**Supplementary Materials and Methods**

***Drosophila* culture, strains and genetics.**

Flies were maintained in standard cornmeal, yeast, dextrose, agar medium, at 25**°**C unless otherwise stated. For selection of staged larvae, 0.05% bromophenol blue was added to the food. Wandering larvae have blue guts, whereas stationary larvae have no blue or very light blue guts. White prepupae were selected by watching larvae, and harvesting them within an hour of pupariation. *w1118* was used as a control. *yw ; Nxt1z2-0488 / CyO Actin-GFP* and *yw ; Nxt1 DG05102 / CyO Actin-GFP* were crossed to generate *Nxt1z2-0488 / Nxt1 DG05102* transheterozygotes for analyses (Caporilli *et al.* 2013). The *Nxt1* mutant larvae were selected by the absence of GFP.

VDRC RNAi lines KK107745 (*Nxt1*), P{GD17336}v52631 (*Nxt1*) KK102231 (*thoc5*), KK100882 (*Nxf1*) and KK109076 (*Ref1*) (Dietzl *et al.* 2007), gave phenotypes and knocked down their respective target gene expression when expressed in spermatocytes in testes, and were therefore used to analyse muscle integrity in the third instar larvae. UAS-eGFP-Nxt1 was described in (Caporilli *et al.* 2013).

w;UAS-*abbafl*, kindly provided by Hanh T. Nguyen, uses the EST GH06739 (GenBank accession number AY121620) as template to amplify and clone the entire *abba* coding or defined regions of *abba* into the pUAST vector (Domsch *et al.* 2013). The 3’ end of the construct encodes an HA tag. UAS-stock is in heterozygous condition. Mef2-Gal4 (Ranganayakulu *et al.* 1996) was kindly provided by Michael Taylor (Cardiff University); Arm-Gal4 and UAS-dicer were from Bloomington stock centre.

**RNA extraction, cDNA synthesis and Quantitative PCR**

Total RNA was extracted using Trizol (ThermoFisher Scientific) and then further cleaned up with the RNeasy Mini Kit (Qiagen) according to manufacturer’s protocol. The DNaseI step was included. RNA was quantified with a nanodrop ND-1000 (ThermoFisher Scientific) and stored at -80oC. For larval carcass RNA sequencing, 30 carcasses were used per replicate in triplicates. For qRT-PCR, either a single carcass or a mix of up to 10 carcasses were used as detailed in the results. cDNA was generated using 100ng total RNA and oligo dT or random primers (for nascent transcript and circRNA analysis) with the Superscript III kit (Invitrogen). The cDNA reaction was diluted to 60 or 120μl with dH2O, and 1μl of this cDNA was further diluted with 7μl dH2O to use as a template in the qRT-PCR reactions. 10μl PowerSybr reagent (ABI) with 1μl of a 10μM solution of each target-specific forward and reverse primers (primer sequences on request) were added for a total reaction of 20μl. For circRNA analysis, primers were designed to flank the back-spliced exon-exon junction most frequently detected in the RNAseR treated circRNA sequence data. All the circRNA structures were validated by sequencing of the RT-PCR product. The qRT-PCR was performed on a Chromo4 instrument (MJR) using the Opticon Monitor 3 software, except for circRNA which was conducted using a Roche lightcycler 96. *Rp49* was used as a control gene for normalization, fold changes were determined by CT. All reactions were performed in triplicate.

**Scanning Electron Microscopy**

Young pupae (<4h) were picked, cleaned with water, and air dried overnight. Up to 6 pupae were put on an aluminium backing plate of a SC500 sputter target. A BIO-RAD SC500 sputter coater was used to coat the non-conducting pupae with ~20nm thick layer of Au90Pd10 according to manufacturer’s protocol (Quorum Technologies). After coating, pins were put in a FEI-XL30 Field Emission Gun Environmental Scanning Electron Microscope. For imaging, a 30 μm diameter final aperture with a beam current <1nA was used to take pictures from each pupa at the posterior/anterior end and the middle section.

**Magnetic Chamber Dissection for larval muscle staining**

For larval dissections, a magnetic chamber was used (Budnik *et al.* 2006). A magnetic strip, with a 30 mm diameter hole in the middle, was glued on a 76 mm x 51 mm slide with 10 mm of the strip sticking out at all sides. Each magnetic chamber uses 2 centre and 4 corner pins. For constructing all pins, see Figure 1 in (Budnik *et al.* 2006). All pins were glued to a vintage metal index tab. A larva is put in the middle of the hole with a drop oflow Ca2+ saline, HL-3 (Pesavento *et al.* 1994). Larva are put ventral side and two centre pins are used to prevent the larva from moving by pinning at the most anterior and posterior side. Vannas Spring Scissors – 3mm blades (RS-5618; Fine Science Tools) were used for cutting. A dorsal incision cut was made at the posterior end with short shallow cuts between the two tracheas until the anterior end was reached.

**Phalloidin staining of larval muscles**

The larva was cleaned on the magnetic chamber by removing all internal organs carefully. All Ca2+ saline, HL-3 solution was removed with a P-100 pipette and fresh drops were added two more times while cleaning the larval carcass. The larval carcass was fixated by adding fresh 4% paraformaldehyde in PBS for 2 minutes. All the pins were removed from the larval carcass and the sample was transferred to a glass well with 100μl 4% paraformaldehyde for another hour. The fixative was removed, and the larval carcass was washed twice with 100μl PBS-T (0.1% Triton X-100) for 5 minutes each. Two drops of Alexa FluorTM 488 Phalloidin (ThermoFisher) was added to 1ml of PBS-T. 100μl Alexa FluorTM 488 Phalloidin mix was added to the well to stain F-actin in larval muscles for 1 hour. Alternatively, FITC phalloidin was used at a final concentration of 1 g/ml in PBS-T. Phalloidin solution was removed and the larval carcass was washed with PBS-T two times for 5 minutes each. The larva carcass was put on a microscope slide and mounted in 85% glycerol + 2.5% n-propyl gallate. Images were made on an Olympus BX 50 (Olympus) microscope and Hamamatsu ORCA-05G digital camera. For a close-up of the larval muscle sarcomeres pattern, images were taken with a Leica DM6000B upright microscope with HC PL Fluotar 20x/0.50 and HCX PL APO 40x/1.25 oil objectives.

**Criteria for scoring muscle defects in larvae**

The integrity of larva muscles was calculated as a percentage from a total of 8 hemisegments. The hemisegments A2-A5 were in the abdominal area and were not damaged by any of the pins. Each hemisegment contained 30 different muscles, so a total of 240 muscles were inspected. The integrity of a muscle is compromised if the muscle is damaged in any way, such as being torn, thin, loss of sarcomeric structure or missing. The muscle damage percentage is the total number of damaged muscles divided by 240 and multiplied by 100.

**Starvation of larvae**

Larvae were fed up to 70, 71, 72 and 73 hours after egg laying (AEL) before removal from the food. The larvae were transferred to a petri dish with a moist filter paper in it to prevent desiccation. Larvae in the petri dish were checked at regular intervals to ensure the filter paper was moist and keep track of the progress of the metamorphosis. Larvae deprived of food 70 hours AEL that survived for four days were analysed for muscle integrity.

**Mobility Assay**

First, second and third instar larvae were used for mobility analysis. Larvae (up to 10) were put in the middle of a 1% agar dish with 1% paraffin oil on one side and an odour (1% 2-propanol) on the other side. Larvae were filmed with a Samsung SDN-550 camera using the micro manager 1.4 program for 200 frames, with 1 frame per second (fps) in the dark under red light. Movies were analysed with MtrackJ (Meijering *et al.* 2012) via Fiji. Larvae were tracked per 20 frames (1st instar) or 10 frames (2nd and 3rd instars) for at least 100 seconds (1st instar) or 50 seconds (2nd and 3rd instar) while larvae were continuously crawling on the plate. Mean speed is calculated in mm/sec.

**Viability Assay**

Third instar larvae were taken from the standard vial and put in a new vial with a low quantity of fresh food for the viability assay. After 5 days of incubation at 25oC vials were taken out and pupa lethality was backtracked by looking at 6 different points: 1) larvae that did not pupate (0h APF), 2) no visible development in pupae (24h APF), 3) head eversion and development of the eye (48h APF), 4) bristles on dorsal thorax (72h APF), 5) complete fly development in pupa (72h APF) and 6) emerging adults.

**Whole larvae mRNA sequencing and analysis**

Eighteen libraries were prepared from total RNA of *w1118* and Nxt1 trans-heterozygotes genotypes. Each genotype had samples of three different stages (wandering larvae, stationary larvae and white prepupae) in triplicates. RNA extraction was performed using Trizol (ThermoFisher Scientific) followed by the RNeasy Mini Kit (Qiagen). The samples were sent to the University of Exeter to perform the library preparation (ScriptSeq RNA-Seq Library Preparation Kit (Illumina)) and sequencing. All samples were 100bp paired-end sequenced on an Illumina HiSeq 2500 using standard mode.

Each sample was prepared in triplicate, and a sequence depth between 6.4M – 11.1M was achieved with sequencing all libraries. The data was analysed with the Tuxedo suite (Trapnell *et al.* 2012) via GenePattern browser (Reich *et al.* 2006). The lists of genes and their Fragments per Kilobase of Exon per Million Mapped Fragments (FPKM) values were compared and statistically tested between the genotypes with Cuffdiff and were imported to excel and divided into more than 2-fold down regulated, more than 2-fold up regulated and non-differentially expressed lists. The differentially expressed genes that were assigned statistically significant had a p-value of <0.05. For analysis of ecdysone-responsive genes we looked for differential expression between mutant and control, minimum of 2-fold up or down regulated and a minimum FPKM value of 10 in at least one condition

**Larval carcass mRNA sequencing**

Six libraries were prepared from total RNA of *w1118* and Nxt1 trans-heterozygotes genotypes. Stationary larvae from each genotype were taken and samples were generated in triplicates. RNA extraction was performed using a combination of Trizol (ThermoFisher Scientific) and the RNeasy Mini Kit (Qiagen). Libraries were generated using the TruSeq Stranded mRNA Library Prep (Illumina) according to manufacturer’s protocol. This library preparation method includes an oligo dT selection to sequence only mature, polyadenylated mRNAs. Samples were sequenced using 2x75bp paired-end with the NextSeq 500/550 Mid Output v2 kit (150 cycles; Cat. No. FC-404-2001) on an Illumina NextSeq500 Sequencer (Illumina). Sequencing was performed by the Genome Research Hub at Cardiff University, School of Biosciences

Each library achieved a read depth between 21M to 32M reads and the data was analysed through the Tuxedo suite (Trapnell *et al.* 2012). Cuffdiff, within this suite, identified significantly differentially expressed genes, alternative splicing and alternative isoform expression. The list of genes with FPKM values provided by Cuffdiff was used to separate the lists in more than 1.5-, 2-, 4- and 16-fold down regulated, more than 1.5-, 2-, 4- and 16-fold up regulated and non-differentially expressed. The differentially expressed genes that were assigned statistically significant had a p-value of <0.05. For graphical display using Excel on log axes, 0.001 was added to all FPKM values.

**Larval carcass total RNA and circRNA sequencing**

For total RNA and circRNA sequencing from larval carcass, a total of 30 larval muscle preparations were used for each replicate. Stationary stage third instar larvae were dissected in PBS using the magnetic chamber before being washed twice in PBS, and then transferred directly into an Eppendorf containing 100l lysis buffer. Batches of 5-7 larval carcasses were pooled in each Eppendorf before being transferred to -20 ͦC until all sample collection was complete. No more than 1hr elapsed between dissection and lysis. *yw; Nxt1z2-0488 / Nxt1DG0510* larvae were used for mutant analysis, with *w1118* controls for comparison.

RNA was extracted using the RNAeasy Mini Kit (Qiagen) according to manufacturer’s recommendation with on-column DNAseI digestion. RNA quality was checked on a High Sensitivity tape on the TapeStation 2200 (Agilent Technologies) and quantity was checked by Qubit. 1 g of RNA was treated with 20 units of RNAseR (Epicentre) in a 25l reaction at 37°C for 30 minutes. Mock samples were treated with the same volume of water. All samples were then cleaned up with 2 volumes of RNA Clean XP beads (Agencourt) and eluted with 13l of water. 10l of mock and RNAseR treated samples were used to make RNA libraries using the TruSeq Stranded total RNA-seq with Ribozero gold LT kit (Illumina) following manufacturers recommendations. Libraries were pooled in equimolar proportions and sequenced on a 2x75bp High-output NextSeq500 cartridge, yielding on average of 40 million reads per sample (Suppl Table xx: CircRNA\_data.xlxs). All sequencing was performed by the Genome Research Hub at Cardiff University, School of Biosciences.

**Total RNA-seq Analysis.**

Quality control analyses were performed on all samples using FastQC v 0.10 ([http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/%29). Raw reads were trimmed with trimmomatic v.0.36 [Bolger et al, 2014] (HEADCROP:15 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36). Sequence reads were aligned to the *Drosophila melanogaster* genome (Dm6.23) using STAR v.2.5.3a (Dobin *et al.* 2012) with options: --outMultimapperOrder Random --outSAMmultNmax 1. Flybase r6.23 annotations were supplied to index the genome but annotations were not provided for mapping. A second QC analysis was run using Bamtools v.2.4.1 (Barnett *et al.* 2011) and the MarkDuplicate tool from the Picard suite v.2.18.14 (https://broadinstitute.github.io/picard/).

High rates of duplicated and multi-mapping reads were observed in all samples. Upon thorough investigation, we concluded that those were not artificial. Duplication rates were due to a very high abundance of a small number of genes, and was linked to multi-mappers (i.e., multi-mapping reads were also duplicated). Multi-mapping was due to a small rRNA contamination, and highly expressed small RNA (e.g. cuticle gene Cpr49Ac) or RNA with repeated conserved domains (e.g. cuticle genes lcp1 and lcp2). Differential expression analyses were also run with and without deduplication and led to the same conclusions. Multi-mappers were excluded from further analyses but duplicated reads were kept.

Read counts were generated using Subread package FeatureCounts v. 1.6.2 (Liao *et al.* 2014); reads mapping to one location were assigned to exons (for mRNA) or introns (for pre-mRNA) and reported by gene\_id. (metafeature). DESeq2 (Love et al, 2014) was used to normalise reads counts and calculate fpkm values. For visualization in genome browsers (i.e. Integrative Genomic Browser), strand-specific genome-wide normalised read coverage files were generated with the bamCoverage tool of the deepyools suite v.3.0.1 (--ignoreDuplicates --effectiveGenomeSize 142573017 --normalizeUsing RPKM).

**CircRNA Identification.**

Paired-end reads were merged prior to analysis. CircRNAs were identified using PTESFinder v.1 (Izuogu *et al.* 2016) (parameters -s 65 -u) with alignments to the *Drosophila* genome (Dm6.23) and the UCSC Refseq transcriptome. This software has been used in similar work (Izuogu *et al.* 2018; Mellough *et al.* 2019; Haque *et al.* 2020a; Haque *et al.* 2020b) and has been shown to reliably identify circRNA structures (Zeng *et al.* 2017). We imposed a detection threshold on each set of replicate samples: circRNAs structures that were identified by at least 10 back-spliced junction spanning reads, derived from at least two RNAseR-treated biological replicates were included. This is a highly stringent criterion given the high variability between RNase R-treated replicates and the sequencing depth (Szabo and Salzman 2016).

As a measure of circRNA abundance, we used the number of back-splice per million mapped reads (BPM), as defined by [Haque et al 2020a]:

where *Ji i*s the number of reads mapped to the back-spliced junction of the circular RNA, *c* is the number of reads mapped to canonical sites of the gene with the circular RNA and *n* is the number of circular RNAs identified. This measure is conceptually similar to fpkm values from RNA-seq and comparable between samples. We averaged the three biological replicates, and compared the mean BPM for each structure between wild type and mutant. Expression values for each replicate are available in Suppl data file: CircRNA\_data.xls. All structures that passed the expression threshold had a BPM>40. The mock (non RNAseR treated) replicates were used to quantify the spliced RNA (but not necessarily poly adenylated) and pre-mRNA of the host transcript. Plots were generated in R with the ggplot2 package.



**Supplementary figure 1**

**Air bubble movement fails in Nxt1 trans-heterozygote pupae.** Still images of wild type (upper) and Nxt1 trans-heterozygotes (lower) for 15 hours. A) Faint visibility of air bubble in both genotypes. B-D) Air bubble becomes more prominent. E-G) Air bubble is clearly visible halfway along the pupa. I-L) Air bubble expands further. M) Air bubble has disappeared and permitted head eversion for wild type, whereas Nxt1 mutant still show air bubble. N) Larva body in Nxt1 mutant released from cuticle at posterior end and retracts to anterior. O-P) No further changes observed in both genotypes.



**Supplementary figure 2**

**Timing of pupal lethality in Nxt1 trans-heterozygotes.** Pupae were examined at six different time points. Many Nxt1 mutants had defects in head eversion, which becomes visible 48 hours into the pupa phase. Nxt1 hypomorph/+ heterozygotes showed a slight reduction in pupa viability.



**Supplementary figure 3**

**Q-RT-PCR validation of under-expression of genes with long introns in *Nxt1* trans-heterozygote larval carcass.** Genes with an average total intron length of 5700, 7500, 10000 and 15000 were examined. Expression of *abba* and *Nxt1* are also shown. CG1307 and CG9330 were used as control genes with low total intron length. All genes examined have expression in the carcass. Rp49 was used for normalization.

****

****

**Supplementary figure 4**

Examples of genes that produce circRNAs**.**

A-D, Genomic windows show normalised coverage for the total RNA sequencing (dark and light pink tracks), circRNA sequencing (with RNAseR, black and red tracks) and mRNA (polyA, dark and light blue), showing examples of patterns of expression changes between WT and *Nxt1* transheterozygotes for the four loci (*CG7378, Hs6st, Src42A* and *tn*) whose circRNA expression was tested by q-RT-PCR. The sum of all three normalised replicates is shown. The FlyBase annotation reveals the range of mRNA structures produced for each locus, and the exons that contribute to the most abundant circRNA structure for each gene are indicted by brackets. In all cases the circRNA does not contain any intron sequences. Interestingly, the exon that produces *Src42A* circRNA is one that shows tissue-specific alternative splicing, and it is mostly absent from the mRNA expressed in muscles.

**Supplementary Table 1**

**Transcript lengths from down and up regulated genes.** Genes that are down regulated also have a longer transcript length compared to up regulated and non-differentially expressed genes. For down regulated genes, the number of transcripts is higher and their shortest/longest transcripts are larger.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Median | No. of genes | Gene length | No. of transcripts | Shortest transcript | Longest transcript |
| >1.5x down | 1,821 | 8,440 | 3 | 3,220 | 4,263 |
| >2x down | 1,340 | 9,706 | 3 | 3,301 | 4,533 |
| >4x down | 567 | 12,162 | 3 | 3,722 | 5,168 |
| >16x down | 32 | 14,213 | 4 | 3,486 | 5,761 |
| Median | No. of genes | Gene length | No. of transcripts | Shortest transcript | Longest transcript |
| >1.5x up | 1,339 | 1,584 | 2 | 1.034 | 1,186 |
| >2x up | 572 | 1,411 | 2 | 901 | 1,027 |
| >4x up | 89 | 1,353 | 1 | 885 | 994 |
| >16x up | 15 | 1,411 | 1 | 917 | 1,020 |
| Median | No. of genes | Gene length | No. of transcripts | Shortest transcript | Longest transcript |
| Unchanged | 4,754 | 2,496 | 2 | 1,642 | 1,856 |

Supplementary table 2

Sequences of primers used in this manuscript

|  |  |  |
| --- | --- | --- |
| Primer name | Primer sequence | Primer purpose |
| CG7378-exon2-5R | TGATGGTTTGATGGACTGATG | circRNA analysis |
| CG7378-Exon5-3F | GGGATGTCGGTATGGTCAG | circRNA analysis |
| Hs6st-exon\_3-5R | TAAAGCGAGGATATCCGAAGAC | circRNA analysis |
| Hs6st-exon\_5-3F | ACAGTCTGGACATTGAGCTG | circRNA analysis |
| Src42A-exon2-5R | CTGTGGGCTTATTGTTCACC | circRNA analysis |
| Src42A-exon2-3F | GACGACGGCCAAGAAAGGC | circRNA analysis |
| tn-exon\_7-5R | GCAACTTGCTGAGCTCTCG | circRNA analysis |
| tn-exon\_7-3F | GGTACCAGCGCCACATCAC | circRNA analysis |
| rp49\_F1 | ATCCGCCACCAGTCGGATCGATATGCTAAG | Normalisation of Q-RT-PCR and circRNA |
| rp49\_R2 | TCTTGAGAACGCAGGCGACCGTTGGGGTTG | Normalisation of Q-RT-PCR and circRNA |
| CG9297\_F | AACAGATGGTGTCAGGGATGG | Q-RT-PCR |
| CG9297\_R | ATCAACCGCAGAATGCCCAC | Q-RT-PCR |
| CG2924\_F | GCCACTCTCGTCAAGGGAAA | Q-RT-PCR |
| CG2924\_R | CGGGGTGAACCATCCATTCT | Q-RT-PCR |
| verm\_F | GTGAAGTCCTGCGATGTGCT | Q-RT-PCR |
| verm\_R | GACTTGCCGTTGCAGAAGAG | Q-RT-PCR |
| r\_F | CGGTCTCCCATCGTTTGTCA | Q-RT-PCR |
| r\_R | TCCTTGGCGGAACGAATCAA | Q-RT-PCR |
| ci\_F | GTACACGCGGTGAAAAACCA | Q-RT-PCR |
| ci\_R | AGCAGCCTTCAAACGTGCAT | Q-RT-PCR |
| lid\_F | ACTTGGTGCCCACTGAGATG | Q-RT-PCR |
| lid\_R | CCAGAGCCGTGATCCATTGT | Q-RT-PCR |
| CG31999\_F | CCCAAATATGTCCCAGCGGT | Q-RT-PCR |
| CG31999\_R | AGCAGTGGTATCCGCCATTC | Q-RT-PCR |
| wnd\_F | ACCGACAATCTGCGAAGGAG | Q-RT-PCR |
| wnd\_R | GAGAGTCCAGGTCGGAGGAT | Q-RT-PCR |
| Nak\_F | CTGCCGCTGTGTCTCCTTAC | Q-RT-PCR |
| Nak\_R | CACCAGATGAAACGCCGATG | Q-RT-PCR |
| CG5522\_F | CAACAAGCGTTGCAGATGCC | Q-RT-PCR |
| CG5522\_R | TCTAGCAGGTGGTGACGAGG | Q-RT-PCR |
| CG2865\_F | TAACCGAAAACGCCACCTGT | Q-RT-PCR |
| CG2865\_R | TAGCAGCATGGGTGCATTGA | Q-RT-PCR |
| PRL-1\_F | AGGCATTACCGTCAAGGACC | Q-RT-PCR |
| PRL-1\_R | CCAGACCAGCCACACAATGA | Q-RT-PCR |
| CG5886\_F | AACGAAAAGGAGCAGACCGT | Q-RT-PCR |
| CG5886\_R | GATTAGGGGAGGCACTTCGG | Q-RT-PCR |
| CG1307\_F | CGCAGGAATCTTCGGTGTCT | Q-RT-PCR |
| CG1307\_R | TTGCCCTTGCCCATCTTGAG | Q-RT-PCR |
| CG9330\_F | GAGACTGTGGTCACGCTCAA | Q-RT-PCR |
| CG9330\_R | GAGTCCGATGTGGCTGAGAG | Q-RT-PCR |
| abba\_F | ACTGGTCCCTTGGGCTTTTC | Q-RT-PCR |
| abba\_R | CGGAAGGTGCAGGAAGTTCA | Q-RT-PCR |
| Nxt1\_F | ACCGACGCCAACAAATTGGA | Q-RT-PCR |
| Nxt1\_R | GTTTGAGGATGGCAGCTCCT | Q-RT-PCR |

Supplementary data file

CircRNA\_data.xlsx contains library sequence depth and mapping information for the circRNA sequencing experiments. BPM tab contains the raw read information and BPM calculations for all of the identified circRNA structures. Each circRNA structure is via the back-splicing event, in the format transcriptID.SpliceDonorExon.SpliceAcceptorExon.