Loss of Cardiac Splicing Regulator RBM20 Is Associated With Early-Onset Atrial Fibrillation

Oliver B. Vad, MD,a,b,* Elisavet Angeli, MSC,b,* Martin Liss, PhD,c Gustav Ahlberg, MD, PhD,a Ingrid E. Christophersen, MD, PhD,a,b Camilla C. Hansen, PhD,1 Sophie Møller, PhD,1 Ylva Hellsten, PhD,1 Stig Haunsoe, MD, DMS,2 Arnljot Tveit, MD, PhD,2,8 Jesper H. Svendsen, MD,3,8 Michael Gotthardt, MD, PhD,1,8,9 Pia R. Lundegaard, PhD,a,b,10 Morten S. Olesen, PhD,a,b,10

HIGHLIGHTS

- Loss of function variants in the key cardiac splicing regulator RBM20 were associated with atrial fibrillation in 2 independent cohorts.
- Loss of RBM20 facilitated structural and transcriptional changes in the atria and affected mitochondrial function.
- Alternative splicing may act as a novel proarrhythmic mechanism in the atria through atrial remodeling and altered mitochondrial energetics.
We showed an association between atrial fibrillation and rare loss-of-function (LOF) variants in the cardiac splicing regulator RBM20 in 2 independent cohorts. In a rat model with loss of RBM20, we demonstrated altered splicing of sarcomere genes (NEXN, TTN, TPM1, MYOM1, and LDB3), and differential expression in key cardiac genes. We identified altered sarcomere and mitochondrial structure on electron microscopy imaging and found compromised mitochondrial function. Finally, we demonstrated that 3 novel LOF variants in RBM20, identified in patients with atrial fibrillation, lead to significantly reduced splicing activity. Our results implicate alternative splicing as a novel proarrhythmic mechanism in the atria. (J Am Coll Cardiol Basic Trans Science 2024;9:163-180) © 2024 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
function. Finally, we observed reduced splicing activity in HEK293-cell models with variants identified in the early-onset AF cohort.

METHODS

EARLY-ONSET AF COHORT. We recruited Scandinavian patients with early-onset AF and no other cardiovascular disease. Early-onset AF was defined as onset of AF at <50 years of age. Danish patients were identified and recruited through the Danish National Patient Registry, with AF defined by International Classification of Diseases, 10th Revision code I48. The validity of the AF diagnosis in the Danish National Health Registry has previously been shown to be approximately 92.5%.12 Individuals fulfilling the inclusion criteria were given written information and were offered the opportunity to participate. We collected clinical information and peripheral blood samples. The study was approved by the scientific ethics committee of the Capital Region of Denmark (protocol number: H-2004862).

Norwegian patients were recruited from the region of Oslo and Vestre Viken, Norway. Patients were identified through clinical practice and the AF registries at Vestre Viken Hospital Trust and Oslo University Hospital. AF was defined as AF for more than 30 seconds on an electrocardiogram. The study was approved by the South-Eastern Regional Ethics Committee (Project no.: 240149) in Norway.

Patients with diabetes, hypertension, hyperthyroidism, or other cardiac disease (congenital heart defects, valvular disease, congestive heart failure, ischemic heart disease) at inclusion were excluded. All participants gave written informed consent.

We extracted genomic DNA from leukocytes in peripheral blood samples from early-onset AF patients. Endonucleases were used to fragment the DNA, and gene-specific probes from the Illumina TruSight Cardio (Illumina Sequencing) sequencing kit were used to create hybridized molecules. We then captured the hybridized molecules using magnetic beads, and following polymerase chain reaction amplification, performed sequencing using Illumina HiSeq 2500 and NextSeq technology. The sequenced reads were trimmed from adapter sequences, filtered for low-quality reads, aligned to the Human Reference Genome (build GRCh37/hg19) using Burrows-Wheeler Aligner software,13 and postprocessed according to Genome Analysis Toolkit best practice guidelines.14 Our sequencing methods and bioinformatics pipeline have previously been described in detail by Ahlberg et al.7

REFERTENT POPULATIONS. We compared the burden of variants predicted to cause LOF in RBM20 in our data with available genetic data from 2 large population-based datasets: the SweGen cohort containing genetic data on 1,000 participants, representing a cross section of the Swedish population,15 and the Genome Aggregation Database (gnomAD) version 3, containing genetic information on 76,156 unrelated individuals.16

Categorical data are presented as counts and percentages and compared between the early-onset AF and control cohorts using the Fisher exact test because of small sample sizes. R version 4.1.2 was used for analysis.17 A P value of <0.05 was considered significant.

UK BIOBANK EXOMES. The UK Biobank is a large population-based cohort study, comprising almost 500,000 individuals aged 40 to 69 years at inclusion. The biobank includes genetic and clinical information, including health records, representative of the general population. All participants in the UK Biobank gave written informed consent. The study protocol and data collection have been described in detail by Bycroft et al.18

We examined the burden of rare LOF variants in RBM20 in whole-exome sequencing data from the UK Biobank.19 We accessed the whole-exome sequencing dataset, of which 404,012 participants were of European ancestry. Detailed information on exome sequencing methodology, quality control, alignment, variant calling, and annotation in the dataset has been described in detail previously.20 Only variant sites where at least 90% of the genotypes had a read depth of >10 were included in the subsequent analyses. We filtered samples based on heterozygosity, missing rates, excess relatedness, and kinship inference, based on UK Biobank resource 531, and excluded samples with disagreement between reported sex and genetically determined sex.

AF was defined by UK Biobank data field 131351, corresponding to International Classification of Diseases, 10th Revision code I48. The AF diagnosis in the UK Biobank was based on in-patient hospital records, death records, primary care records, and self-reports from interviews conducted by qualified nurses. We then conducted rare variant association tests for LOF variants with a minor allele frequency of <0.01 in RBM20. Association tests were conducted using the logistic regression model integrated in REGENIE,21 adjusted for sex, age at inclusion, and 10 first principal components. We tested for association across all rare LOF variants collapsed into a single gene-based test as well as for each rare LOF variant
individually. We applied Firth logistic correction to account for potential bias in maximum likelihood estimates caused by case-control imbalance and small sample size. Results from the logistic regressions are presented as ORs with 95% CIs.

Finally, we tested for association between rare LOF variants in RBM20 and cardiac structure and function in a subset of UK Biobank participants who had undergone cardiac magnetic resonance imaging. Imaging annotation has previously been described in detail for right heart structure,22 left ventricular structure,23 and left atrial structure.24 Genetic association tests were conducted in REGENIE and adjusted for sex, age at imaging, body surface area at imaging, imaging center, and 10 first principal components.

ANIMAL MODEL CHARACTERISTICS. We used atrial tissue from rats carrying a 95-kb deletion, resulting in the removal of exons 2 through 14 in RBM20. The rats were described by Guo et al9 and outcrossed to Sprague Dawley rats. We examined atrial tissue from 8 heterozygote RBM20+/− and 8 homozygote RBM20−/− animals. For controls, we used atrial tissue from 5 unrelated Sprague Dawley rats. The animal studies complied with all institutional and national requirements for the care and use of laboratory animals (Berlin State Office for Health and Social Affairs: Reg Y-9004/14).

TRANSMISSION ELECTRON MICROSCOPY. Atrial tissue from 3-month-old (n = 6) RBM20+/− siblings and unrelated 3-month-old WT (n = 5) rats was dissected out and separated into left and right atria. The samples were fixed in a 2% volume/volume glutaraldehyde in 0.05-mol/L sodium phosphate buffer (pH 7.4) for 24 hours. After a brief rinse in a sodium cacodylate buffer (0.15 mol/L, pH 7.4), a subsequent 2-hour postfixation step followed (1% weight/volume OsO4; 0.05-mol/L potassium ferricyanide in 0.12-mol/L sodium cacodylate buffer; pH 7.4). Samples were then dehydrated in ethanol and transferred to propylene oxide before being embedded in Epon. Samples were cut into sections of approximately 80 nm and collected on copper grids in Formvar supporting membranes. Samples were subsequently stained using uranyl acetate and lead citrate. Samples were imaged using transmission electron microscopy on a Philips CM 100 transmission electron microscope (Philips), operated at an accelerating voltage of 80 kV. Images were captured using an OSIS Veleta digital slow scan 2k × 2k charge-coupled device camera (Olympus).

DIRECT COMPLEMENTARY DNA SEQUENCING. We performed direct complementary DNA (cDNA) sequencing on atrial tissue from 3 RBM20+/− rats and 3 RBM20−/− rats and compared them with 3 WT controls. RNA was isolated from left and right atrial tissue and purified using TRIzol Reagent (Thermo Fisher Scientific) combined with the RNeasy Micro kit (Qiagen).

Reads were aligned to the Norwegian Brown reference transcriptome using fasta and gtf files provided by Ensembl (Rnor 7.2).25 We used the minimap2 software (version 2.24-r1112),26 with arguments -ax map-ont for direct cDNA nanopore sequencing data. Transcript abundance was then estimated using Salmon (version 1.8.0).27

Detailed methods for direct cDNA sequencing, read alignment, and transcript abundance quantification are summarized in the Supplemental Appendix.

DIFFERENTIAL TRANSCRIPT USAGE ANALYSIS. For differential transcript usage (DTU), we based our approach on the one described in the pipeline-transcriptome-de pipeline. Based on expression data from the Protein Atlas 28 we focused on genes that were part of the Heart Specific Proteome. Using the gene symbols in the Ensembl database, we matched the 13,668 genes on this list with 11,936 orthologues in the rat reference genome. The remaining genes were excluded before DTU analysis. Raw counts were filtered (≥1 transcript count in ≥4 samples), and a data frame was constructed using the DRIMseq package (version 1.22.0)29 and analyzed using the DESeq2 package (version 1.34.0).30 We controlled for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) post hoc method. Associations with an FDR of <5% were considered statistically significant. All analyses were conducted in R version 4.1.2.31

DIFFERENTIAL GENE EXPRESSION ANALYSIS. Before differential gene expression analysis, we performed prefiltering to ensure high quality of the data. Only genes with ≥10 counts in ≥4 samples and part of the heart-specific proteome (described earlier) were included in the analysis. Differential gene expression analysis was conducted in R with the DESeq2 package,32 using standard settings, and an FDR cutoff of <0.05 to account for multiple testing (using the Benjamini-Hochberg method). Following analysis, results were shrunk using the lfcshrink function in DESeq2, set to use the apeglm method.32 For visualization of data, data were normalized by the regularized logarithm transformation function in the DESeq2 package, and a principal component plot was constructed according to the instructions of the package vignette. Differentially expressed genes were visualized on volcano plots using the ggplot2 package in R.33 Finally, we conducted a pathway analysis of significantly enriched gene sets in RBM20−/− rats using gene
set enrichment analysis, as described in detail in the Supplemental Appendix.

MITOCHONDRIAL FUNCTION ANALYSIS. Samples for mitochondrial analyses were dissected out and were immediately placed in ice-cold Buffer X as described by Perry et al. Mitochondrial respiratory capacity was analyzed using high-resolution fluororespirometry by Oxygraph-2k technology (Oroboros Instruments) in accordance with that previously optimized for skeletal muscle. Our analyses were conducted with mitochondrial complex I- and II-linked substrates (ie, glutamate, malate, succinate, and adenine diphosphate). Mitochondrial respiration and hydrogen peroxide (H$_2$O$_2$) formation were measured in real time simultaneously. H$_2$O$_2$ levels were determined using Amplex UltraRed (Thermo Fisher Scientific) according to previously published methods. H$_2$O$_2$ is a stable form of reactive oxygen species (ROS) and will from here on be referred to as ROS. We performed statistical analyses in R (version 3.4.1). Residual and Q-Q plots were used to confirm homogeneity of covariance and normal distribution of the data. The main effect was analyzed using analysis of variance. Group differences were analyzed using a linear mixed model with normal distribution of the data. The main effect was used to conduct the Tukey post hoc test and adjusted using the DFR. A value of $P < 0.05$ was considered statistically significant, and values are reported as mean ± SEM.

IDENTIFICATION OF HUMAN RBM20 MUTATIONS WITH REDUCED ACTIVITY. We used site-directed mutagenesis to examine the splicing activity of the included variants from the Scandinavian AF cohort. Mutagenesis was performed using the Q5 proofreading DNA polymerase (New England Biolabs) and a standard protocol. The primers used (Supplemental Table 1) were designed using the QuickChange Primer Design Program and have been described in detail elsewhere. Final constructs were sequence-verified. Detailed methods for analysis of RNA and protein expression are described in the Supplemental Appendix.

RESULTS

IDENTIFICATION OF RARE VARIANTS IN RBM20 IN EARLY-ONSET AF PATIENTS. We included 531 Scandinavian individuals with early-onset AF (age at AF onset: <50 years) Individuals with left ventricular ejection fraction of <55% or structural heart disease as determined by echocardiographic examination were excluded. The median age at AF onset was 30 years (IQR, 25th-75th percentile: 24-36 years), and there was a male predominance (83.4% men).

In the early-onset AF cohort, 3 individuals carried rare LOF variants (~0.56%), all of which were novel. The clinical characteristics of variant carriers have been summarized in Supplemental Table 2. One individual (A-01) carried a variant resulting in a premature stop codon (p.Gln345*) and developed AF at age 47 years without predisposing comorbidity. Another individual (A-11) carried a frameshift variant (p.Glu725Valfs*34). This individual developed AF at age 26 years with no other comorbidities. At the time of inclusion, the individual had AF with monthly episodes. Both electric cardioversion and radiofrequency ablation had been attempted without success. A third individual (A-13) carried a variant resulting in a premature stop codon in RBM20 (p.Glu787*). The individual developed AF at the age of 22 years with no other comorbidities.

LOF VARIANTS IN RBM20 ARE ASSOCIATED WITH AN INCREASED RISK OF AF. We evaluated whether there was an association between an increased burden of rare LOF variants in RBM20 and AF. The number of variants in the Scandinavian early-onset AF cohort was compared to the number of RBM20 variants in the 2 population-based databases of normal variation, gnomAD and SweGen. The burden of LOF variants in RBM20 was significantly higher in patients with AF compared with the gnomAD cohort (0.5% vs 0.04%; Fisher exact test: $P = 0.003$). In a sensitivity analysis, excluding individuals of non-European ancestry from the gnomAD population, we saw consistent results (0.5% vs 0.06%; $P = 0.007$). We performed the same association test using the population-based, whole-genome sequencing data from SweGen representing a cross section of the Swedish population and found similar results (0.5% vs 0%; $P = 0.041$).

REPLICATION IN A POPULATION-BASED COHORT. We then assessed whole-exome sequencing data from 404,012 unrelated individuals of European ancestry from the UK Biobank. The individuals in the cohort had a median age at inclusion of 58 years (IQR, 25th-75th percentile: 51-63 years), and 54.1% were of female sex. We identified 31,843 individuals diagnosed with AF and considered the remaining 372,169 individuals as controls. Characteristics of the cohort are summarized in Supplemental Table 3.

In this population-based cohort, we replicated the association between rare LOF variants in RBM20 and AF (OR: 2.43; 95% CI: 1.48-4.00; $P < 0.001$). We identified 5 unique LOF variants identified in the cohort, of which 3 were associated with an increased OR for AF (Supplemental Table 4). We examined cardiac structure and function in a subset of
individuals who had undergone cardiac magnetic resonance imaging (44,123 and 47,005 individuals for atria and ventricles, respectively). We found that loss of RBM20 was associated with a significant increase in left atrial minimum volume (β = 0.77; P = 0.035). We also observed some suggestive, borderline-significant associations with increased left atrial maximum volume (β = 0.66; P = 0.067) and decreased left atrial total emptying fraction (β = -0.62; P = 0.089). Results are illustrated in Figure 1 and summarized in Supplemental Table 5.

DIFFERENTIAL ISOFORM USAGE IN RBM20-DEPENDENT TRANSCRIPTS. We investigated the transcriptional consequences of loss of RBM20 function using nanopore long-read direct cDNA sequencing on 9 atrial tissue samples from wild-type (WT) rats (n = 3), rats with heterozygous loss of RBM20 (RBM20+/−) (n = 3), and rats with homozygous loss of RBM20 (RBM20−/−) (n = 3). The mean number of quality control passed reads per sample was 2,105,938 (range: 1,398,064-3,708,577). Metrics of read lengths, quality, and mapping parameters are summarized in Supplemental Table 6. All findings were corrected for multiple testing using a Benjamini-Hochberg FDR of <5%.

In an analysis of DTU, we observed a switch in isoform expression in several putative splicing targets of RBM20. In atria from RBM20+/− rats, we observed altered expression of transcripts in the TPM1 gene, encoding the actin-binding tropomyosin 1 protein and in LDB3, which encodes the important sarcomere-stabilizing LIM domain binding 3 protein (Table 1). In RBM20−/− rat atria, DTU was also identified in LDB3 and TPM1 and several additional genes spliced by RBM20, including TTN. We also identified DTU in NEXN, encoding the Z-disk protein nexilin; MYOM1, encoding the M-band protein myomesin 1; and KCNIP2, which encodes the ion-channel

![Volcano plot of results from genetic association test between LOF variants in RBM20 and cardiac magnetic resonance imaging data. The figure shows an association with increased left atrial minimum volume. Associations with P < 0.10 are labeled with beta coefficients (β) and P values. LAMAX = left atrial maximum volume; LAMIN = left atrial minimum volume; LATEF = left atrial total emptying fraction; LOF = loss of function; LVEDV = left ventricular end-diastolic volume; LVESV = left ventricular end-systolic volume.](image-url)
interacting protein Kv channel-interacting protein 2 (Table 2). Results from the DTU analysis are summarized in the Supplemental Appendix. All putative splicing targets of RBM20 are summarized in Supplemental Table 7.

**LONG-READ cDNA SEQUENCING REVEALS DIFFERENTIAL GENE EXPRESSION.** In RBM20−/− rat atria, we identified 444 differentially expressed genes compared with WT rats, (Figures 2E and 2F). The altered gene expression was more pronounced in RBM20−/− rat atria, where we observed 585 genes with differential expression compared with WT rats (Figures 3E and 3F). Among the most significant changes in expression were an up-regulation of the MYL4 gene (log2 fold change: 1.17; $P = 3.27 \times 10^{-22}$) and of RYR2 (log2 fold change: 0.95; $P = 1.47 \times 10^{-15}$) in RBM20−/− rat atria, with similar findings in RBM20−/− rat atria. Gene set enrichment analysis indicated enrichment of altered expression in genes that are involved in mitochondrial and sarcomere function. Among the 10 most up-regulated gene sets in both RBM20−/− and RBM20+/− atria were cytoplasmic translation and cellular respiration (Figures 2F and 3F), whereas gene sets related to cell migration such as actin-cytoskeleton reorganization were down-regulated (Figure 2F and 3F). Notably, we also observed the gene set sarcomere organization among the most up-regulated pathways in RBM20−/− atria (Figure 3F) ($P = 0.001$). Results from differential gene expression and pathway analyses are summarized in the Supplemental Appendix.

**ALTED ATRIAL TISSUE STRUCTURE IN AN RBM20-DEFICIENT RAT MODEL.** We imaged atrial tissue samples from RBM20+/−, RBM20−/−, and WT rats, using transmission electron microscopy (Figure 2). Compared with WT rats, the tissue from RBM20−/− rats showed an overall tissue degradation, with disorganized sarcomere structures showing poorly defined M- and I-bands and thin, fuzzy Z-discs (Figures 2C and 2D). Interestingly, mitochondria appeared to be more abundant in RBM20−/− rats; however, their structure was less well defined, with a high number of the present mitochondria showing altered morphology, appearing bloated with poorly defined cristae and degraded membranes. A similar phenotype was observed in atria from RBM20−/− rats (Figures 3C and 3D).

**IMPARED MITOCHONDRIAL FUNCTION IN RBM20-DEFICIENT RATS.** Based on the transmission electron microscopy imaging, we followed up with analyses on mitochondrial function. Respiration and absolute formation of ROS (measured as hydrogen peroxide) were significantly affected in the RBM20-deficient rats compared with WT rats (group effect, respiration: $P = 0.008$; ROS: $P = 0.01$). The RBM20−/− atria showed significantly lower respiration in complexes I and II than the WT ($P = 0.024$) and the RBM20−/− atria ($P < 0.001$). Respiration with the maximal OXPHOS-linked substrates was also lower in the RBM20−/− atria compared with RBM20+/− atria ($P = 0.042$), although the difference was not statistically significant compared with WT atria ($P = 0.075$) (Figure 4A). We observed no difference in the absolute ROS level in RBM20−/− and RBM20−/− rats compared to WT (complexes I and II: 0.058 vs RBM20−/− 0.16 pmol·s⁻¹·mg w/v⁻¹ and vs RBM20−/−.

| Table 1 Altered Isoform Expression in Splicing Targets (RBM20−/− Atria) |
|--------------------------|------------------|-------------|------------------|------------------|

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Isoform</th>
<th>Transcript ID</th>
<th>Protein Size (aa)</th>
<th>Log2 Fold Change</th>
<th>P Value</th>
<th>Adjusted P Value (FDR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDB3</td>
<td>LIM domain binding 3</td>
<td>Ldb3-206</td>
<td>ENSRNOT000000081631</td>
<td>330</td>
<td>-0.0501</td>
<td>2.54 × 10⁻⁶</td>
<td>0.0003</td>
</tr>
<tr>
<td>Ldb3</td>
<td>LIM domain binding 3</td>
<td>Ldb3-203</td>
<td>ENSRNOT000000083623</td>
<td>729</td>
<td>4.0394</td>
<td>0.008</td>
<td>0.12</td>
</tr>
<tr>
<td>Ldb3</td>
<td>LIM domain binding 3</td>
<td>Ldb3-205</td>
<td>ENSRNOT000000086627</td>
<td>707</td>
<td>-1.016</td>
<td>0.12</td>
<td>0.47</td>
</tr>
<tr>
<td>Ldb3</td>
<td>LIM domain binding 3</td>
<td>Ldb3-202</td>
<td>ENSRNOT000000087521</td>
<td>771</td>
<td>0.7909</td>
<td>0.82</td>
<td>0.97</td>
</tr>
<tr>
<td>Ldb3</td>
<td>LIM domain binding 3</td>
<td>Ldb3-201</td>
<td>ENSRNOT000000085414</td>
<td>679</td>
<td>0.3489</td>
<td>0.88</td>
<td>0.98</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-208</td>
<td>ENSRNOT00000024493</td>
<td>287</td>
<td>1.0037</td>
<td>1.02 × 10⁻⁶</td>
<td>1.76 × 10⁻⁴</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-201</td>
<td>ENSRNOT00000057641</td>
<td>326</td>
<td>-2.6584</td>
<td>6.69 × 10⁻⁴</td>
<td>0.023</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-202</td>
<td>ENSRNOT00000024575</td>
<td>284</td>
<td>0.8738</td>
<td>0.010</td>
<td>0.13</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-207</td>
<td>ENSRNOT00000048044</td>
<td>284</td>
<td>-0.3951</td>
<td>0.011</td>
<td>0.14</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-209</td>
<td>ENSRNOT00000090288</td>
<td>300</td>
<td>1.5946</td>
<td>0.067</td>
<td>0.37</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-210</td>
<td>ENSRNOT00000099012</td>
<td>267</td>
<td>2.2200</td>
<td>0.070</td>
<td>0.38</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-204</td>
<td>ENSRNOT00000040808</td>
<td>281</td>
<td>-1.3919</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-206</td>
<td>ENSRNOT00000112475</td>
<td>257</td>
<td>0.7770</td>
<td>0.10</td>
<td>0.46</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-203</td>
<td>ENSRNOT00000085894</td>
<td>294</td>
<td>0.2243</td>
<td>0.13</td>
<td>0.49</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-205</td>
<td>ENSRNOT00000024617</td>
<td>248</td>
<td>-1.4616</td>
<td>0.13</td>
<td>0.49</td>
</tr>
</tbody>
</table>

aa = amino acids; FDR = false discovery rate.
TABLE 2  Altered Isoform Expression in Splicing Targets (RBM20+/- Atria)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Isoform</th>
<th>Transcript ID</th>
<th>Protein Size (aa)</th>
<th>Log2 Fold Change</th>
<th>P Value</th>
<th>Adjusted P Value (FDR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDB3</td>
<td>Lim domain binding 3</td>
<td>Ldb3-205</td>
<td>ENSRNOT00000086627</td>
<td>707</td>
<td>1.2314</td>
<td>0.002</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ldb3-206</td>
<td>ENSRNOT00000081631</td>
<td>330</td>
<td>-0.1636</td>
<td>0.007</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ldb3-201</td>
<td>ENSRNOT0000085414</td>
<td>679</td>
<td>-1.2065</td>
<td>0.15</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ldb3-203</td>
<td>ENSRNOT0000083623</td>
<td>729</td>
<td>3.8914</td>
<td>0.31</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ldb3-202</td>
<td>ENSRNOT0000087521</td>
<td>771</td>
<td>-0.4304</td>
<td>0.44</td>
<td>0.76</td>
</tr>
<tr>
<td>KCNIP2</td>
<td>Potassium voltage-gated channel interacting protein 2</td>
<td>Kcnip2-201</td>
<td>ENSRNOT0000086164</td>
<td>225</td>
<td>-2.0340</td>
<td>9.55 x 10^-6</td>
<td>4.65 x 10^-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kcnip2-202</td>
<td>ENSRNOT000024709</td>
<td>270</td>
<td>1.5962</td>
<td>1.93 x 10^-5</td>
<td>8.32 x 10^-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kcnip2-203</td>
<td>ENSRNOT000024750</td>
<td>220</td>
<td>2.1061</td>
<td>0.0174</td>
<td>0.13</td>
</tr>
<tr>
<td>MYOM1</td>
<td>Myomesin 1</td>
<td>Myom1-202</td>
<td>ENSRNOT00000112958</td>
<td>1,592</td>
<td>-1.4593</td>
<td>0.002</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myom1-201</td>
<td>ENSRNOT0000085362</td>
<td>1,676</td>
<td>1.6058</td>
<td>0.007</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myom1-203</td>
<td>ENSRNOT00000112958</td>
<td>1,689</td>
<td>1.6058</td>
<td>0.007</td>
<td>0.07</td>
</tr>
<tr>
<td>NEXN</td>
<td>Nexilin</td>
<td>Nexn-201</td>
<td>ENSRNOT00000111795</td>
<td>669</td>
<td>-1.4136</td>
<td>8.23 x 10^-9</td>
<td>8.22 x 10^-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nexn-202</td>
<td>ENSRNOT00000110679</td>
<td>592</td>
<td>2.2290</td>
<td>4.59 x 10^-4</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nexn-203</td>
<td>ENSRNOT0000100634</td>
<td>535</td>
<td>0.2620</td>
<td>8.29 x 10^-4</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nexn-207</td>
<td>ENSRNOT0000099274</td>
<td>656</td>
<td>2.6718</td>
<td>0.001</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nexn-206</td>
<td>ENSRNOT0000066780</td>
<td>670</td>
<td>0.7464</td>
<td>0.026</td>
<td>0.17</td>
</tr>
<tr>
<td>TPM1</td>
<td>Tropomyosin 1</td>
<td>Tpm1-208</td>
<td>ENSRNOT000024493</td>
<td>287</td>
<td>1.0537</td>
<td>4.06 x 10^-5</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-202</td>
<td>ENSRNOT000024575</td>
<td>284</td>
<td>0.9882</td>
<td>0.021</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-205</td>
<td>ENSRNOT000024617</td>
<td>248</td>
<td>-1.9824</td>
<td>0.026</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-204</td>
<td>ENSRNOT000040808</td>
<td>281</td>
<td>-1.8077</td>
<td>0.030</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-206</td>
<td>ENSRNOT0000112475</td>
<td>257</td>
<td>1.0082</td>
<td>0.045</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-201</td>
<td>ENSRNOT0000057641</td>
<td>326</td>
<td>-0.8091</td>
<td>0.080</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-209</td>
<td>ENSRNOT0000075894</td>
<td>294</td>
<td>0.04213</td>
<td>0.67</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-207</td>
<td>ENSRNOT000048044</td>
<td>284</td>
<td>-0.1326</td>
<td>0.21</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tpm1-203</td>
<td>ENSRNOT0000085894</td>
<td>294</td>
<td>0.04213</td>
<td>0.67</td>
<td>0.88</td>
</tr>
<tr>
<td>TTN</td>
<td>Titin</td>
<td>Ttn-202</td>
<td>ENSRNOT0000101577</td>
<td>26,931</td>
<td>-0.9696</td>
<td>0.001</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ttn-201</td>
<td>ENSRNOT0000114553</td>
<td>35,375</td>
<td>1.1073</td>
<td>0.002</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ttn-203</td>
<td>ENSRNOT0000108121</td>
<td>5475</td>
<td>-0.0054</td>
<td>0.20</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Significant P values adjusted by an FDR of <5% are marked in bold.

Abbreviations as in Table 1.

0.064 pmol s^-1 mg w/v^-1; P = 0.058 and P = 0.90, respectively. OXPHOS: RBM20+/- 0.093 and RBM20+/- 0.056 vs 0.086 pmol s^-1 mg w.v^-1; P = 0.90 and P = 0.70, respectively (Figure 4B). However, when the formation was expressed relatively as ROS formation per oxygen used by the mitochondria, ROS/O2, RBM20+/- rats were found to have a significantly higher ROS/O2 ratio than both the WT and RBM20+/- rats with complexes I and II-linked substrates (0.0046 vs WT 0.0011 pmol s^-1 mg w.w^-1 and vs RBM20+/- 0.0016 pmol s^-1 mg w.w^-1; P = 0.001 and P = 0.011, respectively) (Figure 4C). This difference in ROS/O2 was no longer observed at maximal OXPHOS (RBM20+/- 0.0015 vs WT 0.00098 and RBM20+/- 0.00077 pmol s^-1 mg w.w^-1; both P = 0.90).

**REDUCED SPlicing ACTIVITY IN HUMAN RBM20 VARIANTS.** To examine the effects on TTN splicing in human RBM20 variants, we constructed a series of human RBM20 single nucleotide base exchange mutants (point mutants) based on the variants identified in our early-onset AF cohort. Here, we included all rare variants (minor allele frequency: <0.1%) predicted to be deleterious in in silico analyses (combined annotation-dependent depletion score: >20). (Figures 5A and 5B).

Quantitative reverse transcription polymerase chain reaction (analysis showed that all 3 LOF variants (RBM20-E725Vs*34, RBM20-E787*, and RBM20-Q345*) led to significantly reduced splicing activity (Figure 5C). The missense variants did not show a significant impact on protein expression, except for the RBM20-P411L variant, which showed a modest reduction of RBM20 expression. The 4 variants that displayed reduced splicing activity also resulted in significantly reduced protein expression levels. The RBM20-Q345* variant was not detected because of a lack of expression of the C-terminal part of the RBM20 protein, which is required for the anti-RBM20 antibody to bind.
FIGURE 2  Electron Microscopy Imaging and Differential Gene Expression in RBM20+/– Rat Atria

Continued on the next page
DISCUSSION

In this study, we used targeted next-generation sequencing and whole-exome data to uncover a novel association between LOF in RBM20 and AF in 2 independent cohorts. In a cohort of highly selected early-onset AF patients, the prevalence of LOF variants in RBM20 was significantly higher in individuals with AF compared with population-based control individuals. In the UK Biobank, an independent population-based cohort consisting of approximately 400,000 exomes, AF was similarly associated with a higher burden of rare LOF variants in RBM20.

Studies have shown RBM20 to be an important regulator of splicing of numerous cardiac genes, including key sarcomere genes (eg, TTN, TPM1, LDB3) and genes handling intracellular Ca\(^{2+}\) dynamics (eg, CACNA1C, CAMK2D, RYR2).

We hypothesized that altered splicing resulting from a loss of RBM20 function would affect atrial structure and function, resulting in an atrial cardiomyopathy phenotype. To examine this systematically, we conducted electron microscopy imaging and long-read cDNA sequencing of atrial tissue from a rat model with loss of RBM20. Electron microscopy imaging revealed substantial changes in tissue integrity, including sarcomere structure and altered mitochondrial morphology, while long-read cDNA sequencing revealed differential transcript use in several putative splice targets of RBM20.

In RBM20\(^{+/−}\) atria, we observed altered splicing of TPM1 and LDB3. TPM1 encodes the protein tropomyosin 1, which is important in thin filament function and actin-myosin interaction in sarcomeres. Pathogenic variants in the gene have previously been associated with cardiomyopathy and reduced sarcomere content. Furthermore, alternative splicing of TPM1 isoforms has been observed in patients with heart failure and cardiomyopathy. LDB3, which encodes the LIM domain binding 3 protein, ensures the structural integrity and stability of the sarcomere Z-disks, as evidenced by the disruption of Z-disks seen in mice with LDB3 mutations.

Specifically, RBM20 regulates the inclusion of the cardiac isoform-specific exon 4, whereas loss of RBM20 function leads to the inclusion of exons 5 and 6 instead, normally only expressed in LDB3 isoforms of skeletal muscle. In line with these results, our data showed a down-regulation of the short cardiac isoform of LDB3 in RBM20\(^{+/−}\) rat atria (ENSRNOT00000081631).

We found more profound effects on splicing in RBM20\(^{+/−}\) atria, with altered splicing of the sarcomere genes TPM1, LDB3, NEXN, TTN, MYOM1, and KCNIP2, which encodes an ion-channel interacting protein. Numerous isoforms were differentially expressed in NEXN, which encodes the crucial Z-disc protein nexilin. Mutations in the NEXN gene have previously been associated with cardiomyopathy. Myomesin 1, the protein product of MYOM1, is part of the M-band and is important for sarcomere integrity. Alternative splicing of MYOM1 is known to affect the stability of sarcomere M-bands, and an up-regulation of a fetal isoform of myomesin 1 has been linked with dilated cardiomyopathy (DCM).

The key sarcomere protein titin, encoded by TTN, plays an essential role in the sarcomere contractile function. Changes in titin isoform composition modulate cardiac elasticity and stiffness, and protein-truncating variants in titin have been associated with early-onset AF. In RBM20\(^{+/−}\) atria, we observed a down-regulation of the short isoform of titin (ENSRNOT00000101577) and up-regulation of the long isoform (ENSRNOT0000014553). A similar switch from short to long titin isoforms has previously been reported in human DCM hearts.

At the gene expression level, we also found differential expression of several genes involved in sarcomere function, such as MYL4, TNNT2, TCAP, and SYNPO2L. We speculate that the changes in gene expression of these genes may be part of a compensatory mechanism. Collectively, these changes in gene expression and isoform composition may explain the sarcomere degradation found in electron microscopy imaging and long-read cDNA sequencing of atrial tissue from a rat model with loss of RBM20.

**FIGURE 2** Continued

(A, B) TEM images of left atria of a 3-month-old WT rat. Sarcomeres are well organized, with clearly defined Z-disks and I-bands. (C, D) TEM images from left atria of a 3-month-old RBM20\(^{+/−}\) rat. In contrast to WT, Z-disks appear thin and fuzzy, and I-bands and M-bands are blurry and poorly defined. Mitochondria are abundant and show compromised structure. (E) A volcano plot of differentially expressed genes in RBM20\(^{+/−}\) rat atria compared with WT rat atria. Genes shown in red are significantly up-regulated, and genes shown in blue are significantly down-regulated. (F) The 10 most up- and down-regulated gene sets by biological process in RBM20\(^{+/−}\) rat atria compared with WT rat atria. The x-axis denotes normalized enrichment score, and the y-axis lists gene ontology pathways. Data points are colored by adjusted P value. TEM = transmission electron microscope; WT = wild type.
FIGURE 3  Electron Microscopy Imaging and Differential Gene Expression in RBM20−/− Rat Atria

Continued on the next page
microscopy imaging, where we observed unorganized thick and thin filaments, fuzzy Z-discs, and blurred M-lines.

The mitochondrial disarray and differential expression of numerous genes involved in cellular respiration and OXPHOS prompted us to perform analyses of mitochondrial respiration and ROS formation. Interestingly, we identified an increased formation of mitochondrial ROS/O2 in RBM20−/− rats, probably because of altered proton or electron leak and/or impaired antioxidant defense. In RBM20−/− rats, we observed similar evidence of compromised mitochondrial function, indicated by significantly lowered mitochondrial respiration. In a recent study on RNA sequencing data from cardiac tissue in humans with DCM, titin-truncating variants were reported to alter mitochondrial energetics. We propose that altered TTN isoform composition could also influence mitochondrial function in loss of RBM20.

The mitochondrial genes found to be differentially expressed in both RBM20−/− and RBM20+/− atria include the gene that codes for mitochondrial complex I (MT-ND5). Complex I, the largest complex in the respiratory chain, is the first enzyme in the respiratory chain and plays a major role in mitochondrial metabolism. Mutations in MT-ND5 have been shown to lead to lower complex I activity, consequently resulting in elevated ROS. A similar elevated ROS formation has been seen in a cancer model with a mutation in the MT-ND3 gene, which was also found to be differentially expressed in the RBM20+/− atria. Collectively, these factors could account for the significantly elevated ROS/O2 ratio found in the RBM20+/− atria and not in the RBM20−/− atria.

The additional differentially expressed mitochondrial gene found in RBM20−/− atria, COX7A1, could contribute to the observed worsened mitochondrial function compared to RBM20+/− atria. The COX7A1 gene codes for the cytochrome c oxidase (complex IV), the terminal component of the mitochondrial respiratory chain, and is generally accepted as the major regulator for OXPHOS. Thus, our observed differential expression of mitochondrial genes would have a definite impact on the electron transfer processes and the electron transport chain integrity, resulting in proton and/or electron leak, and may explain the subsequent overwhelming of the mitochondrial antioxidant defense system. Recent studies have implicated mitochondrial dysfunction and oxidative stress in the pathogenesis of heart disease and AF and found that increased oxidative stress may affect cardiomyocyte function through ROS-mediated modifications of sarcomere proteins, compromising cardiac contractility.

One of the most significant changes in gene expression we observed was an up-regulation of RYR2 expression in both RBM20−/− and RBM20+/− rats. The RYR2 gene encodes the ryanodine receptor 2 channels (RYR2 channels). We speculated that this may be a compensatory mechanism in response to altered Ca2+ dynamics in dysfunctional sarcomeres. However, this could also be caused by increased oxidation of RYR2 channels by ROS. Previous studies have reported that RYR2 channels are susceptible to oxidation, inducing a conformational change that leads to a higher open probability of the channel and, thus, increased Ca2+ leak. Such Ca2+ leak and altered intracellular Ca2+ dynamics have been implicated in the pathogenesis of early afterdepolarizations, an important mechanism in the initiation of AF and other arrhythmias. Our results indicated differential expression of several genes involved in intracellular calcium handling (Figures 6C and 6D). Calcium sparks and calcium waves initiated by disturbed Ca2+ handling may facilitate early afterdepolarizations, mainly through an increase in the sodium-calcium exchanger current (I_{NCX}), which acts as a depolarizing current that can facilitate arrhythmia (Figure 6E). This could be compounded by the alternative splicing observed in KCNIP2, because the protein encoded by the gene has been reported to play an important role in cardiac electrical stability and arrhythmogenesis.

Finally, we explored the effects of splicing in all potentially deleterious RBM20 variants detected in AF patients. Using point mutants in cell models, we observed that all LOF variants significantly reduced splicing activity. Most missense variants, except...
RBM20-P411L did not seem to significantly affect splicing. Based on our findings, we hypothesize that rare LOF variants in RBM20 may lead to changes in the splicing of key sarcomere genes, resulting in a dysfunctional contractile function of cardiomyocytes. This may, in turn, increase stress on the metabolic systems of the cardiomyocyte such as the mitochondria, as indicated by the increased formation of ROS/O2. Finally, altered intracellular Ca^{2+} dynamics, facilitated by altered calcium handling and leak through oxidized and up-regulated RYR2 channels, may induce a proarrhythmic condition, which, in turn, may lead to a vicious, self-reinforcing cycle of AF and atrial dysfunction. Our findings are in line with previous studies that report a higher incidence of atrial arrhythmias in human patients and mouse models with DCM caused by missense variants in RBM20.62,63

**STUDY LIMITATIONS.** Whereas our study provided novel insights into plausible mechanisms of arrhythmia and atrial cardiomyopathy in individuals...
FIGURE 5 Rare RBM20 Variants Identified in an Early-Onset AF Cohort Affect Splicing Activity

(A) Schematic representation of (human) RBM20 and the single nucleotide base exchange mutants (point mutants). The 2 zinc finger (ZnF) domains are depicted in blue, the RNA recognition motif (RRM) is in green, and the serine-rich domain-like domain is in orange. Unstructured regions are indicated as small gray boxes. RBM20 amino acid residues are indicated underneath the horizontal line in the scheme, and the vertical lines show changes in amino acid residue and their approximate position in the protein sequence.

(B) Table of rare variants in RBM20 identified in our early-onset AF cohort. (C) Real-time reverse transcription polymerase chain reaction (SYBR Green, Thermo Fisher Scientific) analysis of HEK293 cells transfected with the TTN241-3 splicing reporter and the mutant RBM20 expression constructs; n = 3; ***P < 0.001. (D) WesTM (ProteinSimple) analysis of RBM20 mutants, transfected into HEK293 cells, using anti-RBM20 antibody and anti-β-tubulin as the loading control. Group comparisons were analyzed by 1-way analysis of variance. AF = atrial fibrillation; HEK293 = human embryonic kidney 293 cells; NA = not available; refSNV (formerly SNP) = reference single-nucleotide variation.
with rare variants in RBM20, our approach also had some limitations.

First, our results may have limited generalizability across ancestry groups because of the inclusion of individuals of European ancestry only. Second, we do not have phenotypic information on the referent populations from gnomAD and SweGen. Thus, some of the control individuals may have AF, although this would lead to bias toward no association. The in-house cohort of early-onset AF patients, although highly selected and well defined, was also of medium size and not representative of the general population. Moreover, the number of individuals identified in the in-house cohort with RBM20 LOF mutations is small. Replicating our findings in the UK Biobank may have broadened the implications of our findings. Furthermore, because RBM20 is a known splicing regulator, we performed long-read sequencing of cDNA to enable the study of the entire transcriptome, which also affected read depth compared with short-read sequencing. In turn, this could negatively influence our power to detect differentially expressed genes and transcripts, and our results are therefore not exhaustive. Finally, imaging, cDNA sequencing, pathway analyses, and mitochondrial analyses were conducted on a rat model, in which we had only a very limited number of animals available, which may also have negatively affected power.

CONCLUSIONS

To our knowledge, our results provide the first comprehensive evidence of a link between rare LOF variants in the cardiac splicing regulator RBM20 and AF. Our experimental data suggest that loss of RBM20 function in the atria may lead to sarcomere and mitochondrial dysfunction as mechanisms underlying the development of atrial arrhythmia in individuals with such variants. Our findings contribute to a more complete understanding of the complex genetics of AF and offer new insights into the role of alternative splicing in cardiac arrhythmia.

ACKNOWLEDGMENTS This research has been conducted using the UK Biobank resource under application number 43247. The authors thank all individuals for their participation in this study. The Visual Abstract and Figure 6E were created using BioRender.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This work was supported by the Novo Nordisk Foundation, John and Birthe Meyer Foundation, Research Foundation of the Heart Centre Rigshospitalet, Villadsen Family Foundation, Arvid Nilsson Foundation, Danish Council for Independent Research, Research Foundation at Rigshospitalet, Skibsrer Per Henriksens R., og hustrus fond, and Department of Clinical Medicine (University of Copenhagen). Dr Ahlberg has received a grant from the Novo Nordisk Foundation, BRIDGE-Translational Excellence Programme. Dr Gotthardt was supported by the Deutsche Forschungsgemeinschaft and the Leducq Foundation. Dr Svendsen has received a fee for participating in the Advisory Board (Medtronic); and speaker fees from Medtronic not related to this work. Dr Gotthardt has a consultancy agreement with River BioMedics; and has received speaker honoraria from Bayer not related to this work. Dr Olesen has received speaker fees from Biosense Webster not related to this work. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Morten Salling Olesen, Department of Biomedical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen N, Denmark. E-mail: mortensol@sund.ku.dk.
PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: AF is a complex phenotype with multifaceted pathophysiological mechanisms. Our results link the RBM20 gene with AF and highlight that atrial cardiomyopathy may be an underlying driver of arrhythmia in some individuals presenting with AF. These findings expand our knowledge of the genetic component of AF and may contribute to possible future genetic risk stratification in individuals with early onset of AF.

TRANSLATIONAL OUTLOOK: Alternative splicing has been reported to play an important role in the form and function of the ventricles. Our study illustrates how changes in isomform expression may also act as a proarrhythmic mechanism in the atria. Future studies into RBM20 and other splicing regulators may identify novel therapeutic targets in atrial arrhythmia and remodeling.

REFERENCES

38. Collins MM, Ahlberg G, Hansen CV, et al. Early sarcormere and metabolic deficits in a zebrafish...

KEY WORDS alternative splicing, atrial cardiomyopathy, atrial fibrillation, genetics, RBM20

APPENDIX For supplemental material and tables, please see the online version of this paper.