**Supplemental Appendix**

## Direct cDNA sequencing, read alignment and transcript quantification

The quality and quantity of the RNA were assessed with Agilent 2100 Bioanalyzer, using the Agilent RNA 6000 Nano kit (Agilent Technologies, USA). An RNA integrity number (RIN) threshold, above which a sample was used for library preparation, was set at 5. The selection of poly(A)+ RNA molecules was performed using Lexogen’s Poly(A) Selection RNA Kit (Lexogen GmbH, Vienna, Austria). The concentration was subsequently determined with the use of RNA HS Assay Kit (Thermo Fischer Scientific) on Qubit Fluorometer (Thermo Fisher Scientific, USA). We then constructed an RNA library based on the Direct cDNA Native Barcoding (SQK-DCS109 with EXP-NBD104 and EXP-NB114 by Oxford Nanopore Technologies Ltd.). After barcoding, each sample was tested for quality and quantity with Agilent 2100 Bioanalyzer, using the Agilent High Sensitivity DNA kit as instructed by the manufacturer (Agilent Technologies, USA). Flowcell loading was performed according to the instructions using the EXP-FLP002 priming mix (Oxford Nanopore Technologies Ltd., United Kingdom). Upon sequencing completion, raw data were basecalled using Guppy Basecalling Software: Version 4.4.0+3a263d4b1 (Oxford Nanopore Technologies, UK). Reads were filtered using parameters --qscore\_filtering –min\_qscore 7, representing quality score above 7, and a basecall accuracy of ~80% or higher.

Before alignment, fastq files from the left and right atria of each sample were merged to obtain higher read depth. Reads were aligned to the Norwegian Brown reference transcriptome using *fasta* and *gtf* files provided by Ensembl (Rnor 7.2)1. We used the minimap2 software (version 2.24-r1112)2, with arguments *-ax map-ont* for direct cDNA nanopore sequencing data. The index was constructed using parameters *-k15 -w5* in minimap2. To make the aligner focus on annotated splice junctions, we added the argument *--junc-bed*, and a bed file was constructed from the Ensembl *gff3* file, using BEDOPS3. Following this, *bam* files were sorted and indexed using SAMtools (version 1.15.1)4.

Transcript abundance estimation was conducted using Salmon (version 1.8.0)5 with the *–noErrorModel -p -l U* arguments. Our pipeline was based on code provided in the Oxford Nanopore Technologies *pipeline-transcriptome-de* pipeline for cDNA transcriptome analyses(available at: [https://github.com/ nanoporetech/pipeline-transcriptome-de](https://github.com/%20nanoporetech/pipeline-transcriptome-de)). Gene counts were merged from raw transcript counts using the python based *merge\_count\_tsvs.py* script in this pipeline.

## Genetic pathway analysis

We used GSEA6, a threshold-free method for gene-set enrichment analysis, to analyze significantly up- and downregulated gene sets. As input, we analyzed differentially expressed genes identified in atrial tissue from the *RBM20*+/- and *RBM20*-/- rat models, and the gene-sets “biological processes” (c5.go.bp.v7.5.1) provided at <https://www.gsea-msigdb.org/>. Genes were ranked based on the test statistic “stat” generated in the DESeq2 analysis.

We used the fast GSEA package7 version 1.20.0 in R version 4.1.2, using default settings. Following analysis, we filtered for gene-sets with an FDR corrected P-value < 0.01. Similar pathways were collapsed using the FGSEA function *collapsePathways*, using default settings, resulting in 74 enriched gene sets in *RBM20*+/- rat atria and 118 enriched gene-sets *RBM20*-/- rat atria. The ten most up- and downregulated gene-sets in RBM20+/- and RBM20-/- are presented in **Figure 2F and 3F,** respectively. All enriched gene sets are summarized in **Supplemental Data 5-6**.

## Identification of human RBM20 mutations with reduced activity

We cultured human embryonic kidney 293 (HEK293) cells in DMEM supplemented with 10% (v/v) FBS, 100 U/ml penicillin and 100 ug/ml streptomycin. We used PEI40 to transfect plasmids in DMEM without antibiotics (10 min complex formation, 48h incubation). To investigate TTN241-3 and *RBM20* variants, we co-transfected with equimolar ratios of the two, and DNA/PEI40 ratio of 1:3 at 440 ng DNA/per well was used for transfection of a well in a 6-well dish. Total RNA was collected 48 h after transfection. Each transfection experiment was performed using three technical replicates and repeated three times. Cells from each transfected well were equally divided for RNA and protein analysis.

To isolate total RNA from cells, we used Trizol (Invitrogen). Preparations of <2 μg of total RNA were treated with DNase I (Thermo Fisher Scientific), according to manufacturer instructions. For first-strand cDNA synthesis, we used High-Capacity RNA-to-cDNA kit (Applied Biosystems).

Quantitative RT-PCR was performed using SYBR Green master mix (Applied Biosystems) in a 7900 HT cycler (Applied Biosystems). The quantification of the gene expression was performed using the ∆∆CT method. Relative levels of splice isoforms are presented as a ratio of mRNAs, with exon 242 included, versus mRNAs, with exon 242 skipped. The fold change in inclusion/skipping ratio was obtained compared to the control (transfection with pcDNA plasmid). N=3 for all samples, all data are expressed as mean ± SEM. Group comparisons were analyzed by one-way ANOVA and Bonferroni posttest.

For immunodetection, HEK293 cells were harvested by rinsing them from the cell culture dish using PBS and snap-frozen in liquid nitrogen. For protein extraction samples were suspended in RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40, 0.25% sodium deoxycholate, 1.0 mM EDTA; 0.1% SDS) and lysed via sonication with 20 bursts at 70% energy using a VialTweeter (Hielscher Ultrasonics GmbH, Teltow, Germany). Protein concentration was estimated after 20’ precipitation of cell debris at 13,000g for 20 min at 4˚C using BCA assay (Life Technologies GmbH, Darmstadt, Germany). Samples were kept cool at all times. 0.4 mg/ml lysates were separated by WesTM according to manufacturer’s instructions. Antigens were probed with rabbit anti-RBM20, (PA5-57404, Thermo Fisher Scientific, USA) and rabbit anti--tubulin (2146S, Cell Signaling Technology, USA).

# Supplemental Table 1. Sequences of oligonucleotides used for site-directed mutagenesis

|  |  |  |
| --- | --- | --- |
| **Mutation name** | **Primer name** | **Sequence** |
| T177I | hRBM20-T177I\_F | 5’-ggcctcgtatctggctggggggtg-3’ |
| T177I | hRBM20-T177I\_R | 5’-caccccccagccagatacgaggcc-3’ |
| Q345\* | hRBM20-Q345stop\_F | 5’-cagctcatagggctagttgtggggtggag-3’ |
| Q345\* | hRBM20-Q345stop\_R | 5’-ctccaccccacaactagccctatgagctg-3’ |
| G404S | hRBM20-G404S\_F | 5’-tgaacgactttcacagtgtggcccccctc-3’ |
| G404S | hRBM20-G404S\_R | 5’-gaggggggccacactgtgaaagtcgttca-3’ |
| P411L | hRBM20-P411L\_F | 5’-atgctacagatatgcagcaagtggagggggg-3’ |
| P411L | hRBM20-P411L\_R | 5’-cccccctccacttgctgcatatctgtagcat-3’ |
| L429P | hRBM20-L429P\_F | 5’-cttccctttcacatgcggctcccagtccttca-3’ |
| L429P | hRBM20-L429P\_R | 5’-ttgaaggactgggagccgcatgtgaaagggaag-3’ |
| E457G | hRBM20-E457G\_F | 5’-aagcacacaatgttccccctgccgaaccaagtata-3’ |
| E457G | hRBM20-E457G\_R | 5’-tatacttggttcggcagggggaacattgtgtgctt-3’ |
| E725Vfs\*34 | hRBM20-E725Vfs\*34\_F | 5’-gtcgctcgtccaacgtccaactcagccttgtccag-3’ |
| E725Vfs\*34 | hRBM20-E725Vfs\*34\_R | 5’-ctggacaaggctgagttggacgttggacgagcgac-3’ |
| E728G | hRBM20-E728G\_F | 5’-acgagcgaccaggaggagggaggc-3’ |
| E728G | hRBM20-E728G\_R | 5’-ggcctccctcctcctggtcgctcgt-3’ |
| E787\* | hRBM20-E787stop\_F | 5’-gcagcctggcctagtctttcctcctggacc-3’ |
| E787\* | hRBM20-E787stop\_R | 5’-ggtccaggaggaaagactaggccaggctgc-3’ |
| E1101K | hRBM20-E1101K\_F | 5’-tccggggtcaacactttttggcaagcttggtttt-3’ |
| E1101K | hRBM20-E1101K\_R | 5’-aaaaccaagcttgccaaaaagtgttgaccccgga-3’ |
| V1155A | hRBM20-V1155A\_F | 5’-agtcctgggaaccgcgaactccaccc-3’ |
| V1155A | hRBM20-V1155A\_R | 5’-ggggtggagttcgcggttcccaggac-3’ |
| S1195Y | hRBM20-S1195Y\_F | 5’-tcggccagctggtacaaatatttctgtaagttcctg-3’ |
| S1195Y | hRBM20-S1195Y\_R | 5’-caggaacttacagaaatatttgtaccagctggccga-3’ |

## **Supplemental Table 2**. Clinical characteristics of individuals with *RBM20* variants.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Study ID | Sex | Variant | | | Age of AF onset (years) | Comorbidities | Type of AF |
| Position  (GRCh37/hg19) | Amino-acid change | Consequence |
| A-01 | Male | 10:112541400 | p.(Gln345\*) | STOP-GAINED | 47 | None | Paroxysmal |
| A-02 | Male | 10:112541577 | p.(Gly404Ser) | MISSENSE | 28 | None | Persistent |
| A-03 | Male | 10:112541599 | p.(Pro411Leu) | MISSENSE | 30 | None | Persistent |
| A-04 | Male | 1 0:112543134 | p.(Leu429Pro) | MISSENSE | 40 | None | Paroxysmal |
| A-05 | Female | 10:112543134 | p.(Leu429Pro) | MISSENSE | 41 | None | Persistent |
| A-06 | Male | 10:112543134 | p.(Leu429Pro) | MISSENSE | 21 | None | Paroxysmal |
| A-07 | Female | 10:112543134 | p.(Leu429Pro) | MISSENSE | 28 | None | Paroxysmal |
| A-08 | Female | 10:112544131 | p.(Glu457Gly) | MISSENSE | 28 | None | Paroxysmal |
| A-09 | Female | 10:112572302 | p.(Arg716Gln) | MISSENSE | 29 | None | Paroxysmal |
| A-10 | Male | 10:112572302 | p.(Arg716Gln) | MISSENSE | 26 | None | Paroxysmal |
| A-11 | Male | 10:112572320 | p.(Glu725Valfs\*34) | FRAMESHIFT | 22 | None | Persistent |
| A-12 | Male | 10:112572338 | p.(Glu728Gly) | MISSENSE | 38 | None | Persistent |
| A-13 | Male | 10:112572514 | p.(Glu787\*) | STOP-GAINED | 26 | None | Persistent |
| A-14 | Male | 10:112581678 | p.(Glu1101Lys) | MISSENSE | 36 | None | NC |
| A-15 | Male | 10:112581678 | p.(Glu1101Lys) | MISSENSE | 37 | None | Paroxysmal |
| A-16 | Male | 10:112590831 | p.(Val1155Ala) | MISSENSE | 28 | None | Persistent |
| A-17 | Male | 10:112595636 | p.(Ser1195Tyr) | MISSENSE | 28 | None | Paroxysmal |

**Abbreviations**: AF, Atrial Fibrillation; NC, Not classified

## **Supplemental Table 3**. Characteristics of the UK Biobank cohort

| UK Biobank cohort (n = 404,012) | |
| --- | --- |
| Age, years, median (Q1-Q3) | 58 (51-63) |
| Female sex, n (%) | 218,753 (54.1%) |
| BMI, kg/m2, mean (SD) | 27.4 (4.7) |
| Systolic BP, mmHg, mean (SD) | 140 (20) |
| Diastolic BP, mmHg, mean (SD) | 82 (11) |
| Atrial fibrillation, n (%) | 31,843 (7.9%) |

Q1-Q3, 25th-75th percentiles; SD, Standard deviation

## **Supplemental Table 4**. RBM20 Loss-of-function variants in the UK Biobank

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Position  (GRCh38/hg38) | DNA-change | Amino-acid change | Consequence | OR (95% CI) | P-value |
| 10:110780967 | c.359delT | p.(L120Rfs\*18) | Frameshift | 3.54 (1.54; 8.13) | 0.003 |
| 10:110821669 | c.3053\_3056delCCCT | p.(S1018Wfs\*85) | Frameshift | 5.76 (1.23; 26.83) | 0.026 |
| 10:110821305 | c.2687delA | p.(E896Gfs\*14) | Frameshift | 26.09 (1.43; 475.53) | 0.028 |
| 10:110784340 | c.1338-1G>T | NA | Splice Acceptor site variant | 2.80 (0.29; 27.05) | 0.37 |
| 10:110821879 | c.3267dupC | p.I1090Hfs\*4 | Frameshift | 1.94 (0.14; 26.79) | 0.62 |

CI, Confidence intervals, GRCh38/hg38, Genome Reference Consortium Human Build 38.

## **Supplemental Table 5**. Results from genetic association analysis between loss of *RBM20* and cardiac dimensions

|  |  |  |  |
| --- | --- | --- | --- |
| **Cardiac Chamber** | **Measurement** | **Beta** | **P-value** |
| Left Atria | Left atrial maximum volume | 0.6586 | 0.067 |
| Left atrial minimum volume | 0.7685 | 0.035 |
| Left atrial total emptying fraction | -0.6210 | 0.089 |
| Left atrial active emptying fraction | -0.5634 | 0.13 |
| Left atrial passive emptying fraction | -0.2872 | 0.39 |
| Right Atria | Right atrial maximum volume | 0.0779 | 0.82 |
| Right atrial minimum volume | 0.0924 | 0.78 |
| Right atrial emptying fraction | 0.1400 | 0.70 |
| Right atrial stroke volume | 0.3110 | 0.39 |
| Left Ventricle | Left ventricular end-systolic volume | 0.5477 | 0.064 |
| Left ventricular end-diastolic volume | 0.5234 | 0.072 |
| Left ventricular ejection fraction | -0.5265 | 0.14 |
| Left ventricular stroke volume | 0.2664 | 0.41 |
| Right Ventricle | Right ventricular end-systolic volume | 0.2150 | 0.44 |
| Right ventricular end-diastolic volume | 0.2523 | 0.37 |
| Right ventricular ejection fraction | -0.1500 | 0.67 |
| Right ventricular stroke volume | 0.2916 | 0.35 |

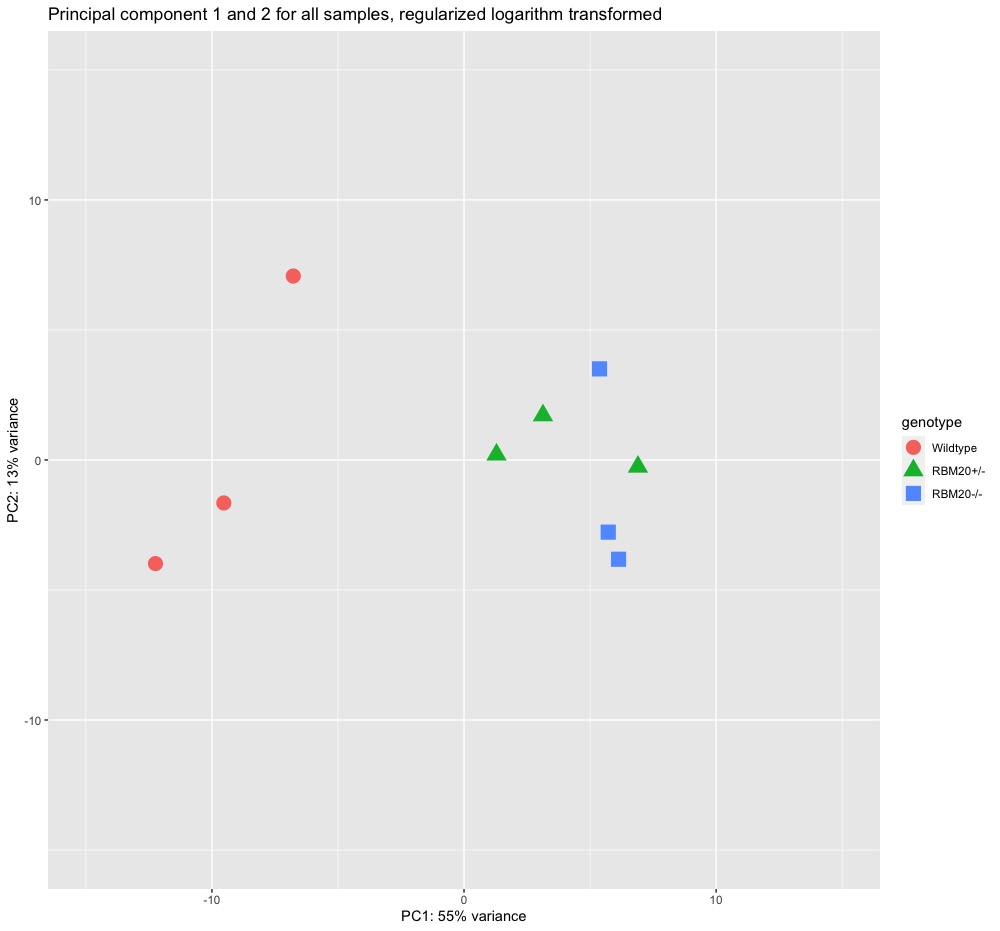
# Supplemental Table 6. Direct cDNA mapping parameters

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sample** | **Genotype** | **QC-passed reads** | **Mean read length [median]** | **Mean read quality** | **Primary alignments** | **Mapped reads (%)** |
| A1 | Wildtype | 1,398,064 | 653.9 [504] | 19.7 | 1,158,623 | 82.87 % |
| A2 | Wildtype | 1,508,077 | 629.4 [481] | 20.3 | 1,217,094 | 80.71 % |
| A3 | Wildtype | 3,708,577 | 633.8 [471] | 20.3 | 3,028,741 | 81.67 % |
| B1 | *RBM20*+/- | 1,722,632 | 725.7 [570] | 19.7 | 1,469,045 | 85.28 % |
| B2 | *RBM20*+/- | 1,415,229 | 719.0 [574] | 19.3 | 1,225,094 | 86.57 % |
| B3 | *RBM20*+/- | 3,130,167 | 779.7 [631] | 19.6 | 2,755,474 | 88.03 % |
| C1 | *RBM20*-/- | 2,292,182 | 742.4 [608] | 19.1 | 2,045,410 | 89.23 % |
| C2 | *RBM20*-/- | 1,497,232 | 754.8 [587] | 20.0 | 1,937,591 | 84.93 % |
| C3 | *RBM20*-/- | 2,281,279 | 797.5 [630] | 19.8 | 1,255,640 | 83.86 % |

# Supplemental Table 7. Putative splicing targets of RBM20

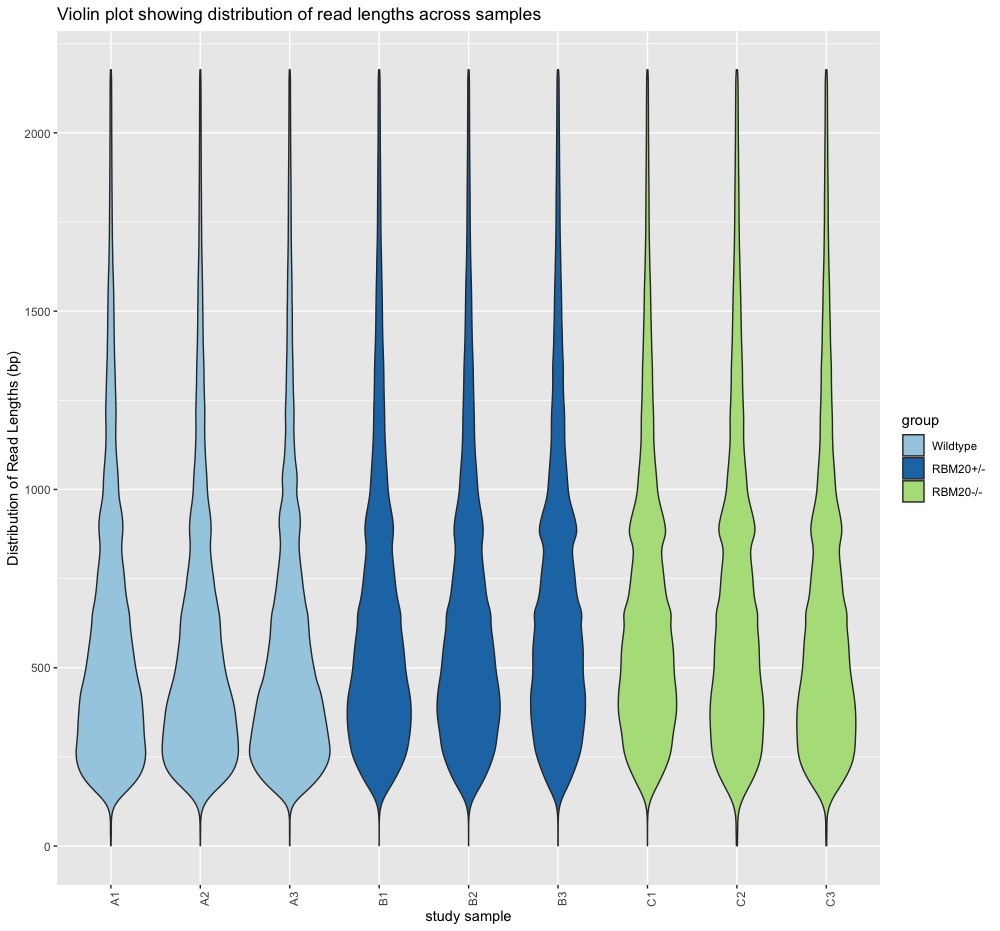
|  |  |  |
| --- | --- | --- |
| **Gene** | **Protein** | **Evidence/Reference** |
| *APTX* | Aprataxin | Guo et al.8 |
| *CACNA1C* | Cav1.2 | Guo et al.8 |
| *CAMK2D* | Calcium/calmodulin-dependent protein kinase type II delta chain | Guo et al.8, Maatz et al.9 |
| *CAMK2G* | Calcium/calmodulin-dependent protein kinase type II gamma chain | Guo et al.8 |
| *DAB1* | Disabled-1 | Guo et al.8 |
| *DNM3* | Dynamin-3 | Guo et al.8 |
| *DST* | Dystonin | Maatz et al.9 |
| *DTNA* | Dystrobrevin Alpha | Guo et al.8 |
| *ENAH* | ENAH Actin Regulator | Maatz et al.9 |
| *FHOD3* | Formin Homology 2 Domain Containing 3 | Guo et al.8 |
| *FNBP1* | Formin-binding protein 1 | Guo et al.8 |
| *GIT2* | ARF GTPase-activating protein GIT2 | Guo et al.8 |
| *KALRN* | Kalirin | Guo et al.8 |
| *KCNIP2* | Kv channel-interacting protein 2 | Guo et al.8 |
| *LDB3* | LIM domain binding 3 | Guo et al.8, Maatz et al.9 |
| *IMMT* | Inner Membrane Mitochondrial Protein | Maatz et al.9 |
| *LMO7* | LIM Domain 7 | Maatz et al.9 |
| *LRRFIP1* | LRR Binding FLII Interacting Protein 1 | Maatz et al.9 |
| *MECP2* | Methyl CpG binding protein 2 | Guo et al.8 |
| *MLIP* | Muscular LMNA Interacting Protein | Maatz et al.9 |
| *MTMR1* | Myotubularin-relateret protein 1 | Guo et al.8 |
| *MYH7* | Beta-myosin heavy chain | Maatz et al.9 |
| *MYOM1* | Myomesin 1 | Maatz et al.9 |
| *NEXN* | Nexilin F-Actin Binding Protein | Maatz et al.9 |
| *NFIA* | Nuclear factor 1 A-type | Guo et al.8 |
| *NPRL3* | Nitrogen permease regulator-like 3 | Guo et al.8 |
| *NTRK3* | Tropomyosin receptor kinase C | Guo et al.8 |
| *OBSCN* | Obscurin | Maatz et al.9 |
| *PDLIM3* | PDZ And LIM Domain 3 | Maatz et al.9 |
| *PDLIM5* | PDZ and LIM domain protein 5 | Guo et al.8 |
| *PLEKHA5* | Pleckstrin homology domain-containing family A member 5 | Guo et al.8 |
| *RALGPS1* | Ral GEF with PH domain and SH3 binding motif 1 | Guo et al.8 |
| *RTN4* | Reticulon 4 | Maatz et al.9 |
| *RYR2* | Ryanodine receptor 2 | Maatz et al.9 |
| *SEMA6D* | Semaphorin 6D | Guo et al.8 |
| *SH3KBP1* | SH3 domain-containing kinase-binding protein 1 | Guo et al.8 |
| *SLC38A10* | Putative sodium-coupled neutral amino acid transporter 10 | Guo et al.8 |
| *SORBS1* | CAP/Ponsin protein | Guo et al.8, Maatz et al.9 |
| *SPEN* | Msx2-interacting protein | Guo et al.8 |
| *TNNT2* | Troponin 2 | Maatz et al.9 |
| *TPM1* | Tropomyosin 1 | Guo et al.8 |
| *TRDN* | Triadin | Maatz et al.9 |
| *TTN* | Titin | Guo et al.8, Maatz et al.9 |
| *UBE2F* | Ubiquitin conjugating enzyme E2 F | Guo et al.8 |
| *ZNF451* | Zinc finger protein 451 | Guo et al.8 |

# Supplemental Figure 1.



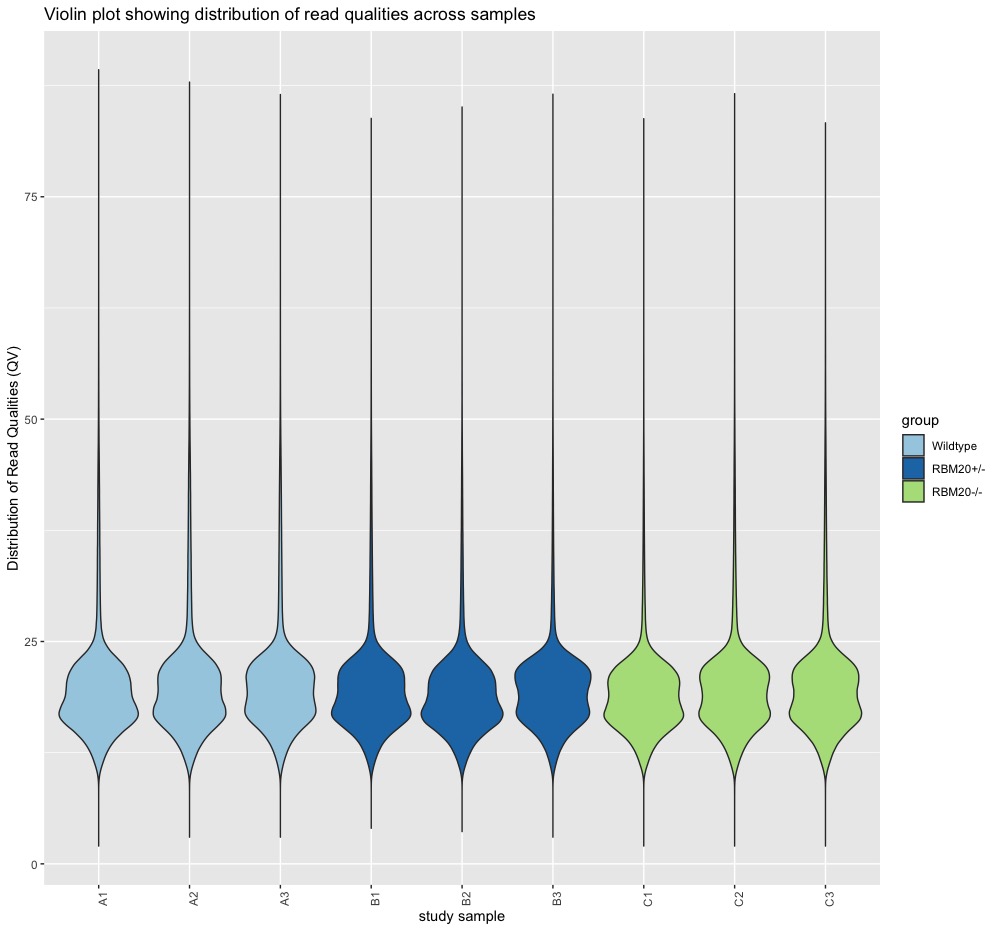
*Supplemental Figure 1:* Principal component analysis plot of all samples. Principal component 1 is on the X-axis and accounts for 55% of variance between samples. Principal component 2 is on the Y-axis and accounts for 13% of variance between samples. Data have been normalized using the regularized transformation function of DESeq2. Samples from wildtype control rats are depicted as red circles, samples from *RBM20*+/- rats are depicted as green triangles, while samples from *RBM20*-/- rats are shown as blue squares.

# Supplemental Figure 2.



*Supplemental Figure 2:* Violin plot of read lengths of cDNA fragments in all samples.

# Supplemental Figure 3.



*Supplemental Figure 3:* Violin plot of read qualities of cDNA fragments in all samples.

# References (Supplemental):

1. Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Research* **50**, D988–D995 (2022).

2. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

3. Neph, S. *et al.* BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**, 1919–1920 (2012).

4. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

5. Salmon provides fast and bias-aware quantification of transcript expression | Nature Methods. https://www.nature.com/articles/nmeth.4197.

6. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545–15550 (2005).

7. Sergushichev, A. A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. 060012 Preprint at https://doi.org/10.1101/060012 (2016).

8. Guo, W. *et al.* RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med.* **18**, 766–773 (2012).

9. Maatz, H. *et al.* RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. *J Clin Invest* **124**, 3419–3430 (2014).