ESC European Society of Cardiology

Cardiovascular Research (2021) **117**, 1760–1775
doi:10.1093/cvr/cvaa307

# Transcriptome and proteome mapping in the sheep atria reveal molecular featurets of atrial fibrillation progression

**Alba Alvarez-Franco [1†], Raquel Rouco [1,‡], Rafael J. Ramirez [2§], Guadalupe Guerrero-Serna [2], Maria Tiana [1], Sara Cogliati [1,4], Kuljeet Kaur[2,¶], Mohammed Saeed[2], Ricardo Magni[1], Jose Antonio Enriquez [1], Fatima Sanchez-Cabo [1], José Jalife [1,2,3]*, and Miguel Manzanares [1,5]***

[1]Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain; [2]Department of Internal Medicine, Center for Arrhythmia Research, University of Michigan, Ann Arbor, MI, USA; [3]Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Spain; [4]Department of Physiology, Institute of Nutrition and Food Technology, Biomedical Research Centre, University of Granada, Granada, Spain; and [5]Centro de Biología Molecular Severo Ochoa, CSIC-UAM, Madrid, Spain

*Received 31 August 2020; editorial decision 23 September 2020; accepted 15 October 2020; online publish-ahead-of-print 29 October 2020*

**Aims** Atrial fibrillation (AF) is a progressive cardiac arrhythmia that increases the risk of hospitalization and adverse cardiovascular events. There is a clear demand for more inclusive and large-scale approaches to understand the molecular drivers responsible for AF, as well as the fundamental mechanisms governing the transition from paroxysmal to persistent and permanent forms. In this study, we aimed to create a molecular map of AF and find the distinct molecular programmes underlying cell type-specific atrial remodelling and AF progression.

**Methods and results** We used a sheep model of long-standing, tachypacing-induced AF, sampled right and left atrial tissue, and isolated cardiomyocytes (CMs) from control, intermediate (transition), and late time points during AF progression, and performed transcriptomic and proteome profiling. We have merged all these layers of information into a meaningful three-component space in which we explored the genes and proteins detected and their common patterns of expression. Our data-driven analysis points at extracellular matrix remodelling, inflammation, ion channel, myofibril structure, mitochondrial complexes, chromatin remodelling, and genes related to neural function, as well as critical regulators of cell proliferation as hallmarks of AF progression. Most important, we prove that these changes occur at early transitional stages of the disease, but not at later stages, and that the left atrium undergoes significantly more profound changes than the right atrium in its expression programme. The pattern of dynamic changes in gene and protein expression replicate the electrical and structural remodelling demonstrated previously in the sheep and in humans, and uncover novel mechanisms potentially relevant for disease treatment.

**Conclusions** Transcriptomic and proteomic analysis of AF progression in a large animal model shows that significant changes occur at early stages, and that among others involve previously undescribed increase in mitochondria, changes to the chromatin of atrial CMs, and genes related to neural function and cell proliferation.

**Keywords** Atrial fibrillation • RNA-seq • Proteomics • Mitochondria • Chromatin



*Corresponding authors. Tel: +34-914 531 200, E-mail: jjalife@cnic.es (J.J.); Tel: +34-914 531 200, E-mail: mmanzanares@cnic.es (M.M.)
†The first two authors contributed equally to this work.
‡Present address: Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland.
§Present address: Pauley Heart Center, Virginia Commonwealth University, Richmond, VA, USA.
¶Present address: Department of Immunology, Tufts University School of Medicine, Boston, MA, USA.

# 1. Introduction

Atrial fibrillation (AF) is the most common arrhythmia. Estimates indicate that it affects over 33 million people worldwide,[1] and US Census projections estimate that by 2050 incidence of AF will more than double.[2] AF typically starts with brief paroxysmal episodes that in many patients evolve, lasting progressively longer until AF becomes persistent and then permanent. Progression from paroxysmal to persistent AF is accompanied by incremental functional and structural modification of the atria that permits its long-term perpetuation.[3] Such changes also increase the risk of hospitalization and subsequent cardiovascular events.[4] However, the fundamental mechanisms governing the conversion from paroxysmal to persistent and permanent forms are not well understood, leading to poor prevention and suboptimal treatment. Thus, there is a clear need for more inclusive and large-scale approaches to understand the molecular mechanisms responsible of the disease.

Obtaining reliable samples from patients is difficult. Therefore, different animal models have been fundamental to understand the mechanism underlying AF onset and progression.[5–7] We have previously shown that in a clinically relevant sheep model of tachypacing induced, long-standing persistent AF, the duration of the AF episodes increase progressively at a rate that predicts the time at which AF stabilizes and becomes persistent.[3,8] In a recent clinical study, remote transmissions of atrial activation frequency in patients with implantable cardioverter-defibrillator ± resynchronization therapy or pacemaker devices enabled the demonstration that the evolution of the AF-induced atrial electrical remodelling follows a remarkably similar general pattern as the sheep model.[9] In the sheep, the changes in AF frequency and duration reflected changes in action potential duration and densities of ion currents.[3] There were also gradual changes in atrial structure including atrial dilatation, hypertrophy, and fibrosis, all of which were mitigated, but not prevented by upstream therapy with antifibrotic agents like eplerenone and a galectin-3 inhibitor.[10,11] The above results in animals and patients highlighted the importance of investigating the mechanisms of AF progression. They also provided strong evidence that monitoring the progression of atrial electrical remodelling can help physicians in the stratification and personalized care of AF patients.[9]

In this work, we have characterized the transcriptomic and proteomic dynamics that result in progressive remodelling of both atria, making the arrhythmia more stable and long-lasting. We used a well-established experimental model of intermittent atrial tachypacing[3,11] and created a user-friendly and interactive web application to enhance and facilitate data sharing in our search for gene and protein expression changes that underlie AF-induced atrial remodelling.

# 2. Methods

## 2.1 Experimental animals

We induced AF using atrial tachypacing as described previously.[3,8] Three groups of male sheep were used: transition (13.75 ± 4.50 days of self-sustained AF without reversal to sinus rhythm), chronic or long-standing persistent AF (289.25 ± 50.29 days of self-sustained AF without reversal to sinus rhythm), and control group (sinus rhythm). Induction of anaesthesia was done with intravenous propofol (4–6 mg/kg), and maintained with isoflurane gas (5–10 mL/kg). Euthanasia was attained by removal of the heart under anaesthesia. Animal procedures were approved by the University of Michigan Committee on Use and Care of Animals and conformed to National Institutes of Health guidelines.

## 2.2 Transcriptomics and proteomics

Hearts were removed by thoracotomy and placed in cold cardioplegic solution. Samples from left atrial appendage (LAA) and right atrial appendage (RAA), or from the posterior left atria (PLA) were dissected and snap-frozen. CM isolation from LAA and RAA was performed as previously described.[3] RNA-seq (LAA and RAA tissue and CMs, $n = 3$ for control, transition and chronic groups; PLA tissue, $n = 6$ for control and transition groups) was performed by CNIC Genomic Unit using an Illumina Hiseq 2500 sequencer, and liquid chromatography tandem–mass spectrometry (LC-MS/MS) by the CNIC Proteomics Unit (LAA and RAA tissue and CMs, $n = 3$ for control, transition and chronic groups).

## 2.3 Data analysis and integration

Only those genes or proteins (which we collectively term as 'features') from sheep with clear human homologues were used, keeping 11 962 unique gene symbols from the human genome. We applied Multiple Co-Inertia Analysis[12] as a method for unsupervised dimensionality reduction, necessary for the analysis of such large databases. Principal components analysis showed that components 1–3 explain ~42% of the total variability in the data, so these were retained for further analysis. Gaussian mixture model (GMM) was used as an unsupervised clustering method to identify in an unbiased way grouping of the selected features.

We performed functional enrichment analysis of the features of these groups using Gene Ontology,[13] in order to identify biological functions or processes statistically over represented among the genes and proteins of each group. Epigenetic modifiers were obtained from the curated database dbEM.[14] Transposable elements (TEs) were annotated and manually curated with the repeat masker track from the UCSC Genome Browser (Oar3.1). Gene set enrichment analysis of genome-wide association studies (GWAS) was performed using gene lists from the public GWAS Catalogue and from meta-analysis of AF-associated genes.[15,16]

RNA-seq data have been deposited at the Gene Expression Omnibus (GEO Database GSE138255) and proteomic data at the Proteome Xchange Consortium (PXD015637). An interactive Shiny app to browse and analyse the full data is available at https://bioinfo.cnic.es/AFibOmics.

## 2.4 Protein analysis

For western blots, cells were harvested in RIPA buffer containing protease inhibitors, resolved in 15% SDS-polyacrilamide gels, and transferred to polyvinylidene difluoride membranes. Bands were quantified using ImageJ and normalized to Troponin T2 (TNNT2). Blue native gel electrophoresis (BNGE) were performed from mitochondria extracts from LAA and RAA isolated CMs (between 2–5 × $10^6$) as described.[17]

# 3. Results

## 3.1 Distinct molecular changes occur rapidly at the transition to early persistent AF

To understand the dynamics of expression that take place during the progression of AF, we took advantage of a well-established model of tachypacing-induced long-standing AF in the sheep.[3,8] We sampled tissue from both the LAA and RAA from three male sheep each as follows: control, transition (1–2 weeks of persistent self-sustained AF in the absence of tachypacing), and chronic (more than 10 months of self-sustained persistent AF) groups (*Figure 1A* and Supplementary material online, *Table S1*). We used LAA and RAA whole-tissue samples for

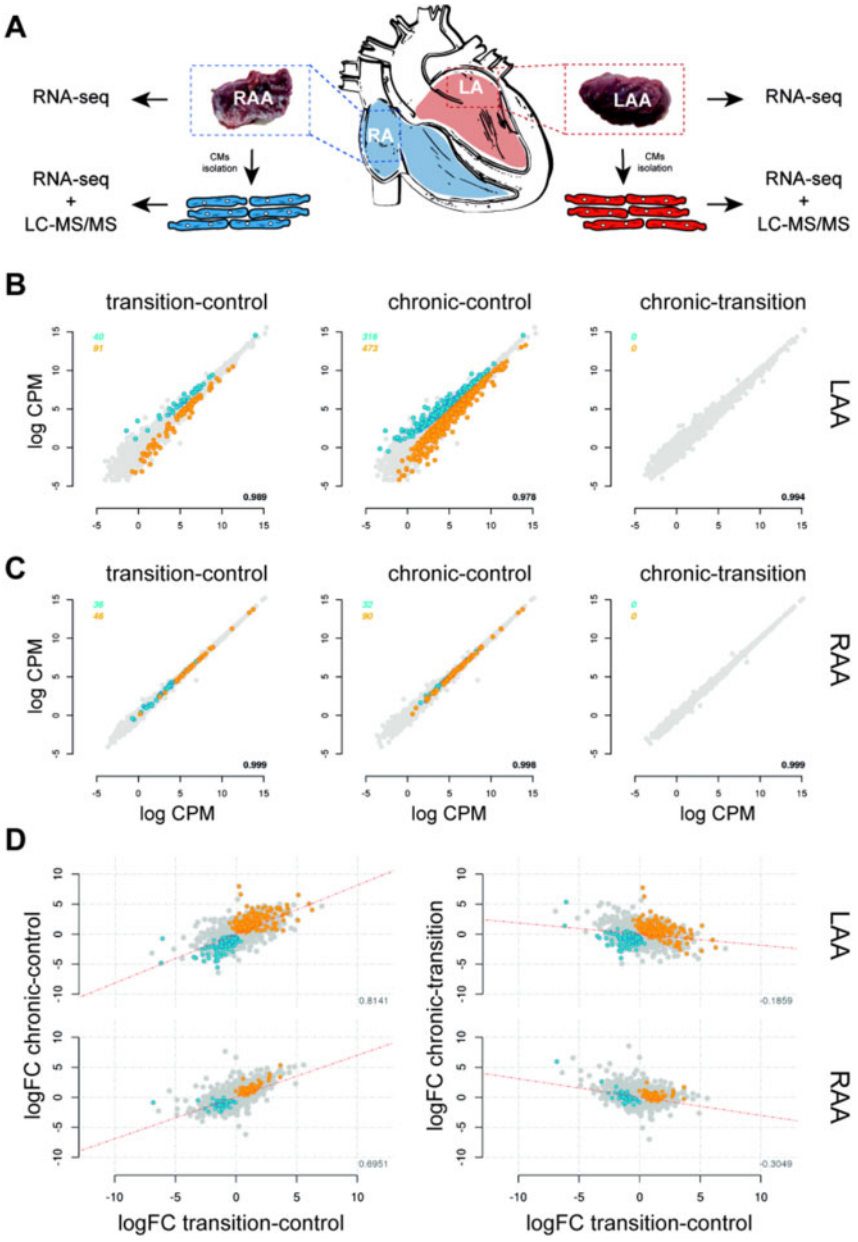


**Figure 1** Transcriptomic profiling of a sheep model of AF progression. (*A*) Schematic diagram of the experimental strategy, the collected samples from the left (LA) and right (RA) atrial appendages and the analyses that were carried out. Three samples were collected in each case and for each analysis. (*B* and *C*) Correlation of mean expression values as log of counts per million (CPM) in atrial tissue RNA-seq between transition and control (left panel), chronic and control (middle panel), and chronic and transition (right panel) from left (*B*, LAA) and right (*C*, RAA) atrial appendages. Orange and blue indicate, respectively, up- and down-regulated genes for each comparison. The Pearson coefficient of correlation is indicated on the lower right corner of each plot. (*D*) Progression of changes in gene expression in atrial tissue along persistent AF. The linear regression adjustment of control-to-transition logFC (fold-change) to those of control-to-chronic (left panels) and transition-to-chronic (right panels), in left (upper panels) and right (lower panels) atrial tissue is shown. Orange and blue indicate up- and down-regulated genes. The R2 value is indicated on the lower right corner of each plot.

transcriptomic profiling by RNA-seq. We also isolated CMs from LAA and RAA samples and performed both RNA-seq and LC-MS/MS proteomic profiling (*Figure 1A*).

To perform this study, we have used three replicates, in line with the normal practice in the field, where statistical methods have been developed specifically to tackle small sample sizes.[18,19] However, we also analysed the intergroup variability by performing unsupervised clustering of the results. This showed how variability among groups (control, transition, or chronic) is lower than between groups and they cluster together (Supplementary material online, *Figure S1A*). Overall, we detected 13 187 and 13 262 transcripts from whole tissue and CMs samples, respectively, of which a total of 1367 and 2479 genes were differentially expressed (5% false discovery rate, FDR) in at least one of the comparisons carried out (Supplementary material online, *Table S2*). On the other hand, we identified 7283 proteins in CMs, from which 581 had significant differential abundances (5% FDR).

To assess how well changes in mRNA and protein expression correlated in our system, we compared log fold-changes between control and transition, as well as control and chronic stages of disease progression from transcriptomic and proteomic data we generated from LAA and RAA CMs. Correlation was relatively low but evident (Supplementary material online, *Figure S1B*), and in line to what has been previously reported.[20,21]

We initially performed comparisons of both transcriptomic and proteomic profiles between control, transition and chronic groups from LAA and RAA. Pairwise correlations of LAA tissue RNA-seq showed that changes to the control condition increase with disease progression (from the transition to the chronic group); however, when we compared transition and chronic groups, we recovered no differentially expressed genes (*Figure 1B*). We observed the same trend for the comparisons of both RNA-seq and LC-MS/MS of LAA CMs (Supplementary material online, *Figure S2A* and *B*). As for the RAA, changes for both tissue and CMs were much lower but followed the same trend and again we detected no changes between transition and chronic states (*Figure 1C* and Supplementary material online, *Figure S2A* and *B*). Therefore, this initial analysis suggests that the LAA undergoes more profound changes during its progression to persistent AF than the RAA. It is also noteworthy that no transcriptomic or proteomic changes are observed from early persistent AF (transition) to long-standing persistent AF (chronic) in LAA or RAA for any of the comparisons.

To explore further the above observation, we compared the degree of change for all expressed genes and proteins (measured as log-fold changes) between control and transition groups to that of control versus chronic or transition versus chronic. In whole atrial appendage tissue RNA-seq, there is a positive linear relationship between the changes occurring from control to transition and control to chronic, indicating that the same trend in expression variation occurs along disease progression (*Figure 1D*, left panels). However, these changes level out and lead to a flat or even slightly descending relation when we compare transition to chronic states (*Figure 1D*, right panels). We observe the same behaviour for both RNA-seq (Supplementary material online, *Figure S2C*) and proteomic data (Supplementary material online, *Figure S2D*) from purified CMs. This analysis thus confirms that the mayor events related to AF progression occur during early phases of the disease and later stabilize as the animal moves from the transition towards the chronic state. We believe this pattern could not be secondary to mayor morphological changes taking place during AF progression, such as the increase in atrial dimensions that occur by dilation of the myocardium.[3] Measurement of the differences in atrial area shows that it increases steadily from control to chronic sheep, and does not level out from transition to chronic as does occur with transcriptional and proteomic changes (Supplementary material online, *Figure S3A*).

Remarkably, the temporal dynamics of the above changes in gene and protein expression corresponded closely with the electrical and structural remodelling we demonstrated previously in the sheep,[3] where dominant frequency measurements in chronic (long-standing persistent AF) sheep did not significantly change compared with the values recorded at transition (early persistent) stage (Supplementary material online, *Figure S3B*).

## 3.2 A three-component model explains molecular variation during AF progression

To understand these early changes in the expression dynamics, we analysed the variability of our multiple datasets and established an integrative data analysis model. Given the complexity of the data and the multiple layers of information (control, transition, and chronic disease states; LAA and RAA samples; whole atrial appendage tissue and isolated CMs; transcriptomics and proteomics; *Figure 1A*), we applied dimensionality reduction methods for data integration.[22] We first analysed each technical dataset (atrial appendage tissue RNA-seq, CM RNA-seq, and CM LC-MS/MS) separately by non-symmetric correspondence analysis, which transforms each dataset into a series of unsupervised lower dimensional units. In this way, we identified 18 principal components that behaved in a similar fashion in the three datasets (Supplementary material online, *Figure S4A*). To integrate the three experiments, we performed co-inertia analysis[23] among datasets obtaining a ranked order of pseudo-eigenvalues that explained the variability of the data (Supplementary material online, *Figure S4B*).

We next explored the biological sources of variation that could underlie each of these components, and found that the progression of the disease from control to chronic explains the first principal component (PC1), accounting for 17.6% of the total variation in the data (disease progression; *Figure* 2A, B and Supplementary material online, *Figure S4C*). The second component (PC2, 15.4% of variation) represents the regional differences between LAA and RAA (left/right identity; *Figure 2A* and Supplementary material online, *Figure S4C*). Interestingly, we observed that the third component (PC3, 8.4% of variation) groups together samples from control and chronic individuals, separating them from transition individuals (transition state; *Figure 2B* and Supplementary material online, *Figure S4C*). This was an unexpected finding, but it reinforced our previous observations (see above) that the mayor changes occurring during disease progression occur during the transition from paroxysmal to persistent AF. We could not identify other possible effects that would account for further components of the variation (Supplementary material online, *Figure S4C*), so we decided to use the first three components, which together explain more than 40% of the total variability, to model and analyse the molecular mechanisms underlying AF progression.

## 3.3 Defining the molecular features responsible for atrial divergence and disease progression

We subsequently characterized the molecular changes occurring during AF progression using two different criteria to select features (genes and/or proteins) from our RNA-seq and proteomic analysis. On one hand, we selected features showing differential expression in any of the pairwise comparisons (Supplementary material online, *Table S2* and *Figure S5A, B*) significantly detected at 5% FDR (Benjamini–Hochberg correction). This resulted in a list of 3278 differentially expressed genes and proteins (DEG/P; Supplementary material online, *Table S2* and *Figure S5D*). On the other hand, we selected features showing extreme values (outer 10%) in our three-component space (Supplementary material online, *Figure S5C*), which totalled 1790 extreme-value genes and proteins (exG/P; Supplementary material online, *Table S2* and *Figure S5D*). Both groups overlapped in 658 features (Supplementary material online, *Figure S5D*), including key factors with atrial-restricted expression (such as *PITX2* or *BMP10*) or related to atrial physiology (*IL6R, KCNN2, NPPA*, or *RCAN1*). In total, we retained 4410 features for further analysis.

To better understand the biological relevance of these features and their behaviour, selected features were analysed by unsupervised clustering using GMM and separated into 31 independent clusters (g_0-g_30; Supplementary material online, *Table S2*) distributed along the three-component space as defined above (*Figure 2C and D*). The majority of

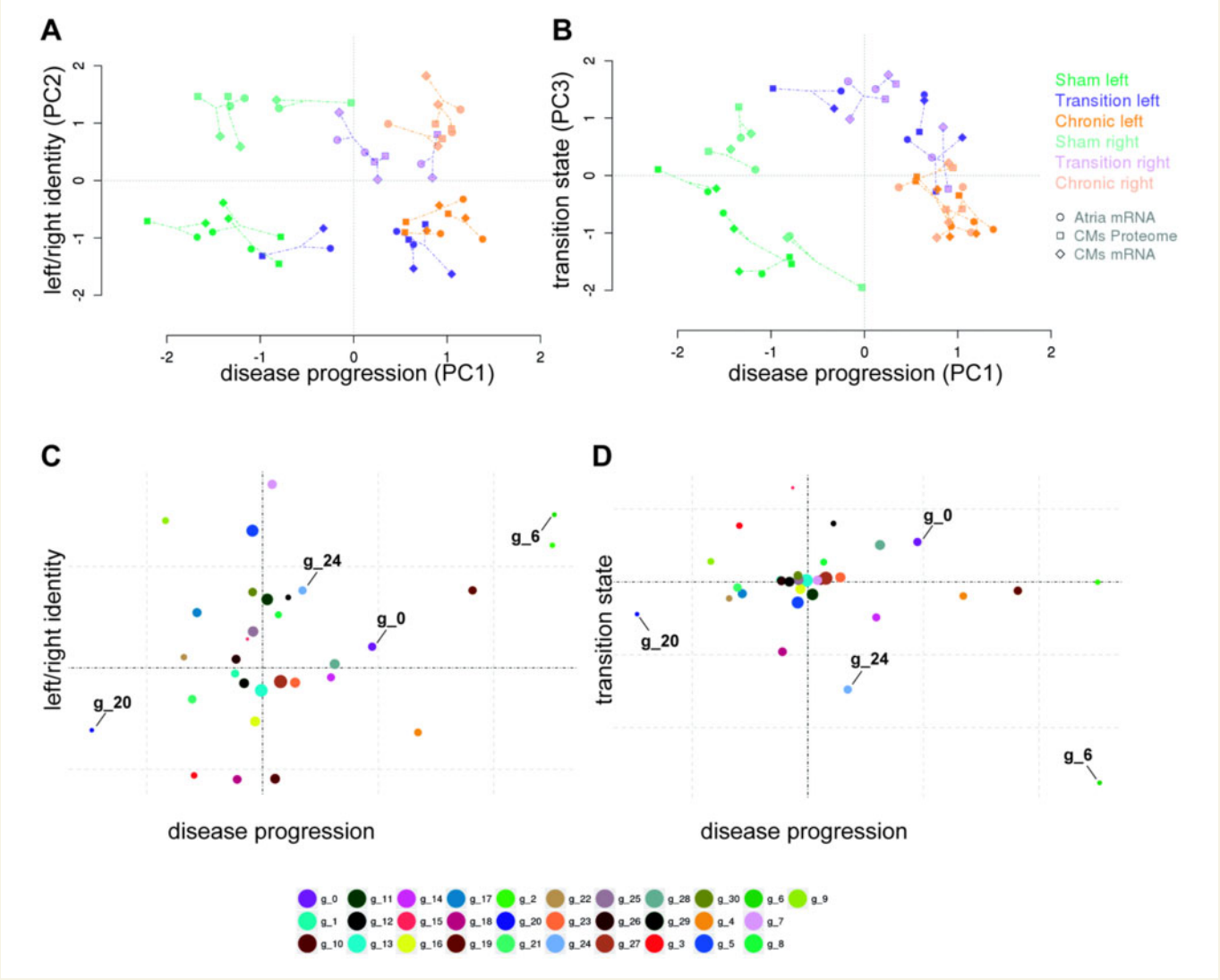


**Figure 2** Co-inertia analysis of multidimensional data identifies the main components that drive variability in the sheep AF model. (*A*) Distribution of transcriptomic and proteomic samples ($n = 3$) in relation to principal components PC1 (disease progression) and PC2 (left/right identity). Lines connect paired samples, obtained from the same individual. Control, green; transition, purple; chronic, orange. LAA samples, dark colours; RAA samples, light colours. Atrial tissue RNA-seq, circles; cardiomyocyte RNA-seq, diamonds; cardiomyocyte LC-MS/MS, squares. (*B*) Distribution of transcriptomic and proteomic samples in relation to components PC1 (disease progression) and PC3 (transition state). Legend as in A. (*C* and *D*) Position of each of the thirty-one clusters identified by GMM unsupervised clustering along the axis that define disease progression and left/right identity (*A*) or transition state (*B*). The size of each cluster represented on the plot correlates with the number of features (genes and proteins) that it includes. Colour legend is shown below. Arrows indicate the position of representative clusters (see *Figure 3*).

clusters showed contribution to more than one component, but those contributing significantly to the third component (PC3, transition state) axis were less abundant (*Figure 2D*). Examination of mean expression of genes or proteins from selected clusters showed direct relationship to the location of the cluster in the three-component space (*Figure 3A–D*). For example, cluster 0 (g_0) that includes among others the genes coding for the calcineurin regulator RCAN1 or Galectin-3 binding protein (*LGALS3BP*), contributes to disease progression and to a lesser extent to left/right identity and to transition state. Expression levels increase in all conditions from control to transition, and remain mainly stable from transition to chronic (*Figure 3A*). Cluster g_6 is interesting in that it shows an extreme position along all three components (*Figure 3B*) and includes *PITX2*, a left atrial marker and the gene most strongly associated to AF by GWAS,[24] or *PCP4*, a calmodulin regulator with specific expression in cardiac Purkinje cells.[25] Cluster g_20 is mostly related to disease progression, showing a trend for down-regulation of expression (*Figure 3C*). This cluster includes receptors for glucagon-like peptide 1 (*GLP1R*) or relaxin (*RXFP1*), both related to the control of insulin secretion. Cluster g_24, where we find key features related with atrial physiology such as the potassium channel gene *KCNH2*, or the nuclear receptor gene *NR4A3*, is strongly associated with the transition state. As such, in CM RNA-seq, we observe diminished expression in transition samples as compared to controls, which does not occur in chronic sheep (*Figure 3D*). Thus, this three-component model, allows us to explore the contribution of significative features to better understand the progression of AF.

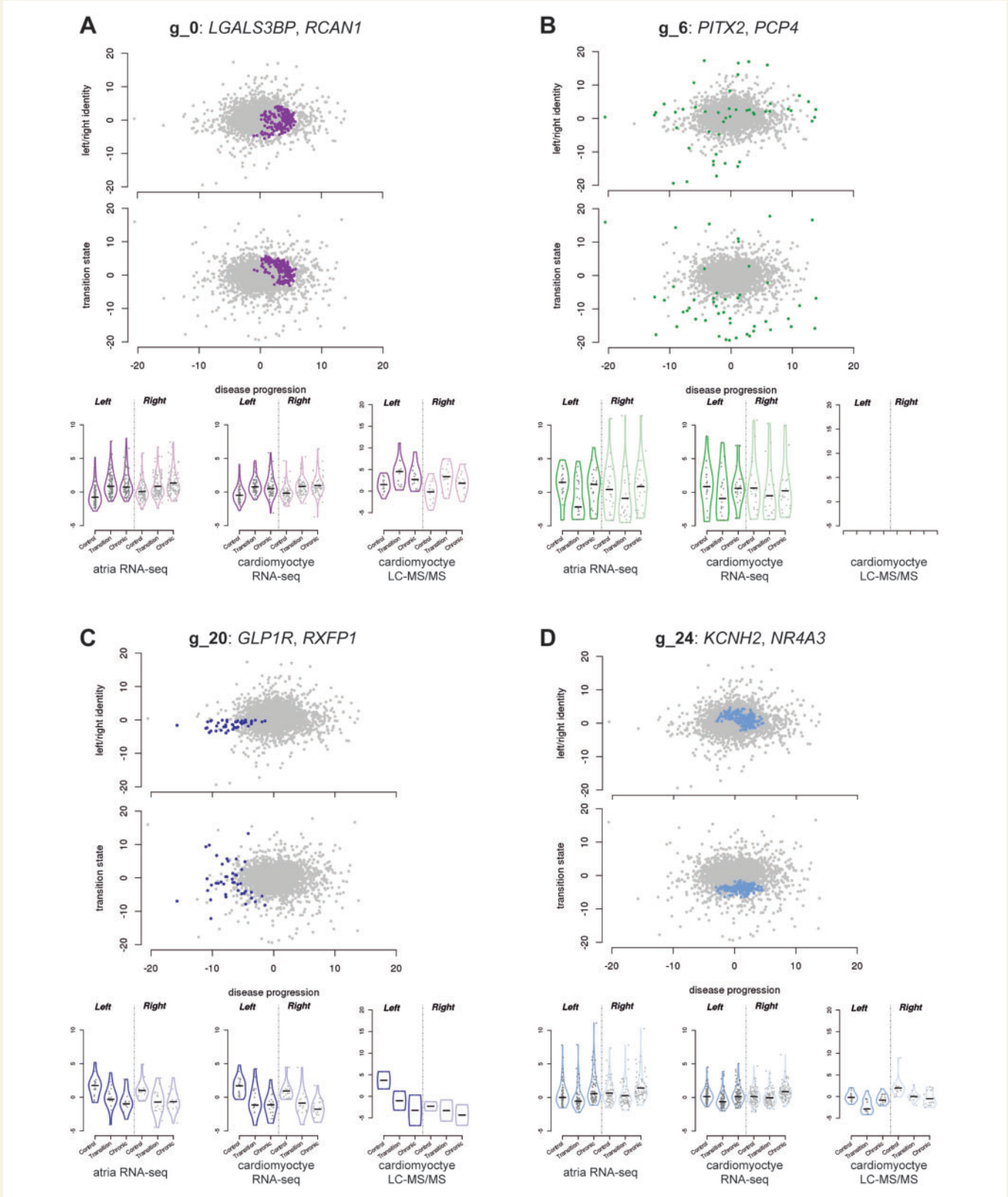


**Figure 3** Distribution and expression of representative GMM clusters in the three-component space of AF progression. The position of all individual features of the specified GMM clusters (*A*, g_0; *B*, g_6; *C*, g_20; *D*, g_24) along the disease progression axis and left/right identity (top) or transition state (middle). Below, violin plots depicting the expression of the features from the specified cluster in each individual experiment (atria RNA-seq, cardiomyocyte RNA-seq, and cardiomyocyte LC-MS/MS), condition (control, transition, and chronic) for both LAA and RAA; mean expression is indicated by a horizontal black line. No proteomic data was available for cluster g_6. Colour legend of the GMM clusters is as in *Figure* 2.

## 3.4 Distinct genetic programmes underlie cell type-specific variation in AF

To gain further insight into the molecular changes occurring during AF progression, we carried out functional enrichment analysis of each of the GMM clusters described above. We searched for statistically enriched Gene Ontology (GO[13]) terms among the features of each cluster, and discerned which of the three different experiments was contributing to this enriched annotation (Supplementary material online, *Figure S6* and *Table S3*). We were able to identify changes in gene or protein expression coming mainly from what we reckon are non-CM cells (those terms enriched in atria RNA-seq data but not in CM data), or post-transcriptional regulatory events (enriched in LC-MS/MS but not RNA-seq data).

As such, terms related to extracellular matrix were enriched in several clusters, showing a trend of increased gene expression during AF progression but mainly in whole-atrial appendage tissue and not in CMs (Supplementary material online, *Figure S6A*). This suggests that these changes occur in non-CM cells, most surely atrial fibroblasts, and relate to the increased fibrosis that has been described during AF progression.[3,26] Genes belonging to other broad categories, such as inflammation or ion channels, change in both atrial tissue and CMs (Supplementary material online, *Figure S6B* and *C*), suggesting a complex interplay between different cell types during disease progression. Ion channels genes show an interesting pattern of expression changes. While various potassium (*KCNJ3, KCNJ5*), calcium (*CACNA1C*) or sodium (*SCN5A*) channel genes show a decrease in expression in CMs in AF, most prominent when comparing control and transition samples (Supplementary material online, *Figure S6C*), other components such as *HCN2* or *KCNH7* are increased. These data are compatible with previous results in sheep atrial tissue showing increased or decreased protein levels of ion channels upon transition to persistent AF.[3]

We also found a large group of annotations related to heart muscle and myofibril structure (Supplementary material online, *Figure S6D*). As expected, changes in these genes were detected almost exclusively in CM samples. Again, here we find genes with different behaviours. Those genes belonging to cluster g_29 and annotated as contractile fibre contain genes such as cardiac muscle alpha actin (*ACTC1*) or myosin heavy chain 2 (*MYH2*), and show a general trend of down-regulation. On the other hand, genes included in cluster g_13 (annotated as myofibril) show specific up-regulation in LAA CMs as compared to RAA in chronic sheep. Genes in this cluster include those coding for Titin (*TTN*), Myomesin (*MYOM1*) or Myosin heavy chain beta (*MYH7*). Titin is the largest protein in humans and is essential for normal myocardial function. GWAS studies have found loss of function variants in TTN to be statistically associated with a diagnosis of early-onset AF.[27,28] *MYH7* encodes the slow molecular motor β-MyHC[29] that expresses only in the atria during cardiac development and not in the adult, but its expression is elevated in atrial myocytes of patients with chronic AF as well as in the ovine model of chronic AF.[30] Thus, our analysis identifies many of the previously reported changes in human and other animal models.

## 3.5 Mitochondria content increases during AF

Our functional enrichment analysis also uncovered novel and unexpected molecular changes associated to AF. For example, we identified a single category enriched in the proteomic but not transcriptomic data, corresponding to mitochondrial respiratory Complex I (Supplementary material online, *Figure S6E*). Interestingly, these proteins showed a tendency to increase along AF progression in LAA but not in RAA. To explore whether the increase of mitochondrial proteins affect the organization of the respiratory chain,[31] we performed Blue-Native Gel Electrophoresis (BNGE) of mitochondrial proteins from isolated atrial CMs. Immunodetection of Complexes I, II, III and IV (*Figure 4A* and Supplementary material online, *Figure S7*) showed a correct and equal distribution (Supplementary material online, *Figure S7A*) among well-characterized supercomplexes as well heart-specific bands for Complex IV[32] (IVn, *Figure 4A*). Interestingly, we observed a new band positive for Complex I (Ix, *Figure 4A*) between Complex I and Supercomplex I+III that increased during AF progression and would merit further characterization.

Despite the fact that BNGE showed an apparent increase of both complexes and supercomplexes through AF progression (*Figure 4A*), normalization using mitochondrial voltage-dependent anion channel, VDAC1 (Supplementary material online, *Figure S7B* and *C*) revealed no changes through AF progression, suggesting that their apparent increase could be due to an increment of the total mitochondrial mass. If so, we would expect that other proteins located in the mitochondria increase their expression during AF progression. In fact, if we analyse proteins that show a significant difference in the LC/MS-MS data, these are highly represented among those that increase their expression in CMs from LAA or RAA of chronic sheep as compared to controls (*Figure 4B* and Supplementary material online, *Table S3*). To confirm this hypothesis, we quantified mitochondrial DNA (mtDNA) copy number versus nuclear DNA as standard indicator of the mitochondrial mass. The results showed that while there is no difference in mtDNA content between left and right, in both LAA and RAA isolated CMs from the transition and chronic samples, mtDNA increases an average of three-fold compared to controls (*Figure 4C*). This strongly suggests that during AF progression, the mitochondrial mass increases together with the mitochondrial electron transport chain subunits without changing their relative distribution.

## 3.6 Chromatin dysregulation in cardiomyocytes is a hallmark of AF

An unexpected observation was the enrichment in terms related to chromatin, that were present in four independent clusters (g_1, g_25, g_26, and g_29; Supplementary material online, *Table S3*). The general trend was for a decrease in expression in both LAA and RAA, mainly in CMs (Supplementary material online, *Figure S6F*). We analysed the expression in CM RNA-seq data of 142 genes encoding chromatin related factors.[14] We identified an overall decrease in expression in transition and chronic sheep as compared to controls (*Figure 5A*). Down-regulated genes included those coding for different histone modifying enzymes (methyl-transferases, de-methylases, acetyl-transferases), related to both active and inactive chromatin and transcription, as well as nucleosome remodellers such as the NuRD complex (*Figure 5A*).

This observation led us to ask if there was a global dysregulation of chromatin in CMs from sheep with induced AF. We first measured the total amount of histones present in CMs by western blot (*Figure 5B* and Supplementary material online, *Figure S8A*), and found that there was an important decrease of both Histone 3 and Histone 4 in transition and chronic individuals. Interestingly, the decrease in histones was much more pronounced in LAA than RAA (*Figure 5B*).

The general decrease in chromatin remodellers together with lower amounts of histones suggested that an overall de-compaction of chromatin could be occurring in CMs during AF. Therefore, we decided to examine the expression of TEs from the sheep genome as a proxy for

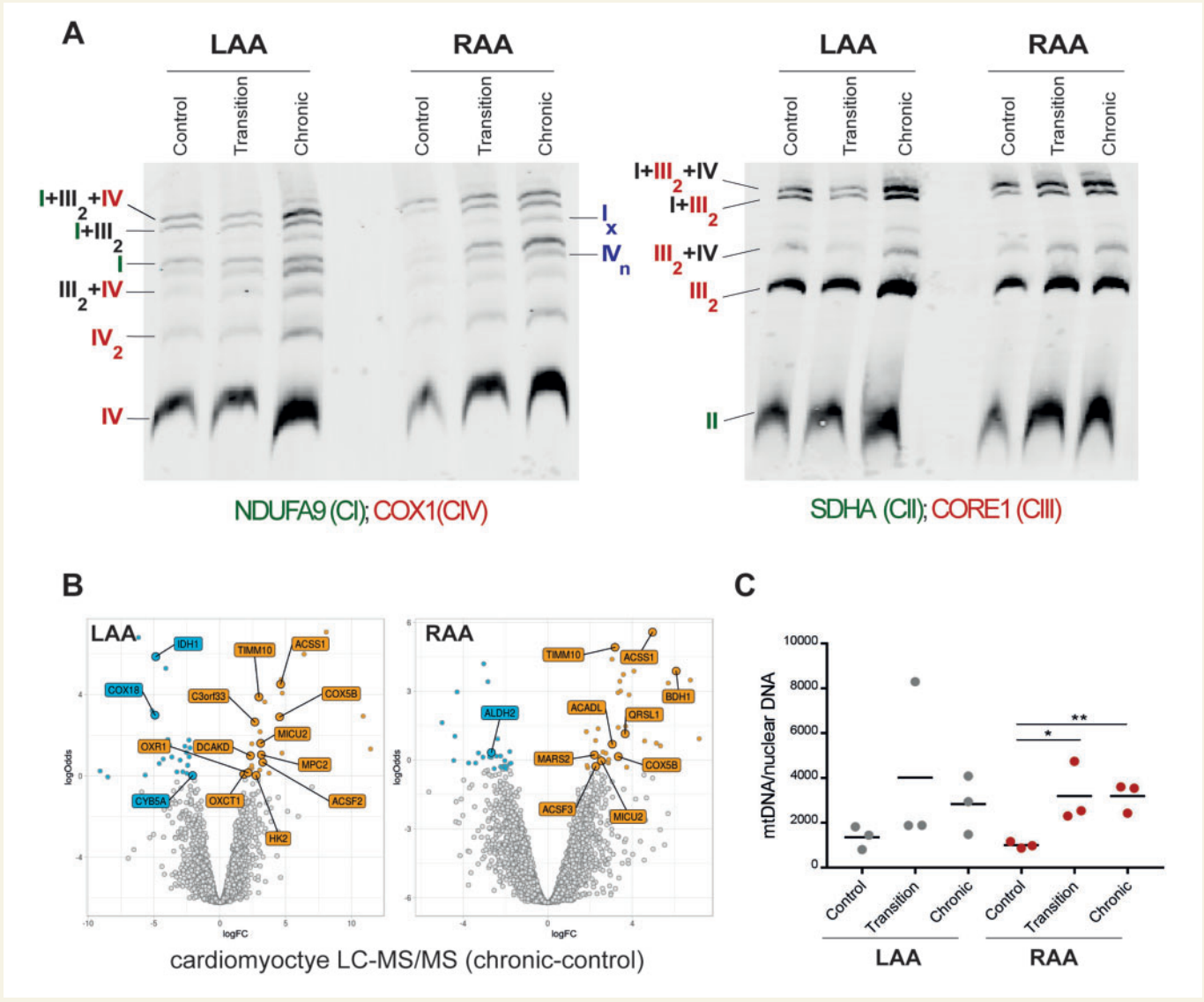


**Figure 4** AF progression increases mitochondrial mass. (*A*) Immunodetection of the respiratory super-complexes after BNGE of digitonin-solubilized mitochondrial proteins. Complexes and super-complexes detected are indicated. (*B*) Volcano plot of LAA and RAA cardiomyocyte LC/MS-MS data indicating significantly up-regulated (orange) and down-regulated (blue) proteins that localize to mitochondria in chronic sheep as compared to controls. (*C*) mtDNA copy number quantification in LAA and RAA by qPCR relative to nuclear DNA. $n = 3$; $^{*}P < 0.05$; $^{**}P < 0.01$; unpaired Student's *t*-test.

chromatin deregulation. Under normal circumstances, TEs are silenced except in very early stages of development.[33] However, it has been recently proposed that TE expression can occur in some pathological states and also during organismal ageing.[34] We reanalysed our transcriptomic data to assess TE expression, as this information is filtered out during standard pre-processing of RNA-seq data. In first place, we annotated the sheep genome's complement of TEs (Supplementary material online, *Figure S8B*), finding that the most abundant was the homologue of bovine BovB long interspersed element (LINE).[35] Expression of BovB LINE was increased in CMs during progression of AF, both in transition and chronic individuals and in left and right atria (*Figure 5C*). Altogether, these results suggest that progression of AF results in a global disorganization of chromatin, with a reduction of histones and remodelling factors, leading to a de-repression of TE expression.

## 3.7 Changes in the posterior left atrium mirror those in the atrial appendage

So far, our analysis was based on tissue and cells from the atrial appendage, but we wished to know if the molecular changes occurring in this tissue were representative of those happening in other anatomical locations of the atria. We took advantage of samples available from control and transition sheep from the PLA, which together with the pulmonary veins are the prime substrate for AF initiation and maintenance. Furthermore, we obtained six samples to analyse by RNA-seq from each condition, allowing us to test how reproducible our transcriptomic analysis was using small sample numbers (see above). We subsampled our dataset taking all possible combinations from 3 to 6 samples for each condition and carrying our differential gene expression analysis. We then compared the degree of overlap between the sets of differentially

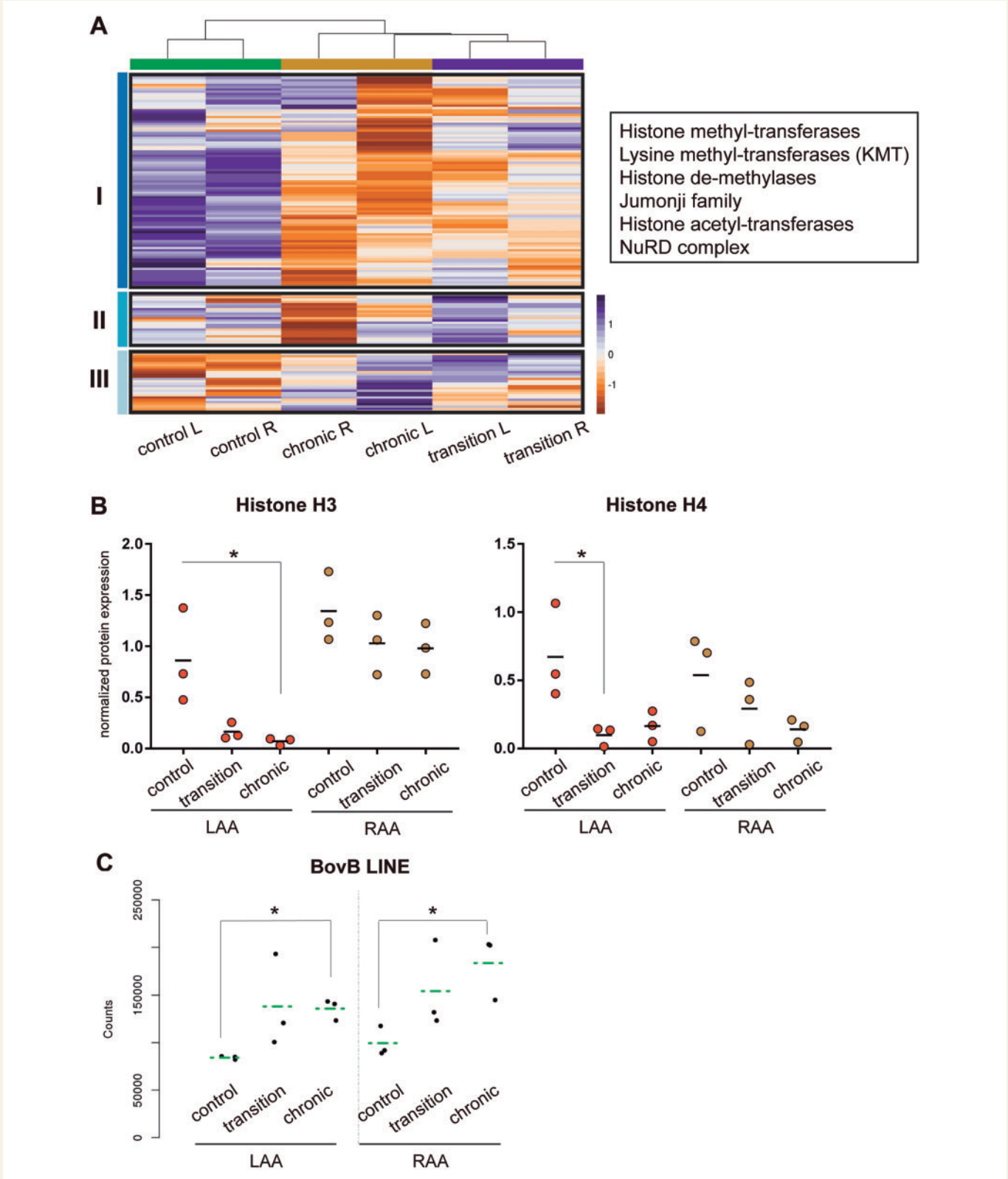


**Figure 5** Cardiomyocyte chromatin is disorganized in AF. (*A*) Heatmap showing the expression (as z-scores) of 142 genes encoding for chromatin remodellers in cardiomyocytes from LAA and RAA of control, transition, and chronic AF sheep. Three main clusters are observed (left), with cluster I showing decreased expression in transition and chronic conditions. This cluster includes mayor histone modifiers and nucleosome remodellers (shown on the right). (*B*) Quantification of the expression of Histone 3 (left) and Histone 4 (right) during AF progression in cardiomyocytes form right and left atria, as measured by western blot. Values were normalized to those of TNNT2 as cardiomyocyte marker. $n = 3$; *$P$-value < 0.05, Student's unpaired $t$-test. (*C*) BovB transposable element transcript abundance (counts) in the RNA-seq data from LAA and RAA cardiomyocytes during AF progression. $n = 3$; *$P$-value < 0.05, DEseq default method.

expressed genes (DEGs) for different sample numbers. We observed that the great majority of DEGs were identified in all conditions, for both up-regulated, down-regulated, and all DEGs (Supplementary material online, *Figure S9A*). Therefore, we can safely conclude that in our analysis, despite being small, a sample number of three as we have used for the atrial appendage, is sufficient to draw robust statistically valid conclusions regarding differential expression.

We first compared the changes in gene expression between control and transition states from the PLA and the LAA. We observed a very high correlation for all genes, both for LAA tissue and isolated CMs (*Figure 6A*). Furthermore, many of the critical genes we had identified as differentially expressed in the LAA, such as *RCAN1, LGALS3*, or *PCP4*, were also changing in the PLA (*Figure 6B* and Supplementary material online, *Table S4*). Unsupervised hierarchical clustering resulted in 12 different clusters showing a clear difference in expression between control and transition PLA (*Figure 6C*). Functional enrichment of GO terms (Supplementary material online, *Table S5*) showed how terms related to heart contraction and the CM were up-regulated in transition samples, while those related to blood vessel development down-regulated (*Figure 6C*). Furthermore, changes in expression in the PLA of genes coding for chromatin related factors (Supplementary material online, *Figure S9B*) showed a very similar trend to that we had previously observed in genes differentially expressed in the atrial appendage along disease progression (*Figure 5A*). Overall, this analysis shows that the molecular changes that take place during AF progression are very similar between different anatomical locations of the atria.

## 3.8 Gene expression identifies differences in the rate of AF progression of individual sheep

An unexpected observation from the clustering analysis was that transition samples separated into two subgroups (top burgundy and blue bars, *Figure 6C*), and that a number of the clusters generated showed differences between these subgroups (clusters 1, 7, 9, 11, and 12; boxed in *Figure 6C*). This subgrouping was also evident when we analysed the expression of chromatin factors (Supplementary material online, *Figure S9B*). When searching for a possible explanation of the differences in the transition samples, we observed that these were perfectly matched by the rate of AF progression in these sheep. This was measured as the days taken for an individual sheep from the start of tachypacing to a period of 7 days in AF with no tachypacing (the criterion used to define transition samples). Three sheep took more than 180 days, and were classified as slow progressing, while other three took less than 70 days and were considered fast (Supplementary material online, *Figure S9C* and *Table S1*).

Those clusters showing differences in expression between slow and fast transition sheep were not enriched for any particular GO term. However, a careful inspection of the genes included in them (Supplementary material online, *Table S4*) showed an unsuspected enrichment in genes related to neural cells, such as those coding for the adhesion molecules NCAM1 or NRCAM, the ion channel HCN2, components of signalling pathways such as TGFB3 or CAMK2B, or members of the Slit/Robo axon guidance signalling pathway such as SLIT1 and ROBO2 (Figure 6C). Finally, it was interesting to find that genes belonging to cluster 1, that show a stronger down-regulation in fast progressing sheep compared to controls than slow sheep, include prominent regulators of cell proliferation such as *KIT, NOTCH1, PRMT8*, or *PLK1* (*Figure 6C*). Future studies will be required to explore if reduced proliferation of cells in the PLA leads to a faster progression of AF to the permanent condition.

## 3.9 Molecular changes that occur during disease progression are enriched for AF risk-associated genes

GWAS have highlighted the genetic basis for a predisposition to AF, and inform in an unbiased manner of possible molecular mechanisms underlying the disease.[24,36,37] On the other hand, our study interrogates the molecular changes that occur as a consequence of the disease, and provides information of the molecular mechanisms responsible for its progression. We asked to what extent did these two processes overlap, and if genes that increase susceptibility to AF were also changing during disease progression.

To do so, we first obtained from the public NHGRI-EBI GWAS Catalogue non-redundant lists of genes associated to electrophysiological cardiovascular diseases (CVD), including AF, and electrophysiological traits such as PR or QT interval, and as a control genes associated to myocardial CVD (such as myocardial infarction or heart failure), obtaining 668 and 212 genes respectively (Supplementary material online, *Table S6*). We also used a third GWAS group of the non-redundant genes listed in two recently published independent GWAS meta-analyses which have extended previous AF associations to hundreds of loci[15,16] (240 genes, Supplementary material online, *Table S6*). We then compared these different lists to the set of selected features (4409 differentially expressed genes and proteins) identified in our study of the LAA and RAA, and as a control, the set of all expressed features (11 960). We found that genes associated to electrophysiological CVD and traits, as well as those identified as associated to AF by meta-GWAS, are enriched in the set of features that change during AF progression in our sheep model, while genes associated with myocardial CVD are not (*Figure 7A*). Among these are genes coding for ion channels (such as *KCNJ5* or *SCN5A*), developmental transcription factors (*PITX2, TBX5*), chromatin regulators (such as *KDM3A* or *MBD5*), structural proteins (*MYOCD, MYOT*), and cell-to-cell communication (*GJA1 or IL6R*) proteins (*Figure 7B*). These results show that gene regulatory networks and molecular pathways that are involved in the genetic predisposition to AF are also altered because of disease progression towards an atrial cardiomyopathy due to external factors, including stress and inflammation.[16]

# 4. Discussion

We have taken advantage of a well-established, clinically relevant large animal model to analyse *in vivo* the molecular changes that occur in the atria during progression of AF from paroxysmal to persistent. We demonstrate that the hallmarks of AF-induced atrial remodelling change only at early transitional stages, but not at later stages of the disease, and that the left atrium undergoes significantly more profound changes in its expression programme than the right atrium.

The temporal dynamics of the above changes in gene and protein expression corresponded closely with the electrical and structural remodelling we demonstrated previously in the sheep,[3] and were remarkably similar to those obtained retrospectively in patients undergoing remote transmissions of AF frequency via implantable cardioverter-defibrillator/cardiac resynchronization therapy with defibrillator or pacemaker devices.[9] In both animal and human studies, the duration of the AF episodes and the electrical activation frequency increased progressively during the transition, and the rate was different for each animal or patient, reaching a maximum at the onset of persistent AF. In addition, the activation frequency did not increase after 1 year of self-sustained persistent AF in the

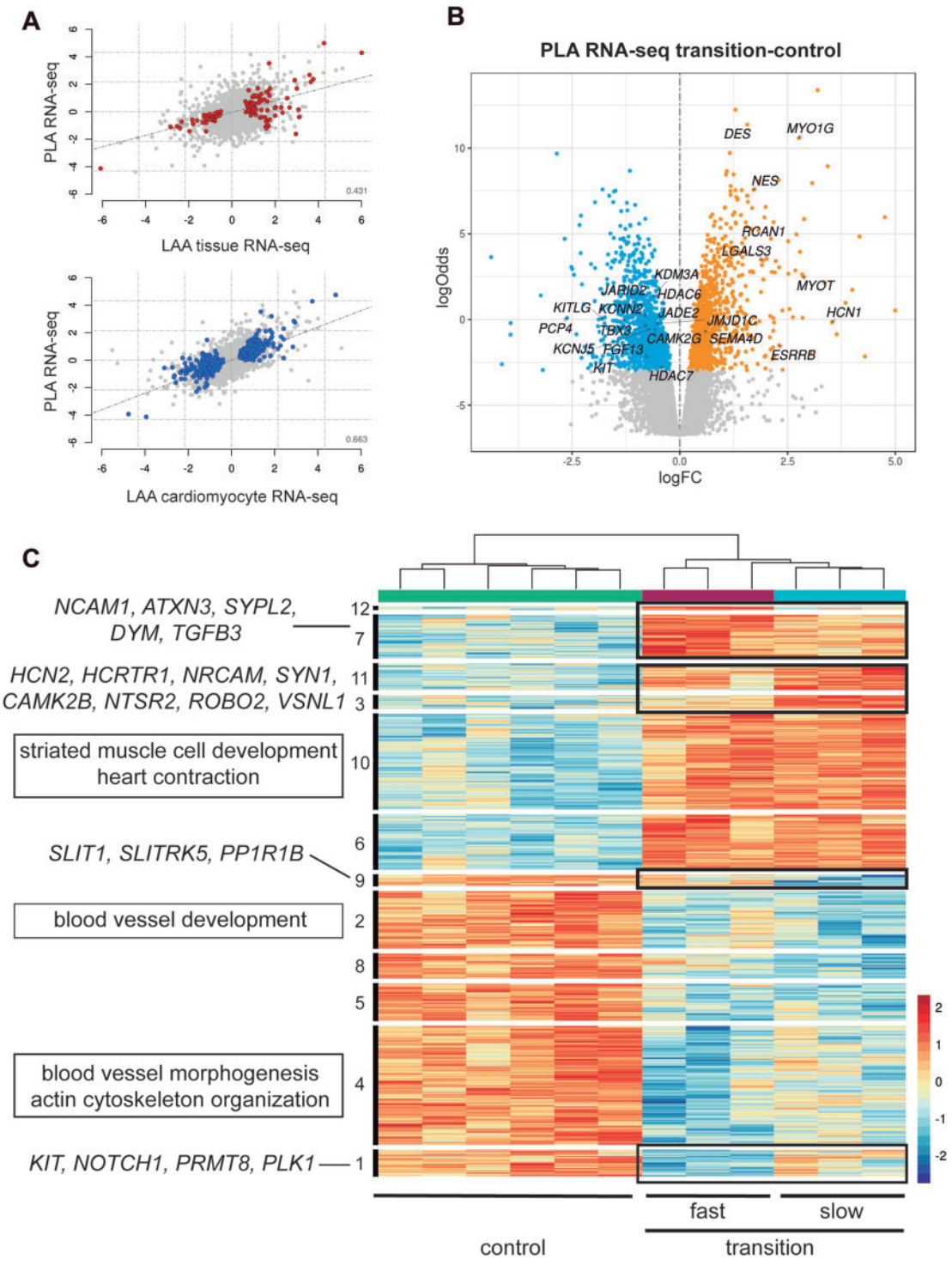


**Figure 6** Transcriptomic profiling of posterior left atria tissue. (*A*) Correlation of the logFC of expression in transition versus control of PLA tissue with LAA tissue (upper panel) and with LAA cardiomyocytes (lower panel). Differentially expressed genes in LAA are shown in red and blue, respectively. Pearson correlation values are indicated in the right bottom corner of each graph. (*B*) Volcano plot of transition vs. control for PLA tissue. Differentially expressed genes at 5% FDR are shown in orange (up-regulated in transition compared to controls) or blue (down-regulated in transition compared to controls). (*C*) Heatmap showing the expression (as z-scores) of the 2185 genes found differentially expressed in the PLA when comparing transition versus control sheep ($n = 6$). Two main branches of the clustering segment the differentially expressed genes into down-regulated and up-regulated for this comparison (transition–control). Various clusters suggest the existence of two different gene expression patterns, for fast and slow sheep to reach persistent AF (indicated as burgundy and blue bars on top of the heatmap, respectively). Genes and GO terms related to individual clusters are indicated on the left.

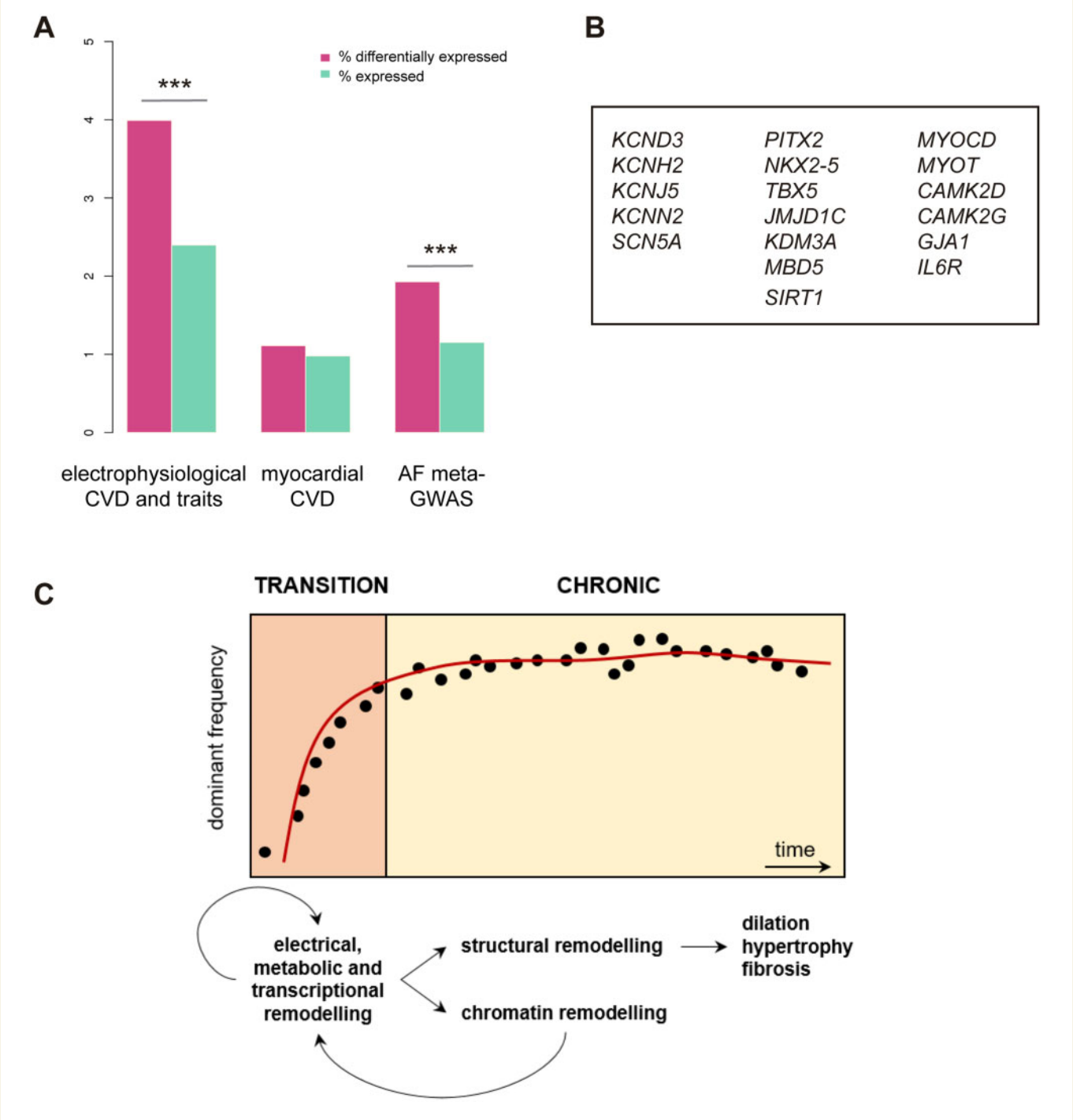


**Figure 7** Overlap of intrinsic genetic determinants and extrinsic genetic changes in AF. (*A*) Graph showing the percentage of differentially expressed (pink) and all expressed (green) features (genes and proteins) in the sheep AF model that are present in the selected list of genes associated by GWAS to electrophysiological CVD and traits, myocardial CVD, and genes associated to AF in two recent meta-analysis loci.[15,16] Differences between both sets was assessed by a hypergeometric test with Benjamini–Hochberg correction for multiple testing. ****P*-value < 1e-04. (*B*) Representative genes included in the overlap between differentially expressed features in the sheep AF model and AF-associated genes, coding for ion channels (left row), developmental transcription factors and chromatin regulators (middle row), and other cellular components (right row). (*C*) Diagram depicting the progression of AF from an early state (transition), where electrical, metabolic, and transcriptional changes take place, to a later long-standing persistent (chronic) state when structural remodelling that leads to dilation and hypertrophy occur as a secondary effect. We also propose that chromatin remodelling is a critical factor in sustaining the disease state.

sheep or 3.4 years in patients. Moreover, at the end of the transition, electrical remodelling in the sheep was accompanied by structural remodelling resulting in atrial dilatation and fibrosis.[3,11]

Different experimental models have been used to understand the pathophysiology of AF.[5,26] But so far, large animal models[8,38,39] have proven to be the most adequate in terms of ease in pacing and similarity to the process of progression in humans.[3,8] Importantly, sheep models have shown their utility to study molecular determinants of disease progression.[10,11] We have performed transcriptomic and proteomic profiling of both atrial tissue and isolated CMs from LAA and RAA of sheep that had been in self-sustained persistent AF for a short period (7 days) and from animals for more than a year in self-sustained persistent AF. We have also analysed the transcriptomic changes driven by AF in the PLA, finding them very similar to those in the LAA. By comparing these two time points of disease progression to control animals, we were able to explore the networks and pathways underlying AF.

The initial global analysis of the data provided an unexpected finding. The changes in gene or protein expression during AF occurred early, during the transition from paroxysmal to persistent, but thereafter were unchanged for up to one year despite AF persistence. This occurs both in whole atrial appendage tissue and isolated CMs, and for either gene or protein expression. Furthermore, when analysing the main drivers of the variability underlying our data, apart from the expected components related to disease progression and left/right identity, we found a third driver that we related to a 'transition state'. We hypothesize the transition state reflects the early changes taking place in the initial period of AF, which in the sheep model are reflected by a rapid increase in dominant frequency (*Figure 7C*). Together, these results suggest that the underlying molecular mechanisms distinguish between paroxysmal and persistent but not between persistent and permanent forms of AF. This outcome makes it necessary to conduct a more detailed analysis of the early stages of disease progression. Another interesting finding from our global analysis is that changes in the LAA are more pronounced than in the RAA, in line with the view that AF is a left atria disease,[40] with subsequent changes occurring in the right atrium.

The functional annotation of the data, and the comparison of whole tissue and CM-specific RNA-seq, allowed us to identify processes and pathways that have been previously associated with AF. For example, fibrosis, inflammation and changes in ion channels have been previously described as part of the mechanisms responsible for perpetuation of AF,[41] and contractile dysfunction in AF (referred to as atrial cardiomyopathy) is emerging as an important contributor to the disease.[42,43] Equally, we observed changes in genes related to calcium-handling, which have a critical role in AF.[44,45] The calcium ion-channel subunit encoding gene *CACNA1C* shows reduced expression in CMs during AF progression, which correlates well with functional studies using the same model.[3,11] On the other hand, *RCAN1*, involved in calcineurin signalling, is one of the most up-regulated genes in our analysis, while *PCP4*, which exerts an opposite effect in this pathway,[25] is strongly down-regulated. It has been shown that rapid atrial activity results in $Ca^{2+}$ loading, that in turn triggers the $Ca^{2+}$-dependent calmodulin–calcineurin–NFAT pathway to cause the down-regulation of $I_{CaL}$ and action potential duration reduction in the atrial cardiomyocytes.[46] RCAN1 has also been implicated in TRPC1/C4 channel-mediated activation of the calcineurin-NFAT pathway,[47] and identified as top hub gene in human AF samples.[48] We also observed an increase in mitochondrial mass in our model that coincides with the observation of higher protein levels of respiratory complexes in patients with persistent AF.[49,50]

Our analysis has revealed changes to chromatin as a major consequence of AF in the sheep. Global down-regulation of chromatin factors, together with a drop in core histone levels and an increase in the expression of TEs, suggestive of chromatin decompaction, all occur during AF progression. Several of these epigenomic-related changes have been associated with the decline that occurs during ageing.[51] Reduction of histones can lead to loss of heterochromatin with derepression of TEs, which would normally be silenced.[52] Hypomethylation of DNA also accompanies chromatin decompaction and has been observed in other CVD such as atherosclerosis,[53,54] and overall epigenomic changes have been argued to be a cause for the progression of common human diseases.[55,56] It is interesting that the analysis of subcellular structures in a goat model of AF revealed dispersed heterochromatin in the nucleus of AF CMs, compared to clustered aggregates found in CMs from sinus rhythm animals, as one of the earliest changes.[38] Therefore, our results suggests that a general decrease in nuclear organization is a hallmark of AF, and could be explored as an early marker of the disease in humans.[57] Furthermore, chromatin remodelling would lead to altered epigenomic states that reinforce the disease-related gene expression programme (*Figure 7C*).

When analysing the transcriptomic data from the PLA, we were able to find differences in gene expression that correlate with sheep atria that transition to persistent AF at different rates. Interestingly, among these we identified many genes related to neural cells, which together with other observations on the role of the neural system in AF,[58,59] suggests a neural input on how quickly AF progresses. We also observed that fast progressing sheep had lower expression of proliferation-related genes. These observations open up novel avenues that could help to identify and treat those patients that will progress more rapidly to permanent, and therefore more adverse, forms of AF.

There is substantial evidence in the literature indicating that AF is heritable. Classical genetics have documented several chromosomal loci and genetic mutations in myocardial sodium, potassium, and potassium-adenosine triphosphate channels.[60,61] In our study, a number of those genes (such as *NPPA, MYL4, PRKAG2, LMNA, SCN5A*, and *KCNH2*) are differentially expressed in the sheep atria during AF progression in our study. In fact, the increase in AF frequency during the progression to persistent AF is a reflection of electrical remodelling in the form of action potential duration abbreviation brought about by differential changes to ion channels, such as decreases in sodium and L-type calcium currents or increase in inward rectifier potassium current.[3] On the other hand, more than 100 new genetic loci have been associated to an increased risk of AF by GWAS, pointing to yet unexplored mechanisms underlying the disease.[62] The analysis of the variants underlying these associations will reveal how they underpin specific atrial substrates or conditions that can modify molecular functions leading to progression of AF.[63] This source of variation might help explain the large variability in the rate of AF remodelling and progression that has been observed in animals and in patients. Nevertheless, research efforts taking together personal genetic profiles, clinical risk factors and monitoring of AF progression[64] along with atrial cell remodelling[63] should help to better understand risk and progression of AF. It may also help to predict the rate of AF progression and the time to completion of atrial electrical remodelling, thus improving stratification and the personalized care of AF patients.

# Supplementary material

Supplementary material is available at *Cardiovascular Research* online.

## Authors' contributions

A.A.-F, R.R., J.J., and M.M. designed research; A.A.-F., R.R., R.J.R., G.G.-S., M.T., S.C., K.K., and M.S. performed research; A.A.-F., R.R., R.M., J.A.E., F.S.-C., J.J., and M.M. analysed data; and A.A.-F., R.R., J.J., and M.M. wrote the paper.

## Acknowledgements

We thank David Filgueiras-Rama (CNIC) for helpful discussions; Luciano di Croce (CRG), and Maria Gomez and Jose Miguel Fernandez-Justel (CBMSO), for discussions and advice on chromatin regulation; Jesus Vazquez and the CNIC Proteomics, Genomics, and Bioinformatics Units for expert assistance; and members of the Manzanares' lab for continued support and encouragement.

**Conflict of interest:** none declared.

## Funding

This work was supported by the Spanish government (BFU2017-84914-P to M.M.; FPI Fellowship to A.A.-F.; FPU Fellowship to R.R.), and in part by grants to J.J. from the National Heart, Lung and Blood Institute (R01 grant HL122352 NIH/NHLBI), the Leducq Foundation (Transatlantic Network of Excellence Program on Structural Alterations in the Myocardium and the Substrate for Cardiac Fibrillation), and the University of Michigan Health System–Peking University Health Science Center Joint Institute for Translational and Clinical Research (UMHS-PUHSC; project: Molecular Mechanisms of Fibrosis and the Progression from Paroxysmal to Persistent Atrial Fibrillation). The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), the Ministerio de Ciencia e Innovación and the Pro CNIC Foundation and is a Severo Ochoa Center of Excellence (SEV-2015-0505).

## Data availability

The data underlying this article are available in Gene Expression Omnibus (GEO) Database at https://www.ncbi.nlm.nih.gov/geo/ (last accesed 7 January 2021) and can be accessed with accession number GSE138255, and at the ProteomeXchange Consortium at http://www.proteomexchange.org (last accesed 7 January 2021) and can be accessed with accession number PXD015637. An interactive Shiny app to browse and analyse the full data is available at https://bioinfo.cnic.es/AFibOmics (last accesed 7 January 2021).

## References

1. Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, Gillum RF, Kim YH, McAnulty JH Jr, Zheng ZJ, Forouzanfar MH, Naghavi M, Mensah GA, Ezzati M, Murray CJ. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* 2014;**129**:837–847.
2. Krijthe BP, Kunst A, Benjamin EJ, Lip GY, Franco OH, Hofman A, Witteman JC, Stricker BH, Heeringa J. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J* 2013;**34**:2746–2751.
3. Martins RP, Kaur K, Hwang E, Ramirez RJ, Willis BC, Filgueiras-Rama D, Ennis SR, Takemoto Y, Ponce-Balbuena D, Zarzoso M, O'Connell RP, Musa H, Guerrero-Serna G, Avula UMR, Swartz MF, Bhushal S, Deo M, Pandit SV, Berenfeld O, Jalife J. Dominant frequency increase rate predicts transition from paroxysmal to long-term persistent atrial fibrillation. *Circulation* 2014;**129**:1472–1482.
4. de Vos CB, Pisters R, Nieuwlaat R, Prins MH, Tieleman RG, Coelen RJ, van den Heijkant AC, Allessie MA, Crijns HJ. Progression from paroxysmal to persistent atrial fibrillation clinical correlates and prognosis. *J Am Coll Cardiol* 2010;**55**:725–731.
5. Nishida K, Michael G, Dobrev D, Nattel S. Animal models for atrial fibrillation: clinical insights and scientific opportunities. *Europace* 2010;**12**:160–172.
6. Nattel S, Harada M. Atrial remodeling and atrial fibrillation: recent advances and translational perspectives. *J Am Coll Cardiol* 2014;**63**:2335–2345.
7. Schotten U, Verheule S, Kirchhof P, Goette A. Pathophysiological mechanisms of atrial fibrillation: a translational appraisal. *Physiol Rev* 2011;**91**:265–325.
8. Filgueiras-Rama D, Price NF, Martins RP, Yamazaki M, Avula UM, Kaur K, Kalifa J, Ennis SR, Hwang E, Devabhaktuni V, Jalife J, Berenfeld O. Long-term frequency gradients during persistent atrial fibrillation in sheep are associated with stable sources in the left atrium. *Circ Arrhythm Electrophysiol* 2012;**5**:1160–1167.
9. Lillo-Castellano JM, González-Ferrer JJ, Marina-Breysse M, Martínez-Ferrer JB, Pérez-Álvarez L, Alzueta J, Martínez JG, Rodríguez A, Rodríguez-Pérez JC, Anguera I, Viñolas X, García-Alberola A, Quintanilla JG, Alfonso-Almazán JM, García J, Borrego L, Cañadas-Godoy V, Pérez-Castellano N, Pérez-Villacastín J, Jiménez-Díaz J, Jalife J, Filgueiras-Rama D. Personalized monitoring of electrical remodelling during atrial fibrillation progression via remote transmissions from implantable devices. *Europace* 2020;**22**:704–715.
10. Takemoto Y, Ramirez RJ, Kaur K, Salvador-Montañés O, Ponce-Balbuena D, Ramos-Mondragón R, Ennis SR, Guerrero-Serna G, Berenfeld O, Jalife J. Eplerenone reduces atrial fibrillation burden without preventing atrial electrical remodeling. *J Am Coll Cardiol* 2017;**70**:2893–2905.
11. Takemoto Y, Ramirez RJ, Yokokawa M, Kaur K, Ponce-Balbuena D, Sinno MC, Willis BC, Ghanbari H, Ennis SR, Guerrero-Serna G, Henzi BC, Latchamsetty R, Ramos-Mondragon R, Musa H, Martins RP, Pandit SV, Noujaim SF, Crawford T, Jongnarangsin K, Pelosi F, Bogun F, Chugh A, Berenfeld O, Morady F, Oral H, Jalife J. Galectin-3 regulates atrial fibrillation remodeling and predicts catheter ablation outcomes. *JACC Basic Transl Sci* 2016;**1**:143–154.
12. Meng C, Kuster B, Culhane AC, Gholami AM. A multivariate approach to the integration of multi-omics datasets. *BMC Bioinformatics* 2014;**15**:162.
13. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;**25**:25–29.
14. Singh Nanda J, Kumar R, Raghava GP. dbEM: a database of epigenetic modifiers curated from cancerous and normal genomes. *Sci Rep* 2016;**6**:19340.
15. Roselli C, Chaffin MD, Weng L-C, Aeschbacher S, Ahlberg G, Albert CM, Almgren P, Alonso A, Anderson CD, Aragam KG, Arking DE, Barnard J, Bartz TM, Benjamin EJ, Bihlmeyer NA, Bis JC, Bloom HL, Boerwinkle E, Bottinger EB, Brody JA, Calkins H, Campbell A, Cappola TP, Carlquist J, Chasman DI, Chen LY, Chen Y-DI, Choi E-K, Choi SH, Christophersen IE, Chung MK, Cole JW, Conen D, Cook J, Crijns HJ, Cutler MJ, Damrauer SM, Daniels BR, Darbar D, Delgado G, Denny JC, Dichgans M, Dörr M, Dudink EA, Dudley SC, Esa N, Esko T, Eskola M, Fatkin D, Felix SB, Ford I, Franco OH, Geelhoed B, Grewal RP, Gudnason V, Guo X, Gupta N, Gustafsson S, Gutmann R, Hamsten A, Harris TB, Hayward C, Heckbert SR, Hernesniemi J, Hocking LJ, Hofman A, Horimoto ARVR, Huang J, Huang PL, Huffman J, Ingelsson E, Ipek EG, Ito K, Jimenez-Conde J, Johnson R, Jukema JW, Kääb S, Kähönen M, Kamatani Y, Kane JP, Kastrati A, Kathiresan S, Katschnig-Winter P, Kavousi M, Kessler T, Kietselaer BL, Kirchhof P, Kleber ME, Knight S, Krieger JE, Kubo M, Launer LJ, Laurikka J, Lehtimäki T, Leineweber K, Lemaitre RN, Li M, Lim HE, Lin HJ, Lin H, Lind L, Lindgren CM, Lokki M-L, London B, Loos RJF, Low S-K, Lu Y, Lyytikäinen L-P, Macfarlane PW, Magnusson PK, Mahajan A, Malik R, Mansur AJ, Marcus GM, Margolin L, Margulies KB, März W, McManus DD, Melander O, Mohanty S, Montgomery JA, Morley MP, Morris AP, Müller-Nurasyid M, Natale A, Nazarian S, Neumann B, Newton-Cheh C, Niemeijer MN, Nikus K, Nilsson P, Noordam R, Oellers H, Olesen MS, Orho-Melander M, Padmanabhan S, Pak H-N, Paré G, Pedersen NL, Pera J, Pereira A, Porteous D, Psaty BM, Pulit SL, Pullinger CR, Rader DJ, Refsgaard L, Ribasés M, Ridker PM, Rienstra M, Risch L, Roden DM, Rosand J, Rosenberg MA, Rost N, Rotter JI, Saba S, Sandhu RK, Schnabel RB, Schramm K, Schunkert H, Schurman C, Scott SA, Seppälä I, Shaffer C, Shah S, Shalaby AA, Shim J, Shoemaker MB, Siland JE, Sinisalo J, Sinner MF, Slowik A, Smith AV, Smith BH, Smith JG, Smith JD, Smith NL, Soliman EZ, Sotoodehnia N, Stricker BH, Sun A, Sun H, Svendsen JH, Tanaka T, Tanriverdi K, Taylor KD, Teder-Laving M, Teumer A, Thériault S, Trompet S, Tucker NR, Tveit A, Uitterlinden AG, Van Der Harst P, Van Gelder IC, Van Wagoner DR, Verweij N, Vlachopoulou E, Völker U, Wang B, Weeke PE, Weijs B, Weiss R, Weiss S, Wells QS, Wiggins KL, Wong JA, Woo D, Worrall BB, Yang P-S, Yao J, Yoneda ZT, Zeller T, Zeng L, Lubitz SA, Lunetta KL, Ellinor PT. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* 2018;**50**:1225–1233.
16. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, Herron TJ, McCarthy S, Schmidt EM, Sveinbjornsson G, Surakka I, Mathis MR, Yamazaki M, Crawford RD, Gabrielsen ME, Skogholt AH, Holmen OL, Lin M, Wolford BN, Dey R, Dalen H, Sulem P, Chung JH, Backman JD, Arnar DO, Thorsteinsdottir U, Baras A, O'Dushlaine C, Holst AG, Wen X, Hornsby W, Dewey FE, Boehnke M, Kheterpal S, Mukherjee B, Lee S, Kang HM, Holm H, Kitzman J, Shavit JA, Jalife J, Brummett CM, Teslovich TM, Carey DJ, Gudbjartsson DF, Stefansson K, Abecasis GR, Hveem K, Willer CJ. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat Genet* 2018;**50**:1234–1239.
17. Wittig I, Braun HP, Schagger H. Blue native PAGE. *Nat Protoc* 2006;**1**:418–428.
18. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–140.

19. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;**43**:e47.
20. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. *Nature* 2011;**473**: 337–342.
21. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* 2012;**13**:227–232.
22. Meng C, Zeleznik OA, Thallinger GG, Kuster B, Gholami AM, Culhane AC. Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief Bioinform* 2016;**17**:628–641.
23. Meng C, Culhane A. Integrative exploratory analysis of two or more genomic datasets. *Methods Mol Biol* 2016;**1418**:19–38.
24. Gudbjartsson DF, Arnar DO, Helgadottir A, Gretarsdottir S, Holm H, Sigurdsson A, Jonasdottir A, Baker A, Thorleifsson G, Kristjansson K, Palsson A, Blondal T, Sulem P, Backman VM, Hardarson GA, Palsdottir E, Helgason A, Sigurjonsdottir R, Sverrisson JT, Kostulas K, Ng MCY, Baum L, So WY, Wong KS, Chan JCN, Furie KL, Greenberg SM, Sale M, Kelly P, MacRae CA, Smith EE, Rosand J, Hillert J, Ma RCW, Ellinor PT, Thorgeirsson G, Gulcher JR, Kong A, Thorsteinsdottir U, Stefansson K. Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* 2007;**448**:353–357.
25. Kim EE, Shekhar A, Lu J, Lin X, Liu FY, Zhang J, Delmar M, Fishman GI. PCP4 regulates Purkinje cell excitability and cardiac rhythmicity. *J Clin Invest* 2014;**124**: 5027–5036.
26. Allessie M, Ausma J, Schotten U. Electrical, contractile and structural remodeling during atrial fibrillation. *Cardiovasc Res* 2002;**54**:230–246.
27. Choi SH, Weng L-C, Roselli C, Lin H, Haggerty CM, Shoemaker MB, Barnard J, Arking DE, Chasman DI, Albert CM, Chaffin M, Tucker NR, Smith JD, Gupta N, Gabriel S, Margolin L, Shea MA, Shaffer CM, Yoneda ZT, Boerwinkle E, Smith NL, Silverman EK, Redline S, Vasan RS, Burchard EG, Gogarten SM, Laurie C, Blackwell TW, Abecasis G, Carey DJ, Fornwalt BK, Smelser DT, Baras A, Dewey FE, Jaquish CE, Papanicolaou GJ, Sotoodehnia N, Van Wagoner DR, Psaty BM, Kathiresan S, Darbar D, Alonso A, Heckbert SR, Chung MK, Roden DM, Benjamin EJ, Murray MF, Lunetta KL, Lubitz SA, Ellinor PT; For the DiscovEHR study and the NHLBI TransOmics for Precision Medicine (TOPMed) Consortium. Association between titin loss-of-function variants and early-onset atrial fibrillation. *JAMA* 2018;**320**:2354–2364.
28. Nielsen JB, Fritsche LG, Zhou W, Teslovich TM, Holmen OL, Gustafsson S, Gabrielsen ME, Schmidt EM, Beaumont R, Wolford BN, Lin M, Brummett CM, Preuss MH, Refsgaard L, Bottinger EP, Graham SE, Surakka I, Chu Y, Skogholt AH, Dalen H, Boyle AP, Oral H, Herron TJ, Kitzman J, Jalife J, Svendsen JH, Olesen MS, Njølstad I, Løchen M-L, Baras A, Gottesman O, Marcketta A, O'Dushlaine C, Ritchie MD, Wilsgaard T, Loos RJF, Frayling TM, Boehnke M, Ingelsson E, Carey DJ, Dewey FE, Kang HM, Abecasis GR, Hveem K, Willer CJ. Genome-wide study of atrial fibrillation identifies seven risk loci and highlights biological pathways and regulatory elements involved in cardiac development. *Am J Hum Genet* 2018;**102**:103–115.
29. Al-Mashhadi RH, Sorensen CB, Kragh PM, Christoffersen C, Mortensen MB, Tolbod LP, Thim T, Du Y, Li J, Liu Y, Moldt B, Schmidt M, Vajta G, Larsen T, Purup S, Bolund L, Nielsen LB, Callesen H, Falk E, Mikkelsen JG, Bentzon JF. Familial hypercholesterolemia and atherosclerosis in cloned minipigs created by DNA transposition of a human PCSK9 gain-of-function mutant. *Sci Transl Med* 2013;**5**:166ra1.
30. Cañón S, Caballero R, Herraiz-Martínez A, Pérez-Hernández M, López B, Atienza F, Jalife J, Hove-Madsen L, Delpón E, Bernad A. miR-208b upregulation interferes with calcium handling in HL-1 atrial myocytes: implications in human chronic atrial fibrillation. *J Mol Cell Cardiol* 2016;**99**:162–173.
31. Cogliati S, Lorenzi I, Rigoni G, Caicci F, Soriano ME. Regulation of mitochondrial electron transport chain assembly. *J Mol Biol* 2018;**430**:4849–4873.
32. Cogliati S, Calvo E, Loureiro M, Guaras AM, Nieto-Arellano R, Garcia-Poyatos C, Ezkurdia I, Mercader N, Vázquez J, Enriquez JA. Mechanism of super-assembly of respiratory complexes III and IV. *Nature* 2016;**539**:579–582.
33. Rodriguez-Terrones D, Torres-Padilla ME. Nimble and ready to mingle: transposon outbursts of early development. *Trends Genet* 2018;**34**:806–820.
34. Wood JG, Helfand SL. Chromatin structure and transposable elements in organismal aging. *Front Genet* 2013;**4**:274.
35. Walsh AM, Kortschak RD, Gardner MG, Bertozzi T, Adelson DL. Widespread horizontal transfer of retrotransposons. *Proc Natl Acad Sci USA* 2013;**110**:1012–1016.
36. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J, Lin H, Arking DE, Smith AV, Albert CM, Chaffin M, Tucker NR, Li M, Klarin D, Bihlmeyer NA, Low S-K, Weeke PE, Müller-Nurasyid M, Smith JG, Brody JA, Niemeijer MN, Dörr M, Trompet S, Huffman J, Gustafsson S, Schurmann C, Kleber ME, Lyytikäinen L-P, Seppälä I, Malik R, Horimoto ARVR, Perez M, Sinisalo J, Aeschbacher S, Thériault S, Yao J, Radmanesh F, Weiss S, Teumer A, Choi SH, Weng L-C, Clauss S, Deo R, Rader DJ, Shah SH, Sun A, Hopewell JC, Debette S, Chauhan G, Yang Q, Worrall BB, Paré G, Kamatani Y, Hagemeijer YP, Verweij N, Siland JE, Kubo M, Smith JD, Van Wagoner DR, Bis JC, Perz S, Psaty BM, Ridker PM, Magnani JW, Harris TB, Launer LJ, Shoemaker MB, Padmanabhan S, Haessler J, Bartz TM, Waldenberger M, Lichtner P, Arendt M, Krieger JE, Kähönen M, Risch L, Mansur AJ, Peters A, Smith BH, Lind L, Scott SA, Lu Y, Bottinger EB, Hernesniemi J, Lindgren CM, Wong JA, Huang J, Eskola M, Morris AP, Ford I, Reiner AP, Delgado G, Chen LY, Chen Y-DI, Sandhu RK, Li M, Boerwinkle E, Eisele L, Lannfelt L, Rost N, Anderson CD, Taylor KD, Campbell A, Magnusson PK, Porteous D, Hocking LJ, Vlachopoulou E, Pedersen NL, Nikus K, Orho-Melander M, Hamsten A, Heeringa J, Denny JC, Kriebel J, Darbar D, Newton-Cheh C, Shaffer C, Macfarlane PW, Heilmann-Heimbach S, Almgren P, Huang PL, Sotoodehnia N, Soliman EZ, Uitterlinden AG, Hofman A, Franco OH, Völker U, Jöckel K-H, Sinner MF, Lin HJ, Guo X, Dichgans M, Ingelsson E, Kooperberg C, Melander O, Loos RJF, Laurikka J, Conen D, Rosand J, van der Harst P, Lokki M-L, Kathiresan S, Pereira A, Jukema JW, Hayward C, Rotter JI, März W, Lehtimäki T, Stricker BH, Chung MK, Felix SB, Gudnason V, Alonso A, Roden DM, Kääb S, Chasman DI, Heckbert SR, Benjamin EJ, Tanaka T, Lunetta KL, Lubitz SA, Ellinor PT; AFGen Consortium. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nat Genet* 2017;**49**:946–952.
37. Ellinor PT, Lunetta KL, Albert CM, Glazer NL, Ritchie MD, Smith AV, Arking DE, Müller-Nurasyid M, Krijthe BP, Lubitz SA, Bis JC, Chung MK, Dörr M, Ozaki K, Roberts JD, Smith JG, Pfeufer A, Sinner MF, Lohman K, Ding J, Smith NL, Smith JD, Rienstra M, Rice KM, Van Wagoner DR, Magnani JW, Wakili R, Clauss S, Rotter JI, Steinbeck G, Launer LJ, Davies RW, Borkovich M, Harris TB, Lin H, Völker U, Völzke H, Milan DJ, Hofman A, Boerwinkle E, Chen LY, Soliman EZ, Voight BF, Li G, Chakravarti A, Kubo M, Tedrow UB, Rose LM, Ridker PM, Conen D, Tsunoda T, Furukawa T, Sotoodehnia N, Xu S, Kamatani N, Levy D, Nakamura Y, Parvez B, Mahida S, Furie KL, Rosand J, Muhammad R, Psaty BM, Meitinger T, Perz S, Wichmann H-E, Witteman JCM, Kao WHL, Kathiresan S, Roden DM, Uitterlinden AG, Rivadeneira F, McKnight B, Sjögren M, Newman AB, Liu Y, Gollob MH, Melander O, Tanaka T, Stricker BHC, Felix SB, Alonso A, Darbar D, Barnard J, Chasman DI, Heckbert SR, Benjamin EJ, Gudnason V, Kääb S. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat Genet* 2012;**44**:670–675.
38. Ausma J, Wijffels M, Thone F, Wouters L, Allessie M, Borgers M. Structural changes of atrial myocardium due to sustained atrial fibrillation in the goat. *Circulation* 1997; **96**:3157–3163.
39. Schuessler RB, Grayson TM, Bromberg BI, Cox JL, Boineau JP. Cholinergically mediated tachyarrhythmias induced by a single extrastimulus in the isolated canine right atrium. *Circ Res* 1992;**71**:1254–1267.
40. Sardana M, Lessard D, Tsao CW, Parikh NI, Barton BA, Nah G, Thomas RC, Cheng S, Schiller NB, Aragam JR, Mitchell GF, Vaze A, Benjamin EJ, Vasan RS, McManus DD. Association of left atrial function index with atrial fibrillation and cardiovascular disease: the Framingham Offspring Study. *J Am Heart Assoc* 2018;**7**:e008435.
41. Jalife J, Kaur K. Atrial remodeling, fibrosis, and atrial fibrillation. *Trends Cardiovasc Med* 2015;**25**:475–484.
42. Shen MJ, Arora R, Jalife J. Atrial myopathy. *JACC Basic Transl Sci* 2019;**4**:640–654.
43. Guichard JB, Nattel S. Atrial cardiomyopathy: a useful notion in cardiac disease management or a passing fad? *J Am Coll Cardiol* 2017;**70**:756–765.
44. Nattel S, Dobrev D. The multidimensional role of calcium in atrial fibrillation pathophysiology: mechanistic insights and therapeutic opportunities. *Eur Heart J* 2012;**33**: 1870–1877.
45. Heijman J, Voigt N, Nattel S, Dobrev D. Cellular and molecular electrophysiology of atrial fibrillation initiation, maintenance, and progression. *Circ Res* 2014;**114**: 1483–1499.
46. Qi XY, Yeh YH, Xiao L, Burstein B, Maguy A, Chartier D, Villeneuve LR, Brundel BJ, Dobrev D, Nattel S. Cellular signaling underlying atrial tachycardia remodeling of L-type calcium current. *Circ Res* 2008;**103**:845–854.
47. Camacho Londono JE, Tian Q, Hammer K, Schroder L, Camacho Londono J, Reil JC, He T, Oberhofer M, Mannebach S, Mathar I, Philipp SE, Tabellion W, Schweda F, Dietrich A, Kaestner L, Laufs U, Birnbaumer L, Flockerzi V, Freichel M, Lipp P. A background Ca2+ entry pathway mediated by TRPC1/TRPC4 is critical for development of pathological cardiac remodelling. *Eur Heart J* 2015;**36**:2257–2266.
48. Tan N, Chung MK, Smith JD, Hsu J, Serre D, Newton DW, Castel L, Soltesz E, Pettersson G, Gillinov AM, Van Wagoner DR, Barnard J. Weighted gene coexpression network analysis of human left atrial tissue identifies gene modules associated with atrial fibrillation. *Circ Cardiovasc Genet* 2013;**6**:362–371.
49. Reilly SN, Jayaram R, Nahar K, Antoniades C, Verheule S, Channon KM, Alp NJ, Schotten U, Casadei B. Atrial sources of reactive oxygen species vary with the duration and substrate of atrial fibrillation: implications for the antiarrhythmic effect of statins. *Circulation* 2011;**124**:1107–1117.
50. Doll S, Dreßen M, Geyer PE, Itzhak DN, Braun C, Doppler SA, Meier F, Deutsch M-A, Lahm H, Lange R, Krane M, Mann M. Region and cell-type resolved quantitative proteomic map of the human heart. *Nat Commun* 2017;**8**:1469.
51. Booth LN, Brunet A. The aging epigenome. *Mol Cell* 2016;**62**:728–744.
52. Song S, Johnson FB. Epigenetic mechanisms impacting aging: a focus on histone levels and telomeres. *Genes* 2018;**9**:201.
53. Zhang W, Song M, Qu J, Liu GH. Epigenetic modifications in cardiovascular aging and diseases. *Circ Res* 2018;**123**:773–786.
54. Baccarelli A, Rienstra M, Benjamin EJ. Cardiovascular epigenetics: basic concepts and results from animal and human studies. *Circ Cardiovasc Genet* 2010;**3**:567–573.
55. Bjornsson HT, Fallin MD, Feinberg AP. An integrated epigenetic and genetic approach to common human disease. *Trends Genet* 2004;**20**:350–358.
56. Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature* 2007;**447**:433–440.
57. Rosa-Garrido M, Chapski DJ, Vondriska TM. Epigenomes in cardiovascular disease. *Circ Res* 2018;**122**:1586–1607.

58. Chen PS, Chen LS, Fishbein MC, Lin SF, Nattel S. Role of the autonomic nervous system in atrial fibrillation: pathophysiology and therapy. *Circ Res* 2014;**114**: 1500–1515.
59. Gussak G, Pfenniger A, Wren L, Gilani M, Zhang W, Yoo S, Johnson DA, Burrell A, Benefield B, Knight G, Knight BP, Passman R, Goldberger JJ, Aistrup G, Wasserstrom JA, Shiferaw Y, Arora R. Region-specific parasympathetic nerve remodeling in the left atrium contributes to creation of a vulnerable substrate for atrial fibrillation. *JCI Insight* 2019;**4**:e130532.
60. Lubitz SA, Yin X, Fontes JD, Magnani JW, Rienstra M, Pai M, Villalon ML, Vasan RS, Pencina MJ, Levy D, Larson MG, Ellinor PT, Benjamin EJ. Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. *JAMA* 2010;**304**: 2263–2269.
61. Judge DP. The complex genetics of atrial fibrillation. *J Am Coll Cardiol* 2012;**60**: 1182–1184.
62. Bapat A, Anderson CD, Ellinor PT, Lubitz SA. Genomic basis of atrial fibrillation. *Heart* 2018;**104**:201–206.
63. Mann SA, Otway R, Guo G, Soka M, Karlsdotter L, Trivedi G, Ohanian M, Zodgekar P, Smith RA, Wouters MA, Subbiah R, Walker B, Kuchar D, Sanders P, Griffiths L, Vandenberg JI, Fatkin D. Epistatic effects of potassium channel variation on cardiac repolarization and atrial fibrillation risk. *J Am Coll Cardiol* 2012;**59**:1017–1025.
64. Weng LC, Preis SR, Hulme OL, Larson MG, Choi SH, Wang B, Trinquart L, McManus DD, Staerk L, Lin H, Lunetta KL, Ellinor PT, Benjamin EJ, Lubitz SA. Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. *Circulation* 2018;**137**:1027–1038.

## Translational perspective

We have generated a detailed molecular map of atrial fibrillation (AF) progression in a clinically relevant large-animal model. Such data would be very difficult if not impossible to obtain from patients. Our results provide a framework for a comprehensive molecular analysis of the disease, pointing to novel avenues of research towards identifying early events that can lead to therapeutically targets to prevent AF-induced atrial remodelling.