryos and absence of wild type ddx39ab transcript in homozygous embryos (Fig 1.E). To confirm that mutation of *ddx39ab* represents the causal event in the homozygous RT-011phenotype, expression level of *ddx39ab*, *ddx39aa* and *ddx39b* was examined, and only expression of *ddx39ab* was found to be absent in homozygous gene trap embryos (Fig. 1F). Furthermore, mRNA encoding wild-type Ddx39ab was injected into embryos from incrosses of heterozygous RT-011 fish. Morphological homozygous RT-011 defects were efficiently rescued by *ddx39ab* RNA (Fig. 1G and H), with injected mutant embryos surviving up to 9 dpf. Taken together, this indicated that developmental abnormalities in RT-011 homozygous embryo resulted from mutation of *ddx39ab*.

To further confirm the endogenous embryonic expression of *ddx39ab*, whole mount RNA *in situ* hybridization was carried out on wild type embryos. *Ddx39ab* demonstrated strong expression during early embryogenesis, with abundant transcript localized to the myotome, heart and eyes. As development proceeded, the expression of *ddx39ab* was constrained to several tissues including pharyngeal arches and liver (Fig. S2). The dynamic expression pattern of *ddx39ab* indicated possible roles in multiple developmental stages and tissues.

Developmental defects in muscular organs and lens in ddx39ab mutants

Based on the expression pattern during embryogenesis and defects displayed in *ddx39ab* homozygous mutants, we examined development of the heart, skeletal muscle and lens in *ddx39ab* mutant embryos.

At 36 hpf, the heart tube in wild type embryos had undergone looping and chamber ballooning, becoming a functional two-chamber pump. In contrast, hearts of *ddx39ab* mutants were morphologically abnormal, with little looping or chamber emergence evident (Fig. 2A left panels). Immunostaining revealed that the cardiac sarcomere was severely dis-organized in *ddx39ab* mutants as compared to control embryos (Fig. 2A mid and right panels). Gene expression analysis of a number of cardiac markers at 26 hpf demonstrated that loss of *ddx39ab* did not cause a readily apparent decrease in expression of many cardiogenic regulatory genes, including *nkx2.5*, *gata5*, *tbx5a* and *bmp4* (Fig. 2B and data not shown*).* Specification of both heart chambers occurred properly in *ddx39ab* mutants, as shown by expression of *myh6* and *myh7* (Fig. 2C). In contrast, *ddx39ab* mutant embryos demonstrated significantly reduced expression of *nppa*, a gene associated with maturation of the heart tube (Auman et al., 2007). Expression of a number of cardiac sarcomere structural genes was also distinctly down-regulated in *ddx39ab* mutant embryos, including *myh6*, *cmlc1*, *acta1b* and *ttn.2* (Fig. 2C). Decreased expression of these cardiac sarcomere components may represent the causal factor for the weak contractility observed in *ddx39ab* mutant hearts.

Next, we examined whether the locomotion defect in *ddx39ab* mutant embryos reflected altered skeletal muscle organization. Immunohistochemical staining for Myosin Heavy Chain (MF20 antibody) and F-Actin revealed myofibrillar protein assembly was severely disrupted in *ddx39ab* mutant embryos when compared with wild type (Fig. 3A). To analyze which step of muscle development was affected, *ddx39ab* mutant embryos were analyzed for myoblast differentiation via RNA *in situ* hybridization for expression of genes encoding myogenic regulatory factors (Bentzinger et al., 2012). Expression of early myogenic specification markers (*myoD* and *myf5*) and late differentiation markers (*myogenein* and *myf6*) appeared normal until 32 hpf (Fig. 3B and data not shown). Similar to the myocardium, at 32 hpf, in *ddx39ab* mutant embryos the expression of a battery of sarcomeric components, including *tnnb*, *myhz2* and *smyhc*, was significantly reduced (Fig. 3 C and D). This indicated that maturation of both fast muscle and slow muscle was compromised in *ddx39ab* mutants. We also found reduction in transcript levels of *casq1a*, which encodes a calcium-binding protein of the skeletal muscle sarcoplasmic reticulum (Yazaki et al., 1990), and *slc25a4*, encoding a muscle cell specific mitochondrial ATP-ADP carrier (Gutierrez-Aguilar and Baines, 2013), in *ddx39ab* mutant embryos (Fig. 3 C). These results suggested that *ddx39ab* plays an important role in skeletal muscle maturation and function.

As *ddx39ab* demonstrated strong expression in developing eyes, we next examined development of retina and lens in *ddx39ab* mutants. Sectioning of retinas at 32 hpf revealed that pigmented epithelium formed and retinal neuroepithelium displayed normal histological features in *ddx3ab* mutants. Further analysis revealed that retinal neuronal production (marked by *atoh7* expression) and ganglion differentiation (marked by and *lhx3* and Alcama expression) were initiated normally in *ddx39ab* mutant retina (Fig. S4). These data indicated that development of the retina in *ddx39ab* mutants is

largely normal. At 28 hpf, the lens of *ddx39ab* mutant embryos displayed no obvious defects at the level of gross morphology, with cell number being comparable in *ddx39ab* mutant and wild type embryos (Fig. 4A). During eye morphogenesis, cells in the center of the lens mass retain spheroidal morphology; differentiating as primary fibers (which originate from the central lens placode) that elongate and wrap around the ovoid-shaped cells in the center, resulting in a crescent shaped layers of fibers (Greiling and Clark, 2009). In contrast to wild type, in *ddx39ab* mutants a majority of lens fiber cells did not form crescent shaped layers, showed relatively irregular and convex shape with significant higher convexity measure (Fig. 4A, lower panels and Fig. S5). We next analyzed expression of lens-specific genes to explore the nature of the disorganization of primary fiber cells in *ddx39ab* mutant embryos. We first examined the expression of a cascade of upstream transcription factors that drive fiber cell differentiation, including *prox1a*, *foxe3* and *pitx3* (Cvekl and Duncan, 2007; Greiling and Clark, 2012; Pillai-Kastoori et al., 2015). Results showed that in *ddx39ab* mutants the transcription of *foxe3* was mildly upregulated, whereas no overt change in *pitx3* and *prox1a* expression were observed (Fig. 4B). During lens fiber cell elongation, soluble proteins known as crystallins are abundantly expressed in lens fibers to increase the refractive index and contribute to transparency (Clark, 2004). We found expression of crystallin genes, including *cryaa*, *cry2d10* and *cryg2d1*, was dramatically down-regulated in *ddx39ab* mutant embryos (Fig. 4C). Differentiated lens fiber cells express a set of cell-cell adhesion molecules required to create refractive index matching of lens membranes and cytoplasm (Bassnett et al., 2011). *In situ* hybridization and qPCR demonstrated that in *ddx39ab* mutant embryos the expression of *lim2.4*, a lens specific receptor for Calmodulin involved in cell junction organization, was notably down-regulated (Fig. 4C). BFSP1 (also known as Filensin) and BFSP2 (also known as Phakinin) are assembly partners of the beaded-chain filament, a type of lens-specific cytoskeletal element.

Significant down-regulation of these two genes in *ddx39ab* lens was also observed (Fig. 4C). These data indicated *ddx39ab* is also indispensable for lens fiber cell terminal differentiation.

The above results indicated *ddx39ab* deletion causes dysregulated expression of structural genes during terminal differentiation of cardiomyocyte, myocyte and lens fiber cells. This hampered cell differentiation in turn leads to multiple defects in the muscular organs and eyes, which likely collectively contribute to the lethal phenotype in *ddx39ab* mutant zebrafish embryos.

Changes in the transcriptomic landscape in ddx39ab mutants

Dead box RNA helicases have been shown to exert their biological function through regulating multiple facets of RNA metabolism (Linder and Jankowsky, 2011). In order to determine how mutation of *ddx39ab* affects the zebrafish embryo transcriptome, we pursued RNA-Seq on wild type and *ddx39ab* mutant embryos at 24 hpf. A total of 878 genes, consisting of 548 with decreased and 330 with increased expression, were significantly altered in the *ddx39ab* mutant embryos (fold change (FC) >2, FDR < 0.05) (Fig. 5A; Table S3). Following gene ontology (GO) analysis, we found that genes down-regulated in *ddx39a* mutants were enriched for GO terms linked to development of muscular tissue and lens (Fig. 5B). Interestingly, consistent with previous observations revealed by *in situ* hybridization, in all three cell types (cardiomyocyte, myocyte and lens fiber cells), the expression of upstream regulatory factors showed no evident change while the expression of many structural constituents was greatly diminished (highlighted in Fig. 5B). These results further suggested that *ddx39ab* plays a common and important role in the terminal differentiation of cardiomyocyte, myocyte and lens fiber cells.

To gain further insight into the impact on transcript levels caused by loss of *ddx39ab*, we examined alternative splicing (AS) by using the OLego program and the Quantas pipeline. 12,236 significant AS events (|dI| > 0.1; Table S4) were scored in *ddx39ab* mutant embryos compared with controls. These events include skipping/inclusion of single cassette exons or tandem cassette exons, intron retention and the use of alternative 5′ or 3′ splice sites (Table S4, number of events shown in Fig. 5C). These data showed that loss of Ddx39ab function leads to extensive intron retention events and cassette skipping, implicating Ddx39ab functions in intron definition. The mixed splicing activities (skipping and inclusion) observed in *ddx39ab* mutants clearly suggest context-dependent splicing defects result from the loss of *ddx39ab*.

Ddx39ab deletion affects pre-mRNA splicing of KMT2 family genes

We proposed that defining the RNA interactome of Ddx39ab would reveal insights into the molecular mechanisms underlying the phenotypes in *ddx39ab* mutant embryos. To systematically identify Ddx39a-associated RNAs, RNA Immunoprecipitation sequencing (RIP-seq) was performed in *ddx39ab* mutant embryos injected with mRNA coding Flag-tagged Ddx39ab (Flag–Ddx39ab). Sequencing results showed Ddx39ab interacts with a diverse set of RNAs (Table S6), of which mRNAs were highly represented (73.1%), with snoRNAs and rRNA contributed 16.8% and 4.5%, respectively, of bound transcripts (Fig. 5D). Comparison of data from RIP-seq and RNA seq revealed that 84% Ddx39a-associated mRNAs showed alternative splicing in *ddx39ab* mutants (which we refer to as high confidence Ddx39ab target transcripts), underscoring the function of Ddx39a in pre-mRNA splicing. Gene ontology term and pathway analysis linked these Ddx39a-associating mRNAs to histone modification. Among these potential Ddx39ab target mRNAs, we noticed members of KMT2 gene family (*kmt2a*, *kmt2ba*, *kmt2bb* and *kmt2cb*)

were prominent. KMT2 family members methylate lysine 4 on the histone H3 tail, a critical regulatory step associated with myogenesis through modulation of chromatin structure and DNA accessibility (Lee et al., 2013; Rao and Dou, 2015).

To confirmed interactions between Ddx39ab and KMT2 family members mRNAs, RNA immunoprecipitation and quantitative reverse transcription PCR (RIP-qPCR) was applied. This further verified that Ddx39ab could bind to these mRNAs (Fig. 5E). We further examined splicing events of KMT2 family genes in control and *ddx39ab* mutant embryos by using RT-PCR analysis (Rios et al., 2011; Rosel et al., 2011). This analysis showed that unspliced mRNAs were retained at higher levels in *ddx39ab* mutants at 24 hpf (Fig. 6A), suggesting that the pre-mRNA splicing of these genes was defective. As a control, we confirmed that in *ddx39ab* mutants the splicing of the housekeeping gene *actb1* was normal. These results suggest that in *ddx39ab* mutants the effect of pre-mRNA splicing may be specific to a certain set of genes.

KMT2 enzymes are the major histone methyltransferases responsible for mono-methylation at lysine 4 of Histone 3 (H3K4me1) at distal enhancers and regions flanking the transcription start site (TSS) (Rao and Dou, 2015). To investigate whether loss of *ddx39ab* affects the presence of H3K4me1 on actively transcribed genes, CHIP-qPCR was performed for a selected set of genes that showed altered transcript levels in control and *ddx39ab* mutant embryos at 24 hpf. Whereas global H3K4me1 levels in *ddx39ab* mutants showed only minor decreases as compared to wild type (Fig. S3), the H3K4me1 occupancy on regions flanking the TSS of myocyte and cardiomyocyte specific genes (*acta1b*, *myhz2*, *nppa*, *myom1a*, *tnnt2d*, *mylpfb* and *smhyc1*) was significantly reduced in mutants (Fig. 6B). Interestingly, the H3K4me1 occupancy level in the TSS region of myogenic regulatory factors (such as *myod*) were comparable to control embryo in *ddx39a* mutants, consistent with analysis of transcript and protein levels via RNA-seq, *in situ* hybridization and Western blot (Fig. 6B and Fig. S3).

These results support a context-dependent requirement of Ddx39ab for proper pre-mRNA splicing of KMT2 family members, with loss of *ddx39ab* leading to a failure in establishment of the epigenetic status required for terminal differentiation.

Discussion

Members of DEAD box RNA helicase family have been shown to be involved in nearly all aspects of RNA metabolism, from transcription to mRNA decay. Beyond the scale of biochemistry and cell biology, recent studies have started to delineate different functions of DEAD box RNA helicase in variant biological scenarios including animal development. For example, *ddx46* has been shown to be expressed in the digestive organs and brain and is required for development of these organs (Hozumi et al., 2012). Researchers have also reported that *ddx18* is essential for hematopoiesis (Payne et al., 2011). The generation of the *ddx39ab* gene-trapping allele provided the opportunity to study the function of this gene in vertebrate development. Here we demonstrated that *ddx39ab* is required for normal gene expression and differentiation of cardiomyocyte, myocyte and lens fiber cells. In previous studies, development and function of heart has proven to be sensitive to defects in RNA metabolism (Ding et al., 2004; Xu et al., 2005). Our observations corroborate these results and suggest that myocyte and lens fiber differentiation similarly bear this cell type specific susceptibility. One possible explanation to this observation is these three type cells exploited a similar, Kmt2 family-depended mechanism to establish epigenetic status during differentiation. An alternative hypothesis is that compared to other cell types (such as neurons), a larger fraction of splicing undergoes Ddx39ab-related regulation in these three types of cell. FACS-based cell sorting followed by RIP-seq may identify cell type specific Ddx39ab-associated transcripts, with this information helping to clarify the underlying mechanisms of cell type specific susceptibility.

RNA-binding proteins (RBP) including DEAD box RNA helicases modulate splicing primarily by positively or negatively regulating splice-site recognition by the spliceosome. The recognition of RBP on pre-mRNAs relies on the distinct regulatory sequences in pre-mRNA that function as splicing enhancers or silencers. Our data showed Ddx39ab could bind to a battery of mRNAs while more sophisticated molecular biology and bioinformatics efforts need to be exploited to unfold the detail mechanism of Ddx39a-mediating splicing events such as how is the specificity defined.

During the development, *ddx39ab* shows a complex and dynamic expression pattern (Fig. 1and Fig. S2). This result prompts future analysis of the role of *ddx39ab* at later developmental events (e.g. formation of pharyngeal arches and development of digestive organs). However, the severe defects observed in *ddx39ab* mutants at early stages of development complicates the analysis of later developmental events. CRIPSR/Cas9-based generation of tissue-specific *ddx39ab* mutants, or generation of a conditional *ddx39ab* allele, will be required to study *ddx39ab* function in other organs/tissues.

Previous studies showed *ddx39ab* acts as a growth-associated factor in cancer cells that is required for genome integrity and telomere protection (Sugiura et al., 2007; Yoo and Chung, 2011). We did not detect abnormal mitoses in *ddx39ab* mutant embryos (Sup Fig 3). However, analysis of later stage *ddx39ab* mutants may be required to observe telomeric defects, after a sufficient number of cell divisions have occurred. It also could be true that the telomere protection function of *ddx39ab* is not conserved between fish and mammals.

It was interesting to note that among a relative small number of genes significantly down-regulated in *ddx39ab* mutant embryos, a large proportion encoded structural constituents including sarcomeric components in muscle cells and crystallin genes in lens fiber cells. In addition to regulating mRNA splicing, it is possible that

Ddx39ab is involved in transactivation of structural components of via an uncharacterized mechanism that is exploited in both muscular and lens fiber cells. Several studies have shown that DEAD box proteins play important roles as regulators of transcription, particularly as co-activators or co-suppressors of transcription (Fuller-Pace and Nicol, 2012; Huang et al., 2015). Further investigation to determine proteins Ddx39a directly binds in various cell types, and the functional consequences of these interactions, should provide important insight into how the specificity of Ddx39ab function is regulated.

Acknowledgement

We thank Qi Xiao for drawing schematic pictures, Peipei Yin for zebrafish husbandry and members of the Scott lab at University of Toronto and Di Chen lab at Nanjing University for feedback and help during this project. This research was undertaken, in part, thanks to grant funding from National Natural Science Foundation of China (NSFC 31471354 and NSFC 31671505 to X.L), and grant funding from the Natural Sciences and Engineering Research Council of Canada (RGPIN 2017-06502 to I.C.S.).

Author contributions

X.L. and I.C.S. conceived and supervised the project and wrote the manuscript. L. Z., Y.Y., B. L. and X.L. performed all experiments.

Competing financial interests

The authors declare no competing financial interests.

References

Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Auman, H.J., Coleman, H., Riley, H.E., Olale, F., Tsai, H.J., Yelon, D., 2007. Functional modulation of cardiac form through regionally confined cell shape changes. PLoS biology 5, e53.

Bassnett, S., Shi, Y., Vrensen, G.F., 2011. Biological glass: structural determinants of eye lens transparency. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 366, 1250-1264.

Bentzinger, C.F., Wang, Y.X., Rudnicki, M.A., 2012. Building muscle: molecular regulation of myogenesis. Cold Spring Harbor perspectives in biology 4.

Bleichert, F., Baserga, S.J., 2007. The long unwinding road of RNA helicases. Molecular cell 27, 339-352.

Bogdanovic, O., Fernandez-Minan, A., Tena, J.J., de la Calle-Mustienes, E., Hidalgo, C., van Kruysbergen, I., van Heeringen, S.J., Veenstra, G.J., Gomez-Skarmeta, J.L., 2012. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. Genome Res 22, 2043-2053.

Calo, E., Flynn, R.A., Martin, L., Spitale, R.C., Chang, H.Y., Wysocka, J., 2015. RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. Nature 518, 249-253.

Clark, J.I., 2004. Order and disorder in the transparent media of the eye. Experimental eye research 78, 427-432.

Clark, K.J., Balciunas, D., Pogoda, H.M., Ding, Y., Westcot, S.E., Bedell, V.M., Greenwood, T.M., Urban, M.D., Skuster, K.J., Petzold, A.M., Ni, J., Nielsen, A.L., Patowary, A., Scaria, V., Sivasubbu, S., Xu, X., Hammerschmidt, M., Ekker, S.C., 2011. In vivo protein trapping produces a functional expression codex of the vertebrate proteome. Nature methods 8, 506-515.

Cvekl, A., Duncan, M.K., 2007. Genetic and epigenetic mechanisms of gene regulation during lens development. Progress in retinal and eye research 26, 555-597.

Ding, J.H., Xu, X., Yang, D., Chu, P.H., Dalton, N.D., Ye, Z., Yeakley, J.M., Cheng, H., Xiao, R.P., Ross, J., Chen, J., Fu, X.D., 2004. Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. The EMBO journal 23, 885-896.

Fleckner, J., Zhang, M., Valcarcel, J., Green, M.R., 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. Genes & development 11, 1864-1872.

Fuller-Pace, F.V., 2013. DEAD box RNA helicase functions in cancer. RNA Biol 10, 121-132.

Fuller-Pace, F.V., Nicol, S.M., 2012. DEAD-box RNA helicases as transcription cofactors. Methods in enzymology 511, 347-367.

Greiling, T.M., Clark, J.I., 2009. Early lens development in the zebrafish: a three-dimensional time-lapse analysis. Developmental dynamics : an official publication of the American Association of Anatomists 238, 2254-2265.

Greiling, T.M., Clark, J.I., 2012. New insights into the mechanism of lens development using zebra fish. International review of cell and molecular biology 296, 1-61.

Gutierrez-Aguilar, M., Baines, C.P., 2013. Physiological and pathological roles of mitochondrial SLC25 carriers. The Biochemical journal 454, 371-386.
Hirabayashi, R., Hozumi, S., Higashijima, S., Kikuchi, Y., 2013. Ddx46 is required for multi-lineage differentiation of hematopoietic stem cells in zebrafish. Stem Cells Dev 22, 2532-2542.
Hou, N., Yang, Y., Scott, I.C., Lou, X., 2017. The Sec domain protein Scfd1 facilitates trafficking of ECM components during chondrogenesis. Dev Biol 421, 8-15.
Hozumi, S., Hirabayashi, R., Yoshizawa, A., Ogata, M., Ishitani, T., Tsutsumi, M., Kuroiwa, A., Itoh, M., Kikuchi, Y., 2012. DEAD-box protein Ddx46 is required for the development of the digestive organs and brain in zebrafish. PloS one 7, e33675.
Huang da, W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44-57.
Huang, W., Thomas, B., Flynn, R.A., Gavzy, S.J., Wu, L., Kim, S.V., Hall, J.A., Miraldi, E.R., Ng, C.P., Rigo, F.W., Meadows, S., Montoya, N.R., Herrera, N.G., Domingos, A.I., Rastinejad, F., Myers, R.M., Fuller-Pace, F.V., Bonneau, R., Chang, H.Y., Acuto, O., Littman, D.R., 2015. DDX5 and its associated lncRNA Rmrp modulate TH17 cell effector functions. Nature 528, 517-522.
Jain, R., Devine, T., George, A.D., Chittur, S.V., Baroni, T.E., Penalva, L.O., Tenenbaum, S.A., 2011. RIP-Chip analysis: RNA-Binding Protein Immunoprecipitation-Microarray (Chip) Profiling. Methods Mol Biol 703, 247-263.
Jarmoskaite, I., Russell, R., 2011. DEAD-box proteins as RNA helicases and chaperones. Wiley interdisciplinary reviews. RNA 2, 135-152.
Jarmoskaite, I., Russell, R., 2014. RNA helicase proteins as chaperones and remodelers. Annual review of biochemistry 83, 697-725.
Kawakami, K., Asakawa, K., Muto, A., Wada, H., 2016. Tol2-mediated transgenesis, gene trapping, enhancer trapping, and Gal4-UAS system. Methods Cell Biol 135, 19-37.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology 14, R36.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9, 357-359.
Lee, J.E., Wang, C., Xu, S., Cho, Y.W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W., Ge, K., 2013. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Elife 2, e01503.
Lindeman, L.C., Vogt-Kielland, L.T., Alestrom, P., Collas, P., 2009. Fish'n ChIPs: chromatin immunoprecipitation in the zebrafish embryo. Methods Mol Biol 567, 75-86.
Linder, P., Jankowsky, E., 2011. From unwinding to clamping - the DEAD box RNA helicase family. Nat Rev Mol Cell Biol 12, 505-516.
Luo, M.L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., Reed, R., 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. Nature 413, 644-647.
Meignin, C., Davis, I., 2008. UAP56 RNA helicase is required for axis specification and cytoplasmic mRNA localization in Drosophila. Dev Biol 315, 89-98.

Payne, E.M., Bolli, N., Rhodes, J., Abdel-Wahab, O.I., Levine, R., Hedvat, C.V., Stone, R., Khanna-Gupta, A., Sun, H., Kanki, J.P., Gazda, H.T., Beggs, A.H., Cotter, F.E., Look, A.T., 2011. Ddx18 is essential for cell-cycle progression in zebrafish hematopoietic cells and is mutated in human AML. Blood 118, 903-915.
Pillai-Kastoori, L., Wen, W., Morris, A.C., 2015. Keeping an eye on SOXC proteins. Developmental dynamics : an official publication of the American Association of Anatomists 244, 367-376.
Rao, R.C., Dou, Y., 2015. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. Nat Rev Cancer 15, 334-346.
Rios, Y., Melmed, S., Lin, S., Liu, N.A., 2011. Zebrafish usp39 mutation leads to rb1 mRNA splicing defect and pituitary lineage expansion. PLoS genetics 7, e1001271.
Rocak, S., Linder, P., 2004. DEAD-box proteins: the driving forces behind RNA metabolism. Nat Rev Mol Cell Biol 5, 232-241.
Rosel, T.D., Hung, L.H., Medenbach, J., Donde, K., Starke, S., Benes, V., Ratsch, G., Bindereif, A., 2011. RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation. The EMBO journal 30, 1965-1976.
Sarkar, M., Ghosh, M.K., 2016. DEAD box RNA helicases: crucial regulators of gene expression and oncogenesis. Front Biosci (Landmark Ed) 21, 225-250.
Schmittgen, T.D., Livak, K.J., 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3, 1101-1108.
Sheets, K.G., Jun, B., Zhou, Y., Winkler, J., Zhu, M., Petasis, N., Gordon, W.C., Bazan, N.G., 2011. Topical Neuroprotectin D1 Attenuates Experimental CNV And Induces Activated Microglia Redistribution. Investigative Ophthalmology & Visual Science 52, 5470-5470.
Shen, H., Zheng, X., Shen, J., Zhang, L., Zhao, R., Green, M.R., 2008. Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. Genes & development 22, 1796-1803.
Shen, J., Zhang, L., Zhao, R., 2007. Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. The Journal of biological chemistry 282, 22544-22550.
Sugiura, T., Sakurai, K., Nagano, Y., 2007. Intracellular characterization of DDX39, a novel growth-associated RNA helicase. Experimental cell research 313, 782-790.
Thisse, C., Thisse, B., 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc 3, 59-69.
Walter, W., Sanchez-Cabo, F., Ricote, M., 2015. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics 31, 2912-2914.
Westerfield, M., 1993. The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish Danio (Brachydanio) rerio. . University of Oregon Press, Oregon.
Wu, J., Anczukow, O., Krainer, A.R., Zhang, M.Q., Zhang, C., 2013. OLego: fast and sensitive mapping of spliced mRNA-Seq reads using small seeds. Nucleic acids research 41, 5149-5163.
Xu, X., Yang, D., Ding, J.H., Wang, W., Chu, P.H., Dalton, N.D., Wang, H.Y., Bermingham, J.R., Jr., Ye, Z., Liu, F., Rosenfeld, M.G., Manley, J.L., Ross, J., Jr., Chen, J., Xiao, R.P., Cheng, H., Fu, X.D., 2005. ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Cell 120, 59-72.

Yazaki, P.J., Salvatori, S., Sabbadini, R.A., Dahms, A.S., 1990. Calsequestrin, an intracellular calcium-binding protein of skeletal muscle sarcoplasmic reticulum, is homologous to aspartactin, a putative laminin-binding protein of the extracellular matrix. Biochemical and biophysical research communications 166, 898-903.
Yoo, H.H., Chung, I.K., 2011. Requirement of DDX39 DEAD box RNA helicase for genome integrity and telomere protection. Aging Cell 10, 557-571.
Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N., Vreven, T., Meignin, C., Davis, I., Zamore, P.D., Weng, Z., Theurkauf, W.E., 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery, Cell, pp. 871-884.

Figures

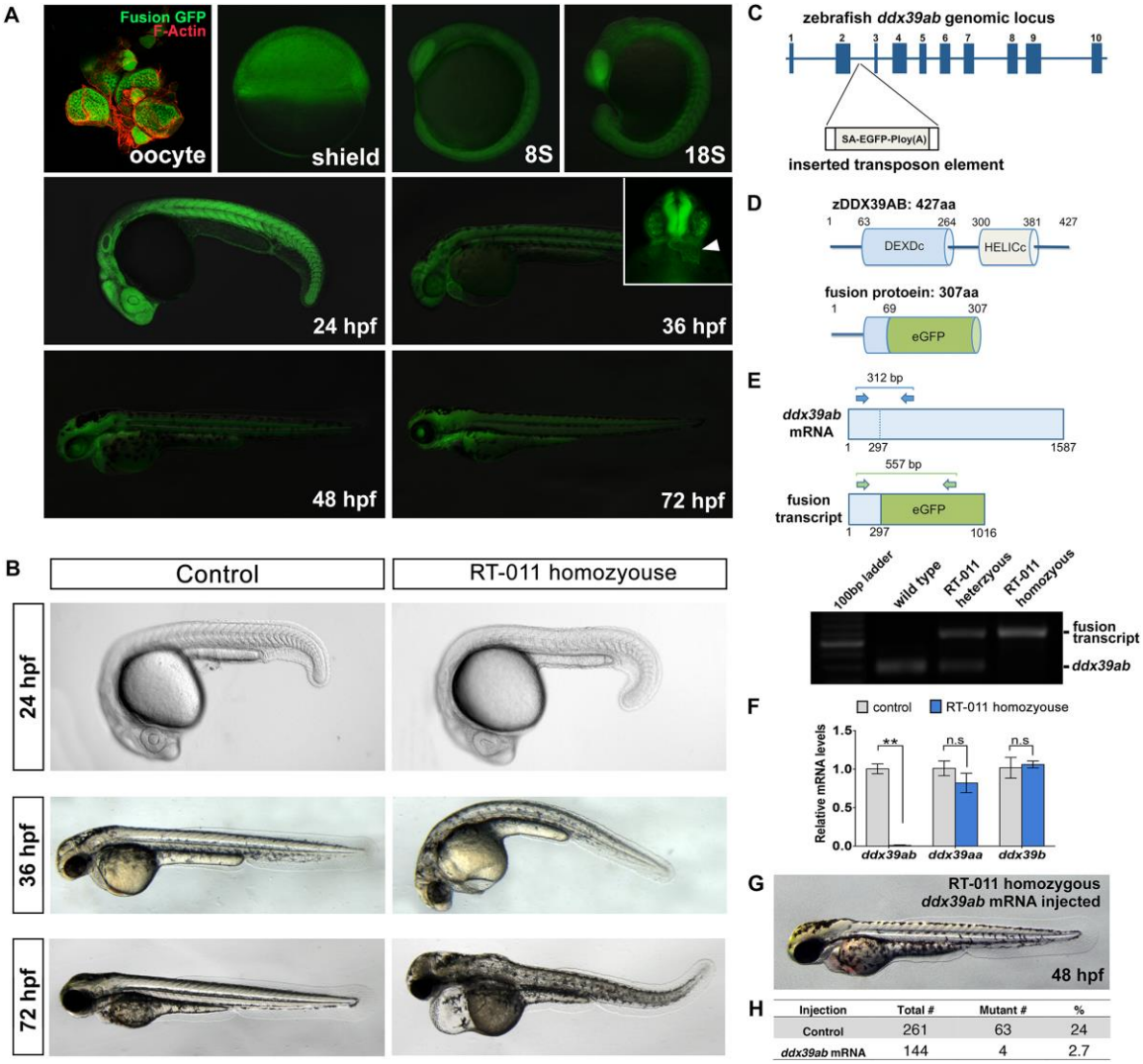


Figure 1. Loss of *ddx39a* leads to an embryonic lethal phenotype in zebrafish. (A): GFP expression pattern in zebrafish trapping line RP-011. Lateral views, hpf, hour post-fertilization. (B): Morphological defects in RP-011 homozygous mutant embryos: RP-011 homozygous embryo did not show morphological defect until 24 phf; at 48 hpf, RP-011 homozygous embryo showed curved body axis; at 72 hpf, RP-011 homozygous embryo showed heart edema, dysmorphic jaw and global degeneration. At least 80 embryos for each genotype were observed and representative samples are showed. (C): The zebrafish *ddx39a* genomic locus. The transposon is inserted after the second intron. (D): Zebrafish DDX39A protein contains DEXDc and HELICc domains. Insertion of the gene trapping element resulted a fusion transcript which will be translated into a fusion protein contain first 69 amino acid from fish DDX39A on the N terminal. (E): RT-PCR results showed the presence of fusion transcript in RP-011 embryo. (F): qPCR results showed absence of *ddx39ab* mRNA in RP-011 homozygous embryo. Data are mean ± SEM. ***: p-value <0.001. ns: not significant. (G): Injection of *ddx39a* mRNA rescue the developmental defects in RP-011 homozygous embryos. (H): Quantification of the percentage of embryo of various genotypes showing development defects in a clutch.

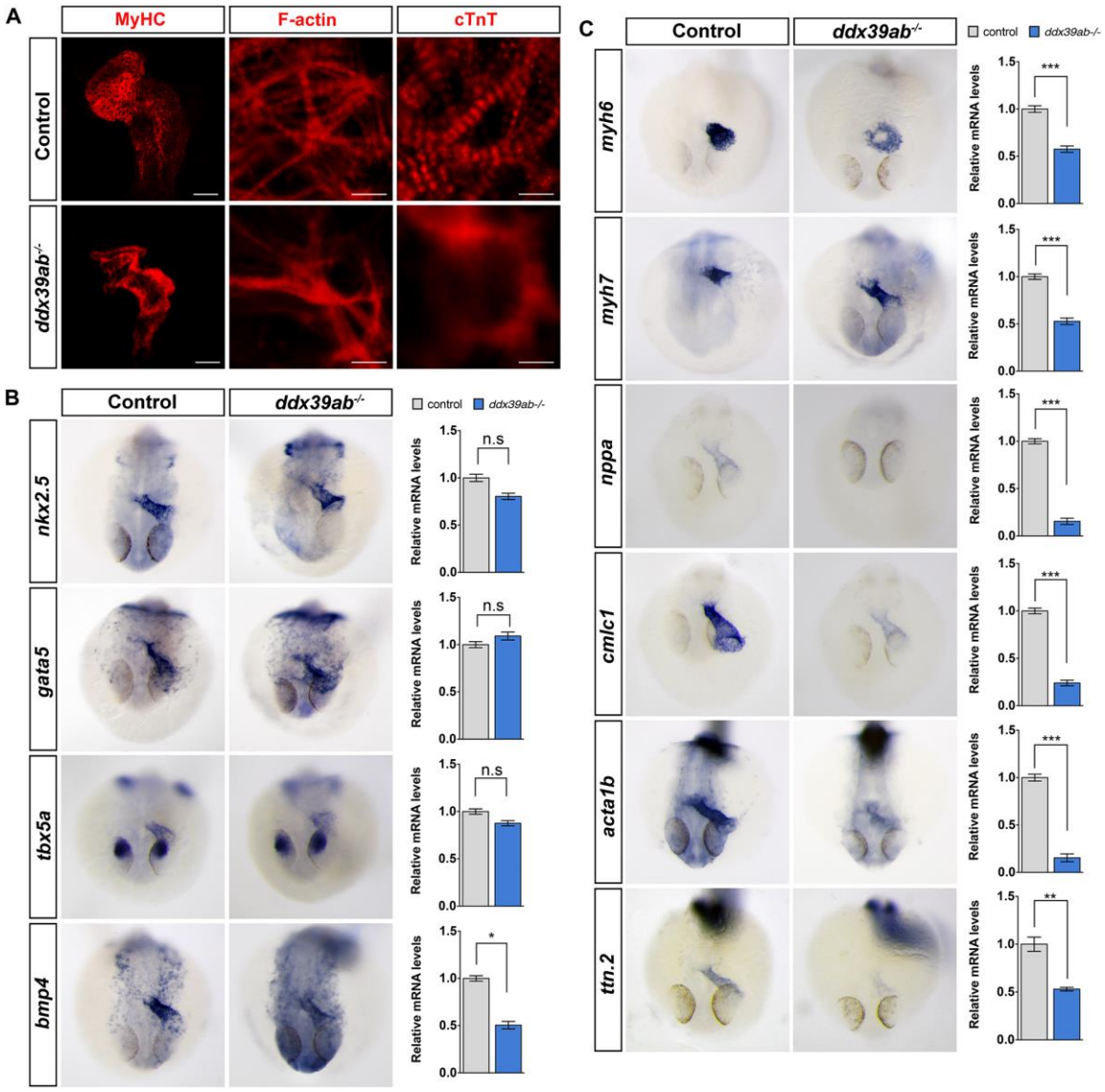


Figure 2. Loss of *ddx39a* leads to cardiomyocyte differentiation defects in zebrafish embryos. (A): Ventral view of heart from whole-mount wide type or *ddx39a* mutant zebrafish embryos at 36 hpf, with cranial to the top. Myocardium was labeled with MF20 antibody. Scale bars: 200um. (B): RNA *in situ* hybridization and qPCR results are shown for cardiogenic regulatory gene expression in WT and *ddx39a* mutant zebrafish embryos. (C): RNA *in situ* hybridization and qPCR results are shown for cardiomyocyte structural gene expression in WT and *ddx39a* mutant zebrafish embryos. B and C, frontal views with dorsal side to the top. For A, at least 15 embryos for each genotype were analyzed and representative samples are shown. For B and C, at least 20 embryos for each genotype were collected and analyzed, representative samples are shown. For qPCR results, data are mean ± SEM. ns: not significant. *: $p<0.05$. **: $p<0.01$. ***: p-value <0.001.

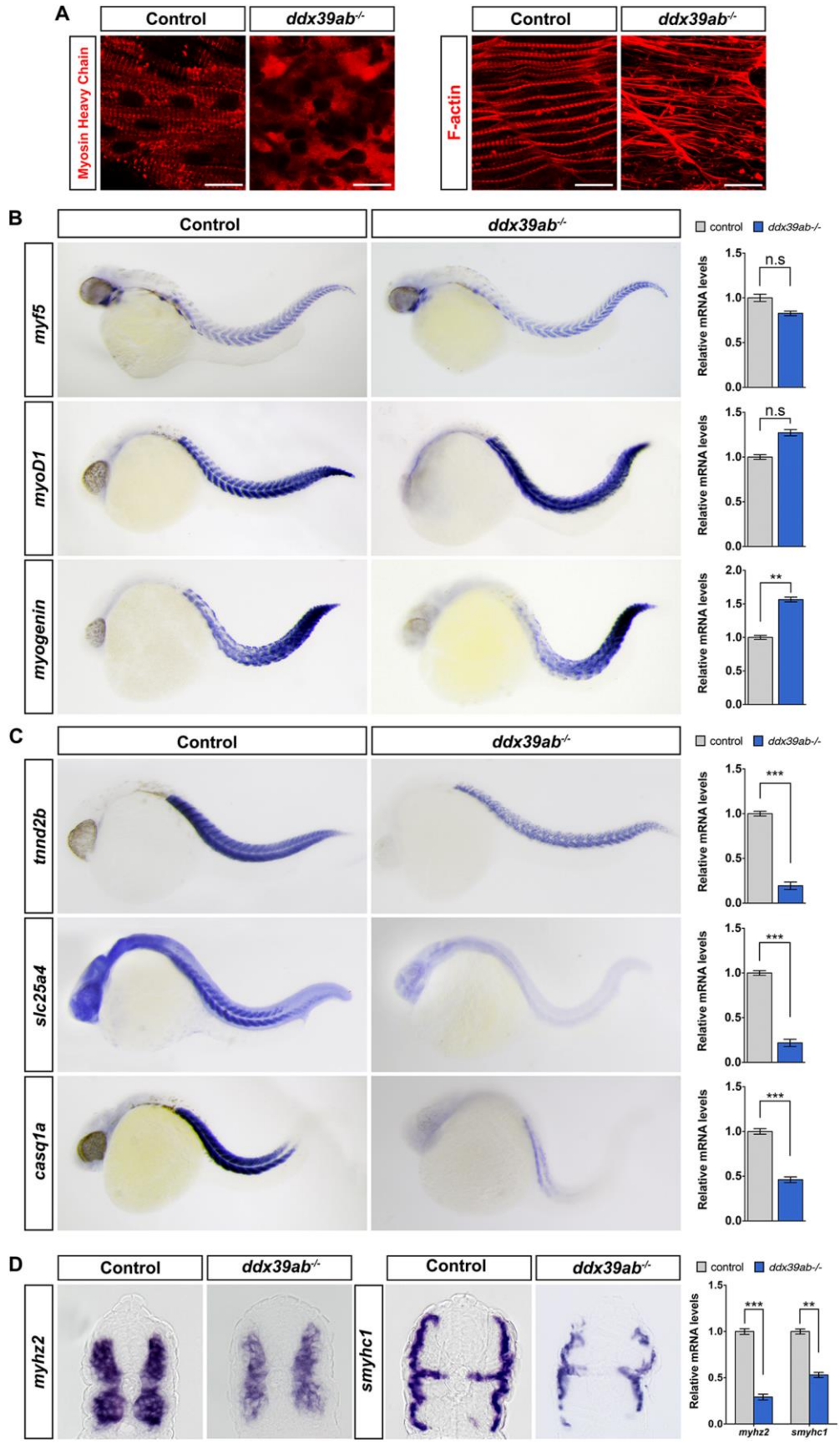


Figure 3. Loss of *ddx39a* results in defective skeletal muscle differentiation in zebrafish. (A): Immunostaining demonstrates the organization of myofilaments in wild type or *ddx39a* mutant embryos at 24 hpf. Anti-MHC antibody (MF20) labeled thick (myosin) filaments and F-actin was visualized by phalloidin staining. Scale bars: 20um. (B): RNA *in situ* hybridization is shown for myogenic regulatory gene expression in WT and *ddx39a* mutant zebrafish embryos. (C and D): RNA *in situ* hybridization and qPCR results are shown for myocyte structural gene expression in WT and *ddx39a* mutant zebrafish embryos. B, C and D, embryos were collected at 32 hpf. B and C, lateral views with anterior to the left. D, cross section with dorsal side to top. For A, at least 15 embryos for each genotype were analyzed and representative samples are shown. For B and C, at least 20 embryos for each genotype were collected and analyzed, representative samples are shown. For qPCR results, data are mean ± SEM. ns: not significant. *: $p<0.05$. **: $p<0.01$. ***: p-value <0.001.

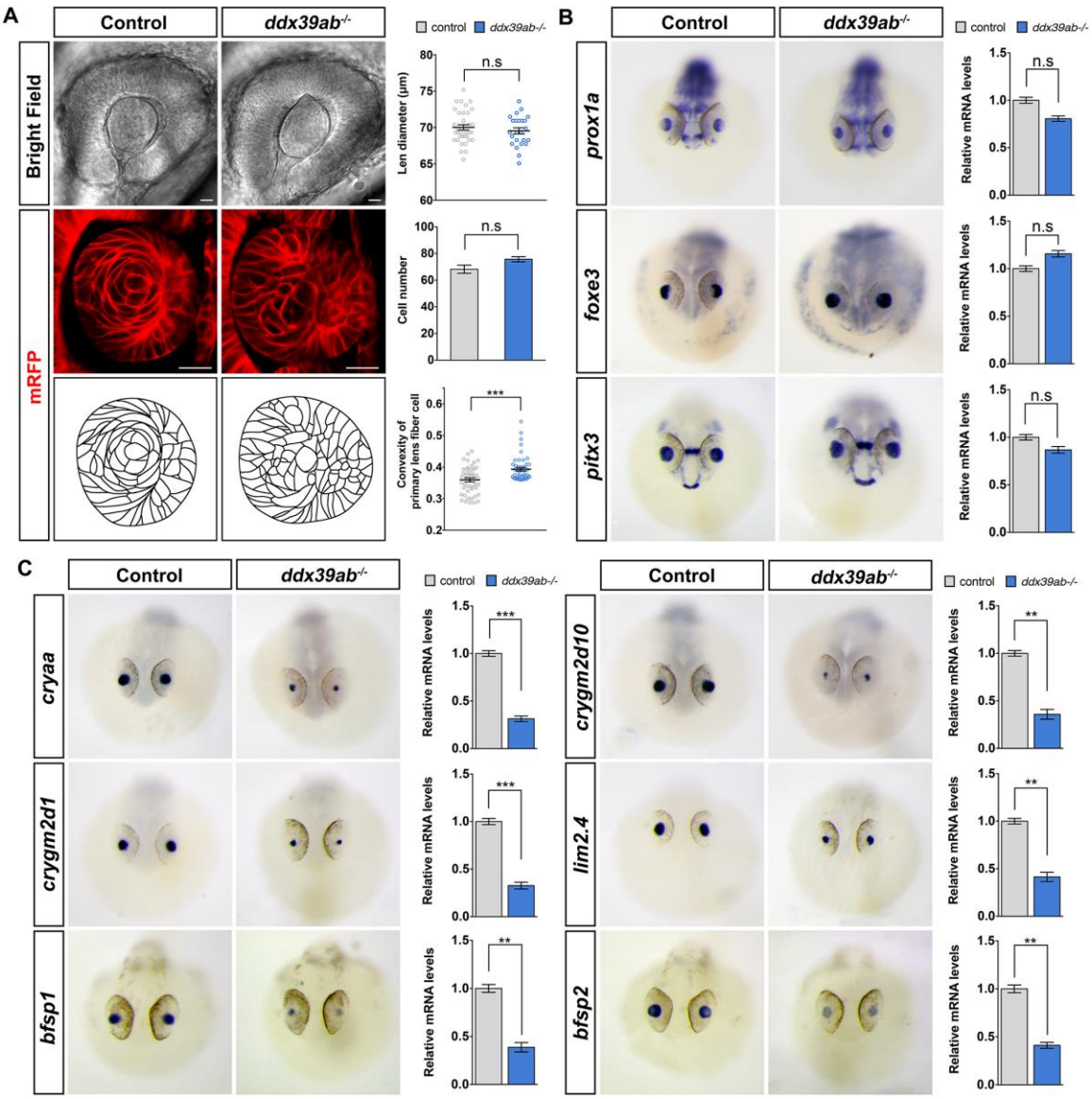


Figure 4. Lens fiber cell differentiation is defective in *ddx39a* mutant embryos. (A): *Ddx39a* mutants possessed defects in lens fiber morphogenesis. Upper left panels, bright field images of the eyes from wild type or *ddx39a* mutant zebrafish embryos at 28 hpf. Upper right panel, scatter plot for diameter of individual lens fibers, mean ± SEM are also shown. ns: not significant. Middle left panels, equatorial sections through the lens center showed organization of lens fibers, cell membrane was labeled with mRFP. Cell number on equatorial section was quantified, data are mean ± SEM. ns: not significant. Bottom panels, drawings schematically show the shape and arrangement of lens fiber cells. Scale bars, 50um. Convexity of lens fiber cell was calculated and plotted, mean ± SEM are also shown. ***: p-value <0.001. (B and C): RNA *in situ* hybridization and qPCR results are shown for lens gene expression in wild type and *ddx39a* mutant zebrafish embryos for indicated transcripts at 32 hpf. B and C, frontal views with dorsal side to the top. For A, B and C, at least 20 embryos for each genotype were collected and analyzed, representative samples are shown. For qPCR results, data are mean ± SEM. ns: not significant. **: p<0.01. ***: p-value <0.001.

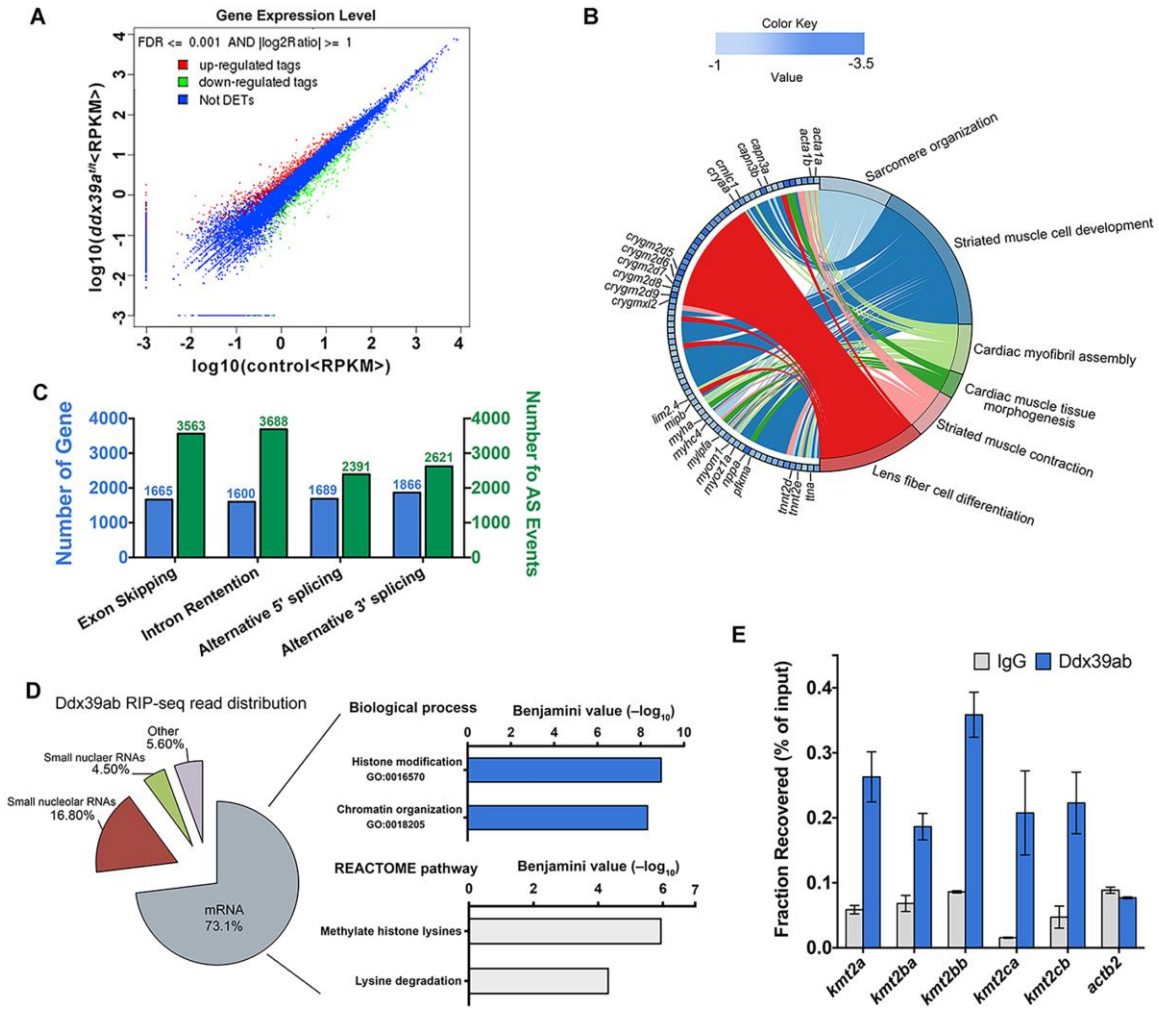


Figure 5. Transcriptome landscape of *ddx39a* mutants and identification of Ddx39a-associating RNAs. (A): Expression levels of differentially expressed genes in control and *ddx39a* mutant embryos. Up-regulated and down-regulated genes are denoted by red and green spots, respectively, while genes not differentially expressed are denoted as blue spots. RPKM, reads per kilobase per million mapped reads. (B): Circular plot showing 77 representative down-regulated genes belonging to enriched functional categories, simultaneously presenting a detailed view of the relationships between expression changes (left semicircle perimeter) and Gene Ontology (right semicircle perimeter). Complete gene lists and GO terms tables are shown in Table S5) (C): Distribution of various alternative splicing events in *ddx39a* mutant embryos. Details are shown in Table S5. (D): Left, Ddx39a RIP-Seq reads annotated to zebrafish genome with percentage of the total RIP-seq reads shown. Right, enriched Gene Ontology and REACTOME pathway terms from Ddx39a-bound mRNAs obtained using the DAVID tool. The x-axis values correspond to the negative Benjamini P value. (E): RIP-qPCR showing binding of Ddx39a to selected epigenetic modulator genes. Results are means±s.e.m. *Actb2* was included as negative control.

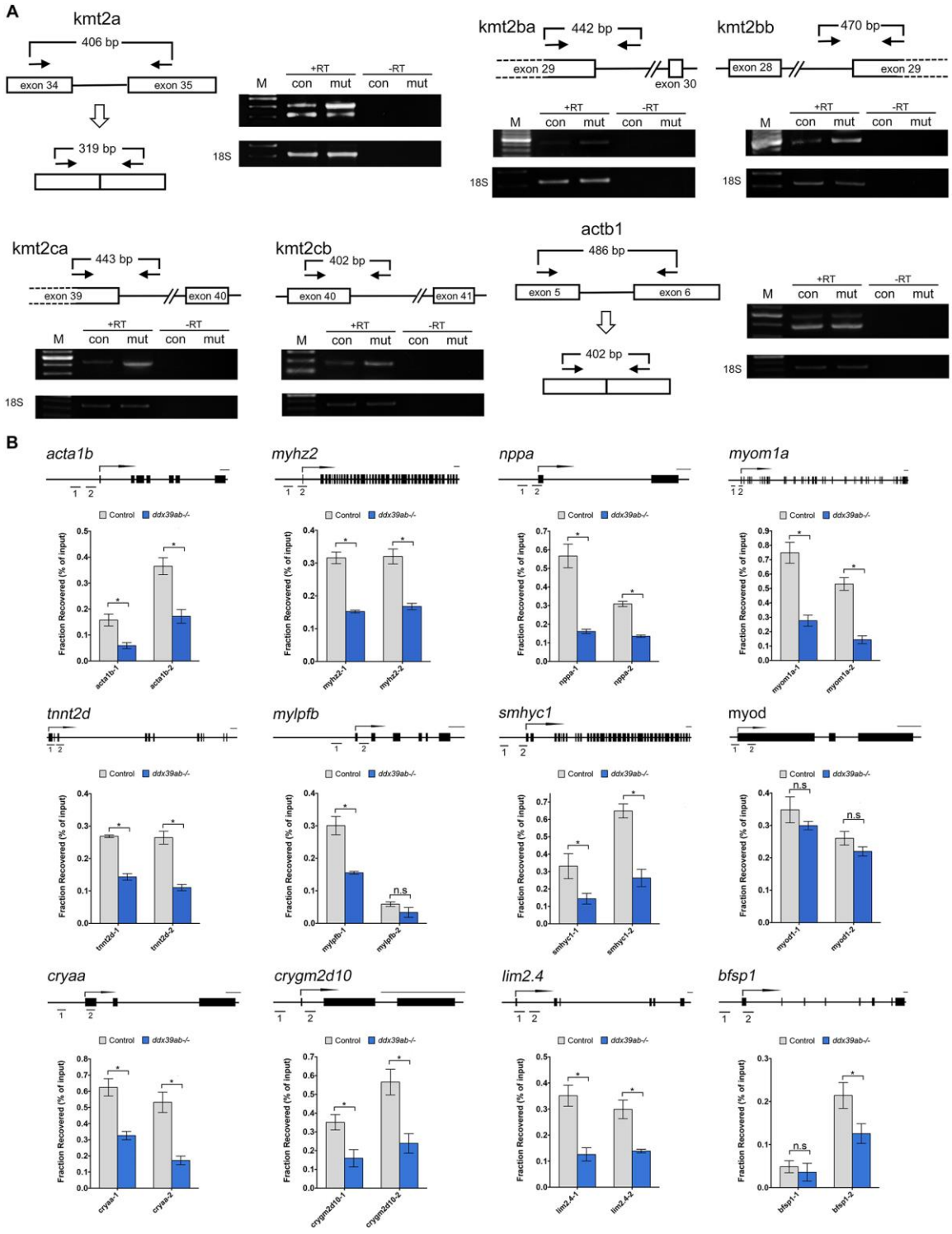


Figure 6. *Ddx39* deficiency affects pre-mRNA splicing of epigenetic modulator genes. (A): The splicing status of *kmt2a, kmt2ba, kmt2bb, kmt2ca, kmt2cb* and *actb1* pre-mRNA was monitored using RT-PCR with the primers indicated in schemes (boxes, exons; lines, introns; arrows, primers). Unspliced *kmt2a, kmt2ba, kmt2bb, kmt2ca and kmt2cb* mRNAs were retained in the *ddx39a* mutant (mut) larvae compared to the control (con) larvae. Unspliced and spliced PCR products were verified by sequencing. +RT refers to the validation reaction itself, and -RT represents the respective control reaction without reverse transcriptase. 18S is a loading control. M, DNA size markers. Control larvae were sibling wild type or *ddx39a*$^{+/-}$ larvae and had normal phenotypes. (B): ChIP-qPCR analyses represent the effect from loss of *Ddx39* on H3K4 methylation levels at promoter and TSS of muscle cell and lens fiber cell genes. Upper schematics showed the gene structure (black rectangles, exons; arrow, TSS) and indicated the DNA segments amplified by PCR. Sample size = 3. Results are means±s.e.m. ns: not significant. **p<0.01, *p<0.05.

Supplementary material

Table S1. Primers sequences and PCR conditions.

Click here to Download Table S1

Table S2. List of antibodies used for western, CHIP, RIP and immunofluorescence studies.

Antibody name	Company	Dilution	Application
Anti-α tubulin	Santa Cruz Biotechnology (sc-23948)	1:2000	Western blotting
Anti-Nkx2.5	Cell Signaling Technology (8792)	1:1000	Western blotting
Anti-Gata4	Santa Cruz Biotechnology (sc-25310)	1:1000	Western blotting
Anti-Myf5	GeneTex (GTX87746)	1:1000	Western blotting
Anti-MyoD	Santa Cruz Biotechnology (sc-377460)	1:500	Western blotting
Anti-Mef2C	Cell Signaling Technology (5030)	1:1000	Western blotting
Anti-Histone 3	Cell Signaling Technology (9715)	1:1000	Western blotting
Anti-H3K4Me1	Abcam (ab8895)	1:500	Western blotting
		2 µg for 25 µg of chromatin	CHIP
Anti-Actin	SIGMA (A2172)	1:200	Immunostaining
Anti-Myosin	DSHB (MF20)	1:200	Immunostaining
ANTI-FLAG M2 Affinity Gel	SIGMA (A2220)	N.A	RIP
Anti-MYH1A	DSHB (F59)	1:50	Immunostaining
Anti-CD166	DSHB (zn8)	1:50	Immunostaining
Anti-TNNT2	Santa Cruz Biotechnology (sc-20025)	1:50	Immunostaining
Anti-mouse IgG (H+L),-Alexa Fluor 594 Conjugate	Cell Signaling Technology (8890)	1:200	Immunostaining
Anti-Rabbit IgG (H+L)-Alexa Fluor 568 Conjugate	Thermo Fisher (A10042)	1:200	Immunostaining

Table S3. Significant different expression gene in *ddx39ab* mutant. This file includes Genes that that have FDR≤0.001 and Fold Change >2.

Click here to Download Table S3

Table S4. Significant alternative splicing events in *ddx39ab* mutant embryo. Four types of alternative splicing events, including exon skipping, intron retention, 5' and 3' splicing sites usage, were calculated from the RNA-seq data of control and *ddx39ab* mutant embryo (24hpf) by the OLego and Quantas software packages. The spread sheet for each type of events contains 7 columns, including gene accession number, chromosome, strand, constitutive exon, alternative events.

Click here to Download Table S4

Table S5. Gene Ontology (GO) enrichment analysis of significant different expression gene in *ddx39ab* mutant. P<0.01, FDR<0.05.

Click here to Download Table S5

Table S6. List of Ddx39ab-interacting RNAs identified by RIP-Seq analysis. q value≤0.1.

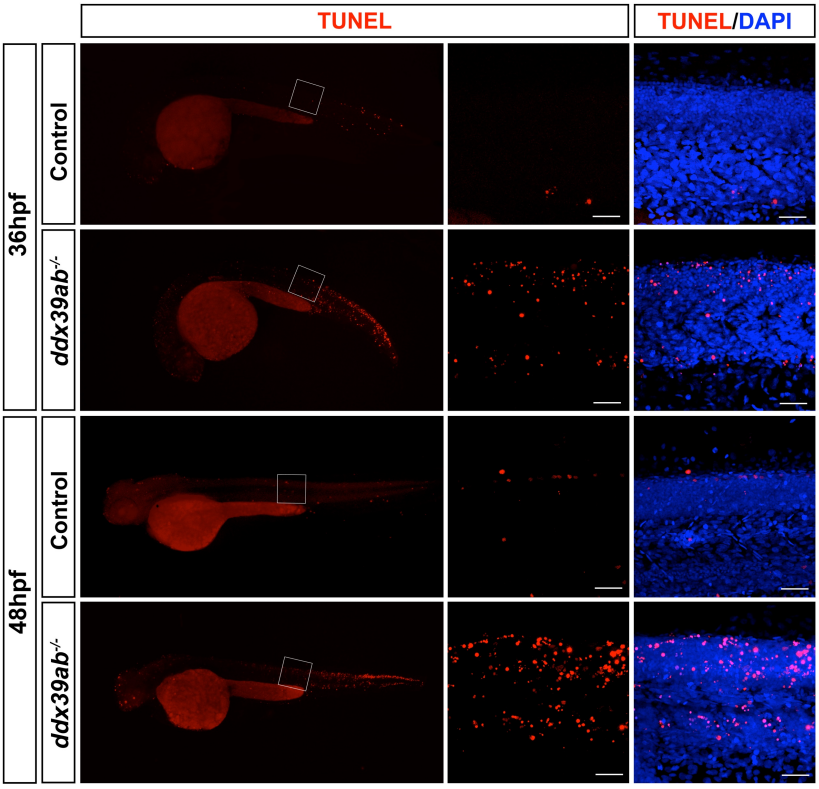
Click here to Download Table S6

Movie S1. Contraction of the definitive heart tube in RP-011 homozygous and control embryo.

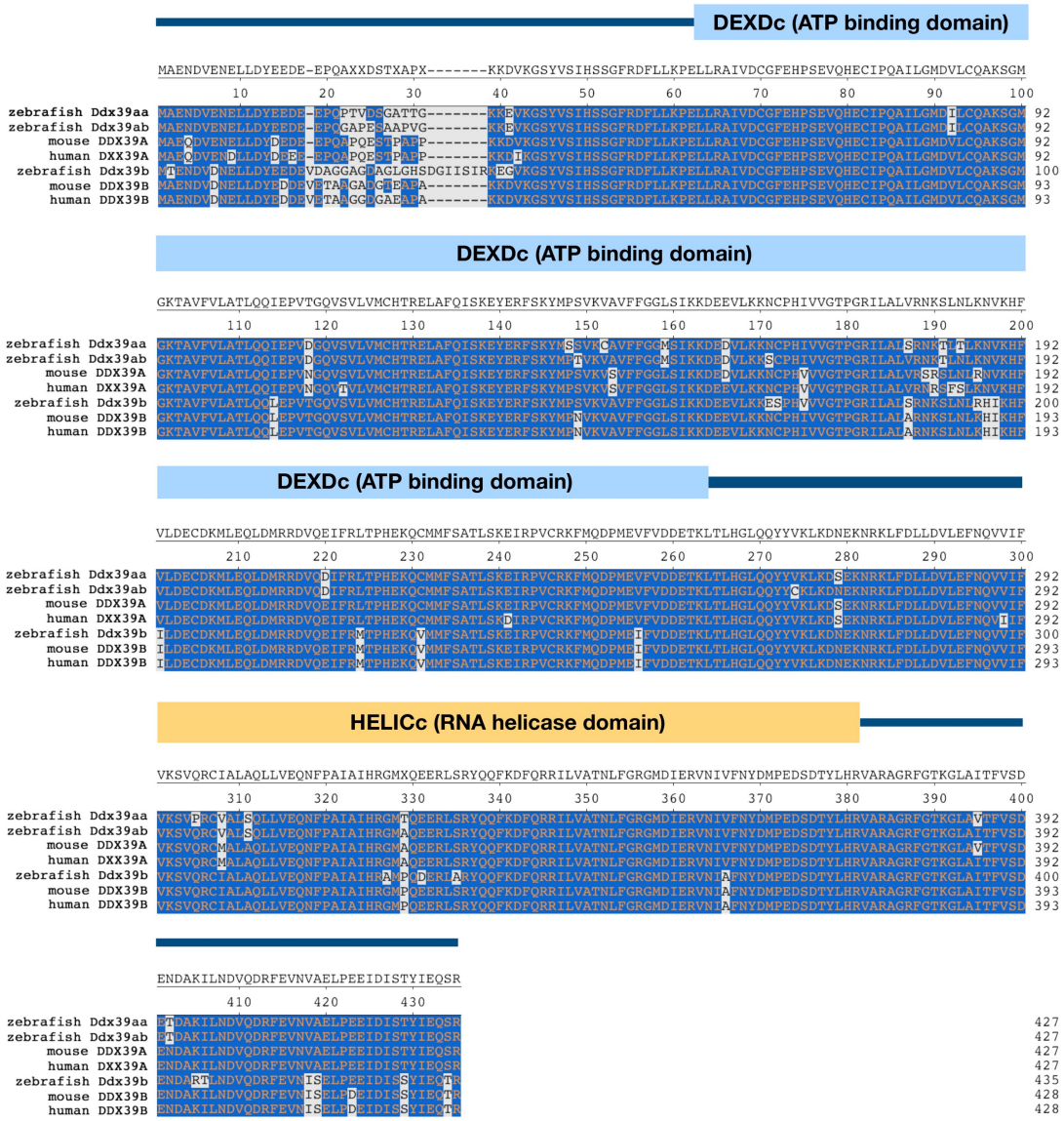
Movie S2. Spontaneous tail movements in RP-011 homozygous and control embryo.

Movie S3. Injection of FLAG-ddx39ab mRNA could rescue the paralysis phenotype in ddx39ab mutants.

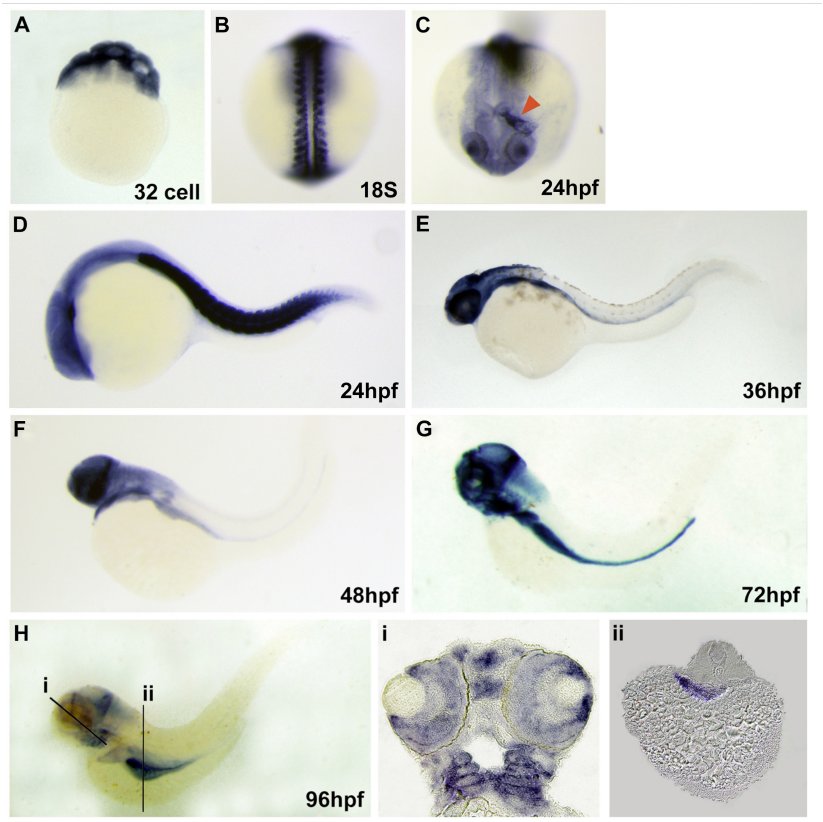
Supplementary figures



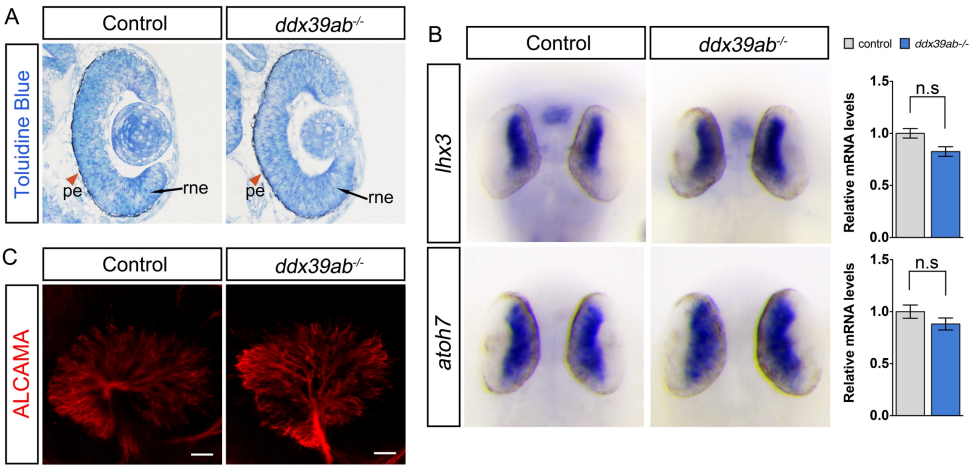
Supplemental Figure 1. **Cell death in wild type or *ddx39ab* mutant embryo.** Right panels, magnified view of the boxed region. Scale bars, 20um. For each stage, at least 10 embryos for each genotype were analyzed and representative samples are shown.



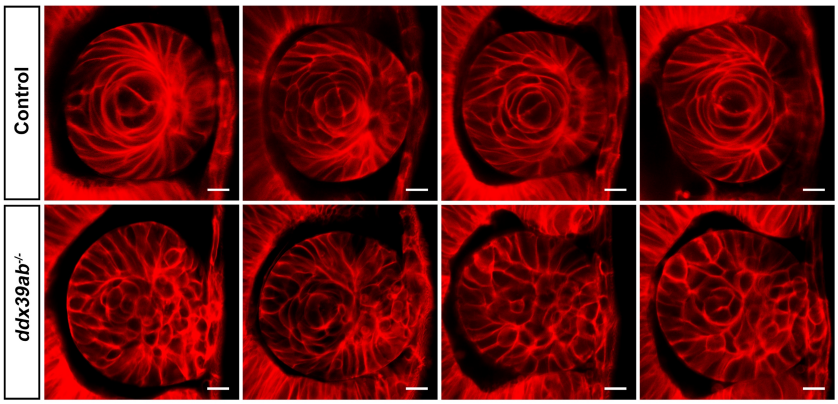
Supplemental Figure 2. DDX39 protein sequences alignment. Functional domains are labeled. The orthologs and paralogs of DDX39 in zebrafish, mouse and human share 94% amino acid identity, differing in only 21 residues.



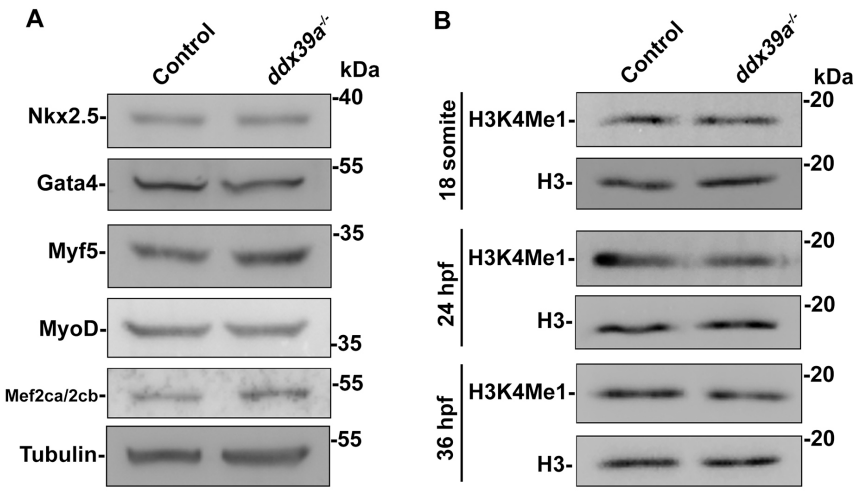
Supplemental Figure 3. Zebrafish *ddx39ab* embryonic expression pattern. (A): In situ hybridization experiment revealed there is maternal deposit ddx39ab mRNA in early stage embryo. (B): Strong expression of *ddx39ab* could be observed in myotome at 18 somite stage. (C-D): At 24hpf enriched *ddx39ab* mRNA could be observed in lens, heart tube and trunk muscle. Triangle indicated expression in heart tube. (E-H): Later on, the expression was restrained to specific regions in brain, retina (i), pharyngeal arches (i) and endoderm derived organs (ii). A, lateral view. B, dorsal view. C, frontal view. D to H, head to left. i and ii are sections of H as indicated. hpf, hour past fertilization.



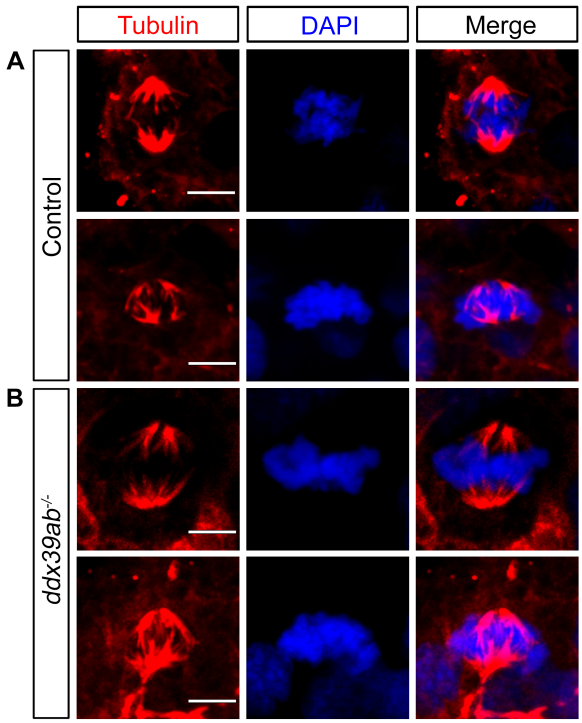
Supplemental Figure 4. Retina development in *ddx39ab* mutants. (A): Frontal sections of retina at 32 hpf. (B): RNA *in situ* hybridization and qPCR results are shown for retinal ganglion marker gene expression in control and *ddx39a* mutant zebrafish embryos. Dorsal views with anterior to the top. For qPCR results, data are mean ± SEM. ns: not significant. (C): Confocal images of retinal ganglion cell (labeled with zn8 antibody) from wild type and *ddx39ab* mutant embryo at 48 hpf. Scale bars, 20um. A, B and C, at least 15 embryos for each genotype were analyzed and representative samples are shown.



Supplemental Figure 5. Disorganization of primary lens fiber cell in *ddx39a* embryo. Cell membrane was labeled with mRFP; living images were taken at 28 hpf. Scale bars, 25 μM.



Supplemental Figure 6. Western blot analysis of the indicated proteins and histone PTMs in control and *ddx39ab* mutant embryo. (A): protein level of myocyte or cardiomyocyte key transcription regulators showed no evident change between control and *ddx39a* mutant at 24 hpf. (B): In *ddx39ab* mutant, minor decrease on H3K4Me1 level could be detected from 24 hpf.



Supplemental Figure 7. *Ddx39ab* mutant embryos exhibit normal mitoses. Mitotic cells stained with DAPI (blue) and α-tubulin (red). Scale bars, 5um.

Supplementary Materials and Methods

Western blotting

Proteins were isolated by homogenizing embryos in RIPA buffer containing protease inhibitor (Sigma). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). For immunoblotting, proteins were transferred to polyvinylidene fluoride (PVDF) membrane using an electrophoretic transfer apparatus (Bio-Rad). The membrane was blocked with 1% non-fat milk (Bio-Rad) and incubated with primary antibody followed by 1:10,000 HRP-conjugated secondary antibodies (Sigma). Signal detection was performed using Pico West Chemiluminescent Substrate (Thermo Scientific). Information for primary antibodies are listed in Supplementary Table 3.

Supplementary material

Table S1. Primers sequences and PCR conditions.

Click here to Download Table S1

Table S2. List of antibodies used for western, CHIP, RIP and immunofluorescence studies.

Antibody name	Company	Dilution	Application
Anti-α tubulin	Santa Cruz Biotechnology (sc-23948)	1:2000	Western blotting
Anti-Nkx2.5	Cell Signaling Technology (8792)	1:1000	Western blotting
Anti-Gata4	Santa Cruz Biotechnology (sc-25310)	1:1000	Western blotting
Anti-Myf5	GeneTex (GTX87746)	1:1000	Western blotting
Anti-MyoD	Santa Cruz Biotechnology (sc-377460)	1:500	Western blotting
Anti-Mef2C	Cell Signaling Technology (5030)	1:1000	Western blotting
Anti-Histone 3	Cell Signaling Technology (9715)	1:1000	Western blotting
Anti-H3K4Me1	Abcam (ab8895)	1:500	Western blotting
		2 µg for 25 µg of chromatin	CHIP
Anti-Actin	SIGMA (A2172)	1:200	Immunostaining
Anti-Myosin	DSHB (MF20)	1:200	Immunostaining
ANTI-FLAG M2 Affinity Gel	SIGMA (A2220)	N.A	RIP
Anti-MYH1A	DSHB (F59)	1:50	Immunostaining
Anti-CD166	DSHB (zn8)	1:50	Immunostaining
Anti-TNNT2	Santa Cruz Biotechnology (sc-20025)	1:50	Immunostaining
Anti-mouse IgG (H+L),-Alexa Fluor 594 Conjugate	Cell Signaling Technology (8890)	1:200	Immunostaining
Anti-Rabbit IgG (H+L)-Alexa Fluor 568 Conjugate	Thermo Fisher (A10042)	1:200	Immunostaining

Table S3. Significant different expression gene in *ddx39ab* mutant. This file includes Genes that that have FDR≤0.001 and Fold Change >2.

Click here to Download Table S3

Table S4. Significant alternative splicing events in *ddx39ab* mutant embryo. Four types of alternative splicing events, including exon skipping, intron retention, 5' and 3' splicing sites usage, were calculated from the RNA-seq data of control and *ddx39ab* mutant embryo (24hpf) by the OLego and Quantas software packages. The spread sheet for each type of events contains 7 columns, including gene accession number, chromosome, strand, constitutive exon, alternative events.

Click here to Download Table S4

Table S5. Gene Ontology (GO) enrichment analysis of significant different expression gene in *ddx39ab* mutant. P<0.01, FDR<0.05.

Click here to Download Table S5

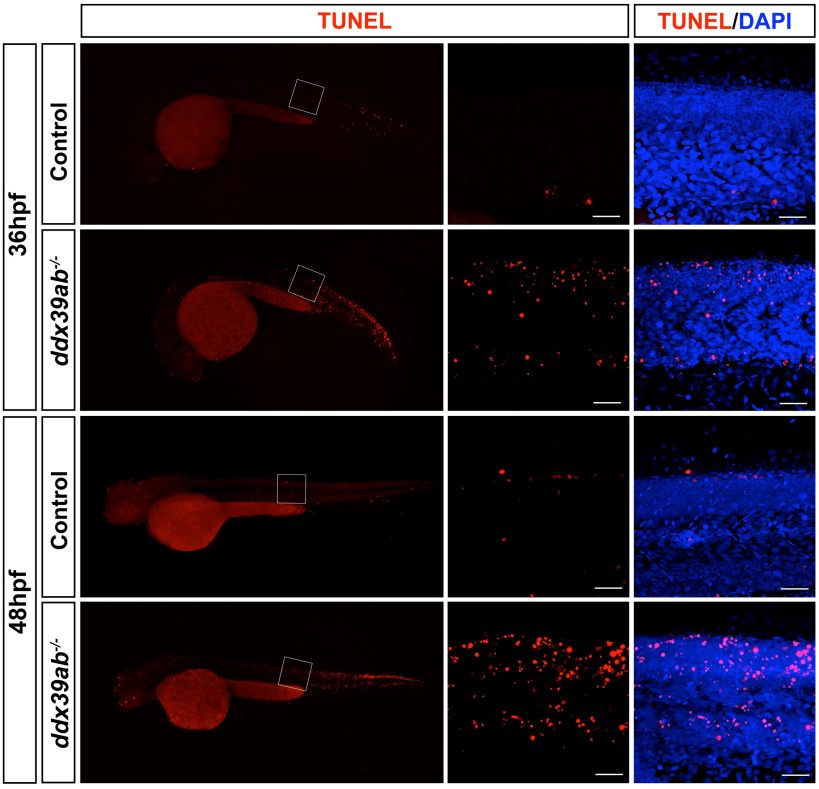
Table S6. List of Ddx39ab-interacting RNAs identified by RIP-Seq analysis. q value≤0.1.

Click here to Download Table S6

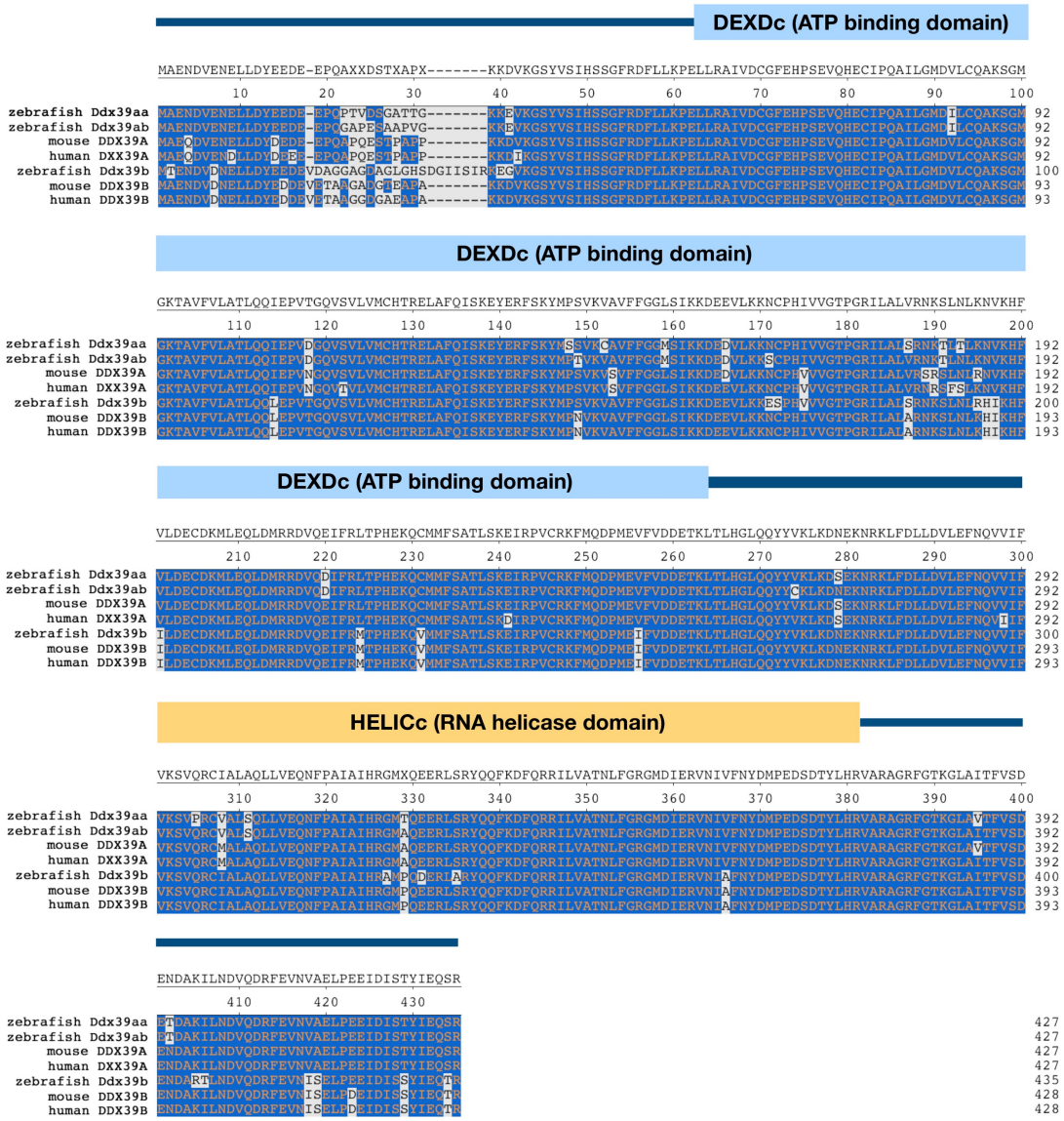
Movie 1. Contraction of the definitive heart tube in RP -011 homozygous and control embryo.

Movie 2. Spontaneous tail movements in RP -011 homozygous and control embryo.

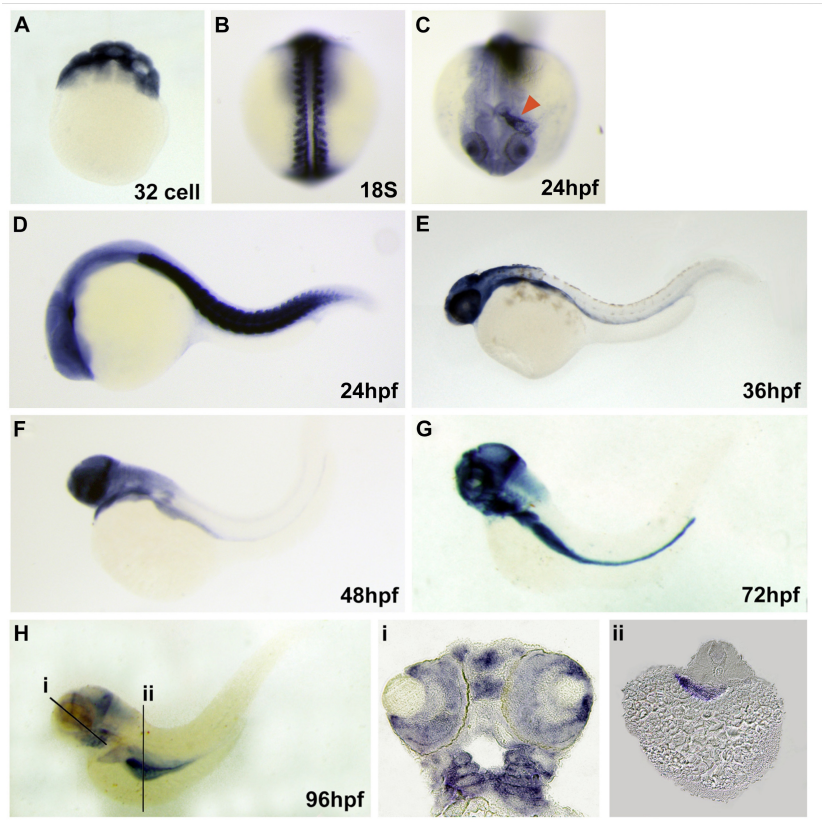
Supplementary figures



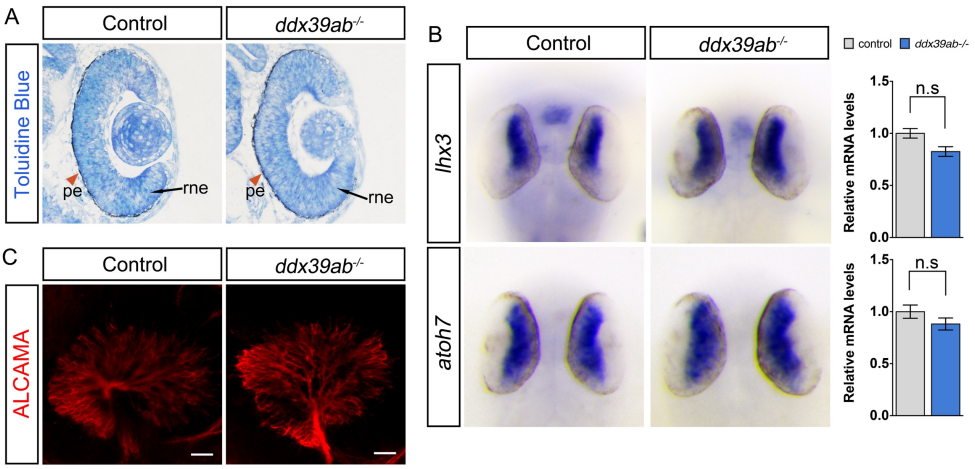
Supplemental Figure 1. **Cell death in wild type or *ddx39ab* mutant embryo.** Right panels, magnified view of the boxed region. Scale bars, 20um. For each stage, at least 10 embryos for each genotype were analyzed and representative samples are shown.



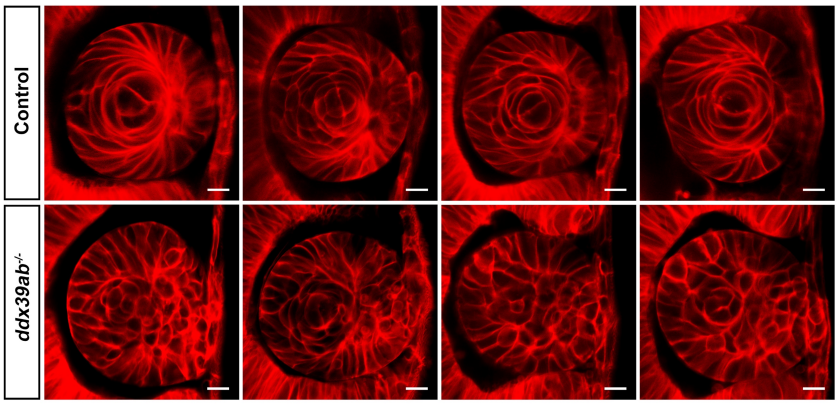
Supplemental Figure 2. DDX39 protein sequences alignment. Functional domains are labeled. The orthologs and paralogs of DDX39 in zebrafish, mouse and human share 94% amino acid identity, differing in only 21 residues.



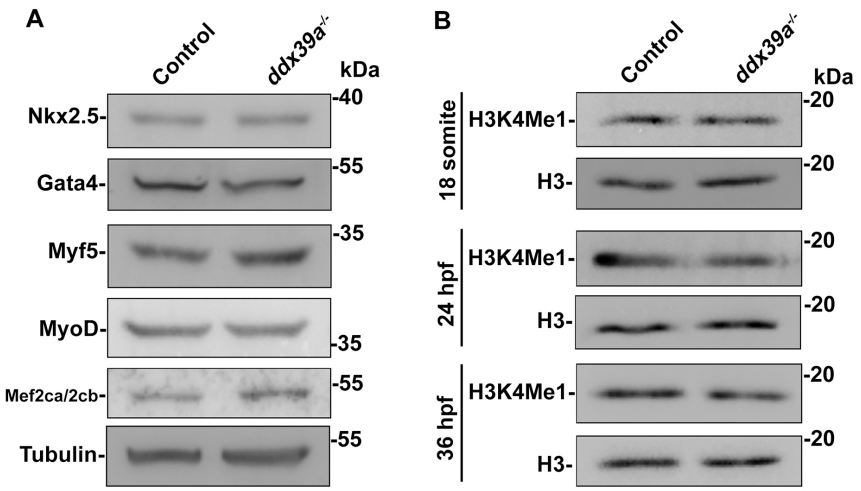
Supplemental Figure 3. Zebrafish *ddx39ab* embryonic expression pattern. (A): In situ hybridization experiment revealed there is maternal deposit ddx39ab mRNA in early stage embryo. (B): Strong expression of *ddx39ab* could be observed in myotome at 18 somite stage. (C-D): At 24hpf enriched *ddx39ab* mRNA could be observed in lens, heart tube and trunk muscle. Triangle indicated expression in heart tube. (E-H): Later on, the expression was restrained to specific regions in brain, retina (i), pharyngeal arches (i) and endoderm derived organs (ii). A, lateral view. B, dorsal view. C, frontal view. D to H, head to left. i and ii are sections of H as indicated. hpf, hour past fertilization.



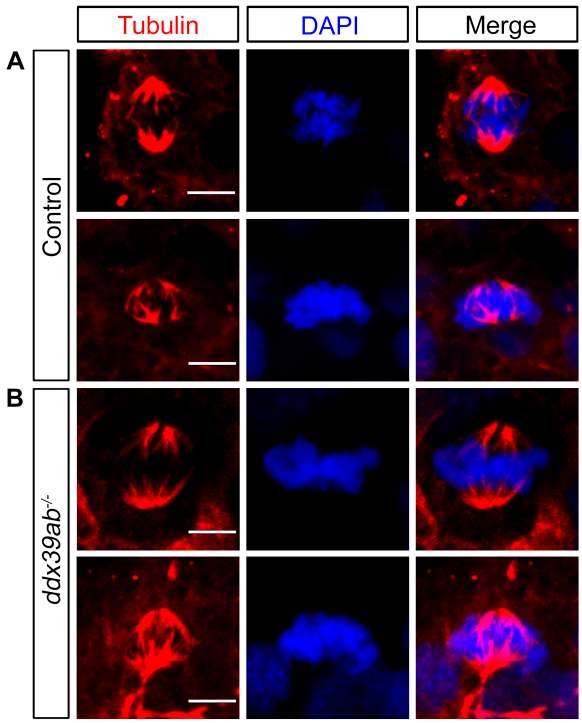
Supplemental Figure 4. Retina development in *ddx39ab* mutants. (A): Frontal sections of retina at 32 hpf. (B): RNA *in situ* hybridization and qPCR results are shown for retinal ganglion marker gene expression in control and *ddx39a* mutant zebrafish embryos. Dorsal views with anterior to the top. For qPCR results, data are mean ± SEM. ns: not significant. (C): Confocal images of retinal ganglion cell (labeled with zn8 antibody) from wild type and *ddx39ab* mutant embryo at 48 hpf. Scale bars, 20um. A, B and C, at least 15 embryos for each genotype were analyzed and representative samples are shown.



Supplemental Figure 5. Disorganization of primary lens fiber cell in *ddx39a* embryo. Cell membrane was labeled with mRFP; living images were taken at 28 hpf. Scale bars, 25 μM.



Supplemental Figure 6. Western blot analysis of the indicated proteins and histone PTMs in control and *ddx39ab* mutant embryo. (A): protein level of myocyte or cardiomyocyte key transcription regulators showed no evident change between control and *ddx39a* mutant at 24 hpf. (B): In *ddx39ab* mutant, minor decrease on H3K4Me1 level could be detected from 24 hpf.



Supplemental Figure 7. *Ddx39ab* mutant embryos exhibit normal mitoses. Mitotic cells stained with DAPI (blue) and α-tubulin (red). Scale bars, 5um.

Supplementary Materials and Methods

Western blotting

Proteins were isolated by homogenizing embryos in RIPA buffer containing protease inhibitor (Sigma). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). For immunoblotting, proteins were transferred to polyvinylidene fluoride (PVDF) membrane using an electrophoretic transfer apparatus (Bio-Rad). The membrane was blocked with 1% non-fat milk (Bio-Rad) and incubated with primary antibody followed by 1:10,000 HRP-conjugated secondary antibodies (Sigma). Signal detection was performed using Pico West Chemiluminescent Substrate (Thermo Scientific). Information for primary antibodies are listed in Supplementary Table 2.