RESEARCH ARTICLE

Investigating gene-microRNA networks in atrial fibrillation patients with mitral valve regurgitation

Joana Larupa Santos1, Ismael Rodríguez1, Morten S. Olesen1,2, Bo Hjorth Bentzen1*, Nicole Schmitt1*,

1 Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen N, Denmark, 2 Department of Cardiology, Laboratory for Molecular Cardiology, The Heart Centre, Rigshospitalet, University Hospital of Copenhagen, Copenhagen Ø, Denmark

☯ These authors contributed equally to this work.
* bobe@sund.ku.dk (BHB); nschmitt@sund.ku.dk (NS)

Abstract

Background
Atrial fibrillation (AF) is predicted to affect around 17.9 million individuals in Europe by 2060. The disease is associated with severe electrical and structural remodelling of the heart, and increased the risk of stroke and heart failure. In order to improve treatment and find new drug targets, the field needs to better comprehend the exact molecular mechanisms in these remodelling processes.

Objectives
This study aims to identify gene and miRNA networks involved in the remodelling of AF hearts in AF patients with mitral valve regurgitation (MVR).

Methods
Total RNA was extracted from right atrial biopsies from patients undergoing surgery for mitral valve replacement or repair with AF and without history of AF to test for differentially expressed genes and miRNAs using RNA-sequencing and miRNA microarray. In silico predictions were used to construct a mRNA-miRNA network including differentially expressed mRNAs and miRNAs. Gene and chromosome enrichment analysis were used to identify molecular pathways and high-density AF loci.

Results
We found 644 genes and 43 miRNAs differentially expressed in AF patients compared to controls. From these lists, we identified 905 pairs of putative miRNA-mRNA interactions, including 37 miRNAs and 295 genes. Of particular note, AF-associated miR-130b-3p, miR-338-5p and miR-208a-3p were differentially expressed in our AF tissue samples. These miRNAs are predicted regulators of several differentially expressed genes associated with...
cardiac conduction and fibrosis. We identified two high-density AF loci in chromosomes 14q11.2 and 6p21.3.

Conclusions
AF in MVR patients is associated with down-regulation of ion channel genes and up-regulation of extracellular matrix genes. Other AF related genes are dysregulated and several are predicted to be targeted by miRNAs. Our novel miRNA-mRNA regulatory network provides new insights into the mechanisms of AF.

Introduction
Atrial Fibrillation (AF) is the most common type of cardiac arrhythmia. The disease represents a major economic burden and prevalence is predicted to increase over the next decades both in Europe and the United States [1]. Genetic variants are well documented in cases of early-onset and ‘lone’ AF [2,3]. Additionally, more than one hundred susceptibility loci were identified by population-based, genome-wide association studies (GWAS) [4,5].

The vast majority of AF patients are over 75 years old indicating that AF is an age-related disease. The disease often starts by short self-terminating episodes originating from a trigger in a vulnerable substrate that leads to rapid focal ectopic firing and re-entry of electrical signals in the atria. In some cases, AF can gradually develop to a permanent condition through a process called remodelling [6]. Electrical remodelling results from ion channel dysregulation while structural remodelling is characterized by increased fibrosis, atrial dilation and conduction abnormalities [7,8]. The available therapeutic options in clinical practice have limited efficiency, especially in those patients with more progressed AF, likely due to higher severity of tissue remodelling [9].

Valvular heart disease (VHD) is an established risk factor for AF [10]. However, not all patients with valve disorders develop AF, suggesting that some VHD patients are more predisposed to develop AF. Differences in the remodelling of affected hearts might be a secondary factor predisposing to AF in these patients.

MicroRNAs (miRNAs) are a class of highly conserved short non-protein-coding RNAs. These 18–24 nucleotides RNA molecules are known to negatively regulate the expression of complementary target genes by binding to the 3'-untranslated region (UTR) of messenger RNAs (mRNAs) in the cytosol, promoting either mRNA degradation or inhibiting translation [11]. MiRNAs can be classified as intergenic, intronic and exonic, and the expression of intronic and exonic miRNAs is mainly controlled by the promoter of the host gene [12].

Original reports on the role of miR-1 in cardiogenesis have triggered a large body of research dealing with the role of miRNAs in cardiac development and pathological processes [13]. Following this report, abnormal expression of many other miRNAs has been described in various types of cardiac disorders [14]. In the context of AF, different studies have shown a role of miRNAs in the deregulation of ion channels based on gene expression analyses of tissue samples from AF patients and additional functional studies [15–17]. Cardiac fibrosis was also shown to be influenced by a number of miRNAs, such as miR-133a [18,19]. However, much is yet to be learned about the complete role of miRNAs in arrhythmogenesis. Hence, the analysis of miRNAs expression can give important insight into regulatory mechanisms involved in disease, especially when combined with gene expression profiling studies [20,21].
The present study identifies several miRNAs and genes that may contribute to the development of AF-associated with VHD and illustrates a miRNA-mRNA network with hundreds of putative regulatory interactions from studying tissue samples from AF patients and controls with mitral valve regurgitation (MVR).

Materials and methods

Patients and tissue samples

Right atrial (RA) posterior wall tissue biopsies were obtained from unrelated Caucasian patients undergoing open-heart surgery for mitral valve replacement or repair at the University Hospital of Copenhagen (Rigshospitalet), as previously described [22]. From these, six patients with AF sustained for at least two months were selected. The control group included six patients in sinus rhythm (SR) and with no previous history of AF who were selected according to age and gender match. Cardiac disease, hypertension and type 2 diabetes mellitus were used as exclusion criteria.

The study was approved by the Ethics Committee of the Capital Region of Copenhagen (protocol reference no. 16238), in accordance with the Declaration of Helsinki. Written informed consent for the use of clinical information and biological samples was provided by all participants. Clinical data from all subjects was obtained through questionnaires and clinical records from the Danish health care system.

RNA preparation

Total RNA including small RNAs was extracted from RA biopsies of six AF patients and six control patients using the miRNeasy kit (QIAGEN, Hilden, Germany) according to manufacturer's instructions. RNA concentration was measured in a NanoDrop 2000 (ThermoScientific, Wilmington, USA) and quality was assessed in a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA samples with a RIN > 6 were used in further experiments. For details see Supplementary methods in S1 File.

RNA-sequencing

Libraries were prepared from ribosomal RNA depleted RNA samples using the TruSeq Stranded Total RNA Library Prep Kit (Illumina, San Diego, California, USA). Libraries were sequenced on an Illumina HiSeq 2500. The control sample HS24 was not used due to low amount of input RNA. The raw sequencing data is available in the Electronic Research Data Archive (University of Copenhagen), see Data Availability section.

Transcriptome sequencing analysis

Raw paired-end reads were aligned to the human GRCh38 reference transcriptome using the Kallisto pseudoaligner with default options [23]. Kallisto quantifies the abundance of reads on transcript level. The resulting quantification files were imported to R for downstream analysis. Transcript counts were collapsed to gene level. Differentially expressed genes (DEGs) in AF compared to controls were obtained using the Bioconductor package DEseq2 [24]. Genes with false discovery rate < 0.05 and log2fold-change (FC) above 1 or below -1 were considered DEGs. Principal component analysis (PCA) and unsupervised hierarchical clustering of sample-to-sample distance matrices were used to analyse sample clustering according to transcriptomic similarities.

DEGs were compared with genes of interest including 1) AF-associated genes from genomic, transcriptomic and proteomic studies from a broad literature search and 2) genes
involved in mechanisms of interest, such as fibrosis, cardiac contraction and ion channel currents, from GO terms and HUGO Gene Nomenclature Committee.

**MiRNAs microarray**

Microarrays were performed with the exact same RNA samples used for the RNA-seq experiments. Mature miRNA transcripts were hybridized to the GeneChipS1 miRNA 4.0 Array (Affymetrix, Santa Clara, USA). Raw CEL data files are available at the Electronic Research Data Archive (University of Copenhagen), see Data Availability section.

Data generated by the Affymetrix’s miRNA array was normalized using the robust multi-array average method [25]. Quality metrics showed one control sample (HS24) as an outlier (S1 Fig in S1 File). FC values and p-values \((p)\) of expression changes were calculated using the Limma package in R/Bioconductor project [26]. A cut-off \(p < 0.01\) was used for selection of differentially expressed miRNAs. No FC cut-off was applied. Samples and differentially expressed miRNAs were subjected to unsupervised hierarchical clustering and plotted as a heat map.

**Validation of RNA-seq and microarray by quantitative PCR**

Qualitative polymerase chain reaction (qPCR) was performed to confirm the reliability of RNA-seq and microarray data. cDNA was synthesized using the Precision nanoScript2 Reverse Transcription kit (PrimerDesign, Southampton, United Kingdom) for mRNAs and the Universal cDNA Synthesis kit II (Exiqon, Wobrun, Massachusetts, USA) for miRNAs. The expression of selected genes was measured using Taqman double dye probes and Precision-PLUS MasterMix with ROX (PrimerDesign, Southampton, United Kingdom). The miRNAs qPCR reactions were performed using commercial miRCURY LNA™ Universal RT microRNA PCR primers and the ExiLENT SYBR® Green master mix (Exiqon, Wobrun, Massachusetts, USA). All experiments were performed in a light cycler CFX Connect Real-Time System (BIO-RAD, Hertfordshire, UK). The following genes—\(\text{KCNA4}, \text{ACTN2}, \text{KCNK3}, \text{NANOG}, \text{TNNT2}, \text{KCNQ5}, \text{KCNJ5}\) and \(\text{KCNB1}\), and miRNAs—miR-143-5p, miR-192-5p, miR-187-3p, miR-208b-3p, miR-338-5p, miR-335-5p, miR-432-5p, miR-490-5p, miR-499a-5p, miR-503-5p and miR-92b-3p, were investigated. \(\text{YWHAZ}\) and \(\text{RPL13A}\) were used as references to normalize the results, while miRNA-16-5p and miR-103a-3p were used to normalize the miRNAs results. Relative expression and FC values were calculated using the \(2^{-\Delta\Delta C_t}\) method.

**Statistical analysis**

Statistical analysis was performed on GraphPad Prism 7 (GraphPad Software Inc., San Diego, USA). Mann-Whitney U-test and was applied for comparison of continuous variables. Patient data is presented as mean ± standard deviation (SD) and percentage. \(P < 0.05\) was considered statistically significant.

**miRNA target genes prediction**

Candidate target genes of the differentially expressed miRNAs were obtained using the miR-Walk3 target prediction database [27,28]. The programme output includes predictions from miRWalk’s machine learning approach (TarPmiR) and experimentally validated interaction from miRTarBase. The 3′-UTR of genes was selected as putative binding site and only predictions with a binding probability above 0.95 were accepted. Afterwards, miRNA–gene interactions in which the target gene was included in our DEGs list were kept and all the others were filtered out. Normalized expression levels from microarray and RNA-seq were used to evaluate
the correlation between each miRNA-gene pair using Spearman’s correlation coefficient in order to find monotonic relations, rather than just linear, given the different magnitudes used for the RNA-seq and microarray experiments. Subsequently, interaction with a negative correlation below -0.5 were used to construct a miRNA-gene regulatory network in Cytoscape version 3.6.1 [29]. Filtering criteria was applied to the networks in order to help visualize and identify relevant interactions. Secondary miRNA-gene networks were built for up- and down-regulated miRNAs previously association with AF in expression studies [30–32] and respective predicted target genes.

**Functional enrichment analysis**

To better understand the functional role of differentially expressed miRNAs and target genes, a functional enrichment analysis of the targeted genes after filtering was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) [33]. The same analysis was performed with the full list of DEGs from RNA-seq experiments. Our enrichment analysis was focused on gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways using an enrichment cut-off of \( p < 0.05 \) after Benjamini-Hochberg correction.

**Positional enrichment of genetic elements**

Up- and down-regulated genes and miRNAs were plotted together with AF-associated single nucleotide polymorphisms (SNPs), according to chromosome position, to visualize possible clusters of these genetic elements. AF SNPs were retrieved from published data and the Human Short Variants dataset from Ensembl Variation 96 [34]. The location of the published SNPs had to be converted from Hg19 assembly to Hg38 using the UCSC liftOver tool. For the rest of the elements, genomic location and cytogenetic band information was retrieved from Biomart (Bioconductor). The nearest gene(s) to each SNP genome coordinate were obtained to plot together with the matching SNP. For 81 SNPs retrieved, no associated gene was found. A cluster was defined as two or more consecutive elements overlapping or not further than 20 kbp. Clusters including five or more co-localized elements of any type were considered as high-density regions. Subsequently, high-density regions comprising exclusively SNPs were excluded from our analysis. The total number of elements per chromosome was normalized to chromosome size and/or gene density. Lastly, the genomic location of DE miRNAs and genes was used to assess co-expression patterns between intronic or exonic miRNAs and host genes.

**Results**

**Patient characteristics**

RA posterior wall biopsies were obtained from MVR patients with (n = 6) and without (n = 6) AF undergoing surgery for valve replacement or repair. Clinical characteristics of all subjects are listed in Table 1. There were no significant differences between the groups apart from AF history. Left atrial (LA) size was increased in both groups. Ejection fraction (EF) was lower in AF subjects but the difference was not significant. EF values were unavailable in one subject from each group, while one LA size measurement was unavailable in the AF group.

**Gene expression signature of AF tissues**

We investigated differences in gene expression between AF and control patients using RNA-sequencing. PCA showed that males and females in the AF group were separated by PC1 (Fig 1A). No other characteristics reported in Table 1 contribute to the clustering of samples. S2 Fig
in S1 File represents clustering of samples according to sample-to-sample matrix distances.

From the 44,852 genes detected in our samples, 644 were differentially expressed in the AF atrial samples with an adj. \( p < 0.05 \) and log\(_2\)FC > 1 or < -1 (Fig 1B). Of these, 265 genes were up-regulated and 379 were down-regulated. A complete list of DEGs is provided in D1 in S1 Dataset.

Pathway and gene ontology enrichment analysis was performed separately in up- and down-regulated genes (Fig 1C and 1D). We found that down-regulated genes were enriched for “integral component of plasma membrane”, “regulation of ion transmembrane transport” and “extracellular region”, while up-regulated genes were greatly involved in “extracellular space”, “nucleosome” and “mitotic nuclear division”. A full list of GO terms and KEGG pathways enriched in up- and down-regulated genes is available in D2 in S1 Dataset.

We validated the RNA-seq results by testing the expression of eight genes in five AF and five control samples using qPCR quantification (S3 Fig in S1 File). The correlation between qPCR and RNA-seq results was high for all the genes tested.

We analysed the dataset to prioritize genes that may be involved in electrical remodelling, structural remodelling or genetic variants associated with AF in GWAS and familial studies. Comparative analysis of DEGs with AF-associated genes and genes involved in mechanisms of interest resulted in 57 up-regulated and 94 down-regulated genes, marked in D1 in S1 Dataset.

Table 1. Clinical characteristics of the study population.

<table>
<thead>
<tr>
<th></th>
<th>AF patients (N = 6)</th>
<th>Controls (N = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>3 males (50)</td>
<td>3 males (50)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>59.3 ± 9.2</td>
<td>59.3 ± 8.7</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>24.6 ± 1.6</td>
<td>22.5 ± 3.3</td>
</tr>
<tr>
<td>AF Type: Paroxysmal</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>2 (33.3)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>4 (66.7)</td>
<td>NA</td>
</tr>
<tr>
<td>AF duration (months)</td>
<td>2–14</td>
<td>NA</td>
</tr>
<tr>
<td>Smoking</td>
<td>3 (50)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>Alcohol consumption (units/week)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \leq 7 )</td>
<td>4 (66.7)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>8–14</td>
<td>1 (16.7)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>15–21</td>
<td>1 (16.7)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>&gt;21</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Type II diabetes mellitus</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Heart failure symptoms (NYHA):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1 (16.7)</td>
<td>1 (16.7)</td>
</tr>
<tr>
<td>II</td>
<td>3 (50)</td>
<td>4 (66.7)</td>
</tr>
<tr>
<td>III</td>
<td>2 (33.3)</td>
<td>1 (16.7)</td>
</tr>
<tr>
<td>IV</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Left atrial dilation(^1)(^2) [35]:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Moderate</td>
<td>1 (16.7)</td>
<td>1 (16.7)</td>
</tr>
<tr>
<td>Severe</td>
<td>4 (66.7)</td>
<td>5 (83.3)</td>
</tr>
<tr>
<td>Ejection Fraction(^1)</td>
<td>56 ± 6.5</td>
<td>67 ± 9.7</td>
</tr>
</tbody>
</table>

\(^1\) One or two unavailable values

\(^2\) Determined using biplane Simpson model; AF = atrial fibrillation; BMI = body mass index; NA–not applicable; NYHA = New York Heart Association classification; Values are shown as n (%) or mean ± SD.

https://doi.org/10.1371/journal.pone.0232719.t001
Specifically, the analysis showed nine genes affected by AF-associated genetic variants, 14 down-regulated and three up-regulated ion channels/subunits, and nine upregulated genes associated with cardiac fibrosis and extracellular matrix components. Genes are listed in Table 2, including reference to published data reporting similar changes in AF patients.

**miRNAs are involved in remodelling of the right atrium**

Microarray experiments allowed us to investigate miRNAs expression profiles in RA biopsies. Different factors were expected to contribute to sample variability due to the heterogeneous
profile of the study participants (AF duration, gender, age, alcohol consumption). We used PCA and sample-to-sample distance matrix to evaluate miRNAs expression variations between samples (S4 Fig in S1 File). We were able to identify a total of 43 miRNAs differentially expressed between AF and control subjects following a cut-off $p < 0.01$. Of those, 21 were down-regulated and 22 were up-regulated in AF patients, as shown in Fig 2. We validated the microarray experiments by qPCR quantification of eleven miRNAs in three AF and three

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene ID</th>
<th>log2FC</th>
<th>p-value</th>
<th>p-adjusted</th>
<th>Reference</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHLDA1</td>
<td>ENSG00000139289</td>
<td>2.94</td>
<td>5.79E-19</td>
<td>5.43E-16</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>MYH7</td>
<td>ENSG00000092054</td>
<td>1.97</td>
<td>6.01E-06</td>
<td>3.28E-04</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>NAV2</td>
<td>ENSG00000166833</td>
<td>1.18</td>
<td>2.68E-13</td>
<td>1.23E-10</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>RPL3L</td>
<td>ENSG00000140986</td>
<td>1.16</td>
<td>2.48E-15</td>
<td>1.70E-12</td>
<td>EV</td>
<td>-</td>
</tr>
<tr>
<td>REC114</td>
<td>ENSG00000183324</td>
<td>-3.01</td>
<td>5.71E-24</td>
<td>1.31E-20</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>HCN4</td>
<td>ENSG00000138622</td>
<td>-1.71</td>
<td>5.15E-11</td>
<td>1.50E-08</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>KCN5</td>
<td>ENSG00000120457</td>
<td>-1.10</td>
<td>1.85E-12</td>
<td>7.56E-10</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>KCNN2</td>
<td>ENSG00000080709</td>
<td>-1.06</td>
<td>1.86E-10</td>
<td>4.62E-08</td>
<td>[34]</td>
<td>-</td>
</tr>
<tr>
<td>MYH6</td>
<td>ENSG00000197616</td>
<td>-1.06</td>
<td>3.89E-05</td>
<td>0.0015</td>
<td>EV</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene ID</th>
<th>log2FC</th>
<th>p-value</th>
<th>p-adjusted</th>
<th>Reference</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>KCNQ4</td>
<td>ENSG00000117013</td>
<td>1.09</td>
<td>0.0021</td>
<td>0.0343</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNQ3</td>
<td>ENSG00000184156</td>
<td>1.36</td>
<td>1.58E-05</td>
<td>7.49E-04</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNA4</td>
<td>ENSG00000182255</td>
<td>1.63</td>
<td>2.49E-09</td>
<td>4.50E-07</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNH7</td>
<td>ENSG00000184611</td>
<td>-2.71</td>
<td>2.29E-07</td>
<td>2.16E-05</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNQ5</td>
<td>ENSG00000185760</td>
<td>-1.79</td>
<td>1.24E-05</td>
<td>6.13E-04</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HCN4</td>
<td>ENSG00000138622</td>
<td>-1.71</td>
<td>5.15E-11</td>
<td>1.50E-08</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CACNA2D2</td>
<td>ENSG0000007402</td>
<td>-1.34</td>
<td>3.74E-11</td>
<td>1.13E-08</td>
<td>[36]</td>
<td>RA</td>
</tr>
<tr>
<td>CACNA1G</td>
<td>ENSG0000006283</td>
<td>-1.33</td>
<td>1.71E-07</td>
<td>1.67E-05</td>
<td>[36]</td>
<td>RA</td>
</tr>
<tr>
<td>KCNK5</td>
<td>ENSG00000164626</td>
<td>-1.32</td>
<td>2.45E-04</td>
<td>0.0066</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNK17</td>
<td>ENSG00000124780</td>
<td>-1.30</td>
<td>5.34E-04</td>
<td>0.0120</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SCN2A</td>
<td>ENSG00000136531</td>
<td>-1.29</td>
<td>8.71E-04</td>
<td>0.0174</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNB1</td>
<td>ENSG00000158445</td>
<td>-1.20</td>
<td>3.96E-04</td>
<td>0.0096</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CACNA1D</td>
<td>ENSG00000157388</td>
<td>-1.16</td>
<td>3.29E-05</td>
<td>0.0013</td>
<td>[36]</td>
<td>RA</td>
</tr>
<tr>
<td>KCNJ5</td>
<td>ENSG00000120457</td>
<td>-1.10</td>
<td>1.85E-12</td>
<td>7.56E-10</td>
<td>[36,37]</td>
<td>RA, LAA</td>
</tr>
<tr>
<td>SCN1A</td>
<td>ENSG00000142885</td>
<td>-1.10</td>
<td>0.0029</td>
<td>0.0429</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KCNN2</td>
<td>ENSG00000080709</td>
<td>-1.06</td>
<td>1.86E-10</td>
<td>4.62E-08</td>
<td>[37]</td>
<td>LAA</td>
</tr>
<tr>
<td>SCN3A</td>
<td>ENSG00000153253</td>
<td>-1.05</td>
<td>0.0029</td>
<td>0.0423</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene ID</th>
<th>log2FC</th>
<th>p-value</th>
<th>p-adjusted</th>
<th>Reference</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPPB</td>
<td>ENSG00000120937</td>
<td>6.61</td>
<td>1.07E-49</td>
<td>8.98E-46</td>
<td>[36,38]</td>
<td>RA</td>
</tr>
<tr>
<td>COLQ</td>
<td>ENSG000000206561</td>
<td>2.90</td>
<td>2.30E-42</td>
<td>1.46E-38</td>
<td>[36–38]</td>
<td>RA, LAA</td>
</tr>
<tr>
<td>NPPA</td>
<td>ENSG00000175206</td>
<td>2.00</td>
<td>4.37E-24</td>
<td>1.10E-20</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TNC</td>
<td>ENSG00000041982</td>
<td>1.87</td>
<td>0.0014</td>
<td>0.0245</td>
<td>[36,38]</td>
<td>RA</td>
</tr>
<tr>
<td>COL21A1</td>
<td>ENSG00000124749</td>
<td>1.66</td>
<td>9.52E-12</td>
<td>3.39E-09</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COL3A1</td>
<td>ENSG00000168542</td>
<td>1.57</td>
<td>2.57E-06</td>
<td>1.62E-04</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COL12A1</td>
<td>ENSG00000111799</td>
<td>1.53</td>
<td>3.01E-07</td>
<td>2.76E-05</td>
<td>[36]</td>
<td>RA</td>
</tr>
<tr>
<td>COL1A1</td>
<td>ENSG00000108821</td>
<td>1.52</td>
<td>1.11E-04</td>
<td>0.0036</td>
<td>[36,39]</td>
<td>RA, RA</td>
</tr>
<tr>
<td>ANGPTL2</td>
<td>ENSG00000136859</td>
<td>1.09</td>
<td>5.47E-11</td>
<td>1.57E-08</td>
<td>[36,37,40]</td>
<td>RA, LAA</td>
</tr>
</tbody>
</table>

EV—Ensembl Variation 96; RAA—Right atrial appendage; LAA—left atrial appendage

https://doi.org/10.1371/journal.pone.0232719.t002
control tissue samples (S5 Fig in S1 File). Altogether, qPCR results were in accordance with the microarray results. Average expression data from all subjects revealed that amongst the top 20 abundant miRNAs were miR-26a, miR-125b, let-7a/c and miR-23b-3p, thereby confirming the high expression levels in RA previously shown by others [41].

Amongst the differentially expressed miRNAs identified in this study, ten were reported previously as abnormally expressed in AF atrial tissues. miR-338-5, miR-10a-5p and miR-200b-5p were down-regulated in our AF patients with similar results reported in RA free wall and RA appendage (RAA) tissue samples of AF patients [30,31]. The same was observed with up-regulated miRNAs, namely miR-30d-5p, miR-216a-5p, miR-106b-3p, miR-130b-3p, miR-574-5p, miR187-5p and miR-187-3p, which were found to be up-regulated in the same human AF studies [30,31]. Increasing the p-value cut-off to 0.05, the number of miRNAs dysregulated in AF increased drastically to a total of 223 (D3 in S1 Dataset), including other commonly AF-associated miRNAs such as miR-15b, miR-26a-5p, miR-24-3p, miR-208a-5p and miR-193a-5p [15,30–32,42–44]. Nonetheless, some of the miRNAs most commonly reported dysregulated in AF, such as miR-1, miR-21 and miR-143, were unchanged in our study [45].

**Novel putative miRNA-gene interactions in AF**

The 43 dysregulated miRNAs were predicted to bind 166,790 3’-UTR sites of 15,213 unique target genes. Interaction with genes outside our DEGs list were excluded. miRNAs are expected to negatively regulate mRNA expression and therefore, expected to have opposite
expression results. We analysed the expression correlation in all miRNA-mRNA interaction pairs using Spearman’s correlation coefficient. Negative correlation ranged -1 to -0.5 were selected. A total of 905 miRNA-mRNA pairs, including 37 miRNAs and 295 genes, passed the filtering (D4 in S1 Dataset). Fourteen of the miRNA-mRNA interactions were experimentally validated according to miRTarBase. We used Cytoscape to visualize the predicted regulatory networks [29]. Two networks were created, one with up-regulated miRNAs and corresponding down-regulated target genes (Fig 3) and a second network with down-regulated miRNAs and corresponding up-regulated target genes (S6 Fig in S1 File).

GO enrichment analysis revealed that our miRNA-gene interactions were enriched in eight categories (D5 in S1 Dataset). Down-regulated miRNAs were predicted to target genes

Fig 3. miRNA-gene regulatory networks build using Cytoscape. Network includes up-regulated miRNAs and candidate gene targets that were down-regulated in right atrial biopsies of AF patients compared to control subjects. The network helps to identify miRNAs targeting multiple DEGs and genes targeted by multiple miRNAs. Triangle nodes represent up-regulated miRNAs with red colour intensity according to FC. Rectangle nodes represent down-regulated genes with blue colouring according to FC. The darkness of the node to node edges correlates with the expression correlation value between a miRNA and its target gene. The darker the line, the closer to -1 is the correlation. All miRNA-gene pairs with correlation above -0.5 were excluded from the network. AF–atrial fibrillation; FC–fold-change.

https://doi.org/10.1371/journal.pone.0232719.g003
enriched for “extracellular space” and “extracellular region” cellular components (Fig 4A). In up-regulated miRNAs, the most significant enrichments of target down-regulated genes were related to plasma membrane components, “regulation of ion transmembrane transport” and “calcium ion binding”. KEGG analysis showed enrichment in the “cAMP signalling pathway” (Fig 4B).

Amongst the up-regulated miRNAs targeting the most DEGs, we identified miR-4436b-3p, miR-4728-5p and miR-6124 as putative regulators of 57, 56 and 67 down-regulated genes. These miRNAs are poorly characterized. Enrichment analysis of the target genes of each miRNA, revealed that down-regulated genes targeted by miR-6124 are enriched for “calcium ion binding”. Considering the importance of calcium dysregulation in the triggering of AF and disease progression of AF [46], it renders miRNA-6124 a potential role in AF. Conversely, AQP4 and RGS6 were the down-regulated genes predicted to be regulated by the higher number of overexpressed miRNAs, ten and thirteen respectively.

Lastly, we selected a subset of interactions including exclusively miRNAs that were previously shown to be dysregulated in AF tissues. The AF related regulatory networks are shown in Fig 5.
AF genomic elements co-localize in high-density regions

Genetics factors contributing to AF involve mono- and polygenic mutations and/or dysregulation of genes and non-coding RNAs, resulting in disrupted protein expression. Additionally, expression regulators, such as methylation and chromatin modifications, can regulate expression of entire sets of genes located in proximal genomic regions. To assess a possible chromosomal clustering of our dysregulated miRNAs and DEGs with previously reported AF-associated genetic variants, we analysed their relative genomic localisation. A total of 1008 genetic components were integrated in the analysis, including 43 DE miRNAs, 644 DE genes and 321 SNPs. Fig 6A illustrates a chromosome plot with the location of all genetic components included in our study. Chromosomes 1 and 6 show the highest density of genetic elements after normalization, with 107 and 78 elements respectively. Zooming into clusters with five or more genetic elements, we have identified seven high-density clusters located in chromosomes 2, 6, 11, 14, 15, 16 and 17. For a full list of clusters, genetics elements and genomic locations, see D6 in S1 Dataset.

An example locus is shown in Fig 6B. The figure represents cluster 580 located in chromosome 14 which includes rs422068, rs28631169, MYH6, MYH7, miR-208b and miR-208a. Another cluster, number 298, is located in the p arm of chromosome 6 and includes the highest number of novel genes associated with AF (S7 Fig in S1 File). This cluster is part of the major histocompatibility complex (MHC) class III gene cluster and contains five up-regulated
genes from our RNA-seq experiments. The MHC III cluster is heterogeneous and poorly characterized. The MHC III region is composed by more than 60 genes encoding signalling molecules involved in inflammatory response and a variety of different cell communication processes [47].

Furthermore, we analysed the genomic location of dysregulated miRNAs and genes to assess possible co-expression patterns. 22 of our 42 differentially expressed miRNAs were located in either intronic or exonic regions. The expression of both miRNA and host genes was changed in the same direction in ten of the 22 miRNAs. In four of those cases, changes were significant for both the miRNA and the gene. These findings support previous studies showing that transcription of intronic and exonic miRNAs is regulated by the same promoter as their host genes [48,49].
Discussion

Transcriptome analysis is a powerful tool to help identify new molecules and pathways of disease as it allows the simultaneous analysis of coding and non-coding RNAs. To date, few studies have applied RNA-sequencing to study transcriptome changes in tissue samples from AF patients [50,51]. In this study, we analyse the expression of genes and miRNAs in RA biopsies to investigate differences between in MVR patients with and without AF. The strength of our study is that we assessed miRNA and gene expression from the exact same tissue samples and pool of extracted RNA, thereby increasing the likelihood of selecting relevant miRNA-mRNA pairs.

We detected changes of expression in 14 cardiac ion channels that can generate electrical conductance disturbances. Notably, genes \textit{KCNJ5}, \textit{KCNN2} and \textit{HCN4}, encoding the inward rectifier K$^+$-channel subunit K$_{i}$3.4 (GIRK4), the small conductance calcium activated potassium channel member 2 K$_{Ca}$2.2 (SK2) and the hyperpolarization activated cyclic nucleotide gated potassium channel 4 (HCN4), respectively, were down-regulated in AF. These findings are in line with other transcriptional studies in different atrial samples [36,37]. The three channels have also been reported in AF GWAS studies. Down-regulation of \textit{KCNJ5} is believed to represent a compensatory mechanism to counteract shortening of the atrial effective refractory period seen in persistent AF [37]. In the heart, \textit{KCNN2} is predominantly expressed in the atria. Our findings of reduced \textit{KCNN2} expression in AF are consistent with Skibsbye et al. who reported \textit{KCNN2} down-regulation and reduced functional importance of the channel in chronic AF [52].

Genes associated with increased cardiac fibrosis and extracellular matrix were upregulated in AF patients. This is an indicator of structural remodelling in hearts of AF patients mediated by interstitial fibrosis due to accumulation of collagen fibres and fibroblasts in the extracellular matrix. The fibrotic processes can be induced via activation of the TGF-ß1/Smad3 signalling pathway with increase \textit{COL1A1} and \textit{COL3A1} expression [53]. Unfortunately, we were unable to investigate interstitial fibrosis in our AF patients, due to limited tissue availability. Lastly, we found \textit{NPPB} to be up-regulated by almost 7-fold in our AF patients. \textit{NPPB} encodes the natriuretic peptide B protein secreted by cardiac myocytes. We hypothesize that its up-regulation might be a response to cardiac fibrosis due to its anti-fibrotic function [54]. Overall, the changes in genes expression described here are in line with the electrical, contractile, and structural remodelling characteristic of AF hearts. We hypothesize that the more severe remodelling of the heart in certain MVR patients can be the key factor to create the trigger and subtract for development and maintenance of AF. The contribution of structural remodelling to AF susceptibility was also described by in two previous studies [55,56], both showing that more severe fibrosis in RA and LA was associated with increased susceptibility to develop AF in patients with concomitant heart disease and after cardiac surgery.

As negative regulator of gene expression, miRNAs bind to the 3'-UTR of target genes to inhibit protein production, by preventing translation by ribosomes and promoting mRNA degradation [11]. Importantly, a gene simultaneously targeted by multiple miRNAs is more likely to be regulated by miRNAs. Likewise, overexpressed miRNAs that target multiple genes are more likely to have an important role in a specific tissue [57]. Hence, it is important to study the collective action of miRNAs in disease. In the present study, differentially expressed miRNAs include ten miRNAs previously reported as changed in AF tissues, but with little known about their functional role. We built an AF related regulatory network including 905 miRNA-target gene pairs that can play a role in the development of AF in MVR patients.

Aquaporin 4 (\textit{AQP4}) and regulator of G-protein signalling 6 (\textit{RGS6}) were the down-regulated genes predicted to be targeted by the highest number of DE miRNAs. \textit{AQP4} contributes
to maintenance of water and electrolyte balance in aging hearts and its decreased expression might compromise homeostasis in cardiac cells [58]. RGS6 is involved in modulation of parasympathetic regulation of heart rate. Down-regulation of RGS6 can contribute to electrical remodelling and shortening of the action potential by increasing the activation time of the G protein-coupled inwardly rectifying K⁺ Kᵢ₃.1/3.4 (GIRK) channel, resulting in a large K⁺ current and membrane hyperpolarization [59,60].

MiR-338-5p targeted 14 of our DEGs, including IGFBP2, LBH and NPPA. IGFBP2 and LBH have both been reported as up-regulated in three different AF studies and they are increased by 8 and 3 FC in our RA samples, respectively [36–38,40]. Mutations in NPPA have been associated with familial AF [61]. NPPA encodes the natriuretic peptide A protein, a circulating hormone released from atrial cardiomyocytes upon stretching. The NPPA protein plays a role in cardiac electrophysiological function, where it shortens the atrial effective refractory period and conduction velocity in human hearts by inhibiting sympathetic and parasympathetic activity or through direct regulation of cardiac ion channels [62–64]. MiR-338-5p was also reported as down-regulated in RAA of AF patients with mitral stenosis [30]. This miRNA has also been implicated in tumor progression and reported to alleviate lung fibrosis [65,66].

MiR-130b-3p was upregulated in RA samples in line with a previous study [31]. MiR-130b-3p is predicted to bind 15 of our down-regulated genes, amongst those there are CACNA2D2 and SLC7A11 which have been reported as down-regulated in AF transcriptomics studies [36,38,67]. CACNA2D2 is a subunit of the voltage-gated calcium channel that modulates the calcium signalling response in cardiomyocytes and induces arrhythmia when down-regulated [68].

Genomic co-localization analysis helped mark regions of interest in the long list of DE genes and miRNAs. A high-density cluster in chromosome 14q11.2 marked a genetic locus that includes genes MYH6, MYH7, miR-208a, miR-208b and SNPs rs422068 and rs28631169 within 20 kbp of distance. We find miR-208a down-regulated in AF patients by increasing the array analysis p threshold to 0.05. The same was reported in RAA from AF patients and atria of ventricular tachypaced dogs [30,32,69]. MiR-208a-3p was described by Li et al. to regulate the expression of connexin 40 in the heart, resulting in electrical impulse propagation defects between adjacent cardiac myocytes [70]. Furthermore, miR-208a-3p is believed to generate protective effects in myocardial infarction and it could become a potential treatment for structural remodelling in AF [71]. Genetic deletion of miR-208a in mouse hearts interferes with cardiac conduction, with lack of P waves in ECG recordings and an AF like phenotype [72]. MiR-208a is encoded by an intron of the MYH6, which is also down-regulated in our samples, therefore suggesting that miR-208a is co-expressed with its host gene. MYH6 is an important protein in cardiac muscle fibre composition, involved in muscle contraction mechanisms. Both miR-208a and MYH6 are expressed predominantly in the atrial chambers. MYH7 and miR-208b are up-regulated in our RNA-seq studies by 4 and 3.5 FC, suggesting once again co-expression. MYH7 is mainly expressed in the ventricular chambers and is part of sarcomere formation, therefore involved in cardiac contraction mechanisms. Several MYH7 mutations have been reported in cardiac muscle disorders, including AF [73]. MiR-208b was reported up-regulated in isolated myocytes from chronic AF patients and is involved in calcium homeostasis and contractile remodelling [74]. In conclusion, our study emphasizes the significance of chr14q11.2 as a AF susceptibility locus.

**Study limitations**

The study cohort size is modest and challenges the assessment of the effect of confounding factors in the results. Furthermore, it is important to mention that different studies investigating
miRNA and gene expression in atrial samples from AF patients overlap poorly [45]. These inconsistencies might reflect 1) heterogeneity between tissue samples (location and collection method), 2) AF type and duration, 3) variation in patient cohorts (comorbidities, life-style, age, medication), and 4) the experimental technique selected to study transcript changes. In fact, our RNA-seq results overlap best with gene expression profiling studies from Jiang et al. using similar tissue samples and experiments [38]. A continuous effort from experts to study larger cohorts and collect similar datasets will allow a combined analysis of multiple samples from AF patients and help on accurate selection of target genes and regulatory networks involving miRNAs.

The in silico miRNA-target gene predictions provide qualified results which are generally accepted by research. However, these interactions require further validation and functional studies to clarify the pathophysiological role of such interactions, such as luciferase assays and in vitro miRNA transfections.

Lastly, the functional role of each DEG in AF is difficult to estimate since many genes are not yet characterized and we miss information on cell type specificity. For instance, we see down-regulation of KCNQ5 in our AF samples, a gene known for being mainly expressed in neuronal, skeletal and smooth muscle cells and therefore, not expected to interfere with cardiac electrical conductance [75]. Single-cell gene expression studies would allow us to detect the cell population creating the change, giving a better idea of the mechanism in which each DEG is involved.

Conclusions

By combining miRNA and gene expression data from RA of MVR patients with AF, we were able to create a novel miRNA-mRNA regulatory network, providing new insights into the mechanism predisposing certain MVR patients to develop and maintain AF. Down-regulation of ion channel genes and up-regulation of extracellular matrix genes summarize the major changes creating a subtract for AF. Lastly, we identified a high-density AF loci in chromosome 14q11.2.

Supporting information

S1 File. Supplementary methods and S1–S7 Figs.
(PDF)

S1 Dataset. Excel file with S1–S6 Data.
(XLSX)

Acknowledgments

The authors would like to thank Dr. Jesper Hastrup Svendsen and Dr. Stig Haunsø (Department of Cardiology, University Hospital of Copenhagen, DK) for sharing tissue biopsies and clinical data from AF and control patients. We are also grateful to the Center for Genomic Medicine at the University Hospital of Copenhagen (Rigshospitalet) for performing microarray and RNA-sequencing experiments.

Author Contributions

Data curation: Joana Larupa Santos, Ismael Rodríguez.

Formal analysis: Joana Larupa Santos, Ismael Rodríguez.

Funding acquisition: Morten S. Olesen, Bo Hjorth Bentzen.
Investigation: Joana Larupa Santos.
Methodology: Joana Larupa Santos, Morten S. Olesen, Bo Hjorth Bentzen, Nicole Schmitt.
Project administration: Bo Hjorth Bentzen, Nicole Schmitt.
Supervision: Nicole Schmitt.
Writing – original draft: Joana Larupa Santos.
Writing – review & editing: Ismael Rodríguez, Morten S. Olesen, Bo Hjorth Bentzen, Nicole Schmitt.

References


