



OPEN ACCESS

EDITED BY

Shizuka Uchida,
Aalborg University Copenhagen,
Denmark

REVIEWED BY

Matt Stratton,
The Ohio State University,
United States
Thomas Hund,
The Ohio State University,
United States

*CORRESPONDENCE

Emilia Entcheva
entcheva@gwu.edu

SPECIALTY SECTION

This article was submitted to
Sex and Gender in Cardiovascular
Medicine,
a section of the journal
Frontiers in Cardiovascular Medicine

RECEIVED 11 May 2022

ACCEPTED 29 June 2022

PUBLISHED 22 July 2022

CITATION

Pressler MP, Horvath A and Entcheva E
(2022) Sex-dependent transcription
of cardiac electrophysiology and links
to acetylation modifiers based on
the GTEx database.
Front. Cardiovasc. Med. 9:941890.
doi: 10.3389/fcvm.2022.941890

COPYRIGHT

© 2022 Pressler, Horvath and
Entcheva. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Sex-dependent transcription of cardiac electrophysiology and links to acetylation modifiers based on the GTEx database

Michael P. Pressler¹, Anelia Horvath² and Emilia Entcheva^{1*}

¹Department of Biomedical Engineering, George Washington University, Washington, DC, United States, ²Department of Biochemistry and Molecular Medicine, McCormick Genomics and Proteomics Center, School of Medicine and Health Sciences, The George Washington University, Washington, DC, United States

Development of safer drugs based on epigenetic modifiers, e.g., histone deacetylase inhibitors (HDACi), requires better understanding of their effects on cardiac electrophysiology. Using RNAseq data from the genotype-tissue-expression database (GTEx), we created models that link the abundance of acetylation enzymes (HDAC/SIRT/HATs), and the gene expression of ion channels (IC) via select cardiac transcription factors (TFs) in male and female adult human hearts (left ventricle, LV). Gene expression data (transcripts per million, TPM) from GTEx donors (21–70 y.o.) were filtered, normalized and transformed to Euclidian space to allow quantitative comparisons in 84 female and 158 male LVs. Sex-specific partial least-square (PLS) regression models, linking gene expression data for HDAC/SIRT/HATs to TFs and to ICs gene expression, revealed tight co-regulation of cardiac ion channels by HDAC/SIRT/HATs, with stronger clustering in the male LV. Co-regulation of genes encoding excitatory and inhibitory processes in cardiac tissue by the acetylation modifiers may help explain their predominantly net-neutral effects on cardiac electrophysiology. *ATP1A1*, encoding for the Na/K pump, represented an outlier—with orthogonal regulation by the acetylation modifiers to most of the ICs. The HDAC/SIRT/HAT effects were mediated by strong (+) TF regulators of ICs, e.g., *MEF2A* and *TBX5*, in both sexes. Furthermore, for male hearts, PLS models revealed a stronger (+/-) mediatory role on ICs for *NKX25* and *TGF1B/KLF4*, respectively, while *RUNX1* exhibited larger (-) TF effects on ICs in females. Male-trained PLS models of HDAC/SIRT/HAT effects on ICs underestimated the effects on some ICs in females. Insights from the GTEx dataset about the co-expression and transcriptional co-regulation of acetylation-modifying enzymes, transcription factors and key cardiac ion channels in a sex-specific manner can help inform safer drug design.

KEYWORDS

GTEx, epigenetics, cardiac electrophysiology, left ventricle, sex differences, histone acetylation, PLS

Introduction

Over the last decade, large-scale annotated biomedical data collections, have been assembled, such as the genotype-tissue expression (GTEx) dataset (1–3), combining RNAseq and other technologies to enable discovery of molecular mechanisms of human disease with unprecedented throughput. Transcriptomics data can be linked to subject's demographics and other characteristics, and molecular relationships can be parsed by sex, age, and tissue context. Now, these resources are being leveraged to estimate safety risk for drugs and to predict the outcomes of clinical trials (4–6), in hopes to reduce cost, time for drug development and potential side effects. Such analyses using the GTEx data have brought awareness to tissue specificity (1), cell specificity (7) and the impact of sex on key molecular pathways and transcriptional control mechanisms (7–10). For example, the GTEx dataset has been useful in identifying heart-specific genes implicated in pathologies, and in exploring cardiac tissue-specific relationships and regulators as well as transcriptome-wide differences between the ventricles and the atria (8, 9, 11–16). Such unique cardiac targets can be useful in designing new and safer therapies. Furthermore, transcriptional principles of ion channel co-expression, underlying bioelectrical stability (lower risk of early afterdepolarizations and other arrhythmias), have been uncovered using the GTEx data from the left ventricle (LV) and validated *in vitro* using human induced pluripotent stem-cell-derived cardiomyocytes, iPSC-CMs (14).

Avoiding cardiotoxicity is a critical consideration in drug design. Cardiotoxicity is at least in part due to side effects of drugs on key cardiac ion channels (directly modulating their function, affecting their transcription or trafficking), thus leading to increased arrhythmia risk (17). Prior failed drugs and recent computational data stress the need for addressing diversity in responses. For example, a drug-response risk classifier trained on male data underpredicts arrhythmia risk in females (18). Hormonal effects are now believed to only partially explain sex differences in these functional responses, as the GTEx data analysis has revealed that about 37% of all genes (majority of these autosomal) exhibit sex-differential gene expression in at least one tissue, including in the LV of the heart, where over 1,300 genes have differential sex expression (7). Importantly, among the most sex-differentially expressed genes are those involved in ion transport processes and response to drugs, including ion channels and CYP enzymes (6, 7, 19). Furthermore, sex differences in gene expression in the heart (8, 9, 20) include differences in the expression of epigenetic modifiers, in chromatin accessibility and the related regulatory networks (21, 22). Even when sex differences in transcription are not dramatic, the cumulative effect of these at the system level can be functionally impactful because of the differential effects on multiple key regulators, including cardiac transcription factors (10).

Here we consider the relationship between a class of acetylation modifiers (with histone and non-histone action) and key cardiac ion channels, defining excitation, repolarization, calcium handling, and cell-cell coupling in the heart's LV. These include histone deacetylases (HDACs and sirtuins, SIRT6) acting as “acetylation erasers” and histone acetyltransferases (HATs) acting as “writers,” which operate to balance chromatin acetylation state and accessibility for transcription factors to regulate gene expression (23) (Figure 1). Sirtuins, SIRT6, represent a sub-class of HDACs with NAD⁺ dependence, therefore, linking metabolism to post-translational modifications. In general, decreased action of HDACs and SIRT6 tilts the balance to increased acetylation, a more open chromatin state and increase in transcription of certain genes (apoptosis and other growth control processes being among the most affected ones), of particular importance in anti-cancer strategies. HDACs and HATs are also known to have important non-histone related functions, including acetylation of cytoplasmic proteins and subsequent effects on multiple cellular processes (24, 25). HDAC inhibitors (HDACi) have emerged in the last two decades as a novel anti-cancer therapy, with four of them FDA approved and several hundred clinical trials on the way (26, 27). Early analysis of safety (from clinical trials) (28, 29) and *in vitro* work with human iPSC-CMs (30–32) indicates that overall HDACi are relatively safe, with few adverse events, including arrhythmias, more prevalent for non-selective HDACi, e.g., panobinostat and vorinostat, and less prevalent for selective inhibitors, such as entinostat. Therefore, the identity of the HDACi and their specific targets among the cardiac electrophysiology key players are of high importance in achieving safer therapeutics. Beyond initial animal studies that have established select engagement of cardiac transcription factors (TFs) upon HDACi application and their role in controlling ion channel (IC) gene expression in the heart (32, 33), a more comprehensive model linking HDACs/SIRT6/HATs to ion channel transcription in the human heart *via* the action of cardiac TFs (33–35) is highly desirable.

Our objective was to use the GTEx RNAseq dataset to construct simplified computational models linking these classes of gene regulators, and to investigate if sex-specific differences exist. We chose partial least-square (PLS) regression modeling (36–38) to quantify the relationships in female and male hearts. Projection of the many master regulators into a reduced-dimension latent structures through PLS helps to elucidate co-regulation mechanisms of cardiac electrophysiology by these histone modifiers and to identify potential key mediators among the cardiac TFs. The exponential rise of publicly accessible human transcriptomics datasets can yield continuous refinement of such models for better stratification of drug action. Such models can help understand the specific effects of acetylation modifiers on the heart and guide the design of safer anti-cancer therapeutics, minimizing cardiac effects, or help

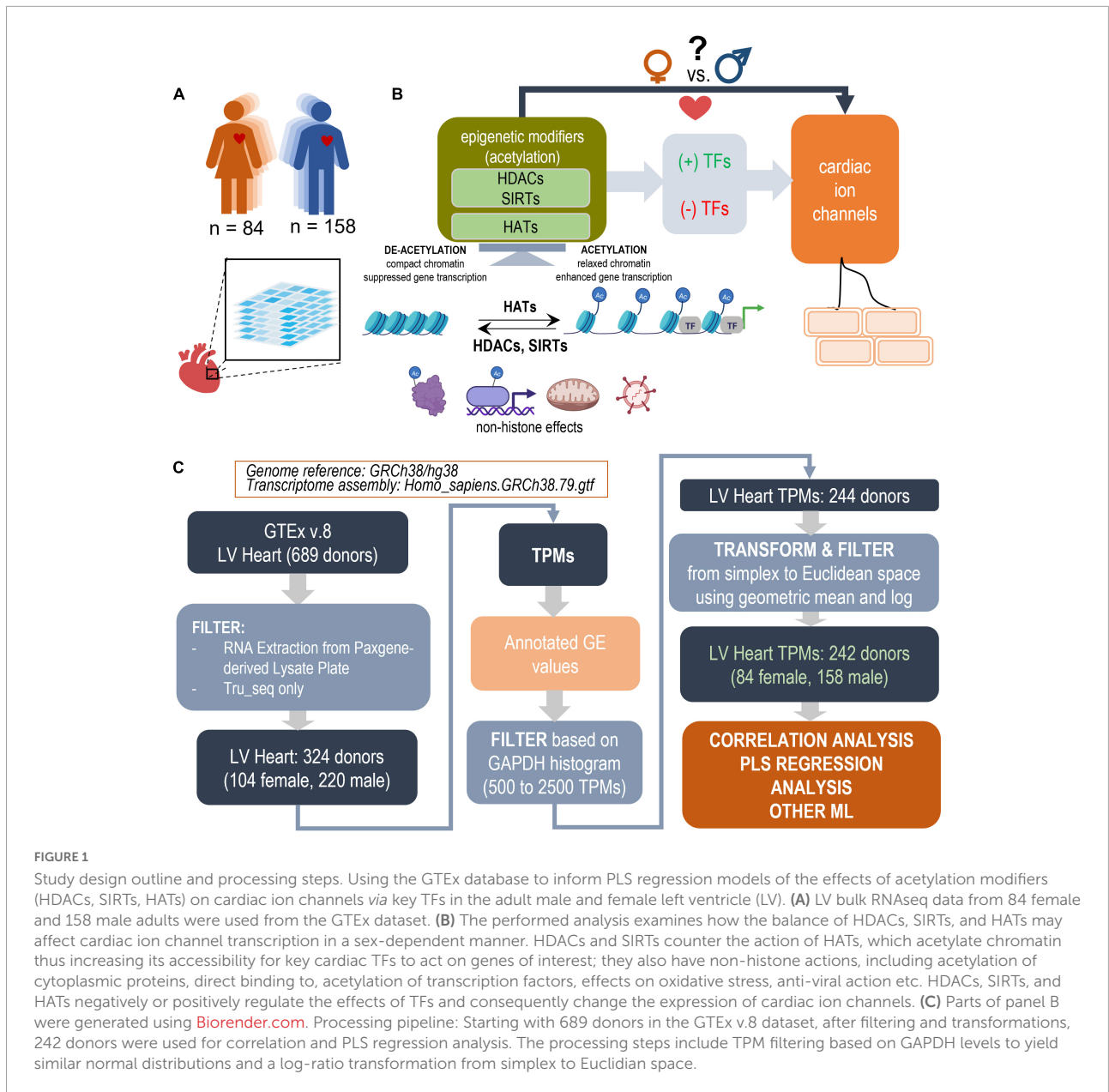


FIGURE 1

Study design outline and processing steps. Using the GTEx database to inform PLS regression models of the effects of acetylation modifiers (HDACs, SIRT6, HATs) on cardiac ion channels via key TFs in the adult male and female left ventricle (LV). (A) LV bulk RNAseq data from 84 female and 158 male adults were used from the GTEx dataset. (B) The performed analysis examines how the balance of HDACs, SIRT6, and HATs may affect cardiac ion channel transcription in a sex-dependent manner. HDACs and SIRT6 counter the action of HATs, which acetylate chromatin thus increasing its accessibility for key cardiac TFs to act on genes of interest; they also have non-histone actions, including acetylation of cytoplasmic proteins, direct binding to, acetylation of transcription factors, effects on oxidative stress, anti-viral action etc. HDACs, SIRT6, and HATs negatively or positively regulate the effects of TFs and consequently change the expression of cardiac ion channels. (C) Parts of panel B were generated using Biorender.com. Processing pipeline: Starting with 689 donors in the GTEx v.8 dataset, after filtering and transformations, 242 donors were used for correlation and PLS regression analysis. The processing steps include TPM filtering based on GAPDH levels to yield similar normal distributions and a log-ratio transformation from simplex to Euclidean space.

targeted HDACi design for cardiac use to purposefully augment cardiac function.

including EP300; 19 cardiac transcription factors (TFs) and 10 genes encoding proteins that form key cardiac ion channels, ICs.

Materials and methods

Classes of human genes in the models

The human genes analyzed here are listed in Table 1. These included the four classes of histone deacetylases, HDACs, for a total of 18: class I (HDAC1, 2, 3, and 8), class IIa (HDAC4, 5, 7, 9), class IIb (HDAC6 and 10), class III (SIRT1 to 7) and class IV (HDAC11); 15 key histone acetyltransferases HATs,

Sourcing the data From GTEx v.8

The analyses are based on the use of study data obtained from the dbGaP web site, under dbGaP accession number phs000424.v8.p2 (GTEx).¹ Expression levels in transcripts per million (TPM) derived from RNA-seq data were downloaded on 08/28/2020 as a single

¹ <https://gtexportal.org/home/datasets>

TABLE 1 Genes considered in the models.

HDACs (11)	<i>HDAC1, HDAC2, HDAC3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9, HDAC10, HDAC11</i>
SIRT5 (7)	<i>SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, SIRT7</i>
HATs (15)	<i>KAT2A, KAT2B, HAT1, ATF2, KAT5, KAT6A, KAT6B, KAT7, EP300, CREBBP, NCOA1, NCOA3, TAF1, GTF3C1, CLOCK</i>
TFs (19)	<i>FOXO1, FOXO3, GATA4, GATA6, HIF1A, KLF4, KLF5, MEF2A, NFAT5, NFKB1, NKX25, NOTCH1, RUNX1, SHMT2, SOD1, TBX5, TGFB1, TRIM28, YY1</i>
ICs (10)	<i>SCN5A, CACNA1C, KCNH2, KCNQ1, KCNJ2, ATP1A1, SLC8A1, ATP2A2, RYR2, GJA1</i>

file combining all body sites and cell types from the GTEx database. From this file, 324 TPM datasets of samples (103 female and 220 male samples) obtained from heart left ventricle (LV) were extracted based on accession numbers listed in the de-identified, open access version of the sample annotations available in dbGaP (GTEx_Analysis_v8_Annotations_SampleAttributesDS, downloaded on 08/28/2020) (see text footnote 1). The subjects age, gender and additional phenotype features were obtained from the de-identified, open access version of the subject phenotypes available in dbGaP (GTEx_Analysis_v8_Annotations_SubjectPhenotypesDS, downloaded on 08/28/2020) (see text footnote 1). Subjects range in age between 20 and 71 y.o. These RNA-seq libraries were generated using non-strand specific, polyA-based Illumina TruSeq protocol and sequenced to a median depth of 78 million 76-bp paired-end reads. Annotated gene expression was obtained using genome reference Homo_sapiens.GRCh38.79.gtf.

Data preprocessing

As indicated in **Figure 1**, TPM values were filtered, normalized and transformed before analysis. The large scale of the GTEx project and the span over a decade implies continued optimization of the sequencing technology and sample handling. Thus, the dataset contains samples of variable quality, mostly influenced by the timing of death with respect to sample processing (39). In addition to the quality control applied by the consortium, we chose to filter the data by *GAPDH* levels (leaving samples with 500–2,500 TPMs in *GAPDH*)- a prominently expressed gene that is usually a good indicator of sample dilution and quality. The aim was to exclude outliers in both the female and male samples and to yield close to normal distribution in both (see **Figure 1** and **Supplementary Figure 1**).

The RNAseq TPM data, used in this study, are compositional data, referenced with respect to the whole transcript/library size. As such, they exist in simplex/Aitchison space (40, 41). To allow derivation of quantitative relationships and obtain meaningful comparisons, correlations and distance measures, a transformation into Euclidian space was necessary and it was done by normalization using the geometric mean and log-ratio representation. Two more outliers were

removed based on a preliminary PLS model, trained within the normal range of *GAPDH*, based on extreme outlier scores (**Supplementary Figure 2**). The remaining 242 samples were normalized by the geometric mean and log scaling (Eqs. 1 and 2) (40, 41) (**Figure 1**) and used in the further analyses.

$$\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} = \sqrt[n]{x_1 x_2 \dots x_n} \quad (1)$$

$$\log_{10} \frac{x_i}{\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}} = \log_{10} \left(\frac{x_i}{\sqrt[n]{x_1 x_2 \dots x_n}}\right) \quad (2)$$

Correlation analysis

The samples were analyzed in 3 blocks of relationships (HDACs, SIRT5, HATs to ion channels), (HDACs, SIRT5, HATs to TFs) and (TFs to ion channels) (**Figure 1**). Pearson's correlations Eq. 3 (x, y representing the correlated variables, n —number of samples) were calculated for each one of these blocks and potential relationships visualized for the female and male samples, as shown in **Figures 2, 3, 4**.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3)$$

Clustering of the genes in **Figures 2, 3** was done using the clustergram function in the python library seaborn. The average and Euclidean were used for the linkage method and distance metric. The algorithm used in the seaborn implementation is described in Bar-Joseph et al. (42).

Partial least squares analysis

Three types of PLS models were trained in MATLAB and Python on the female data and the male data separately. The only differences between the two platforms are the methods employed to execute PLS: MATLAB uses nonlinear iterative PLS (NIPALS) while the sklearn library uses the statistically-inspired modification of PLS (SIMPLS). Both methods produced similar results. When using PLS it is important to use the correct number of components. Once the model was trained, prediction was calculated using Eq. 4, where X and Y are the input and output variable vectors, B is the Beta-coefficient matrix

calculated by PLS. Choosing the correct number of components can be found using cross-validation, percent variance explained, and predicted residual error sum of squares (PRESS). Leave-One-Out Cross Validation (LOOCV) was first used to find the number of components. Next PRESS was performed to confirm this finding. PRESS was calculated as per Eqs 5. The number of components settled on was 4, as it explained the most variance (over 70% for the various models) while keeping the model simple to prevent overfitting (see **Supplementary Table 1**).

$$\hat{Y} = B^T X \quad (4)$$

$$PRESS = ||Y - \tilde{Y}||^2 \quad (5)$$

Visualization by biplots

The results are presented in biplots for the PLS model inputs and model outputs. In all cases, the biplots represent the average results from 1,000 Monte Carlo simulations (see below). The biplots are projections of the model parameters (PLS “loadings”) onto the space of the first two latent variables of the constructed 4-latent-variable PLS models. The biplots also indicate on the axes the % variance explained for each of the first two latent variables. For all models, the first two latent variables explained > 50% of the variance. Proximity in angle between the lines for the model variables indicates co-regulation/similarity of action; the magnitude of the lines signifies the importance of the variable in the shown projection.

Consideration of potential confounding factors and dimension reduction uniform manifold approximation and projection

Using phenotypic information about the GTEx samples, we considered several potential confounders to our female—male analysis, including age, BMI, quality index RIN, and total ischemic time, SMTSISCH. By feeding HDACs, SIRT, HATs, TFs, Ion Channels into uniform manifold approximation and projection (UMAP), a dimension-reduction visualization (43), the samples became embedded in a 2D space and were plotted and colored by the potential confounders. The factor that appeared to systematically influence the clustering was the total ischemic time as shown in **Figure 5A**. Therefore, we performed a histogram matching for the female and male samples based on SMTSISCH, and built a “reduced” male model to compare to the female, as explained in the Results.

Monte Carlo modeling and analysis of male and female differences

Monte Carlo Cross-Validation (MCCV) was used for all PLS models, typically 1,000 runs were done with random assignment of training and testing samples (with 10% holdout for testing). This applies for the results shown in **Figures 2, 3, 4**. MCCV was also used to examine differences in the models trained on male and female samples, as shown in **Figure 5**. By randomly training n numbers of models a possible distribution of models was created. PLS was performed 1,000 times on either the male or female samples. The male and female data was split into training and test sets with a random 10% holdout. After the split, PLS models were trained. The average B-matrix, scores, and loadings were calculated and plotted, and the data used for generation was stored for later use. Subsequently, two more tests were performed using the different inputs and outputs blocks. Each input and output block were analyzed to uncover features of regulation in finer detail.

Three tests to examine the differences between male and female samples were performed. Not all PLS model types were used to analyze the differences. The only relations examined were HDACs, SIRT, HATs predicting ion channels. The first approach involved “male-trained, female tested models.” A model was created using only male samples for training (from the 158 male samples, 84 random samples were selected, of which 10% were used for holdout). With the 76 remaining male samples, a PLS model was trained. The beta matrix was obtained from this model, which is used to calculate the predictions. The errors were calculated as [*prediction* – *actual*] so that underpredictions would generate negative errors and over predictions—positive. Previous steps were repeated 1,000 times with a new random hold out. For the next model a reduced data set was used using the same method. The dataset was reduced to control for the potential error introduced with ischemic time, as seen in **Figure 5A**. Once a bin width was chosen samples were removed randomly from the males within that bin until the number of samples in male and female matched. The third approach involved “mixed-trained, mixed-tested models.” After selecting 84 male samples, they were mixed with the female samples. This left two sets of 84 samples with equal numbers of male and female in each. This model was created to test if the relations are due to random sample bias. Both procedures are visualized in **Figure 5C**.

Beyond individual influence on ion channels, the potential effect on ion channel groupings were examined. These effects were calculated by taking the mean of the beta matrix coefficients for a desired ion channel group. The groupings analyzed were Depolarization (*SCN5A*, *CACNA1C*, *SLC8A1*), Repolarization (*KCNH2*, *KCNQ1*, *KCNJ2*, *ATP1A1*), Net (Depolarization—Repolarization), Resting membrane potential

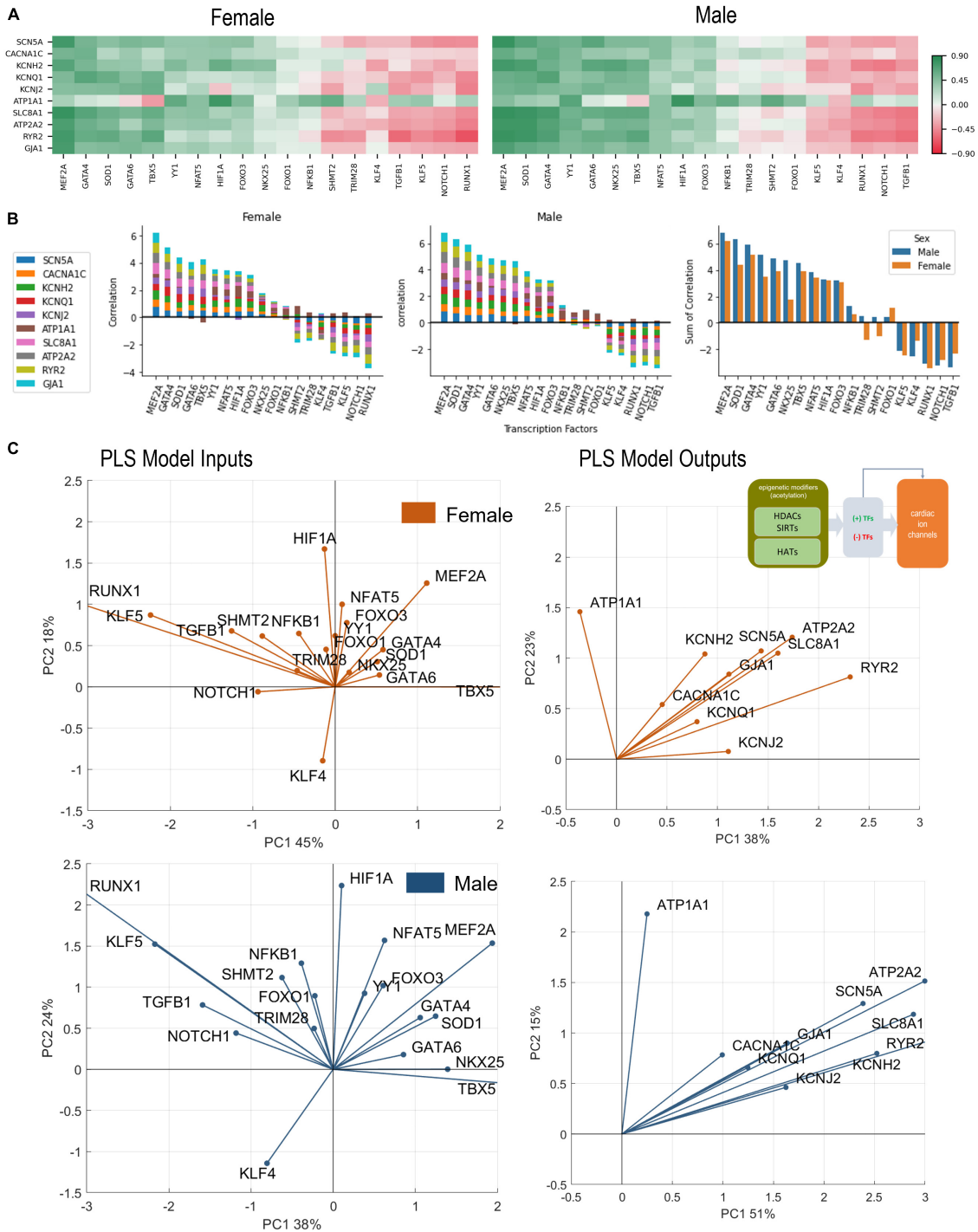


FIGURE 4

Analysis of links between cardiac TFs and key cardiac ion channels. (A) Pearson's correlation of key ion channels with cardiac TFs for female (left) and male (right) LV heart samples. The ranking of the TFs is done based on sum of correlations (from predominantly positive to predominantly negative) from left to right. These ranking yields slightly different order for female and male samples. (B) Bar plots show the cumulative correlation of individual ion channels with each TF for female, male samples, and the final plot shows the overlaid impacts of TFs on ion channels for female and male samples to reveal any differences. (C) PLS regression models trained with the TFs predicting ion channels for female (orange) and male (blue) samples. Inset in B shows the models being investigated. Shown are biplots for the PLS models—constructed as described in Figure 2.

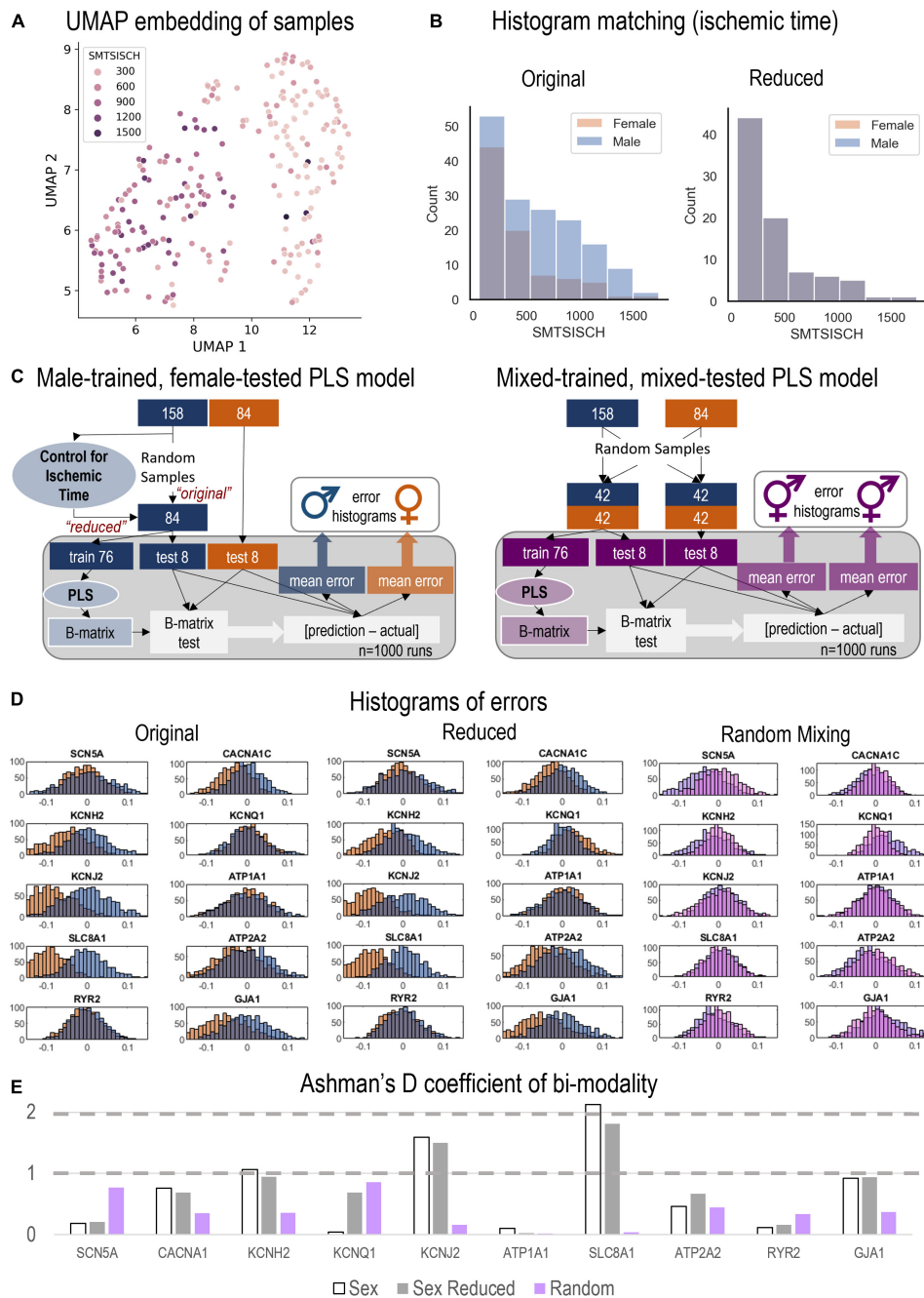


FIGURE 5

Monte Carlo (MC) simulations to examine the predictive power of male-trained PLS models when applied to female data. **(A)** UMAP embedding (dimension reduction) was used to visualize the samples in 2D space, when considering all HDACs, SIRTs, HATs, TFs, and Ion Channels. This representation showed strong role for the total ischemic time (SMTSISCH, colored) in the separation of samples. **(B)** To eliminate the effects of SMTSISCH as a confounder in the sex-difference analysis, a histogram matching was performed—the histograms show the before and after reduction of the set of male samples to match the female samples based on total ischemic time. **(C)** (Left) Details on the MC model runs and estimation of errors for male-trained, female-tested PLS models linking histone acetylation modifiers to cardiac ion channels in the LV. Two types of male models (matching the sample size of the female set, 84) were constructed—“original” with randomly selected 84 male samples and “reduced” (with histogram matching for SMTSISCH). (Right) Details on the construction of MC mixed models—trained on mixed data and tested on mixed female-male data. By mixing male and female, the model will have seen both samples and predict both well, assuming that the training samples represent the actual distribution of male and female samples. **(D)** From the 1,000 runs, the total errors were stored and displayed in histograms for the three types of models depicted in D, respectively. Note systematic shifts in errors from the zero-mean center. Left shift (to more negative error values) indicates underestimation of effects by the created PLS model for the respective variable. **(E)** Ashman's D coefficient (measure of bimodality, or how distinct the two populations of errors are) was calculated for each of the model results and displayed on the bar plot; higher Ashman's D values indicate better separability.

(*KCNJ2*, *ATPIA1*), Calcium Handling (*ATP2A2*, *RYR2*, *SLC8A1*, *CACNA1C*), and Coupling (*GJA1*). The coefficients used in these calculations were from two PLS models using all the samples (HDACs, SIRT, HATs to Ion Channels and TFs to Ion Channels). From these Beta matrices the coefficients for a given group were averaged and then plotted using the seaborn library function barplot. The results of these calculations are seen in **Figure 6**.

Figure 7 was generated using the Sankey figure from the python library plotly. A new PLS model was trained in python using the 4 components. All samples were used to obtain the Beta Matrix used in the plot. Once the models were trained on male and female separately, the Beta Matrices were used for the coloring and generation of the plot. Since the scale of some coefficients in the Beta Matrix were large, it overshadowed the overall effect of most of the HDACs on TFs and TFs on Ion Channels. To mitigate this the MaxAbsScaler function from the library sklearn was used to scale the values of each gene between -1 and 1 without shifting the center, which preserves the sign of the effect. The weight and color of the lines were generated using the scaled data.

Statistical analysis

The Ashman's D coefficient was used to examine if the two distributions of errors are part of the same distribution. For the two distributions to be separable the coefficient must be greater than 2. Ashman's D coefficient is calculated using Eq. 6 (44).

$$D = (2^{\frac{1}{2}}) \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (6)$$

Results

Genes involved in the effects of histone acetylation modifiers on cardiac electrophysiology via key cardiac transcription factors

In this study various acetylation modifiers were considered: all HDACs, including class III, sirtuins, SIRT; as well as key HATs (**Table 1**). From the 1,500+ human transcription factors, previously a subset has been identified as “cardiac TFs,” of which we considered 19 that have been implicated in control of electrophysiological processes in the heart (33–35). From the extended list of genes known to encode proteins relevant to cardiac electrophysiological processes in the LV, sometimes called the “rhythmonome” (14), we focused on the ten key genes controlling the depolarization—repolarization balance, the resting membrane potential, calcium handling and cell-cell coupling in the LV human heart (**Table 2**). The reductionist

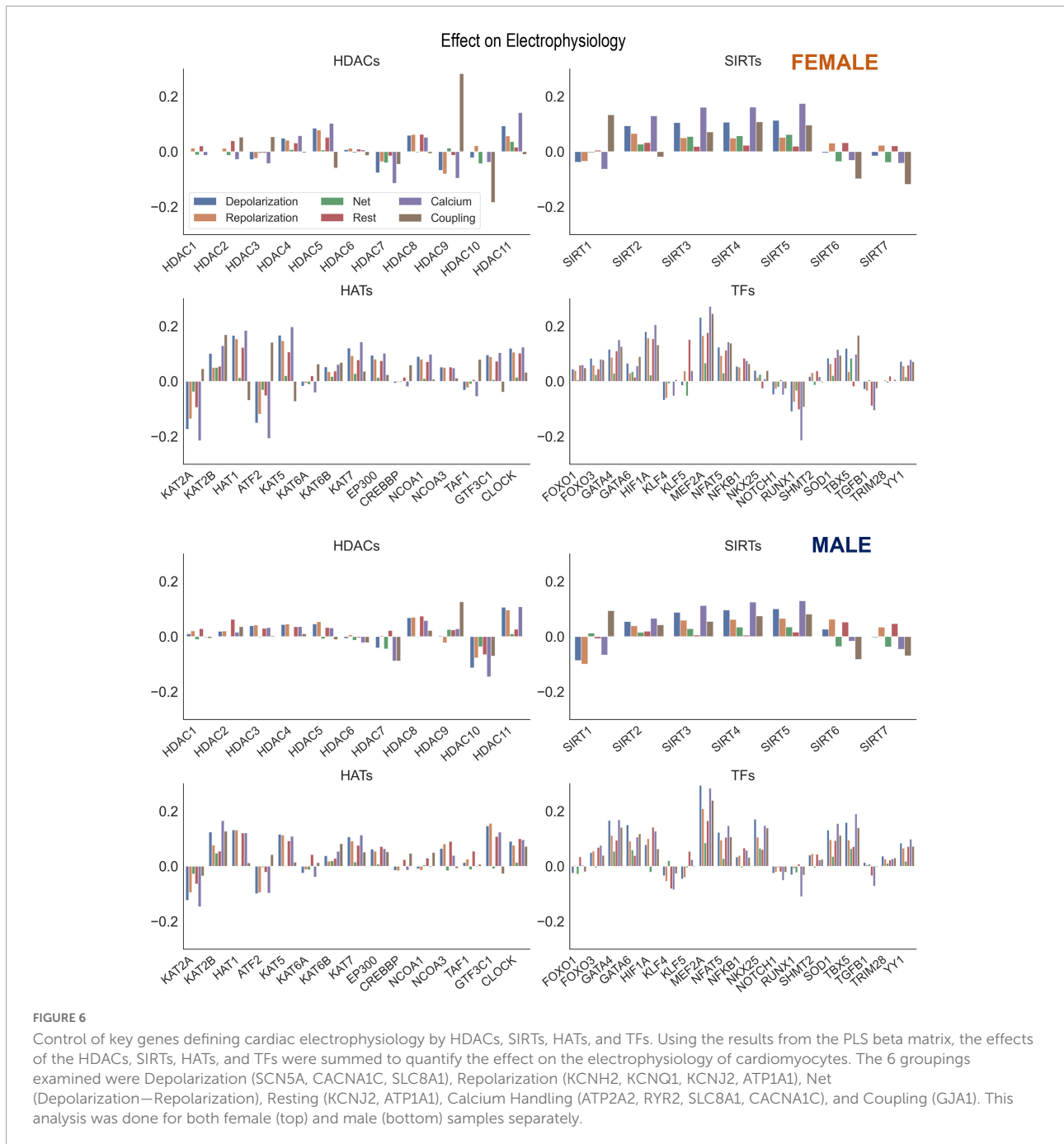
approach here—to choose 19 TFs and 10 ICs out of a bigger set was based on prior knowledge of their importance in cardiac electrophysiology but mostly dictated by the sample size available. Using standard PLS-R, we deliberately kept the number of variables lower than the number of samples. In doing so, important TFs and ICs may have been omitted; a full transcriptome PLS-R would require modified algorithms.

Linking acetylation modifiers and key cardiac ion channels based on transcription

Hierarchical clustering of female and male correlation coefficients (**Figure 2A**), separates approximately 3 groups of negatively, weakly, and positively correlated epigenetic modifiers, i.e., HDACs and HATs, in relation to the select 10 cardiac ion channels (**Table 2**). Many of these correlations are further corroborated by the results of PLS regression modeling on the transcriptomics data from female and male LV tissues (**Figure 2B**).

In female and male LVs, the strongest negative regulators of cardiac ion channels appear to be similar for the two sexes—*HDAC7*, *HDAC10*, *SIRT6*, *SIRT7*, and *KAT2A*, based on correlations of gene expression. However, the strongest positive regulators are different in female and male LV. For male hearts, these positive regulators are mostly HATs (*GTF3C1*, *KAT5*, *HAT1*, *CLOCK*, and *EP300*) plus *HDAC8*. In the female hearts, *SIRT3*, *SIRT4*, and *SIRT5* along with *HDAC9* exhibit the strongest positive regulation, in addition to *KAT7* and *KAT2B*. The mitochondria-localized *SIRT3*, *SIRT4*, *SIRT5* act together, displaying very close (positive) relationships to the ion channel expression in both sexes. In contrast, the nucleus-localized *SIRT6* and *SIRT7* display strong negative correlations with cardiac ion channel expression in both sexes.

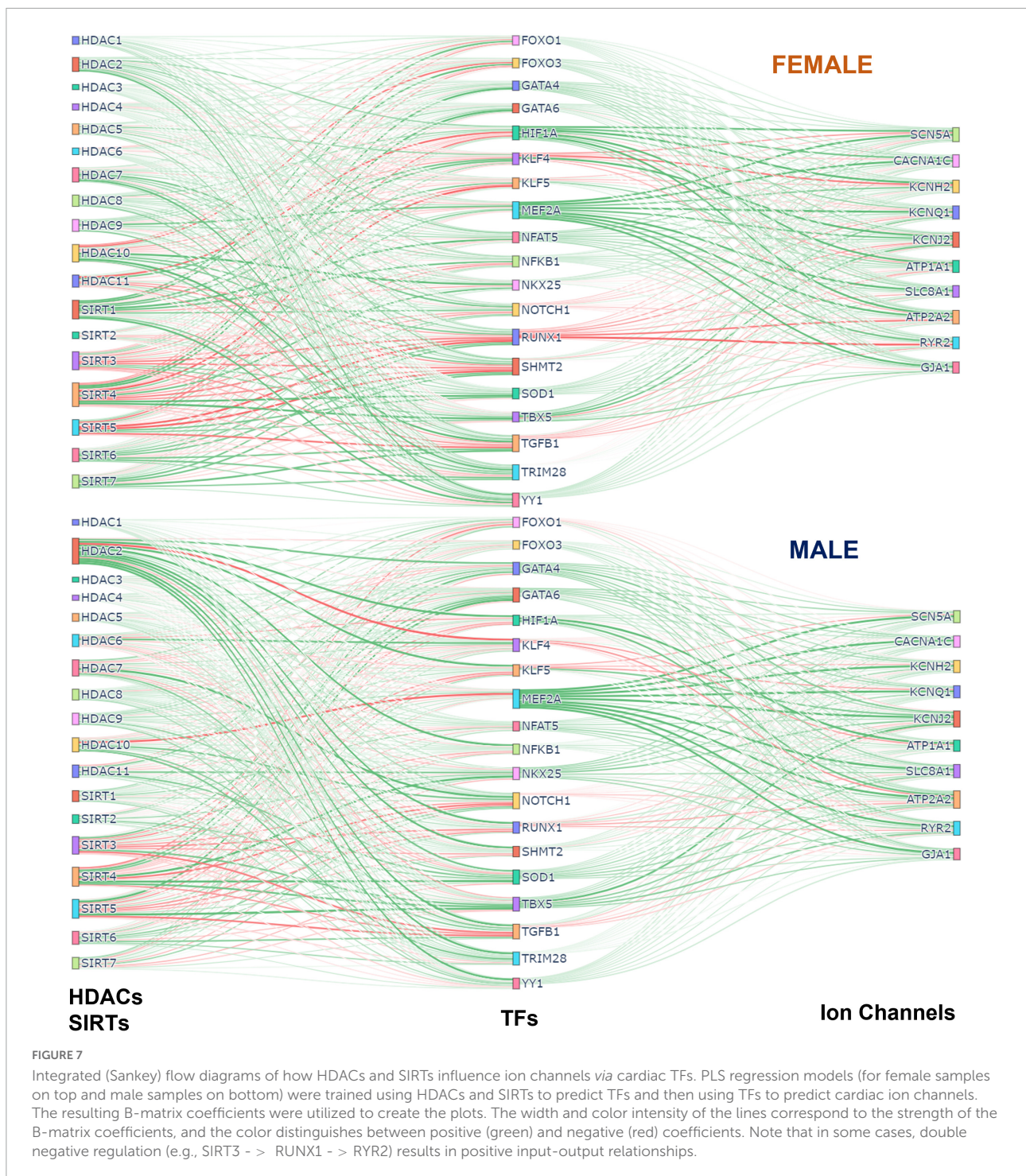
From the correlation analysis one can see that the key cardiac ion channels are largely co-regulated by these acetylation modifiers. An exception to this appears to be *ATPIA1*, which in many cases shows reverse correlation compared to the rest. PLS regression analysis results, presented as biplots in **Figure 2B**, provide a more granular view. For example, the co-regulation of ion channels by HDACs and HATs is more pronounced in the male LV compared to the female LV—tighter clustering and alignment seen in the PLS model output biplots. In both sexes, this key cardiac ion channel set is flanked by two genes regulating the resting membrane potential—*ATPIA1* and *KCNJ2*, with *ATPIA1* being orthogonally regulated compared to the rest. Such distinct pattern may have compensatory role—if the gene expression of all ion channels is upregulated or downregulated simultaneously, there may be undesired changes in the resting membrane potential that need to be countered. Genes involved in calcium handling (*RYR2*, *ATP2A2*, *SLC8A1*) are strongly affected by the acetylation modifiers in both sexes



(see prominent representation lines in the output biplots). According to this PLS model, *SIRT2*, *SIRT3*, *SIRT4*, and *SIRT5* act very closely together to exert positive regulation on the ion channels (exception being *ATP1A1*). Most of the HATs are co-aligned with *ATP1A1*, implying positive regulation of the Na/K pump to maintain the resting membrane potential. *KAT2A* is the exception in both sexes—it has distinct negative effect on all cardiac ion channels in both sexes along with *HDAC10* and to some degree *HDAC7*, *SIRT6*, and *SIRT7*.

Linking acetylation modifiers and key cardiac transcription factors

Most of the gene control by epigenetic modifiers is channeled through transcription factors; and these TFs can act as positive or negative regulators of downstream targets, e.g., ion channels. **Figure 3** shows our results from correlative analysis and clustering of such effects (**Figure 3A**), and from PLS regression modeling linking HDACs and HATs to cardiac



TFs (Figure 3B). From the clustering analysis it is obvious that the effects on TFs are complex and the outcomes are not co-aligned like the ion channels. For some TFs, like *YY1*, most of the effects appear positive, especially for HAT effects on *YY1*. This means that histone acetylation by HATs likely yields increased expression of *YY1*. Similarly, for both sexes, the HATs result in positive regulation of one of the strongest cardiac

TFs – *MEF2A*. Furthermore, *MEF2A* is negatively regulated by the same set of acetylation modifiers that appeared as the most prominent negative modulators of cardiac ion channel expression as per Figure 2: *HDAC7*, *HDAC10*, *SIRT6*, *SIRT7*, and *KAT2A*. This reinforces the idea that the action of HDACs and HATs on the ion channels is likely through potent TFs, as exemplified by *MEF2A*. Prior studies support some of these

TABLE 2 Genes, proteins and ion channels in the LV.

Gene	Protein	Ion current/process
Depolarizing		
SCN5A	Nav1.5	Underlying the main cardiac excitatory (inward) current, I_{Na}
CACNA1C	Cav1.2	Main subunit of the (inward) L-type calcium current I_{CaL}
SLC8A1	NCX1	Main subunit of the sodium-calcium exchanger I_{NCX} , which generates inward current and contributes to excitation in the forward mode
Repolarizing		
KCNH2	Kc11.1 (hERG)	Main subunit of the key repolarizing (outward) current in the LV, the rapid delayed rectifier, I_{Kr}
KCNQ1	Kv7.1	Main subunit of the slow delayed rectifier, I_{Ks} , a repolarizing current in the LV
KCNJ2	Kir2.1	Main subunit of the inward rectifier I_{K1} , which contributes to the maintenance of the resting membrane potential and the late phase of repolarization
ATP1A1	Na ⁺ /K ⁺ ATPase	Main subunit of the sodium-potassium pump I_{NaK} , which provides outward current to maintain the resting membrane potential
Resting membrane potential		
KCNJ2	Kir2.1	Main subunit of the inward rectifier I_{K1} , which contributes to the maintenance of the resting membrane potential and the late phase of repolarization
ATP1A1	Na ⁺ /K ⁺ ATPase	Main subunit of the sodium-potassium pump I_{NaK} , which provides outward current to maintain the resting membrane potential
Calcium handling		
CACNA1C	Cav1.2	Main subunit of the (inward) L-type calcium current I_{CaL} – bringing calcium through the sarcolemma to trigger calcium-induced calcium release (CICR) for contraction
RYR2	RyR2	Main subunit of the ryanodine receptor – key player in the CICR from the sarcoplasmic reticulum (SR)
ATP2A2A	SERCA2	Encodes the main subunit of the SERCA pump which helps terminate the cytosolic calcium transients and refill the SR with Ca ²⁺
SLC8A1	NCX1	Encodes the main subunit of the sodium-calcium exchanger I_{NCX} , which participates in the bi-directional sarcolemmal calcium flux
Cell-cell coupling		
GJA1	Cx43	Encodes the main gap junctional ion channel in the LV, electrically coupling cells

findings, e.g., *HDAC7* serving as a negative regulator of *MEF2A* (33, 45).

The PLS biplots reveal some differences between the female and male LVs. In the male LV, *NKX25* is co-aligned with *TBX5* and strongly influenced by the positive acetylation modifiers *SIRT5* 2-5. In contrast, the response of *NKX25* to acetylation modifiers in the female heart appears minimal. *TBX5* is strongly positively regulated by the mitochondrial SIRT5 (*SIRT3-5*) and strongly negatively regulated by the nuclear SIRT5 (*SIRT6* and *SIRT7*) in both sexes. Similar to *MEF2A*, likely *TBX5* plays an important role in mediating the effect of acetylation modifiers on the ion channels. Other clustered response to histone acetylation is seen for *GATA4*, *GATA6*, and *SOD1*, which show perfect co-regulation in the male LV along an axis similar but distinct from *MEF2A*. In the female LV, these additional TFs are closer to the responses for *MEF2A* yet smaller in magnitude. *GATA4*, *GATA6*, and *SOD1* appear co-aligned with (likely positively regulated by) *KAT2B* and *KAT6B*. A distinct difference between the female and male model outputs is seen in *KAT2B*—this HAT is co-aligned (positive regulation) with *MEF2A* and *GATA4* in the female LV, while it is closer aligned with *GATA4* but not *MEF2A* in the male LV. Other positively regulated TFs, especially by the HATs are *HIF1A*, *NEAT5*, *FOXO1*, and *FOXO3*. The responses of these four TFs are more closely aligned together in the female LV compared to the male. Many of the HATs, including *CLOCK*, *HAT1*, and *EP300*, are

clustered (act together) in their effects on the cardiac TFs and they are most closely aligned with *HIF1A*, *NEAT5*, *YY1*, *FOXO1*, and *FOXO3*, implying positive regulation.

Distinct negative cluster of TFs is formed by *RUNX1*, *KLF5*, *TGFBI*, and *NOTCH1*. For the female LV, these are joined by *TRIM28* and *SHMT2* and together they appear to present a counter-regulator for *TBX5*. Their expression may be positively regulated by *SIRT6*, *SIRT7*, and *HDAC7* (negative regulators of ion channels as seen in Figure 2). In the male LV PLS model, *KLF4* appears as direct counter-regulator of *MEF2A*, presumably exerting negative effects on the ion channels. *HDAC10* and *KAT2A* seem to positively regulate *KLF4* in the male LV. *KLF4*'s role in the female model is less pronounced.

Cardiac transcription factors and control of the gene expression of key cardiac ion channels

The set of 19 TFs was selected because of their known prominent role in regulation of genes underlying cardiac electrophysiology processes (33–35). In Figure 4, we examine closely the role of these “cardiac TFs” on the 10 ion channels of interest. As in Figure 2 for links between the acetylation modifiers and the ion channels, here we also see pronounced co-regulation of ion channels by the TFs. This means that there

are potent “positive TFs” that act to increase the expression of all or almost all key cardiac ion channels; there are also potent “negative TFs” that act to decrease the expression of all or almost all key cardiac ion channels. Among the positive TFs are *MEF2A*, *GATA4*, *GATA5*, *SOD1*, *TBX5*, *YY1*. Their extent of action varies in female and male LV. Perhaps the most striking sex difference is in the stronger role for *NKX25* on ion channel expression in the male LV compared to the female LV (Figure 4B, right panel). Additionally, some of the weaker TFs, like *TRIM28* and *SHMT2*, may show opposite effect on the regulation of the ion channels for the female and male LV. Based on the ranking of correlation in Figures 4A,B and based on the PLS biplots in Figure 4C, the ion channel co-regulation is also driven by a set of exclusively negative cardiac TFs. In this cluster are *RUNX1*, *NOTCH1*, *TGFBI*, *KLF5*, and *KLF4* for both sexes, though their relative ranking is different for female and male LVs. Similar to the results in Figure 2, *ATPIA1* is orthogonally regulated (with respect to the rest of the cardiac ion channels) by many of the positive and negative TFs. This is clearly seen for *TBX5*, which exerts negative effect on *ATPIA1* yet acts as a strong positive regulator of the rest of the ion channels. Similarly, *RUNX1*, *TGFBI*, *NOTCH1*, and *KLF5* exert positive effects on *ATPIA1* yet they act as strong negative regulators of the rest of the ion channels.

The PLS output biplots (Figure 4C), illustrate the co-regulation of cardiac ion channels by TFs. Again, similar to Figure 2B, this co-regulation is much stronger (tighter clustering) in the male LV. And the distinct orthogonal regulation of *ATPIA1* (compared to the rest of the ion channels) is seen in the female and male LV. *HIF1A* and *NFAT5* seem most aligned with *ATPIA1* (positive regulation) for both sexes. *TBX5*, *GATA6* (and *NKX25* for the male LV only) appear to be most aligned with *KCNJ2*. In both sexes, *MEF2A* plays a prominent positive role for the expression of cardiac ion channels, aligned with the cluster of ion channels. *KLF4* directly counters the role of *MEF2A*, especially in the male LV. The PLS input biplots also showcase the extent of clustering of the negative TFs, especially *RUNX1*, *KLF5*, *TGFBI*, *NOTCH1*.

Can male-trained partial least-square models predict female relationships Between acetylation modifiers and cardiac ion channels?

Using the genes of interest in this study (Table 1), we visualized all samples using a dimension-reduction approach, UMAP (Figure 5A). When color-coding various phenotypic variables associated with the samples, including age, BMI etc., we did not observe specific clustering, except for the UMAP plots when total ischemic time was displayed (Figure 5A). SMTSISCH is the total time for handling an LV tissue sample

from the moment of opening the chest to the collection of tissue lysate for transcriptomics in GTEX. To eliminate ischemic time as a potential confounder in our analysis, we looked closely at the SMTSISCH histograms for female and male LV samples used. Figure 5B, left, shows that the larger data set for male LV did also include some longer ischemic times. We performed histogram-matching, by reducing the male dataset, eliminating samples so that the histograms of ischemic time for female and male samples became identical—see “reduced” in Figure 5B. The PLS biplots for the “original” male and female models along with the biplots for the new “reduced” male model are shown in Supplementary Figures 3–5.

We tested if purely male-trained PLS models of influences of HDACs and HATs on cardiac ion channels can accurately predict such influences for female LV samples. This was done using the full male set of 158 samples (Figure 5C, left, “original” model), the “reduced” male data set of 84 samples with histogram-matched ischemic time (Figure 5C, left, “reduced” model) and for a “mixed” model using equal number of female and male samples for training and for testing (Figure 5C, right). Monte Carlo simulations in all three cases yielded prediction errors for various outputs (ion channels) when the models were tested on male and female samples. Histograms of these errors are displayed in Figure 5D for the “original” and “reduced” male-trained models as well as for the “mixed” model. For good predictions, we expect the error histograms to be centered at zero and to be relatively tight. This is mostly the case for the mixed model, where the male and female error histograms overlap. The “original” and “reduced” male-trained models exhibit similar behavior for some of the ion channel genes, such as *SCN5A*, *KCNQ1*, *ATPIA1*, and *RYR2*, implying that predicting the effects of HDACs and HATs on these genes may not be sex dependent. However, other results, especially for *KCNJ2*, *SLC8A1*, *KCNH2*, and *GJA1* had error histograms for female predictions that were left shifted from zero, implying that the extent of regulation of these genes was under-estimated by the male-trained models. This is similar for the “original” and the “reduced” model (after ischemic time has been histogram-matched). Ashman’s D is a metric that can quantify the degree of bi-modality in a lumped distribution that may encompass two distinct populations. In Figure 5E, we computed Ashman’s D for the error histograms obtained in the three models. For the “mixed” model (training on equal mix of male and female samples), Ashman’s D coefficients remained low (< 1) for all output variables. For the “original” male-trained model, Ashman’s D exceeded 1 for *KCNH2*, *KCNJ2*, and *SLC8A1*. The “reduced” model showed similar trends to the “original” male-trained model, though slight decrease was seen in this metrics of bimodality. Ashman’s D was the highest for *SLC8A1*, implying that male-trained models may inaccurately predict the regulation of the Na/Ca²⁺ exchanger by HDACs and HATs in the female LV.

Acetylation modifiers and net effects on cardiac electrophysiology

Our results showed strong co-regulation of the cardiac ion channels by HDACs, HATs (Figure 2) and by the cardiac TFs (Figure 4). We were interested to see the effects of these regulators on groupings of ion channels, as defined in Table 2: genes encoding for depolarizing, repolarizing ion channels, genes for ion channels and processes responsible for the maintenance of the resting membrane potential, for calcium handling or cell-cell coupling. The results of these effects, color-coded by groupings are displayed in Figure 6 for the female (top) and the male (bottom) models. Of particular interest is if a certain perturbation may bias the net effect (positive for net depolarizing effect, negative—for net hyperpolarizing effect). As expected, most of the individual perturbations by HDACs, SIRT, HATs, and cardiac TFs leave minimal (close to zero) net effect. For the mitochondrial SIRT and some of the positive TFs, there is a net depolarizing effect, based on the subset of ion channels considered here, possibly increasing excitability. Similarly, mitochondrial SIRT and some of the HDACs (*HDAC11*, and *HDAC5* for female only) plus some of the HATs (*KAT2B*, *HAT1*, *KAT5*, *KAT7*, *EP300*, *GTF3C1*, *CLOCK*) have prominent positive effect on genes related to calcium handling. Conversely, the strongest negative regulators of calcium handling appear to be *HDAC7*, *KAT2A*, *ATF2*, *SIRT1* (plus *HDAC9* for female LV and *HDAC10* for male LV). The impact of TFs on calcium handling genes directly follows from the results on “positive” and “negative” TFs as seen in the previous figures. The strongest positive regulators of calcium handling genes in the female LV are *MEF2A*, *HIF1A*, *GATA4*, *NFAT5*, and *SOD1*, while in the male LV these are *MEF2A*, *TBX5*, *GATA4*, *NKX25*, and *NFAT5*. The strongest negative regulators of calcium handling genes for both sexes appear to be *RUNX1*, *TGFBI*, and *KLF4*.

Sankey plots visualizing overall relationships Between acetylation modifiers, transcription factors and cardiac ion channels

To visualize the overall effects of acetylation modifiers on cardiac ion channels *via* select transcription factors we used Sankey plots, based on the B-matrix coefficients from the PLS regression models for females (top) and males (bottom) (Figure 7). The sign and the magnitude of the B-matrix coefficients determine the coloring and the thickness of the lines. Among the similarities between female and male, *MEF2A* appears to be the strongest positive TF of ion channels with a mix of positive and negative influences from HDACs and SIRT. The strongest negative TF of ion channels, especially in the female LV, is *RUNX1*, which also gets mixed positive

and negative influences from HDACs and SIRT. For example, for both sexes, the mitochondrial NAD⁺ dependent sirtuins *SIRT3*, *SIRT4*, and *SIRT5* downregulate *RUNX1*, thus exerting ultimately positive effect on the cardiac ion channels, while *HDAC7* and *HDAC10* positively influence *RUNX1*, thus having overall negative effect on ion channel expression. Other positive TFs of ion channels seen here are *GATA4*, *GATA6*, *SOD1*, *NFAT5*, *TBX5*, *YY1*. *HIF1A* is a significantly stronger positive regulator of ion channels in the female LV compared to the male, while *NKX25* is stronger positive TF of ion channels in the male hearts compared to female, in line with previous studies that have identified *NKX25* as a male-biased TF (8). *HDAC2* also appears to play more prominent role in the male LV, regulating various cardiac TFs; *SIRT1* exerts stronger (positive) influences on cardiac TFs in the female LV. *SIRT3*, 4 and 5 are prominent in both sexes, providing downregulation of the negative TFs and upregulation of the positive TFs, thus having an overall positive effect on cardiac ion channels.

Discussion

In this study we considered quantitative links between acetylation modifiers (HDACs and HATs) and the transcription of ion channels in the human heart based on the GTEx dataset. Their canonical role of histone acetylation/deacetylation and controlling the compactness of chromatin is better understood for genes controlling growth or cell death (26, 46–48), often in the context of cancer. Their effect on bioelectricity-related genes is much less studied (33). Cell excitation and calcium handling, as metabolically demanding processes, may be considered positively regulated during growth and hypertrophy, and downregulated during ischemia or other conditions when conservation of resources is needed. The prominent role played by sirtuins, especially the NAD⁺ dependent SIRT3–5 in regulating cardiac ion channels (Figures 2, 3, 4, 7), reflects their role as linkers between metabolic state and post-translational modifications to control transcription according to available resources (49, 50).

Non-histone effects of HDACs and HATs have major impact on cell function through acetylation of cytoplasmic proteins, e.g., alpha-tubulin by HDAC6, or direct acetylation of transcription factors, such as those from the MEF2 family, GATA and others. Among the non-histone effects of HDACs and HATs are impact on protein aggregation, protein clearance/degradation, autophagy, DNA repair capabilities, oxidative stress, anti-inflammatory action, control of chaperone proteins, anti-viral roles, circadian rhythm, and other key cell functions (24, 25, 51). Beyond cancer, these acetylation modifiers have been shown to be impactful through their histone and non-histone action in cardiovascular health, especially related to development, hypertrophy/heart failure, metabolic state, and cardiac aging (52–55).

The models, presented here, linking the gene expression of acetylation modifiers to the expression of cardiac ion channels *via* select cardiac transcription factors, do not aspire to distinguish a mechanism of action, i.e., they are admittedly not mechanistic. A strong effect between an HDAC, HAT and an ion channel based on gene expression alone may result from any of the direct or indirect ways of histone or non-histone action. As the GTEx dataset is still with a limited number of samples (a total of 242 analyzed here), the applicability of some machine learning methods that thrive on large datasets is limited. We chose to use partial least-squares regression (PLS-R) in this study to quantify these relationships and to explore potential sex differences in them. PLS-R was developed by Wold et al. (36), Abdi (37), and Wold (38) to deal with noisy and collinear variables. It is most closely related to Principal Component Regression (PCR). Both methods remove the multicollinearity between samples by projecting them into a latent space. As PCR maximizes the variance explained for the new latent variable, it may overlook the possibility that the latent variable may not explain the predicted variable. This leads to using many more components than needed to predict the outputs. PLS-R, used here, on the other hand, maximizes the covariance between the input and output variables, and therefore maximizes the predictability of the inputs-to-outputs links. This connection between the inputs and outputs makes PLS-R a desirable method for examining RNA-seq data for specific relationships. This method is interpretable, and it holds causal power (beyond simple correlation) if large portion of the variance in the experimental data can be explained by the model (36, 38, 56). When interpreting the results of the PLS-R analysis here, we assume certain simplified and directional input-output structure, which in reality may be more complex. For example, gene regulatory networks, GRNs, may involve bidirectional dependences and regulatory loops, the reconstruction of which would require dedicated new experimental data collection.

In **Figures 2, 3, 4** we show how acetylation modifiers may link to cardiac ion channels *via* cardiac TFs. It is important that some of the key results could be corroborated by prior research, mostly done in rodents, where transgenic models can be used to delete genes of interest and dissect functional impact. One of the main results in this study is the tight co-regulation of genes encoding cardiac ion channels by the acetylation modifiers and by the cardiac TFs. This is seen as clustering and co-alignment of the ion channel genes in the PLS output biplots in **Figures 2, 4**, with the male hearts exhibiting tighter clusters compared to the female. These results are in agreement with a recent report (57), showing highly coordinated transcriptional control of cardiac ion channels in the ventricles, but not in the atria. This may be due, at least in part, to the reported synergistic action of positive TFs in the heart (34, 58)—collaboration between *MEF2A*, *GATA4*, *TBX5*, *NKX25* and the HAT *EP300*, identified based on chromatin co-occupancy through Chip-seq analysis of mouse heart tissue. Furthermore, *GATA4* has been linked to chromatin acetylation state (H3K27ac) in facilitating

gene expression in the (mouse) heart. In the current report, using human RNAseq data, we see similar synergism between these strong positive TFs, including *GATA4*, in the regulation of cardiac ion channels. Furthermore, strong negative TFs mediating the effect of HDACs/SIRT6/HATs on cardiac ion channels, such as identified here *RUNX1*, have been reported in recent studies as well (59, 60). Specifically, these reports showed adverse cardiac effects, including suppressed calcium handling by *RUNX1* after myocardial infarction (59, 61), in line with our results. Another strong negative TF shown here is *KLF5*, closely co-aligned with *RUNX1*. *KLF5* has recently been identified as a key negative regulator in ischemic cardiomyopathy in mouse and human hearts (62, 63) and has been suggested as a potential therapeutic target for drug development.

A recent study (14) combined meta-analysis of human RNAseq datasets, computational modeling of cardiac electrophysiology and *in vitro* experiments with human induced pluripotent stem cell derived cardiomyocytes (iPSC-CMs) to show a strong correlation between *CACNA1C* and *KCNH2* and better stability of cardiac electrical activity (robustness and protection against arrhythmias) for high co-expression/co-regulation of these two genes. In our results, *CACNA1* and *KCNH2* are perfectly co-aligned/co-regulated by HDACs/SIRT6 (Figure 2B) and by cardiac TFs (Figure 4C) in the female LV, and less so in the male LV. If these sex differences are confirmed in further experiments in human cardiac cells and tissue, they could be partially contributing to higher resistance to arrhythmias in female hearts.

The co-regulation of ion channels controlling depolarization and repolarization implies that the HDACs/SIRT6/HATs and cardiac TFs are likely to leave mostly net zero effect in terms of excess depolarization or excess hyperpolarization in the heart, as seen in Figure 6. This is important for overall electrical stability. It may also explain the relatively low cardiotoxicity of HDAC inhibitors (29). Even if the effects of HDACs/SIRT6/HATs and cardiac TFs do not create net depolarization/hyperpolarization, a concerted significant upregulation of ion channel genes can lower cell/tissue impedance; conversely, a concerted downregulation of ion channels will increase tissue impedance. The former will make it harder to excite or for an extra beat to propagate; while the latter may increase the risk for an aberrant beat to propagate. These more subtle effects of HDACs/SIRT6/HATs modulating cardiac ion channels, in tandem, may impact how epigenetic modulation alters arrhythmia propensity (30–33).

A distinct orthogonal regulation of *ATP1A1* (with respect to the rest of the cardiac ion channels) was seen by the HDACs/SIRT6/HATs and by the cardiac TFs (Figures 2, 4). This may be a way to compensate for altered impedance (see above), as *ATP1A1* encodes for the Na/K pump, a major determinant of the resting membrane potential. It has been shown previously that higher level of excitatory activity and higher calcium release in neurons regulate *ATP1A1* transcription toward such compensatory effect (64). *HIF1A* has been identified as a

TF closely linked to *ATPIA1* regulation (64), and our data corroborate that, as seen in **Figure 4C** for both female and male.

Sex differences in transcriptional processes related to membrane ion transport, excitability and interaction with drugs (7), as well as in chromatin regulation and epigenetic influences (21, 65) have been previously reported for various organs. In the heart, the consequences of such sex differences may be of prime importance when developing and testing new drug compounds, including HDAC inhibitors. Recently, cardiac simulations showed that models integrating male-based data underestimated the risk for lethal arrhythmias, such as Torsade de Pointes, in females (18). To aid the comprehensive preclinical analysis of drug action, transcriptional and other relevant data need to be viewed in a stratified manner, with considerations of the target demographics, including sex (6). In the current study, we show several distinct sex-dependent patterns of regulation of the acetylation modifiers and their action on cardiac ion channels. Specifically, we showed that male-trained PLS models may underestimate the effects of HDACs/SIRT/HATs on certain ion channel related genes, including *SLC8A1*, *KCNJ2*, and *KCNH2* (**Figure 5**). Other insights from the PLS analysis, including sex differences in the tightness of cardiac ion channel co-regulation, the more prominent role of *NKX25* in the male LV, and stronger effects of *RUNX1* in the female LV, need to be confirmed as larger human datasets become available, as well as tested explicitly experimentally.

The analysis provided here has certain limitations—we worked only with gene expression data and without the ability to induce controlled perturbations in the effectors of interest (acetylation modifiers). Mechanistic understanding of the complex action of acetylation modulators will require information from additional sources, including chromatin accessibility assays and quantification of protein acetylation at large. Such data can help distinguish between histone and non-histone action of the HDACs and HATs on the ion channels. Furthermore, while PLS-R may suggest input-output causality, ultimately such relationships need to be tested experimentally, with proper input functions. For example, Perturb-seq (66, 67), using interference CRISPR to perturb the genes of interest (HDACs, HATs) and follow up scRNA-seq to quantify the response, including all TFs and ICs of interest, would be a way to experimentally test some of the predictions here and to establish causality, in general. In its current simplified form, our model data corroborate some previously reported relationships, but also show new interactions that need to be tested experimentally. Larger datasets are likely to help better model development and to increase the confidence in model predictions. Our current analysis was done on bulk tissue RNAseq data and cannot differentiate if sex differences may be due to differences in cell type composition. Recent studies have suggested sex-distinct cell type composition, including a larger fraction of myocytes in the female LV compared to the male LV (68), which may explain certain sex differences in transcriptional regulation. Ideally,

single-cell or spatial transcriptomics data (69), combined with the framework developed here, can improve the mechanistic understanding of the action of histone acetylation modifiers in the heart. Moreover, a small selection of genes was analyzed (number of genes smaller than the number of samples) in this study, however, PLS can further be utilized with the full transcriptome. For cases where much longer list of genes is considered—longer than the number of samples, a “sparse PLS” can be used as a dimension-reduction and a parameter selection tool (70).

Conclusion

In conclusion, we present a framework that aims to help the better understanding of how acetylation modifiers may alter cardiac electrophysiology based on RNAseq data from human hearts. With the rapid growth and availability of such data it is possible to refine computational models to reliably predict responses to various epigenetic modifiers in a stratified, personalized manner.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The Python and MATLAB scripts used to generate the results in this study are made available through github: <https://github.com/pressm/PLSGTEEx>. The RNAseq dataset analyzed during the current study are available in the GTEx repository, <https://gtexportal.org/home/dataset>, but restrictions apply to the availability of the phenotypic data, which can be accessed by following the instructions at <https://gtexportal.org/home/protectedDataAccess>.

Author contributions

MP, AH, and EE designed the study. AH accessed the GTEx database, extracted the data, and performed initial filtering and checks. MP and EE constructed the models and interpreted the results. MP wrote all codes related to the project, performed the analysis, and created the figures. EE supervised the project. MP and EE wrote the manuscript with input from AH. All authors reviewed and revised the manuscript.

Funding

This work was supported by the National Science Foundation (EFRI 1830941) and the National Institutes of Health (R01HL144157) to EE.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2022.941890/full#supplementary-material>

References

1. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. (2020) 369:1318–30. doi: 10.1126/science.aaz1776
2. GTEx Consortium. Human genomics. the Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. (2015) 348:648–60. doi: 10.1126/science.1262110
3. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. *Nature*. (2017) 550:204–13.
4. Duffy A, Verbanck M, Dobbyn A, Won HH, Rein JL, Forrest IS, et al. Tissue-specific genetic features inform prediction of drug side effects in clinical trials. *Sci Adv*. (2020) 6:eabb6242. doi: 10.1126/sciadv.abb6242
5. Nguyen PA, Born DA, Deaton AM, Nioi P, Ward LD. Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *Nat Commun*. (2019) 10:1579.
6. Idda ML, Campesi I, Fiorito G, Vecchiotti A, Urru SAM, Solinas MG, et al. Sex-biased expression of pharmacogenes across human tissues. *Biomolecules*. (2021) 11:1206. doi: 10.3390/biom11081206
7. Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, Wucher V, Gewirtz ADH, Cotter DJ, et al. The impact of sex on gene expression across human tissues. *Science*. (2020) 369:eaba3066.
8. InanlooRahatloo K, Liang G, Vo D, Ebert A, Nguyen I, Nguyen PK. Sex-based differences in myocardial gene expression in recently deceased organ donors with no prior cardiovascular disease. *PLoS One*. (2017) 12:e0183874. doi: 10.1371/journal.pone.0183874
9. Deegan DF, Karbalaee R, Madzo J, Kulathinal RJ, Engel N. The developmental origins of sex-biased expression in cardiac development. *Biol Sex Differ*. (2019) 10:46.
10. Lopes-Ramos CM, Chen C-Y, Kuijjer ML, Paulson JN, Sonawane AR, Fagny M, et al. Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell Rep*. (2020) 31:107795. doi: 10.1016/j.celrep.2020.107795
11. Ahn J, Wu H, Lee K. Integrative analysis revealing human heart-specific genes and consolidating heart-related phenotypes. *Front Genet*. (2020) 11:777. doi: 10.3389/fgene.2020.00777
12. Rahman R, Zatorski N, Hansen J, Xiong Y, van Hasselt JGC, Sobie EA, et al. Protein structure-based gene expression signatures. *Proc Natl Acad Sci U S A*. (2021) 118:e2014866118.
13. Pérez-Hernández M, Marrón-Liñares GM, Schlamp F, Heguy A, van Opbergen CJM, Mezzano V, et al. Transcriptomic coupling of PKP2 with inflammatory and immune pathways endogenous to adult cardiac myocytes. *Front Physiol*. (2020) 11:623190. doi: 10.3389/fphys.2020.623190
14. Ballouz S, Mangala MM, Perry MD, Heitmann S, Gillis JA, Hill AP, et al. Co-expression of calcium and hERG potassium channels reduces the incidence of proarrhythmic events. *Cardiovasc Res*. (2021) 117:2216–27. doi: 10.1093/cvr/cvaa280
15. Veerman CC, Podliesna S, Tadros R, Lodder EM, Mengarelli I, de Jonge B, et al. The brugada syndrome susceptibility gene HEY2 modulates cardiac transmembrane ion channel patterning and electrical heterogeneity. *Circ Res*. (2017) 121:537–48.
16. Tan WLW, Anene-Nzulu CG, Wong E, Lee CJM, Tan HS, Tang SJ, et al. Epigenomes of human hearts reveal new genetic variants relevant for cardiac disease and phenotype. *Circ Res*. (2020) 127:761–77.
17. Mamoshina P, Rodriguez B, Bueno-Orovio A. Toward a broader view of mechanisms of drug cardiotoxicity. *Cell Rep Med*. (2021) 2:100216. doi: 10.1016/j.xcrm.2021.100216
18. Fogli Iseppa A, Ni H, Zhu S, Zhang X, Coppini R, Yang P-C, et al. Sex-specific classification of drug-induced torsade de pointes susceptibility using cardiac simulations and machine learning. *Clin Pharmacol Ther*. (2021) 110:380–91. doi: 10.1002/cpt.2240
19. Kim-Hellmuth S, Aguet F, Oliva M, Muñoz-Aguirre M, Kasela S, Wucher V, et al. Cell type-specific genetic regulation of gene expression across human tissues. *Science*. (2020) 369:eaaz8528.
20. Deegan DF, Nigam P, Engel N. Sexual dimorphism of the heart: genetics, epigenetics, and development. *Front Cardiovasc Med*. (2021) 8:668252. doi: 10.3389/fcvm.2021.668252
21. Sim CB, Phipson B, Ziemann M, Rafahi H, Mills RJ, Watt KI, et al. Sex-specific control of human heart maturation by the progesterone receptor. *Circulation*. (2021) 143:1614–28. doi: 10.1161/CIRCULATIONAHA.120.051921
22. Hartman RJG, Huisman SE, den Ruijter HM. Sex differences in cardiovascular epigenetics—a systematic review. *Biol Sex Differ*. (2018) 9:19. doi: 10.1186/s13293-018-0180-z
23. Gillette TG, Hill JA. Readers, writers, and erasers: chromatin as the whiteboard of heart disease. *Circ Res*. (2015) 116:1245–53. doi: 10.1161/CIRCRESAHA.116.303630
24. Narita T, Weinert BT, Choudhary C. Functions and mechanisms of non-histone protein acetylation. *Nat Rev Mol Cell Biol*. (2019) 20:156–74.
25. Drazic A, Myklebust LM, Ree R, Arnesen T. The world of protein acetylation. *Biochim Biophys Acta (BBA) Proteins Proteomics*. (2016) 1864:1372–401.
26. McKinsey TA. Therapeutic potential for HDAC inhibitors in the heart. *Annu Rev Pharmacol Toxicol*. (2012) 52:303–19.
27. Banik D, Moufarrij S, Villagra A. Immunoepigenetics combination therapies: an overview of the role of HDACs in cancer immunotherapy. *Int J Mol Sci*. (2019) 20:2241. doi: 10.3390/ijms20092241
28. Albin A, Pennesi G, Donatelli F, Cammarota R, De Flora S, Noonan DM. Cardiotoxicity of anticancer drugs: the need for cardio-oncology and cardio-oncological prevention. *J Natl Cancer Inst*. (2010) 102:14–25. doi: 10.1093/jnci/djp440
29. Schiattarella GG, Sannino A, Toscano E, Cattaneo F, Trimarco B, Esposito G, et al. Cardiovascular effects of histone deacetylase inhibitors epigenetic therapies: systematic review of 62 studies and new hypotheses for future research. *Int J Cardiol*. (2016) 219:396–403. doi: 10.1016/j.ijcard.2016.06.012
30. Kopljar I, De Bondt A, Vinken P, Teisman A, Damiano B, Goeminne N, et al. Chronic drug-induced effects on contractile motion properties and cardiac biomarkers in human induced pluripotent stem cell-derived cardiomyocytes. *Br J Pharmacol*. (2017) 174:3766–79. doi: 10.1111/bph.13713

31. Kopljar I, Gallacher DJ, De Bondt A, Cougnaud L, Vlaminckx E, Van den Wyngaert I, et al. Functional and transcriptional characterization of histone deacetylase inhibitor-mediated cardiac adverse effects in human induced pluripotent stem cell-derived cardiomyocytes. *Stem Cells Transl Med.* (2016) 5:602–12. doi: 10.5966/sctm.2015-0279
32. Pozo MR, Meredith GW, Entcheva E. Human iPSC-cardiomyocytes as an experimental model to study epigenetic modifiers of electrophysiology. *Cells.* (2022) 11:200. doi: 10.3390/cells11020200
33. Brundel B, Li J, Zhang D. Role of HDACs in cardiac electropathology: therapeutic implications for atrial fibrillation. *Biochim Biophys Acta Mol Cell Res.* (2020) 1867:118459. doi: 10.1016/j.bbamcr.2019.03.006
34. He A, Kong SW, Ma Q, Pu WT. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc Natl Acad Sci U S A.* (2011) 108:5632–7. doi: 10.1073/pnas.1016959108
35. McCulley DJ, Black BL. Transcription factor pathways and congenital heart disease. *Curr Top Dev Biol.* (2012) 100:253–77.
36. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemometrics Intell Lab Syst.* (2001) 58:109–30.
37. Abdi H. Partial least squares (PLS) regression. In: Lewis-Beck M, Bryman A, Futing T editors. *Encyclopedia of Social Sciences Research Methods.* Thousand Oaks, CA: Sage (2003). p. 792–5.
38. Wold H. 11 - Path models with latent variables: the NIPALS Approach**NIPALS = Nonlinear iterative partial least squares. In: Blalock HM, Aganbegian A, Borodkin FM, Boudon R, Capecci V editors. *Quantitative Sociology.* Cambridge, MA: Academic Press (1975). p. 307–57.
39. Ferreira PG, Muñoz-Aguirre M, Reverter F, Cp SG, Sousa A, Amadoz A, et al. The effects of death and post-mortem cold ischemia on human tissue transcriptomes. *Nat Commun.* (2018) 9:490. doi: 10.1038/s41467-017-02772-x
40. Quinn TP, Erb I, Richardson MF, Crowley TM. Understanding sequencing data as compositions: an outlook and review. *Bioinformatics.* (2018) 34:2870–8.
41. Aitchison J, Barceló-Vidal C, Martín-Fernández JA, Pawłowsky-Glahn V. Logratio analysis and compositional distance. *Math Geol.* (2000) 32:271–5.
42. Bar-Joseph Z, Gifford DK, Jaakkola TS. Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics.* (2001) 17(Suppl. 1):S22–9. doi: 10.1093/bioinformatics/17.suppl_1.s22
43. McInnes L, Healy J, Melville J. Umap: uniform manifold approximation and projection for dimension reduction. *arXiv preprint.* (2018): doi: <PMID>ISBN: 1802.03426,
44. Ashman KM, Bird CM, Zepf SE. Detecting bimodality in astrometrical datasets. *Astronomical J.* (1994) 108:2348.
45. Dressel U, Bailey PJ, Wang SCM, Downes M, Evans RM, Muscat GEO. A dynamic role for HDAC7 in MEF2-mediated muscle differentiation*. *J Biol Chem.* (2001) 276:17007–13. doi: 10.1074/jbc.M101508200
46. McKinsey TA, Olson EN. Dual roles of histone deacetylases in the control of cardiac growth. *Novartis Found Symp.* (2004) 259:132–41; discussion 141–5, 163–169.
47. McKinsey TA, Zhang CL, Olson EN. Signaling chromatin to make muscle. *Curr Opin Cell Biol.* (2002) 14:763–72.
48. Backs J, Olson EN. Control of cardiac growth by histone acetylation/deacetylation. *Circ Res.* (2006) 98:15–24.
49. Boon R. Metabolic fuel for epigenetic: nuclear production meets local consumption. *Front Genet.* (2021) 12:768996. doi: 10.3389/fgene.2021.768996
50. Nirello VD, Rodrigues de Paula D, Araújo NVP, Varga-Weisz PD. Does chromatin function as a metabolite reservoir? *Trends Biochem Sci.* (2022): [Online ahead of print], doi: 10.1016/j.tbs.2022.03.016
51. Kouzarides T. Acetylation: a regulatory modification to rival phosphorylation? *Embo J.* (2000) 19:1176–9.
52. Kane AE, Sinclair DA. Sirtuins and NAD⁺ in the development and treatment of metabolic and cardiovascular diseases. *Circ Res.* (2018) 123:868–85.
53. Soni SK, Basu P, Singaravel M, Sharma R, Pandi-Perumal SR, Cardinali DP, et al. Sirtuins and the circadian clock interplay in cardioprotection: focus on sirtuin 1. *Cell Mol Life Sci CMLS.* (2021) 78:2503–15. doi: 10.1007/s00018-020-03713-6
54. Cakir B, Tanaka Y, Kiral FR, Xiang Y, Dagliyan O, Wang J, et al. Expression of the transcription factor PU.1 induces the generation of microglia-like cells in human cortical organoids. *Nat Commun.* (2022) 13:430. doi: 10.1038/s41467-022-28043-y
55. Francois A, Canella A, Marcho LM, Stratton MS. Protein acetylation in cardiac aging. *J Mol Cell Cardiol.* (2021) 157:90–7.
56. Wright S. *Correlation and Causation.* Washington, DC: USDA (1921). p. 557–85.
57. Iacobas S, Amuzescu B, Iacobas DA. Transcriptomic uniqueness and commonality of the ion channels and transporters in the four heart chambers. *Sci Rep.* (2021) 11:2743. doi: 10.1038/s41598-021-82383-1
58. He A, Gu F, Hu Y, Ma Q, Ye LY, Akiyama JA, et al. Dynamic GATA4 enhancers shape the chromatin landscape central to heart development and disease. *Nat Commun.* (2014) 5:4907. doi: 10.1038/ncomms5907
59. Riddell A, McBride M, Braun T, Nicklin SA, Cameron E, Loughrey CM, et al. RUNX1: an emerging therapeutic target for cardiovascular disease. *Cardiovasc Res.* (2020) 116:1410–23.
60. Zhang D, Liang C, Li P, Yang L, Hao Z, Kong L, et al. Runt-related transcription factor 1 (Runx1) aggravates pathological cardiac hypertrophy by promoting p53 expression. *J Cell Mol Med.* (2021) 25:7867–77. doi: 10.1111/jcmm.16704
61. McCarroll CS, He W, Foote K, Bradley A, McGlynn K, Vidler F, et al. Runx1 deficiency protects against adverse cardiac remodeling after myocardial infarction. *Circulation.* (2018) 137:57–70. doi: 10.1161/CIRCULATIONAHA.117.028911
62. Palioura D, Lazou A, Drosatos K. Krüppel-like factor (KLF)5: an emerging foe of cardiovascular health. *J Mol Cell Cardiol.* (2022) 163:56–66. doi: 10.1016/j.jmcc.2021.10.002
63. Hoffman M, Palioura D I, Kyriazis D, Cimini M, Badolia R, Rajan S, et al. Cardiomyocyte krüppel-like factor 5 promotes de novo ceramide biosynthesis and contributes to eccentric remodeling in ischemic cardiomyopathy. *Circulation.* (2021) 143:1139–56. doi: 10.1161/CIRCULATIONAHA.120.047420
64. Li Z, Langhans SA. Transcriptional regulators of Na,K-ATPase subunits. *Front Cell Dev Biol.* (2015) 3:66. doi: 10.3389/fcell.2015.00066
65. Kfoury N, Qi Z, Prager BC, Wilkinson MN, Broestl L, Berrett KC, et al. Brd4-bound enhancers drive cell-intrinsic sex differences in glioblastoma. *Proc Natl Acad Sci U S A.* (2021) 118:e2017148118. doi: 10.1073/pnas.2017148118
66. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell.* (2016) 167:1853–66.e17. doi: 10.1016/j.cell.2016.11.038
67. Replogle JM, Saunders RA, Pogson AN, Hussmann JA, Lenail A, Guna A, et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell.* (2022): [Online ahead of print], doi: 10.1016/j.cell.2022.05.013
68. Walker CJ, Schroeder ME, Aguado BA, Anseth KS, Leinwand LA. Matters of the heart: cellular sex differences. *J Mol Cell Cardiol.* (2021) 160:42–55.
69. Litviouková M, Talavera-López C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. *Nature.* (2020) 588:466–72.
70. Chun H, Keles S. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *J Royal Stat Soc Series B (Stat Methodol).* (2010) 72:3–25.