Co-expression of calcium channels and delayed rectifier potassium channels protects the heart from proarrhythmic events

Sara Ballouz[a,1], Melissa M Mangala[b,1], Matthew D Perry[b,c,1], Stewart Heitmann[b], Jesse A Gillis[a], Adam P Hill[b,c,2], Jamie I Vandenberg[b,c,2]

Author Affiliation:

a. Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Woodbury, NY 11797, USA.
b. Victor Chang Cardiac Research Institute, Darlinghurst, NSW 2010, Australia
c. University of NSW, Sydney. Kensington, NSW 2052, Australia

1. These authors contributed equally to this study

2. Corresponding Authors:

Jamie I Vandenberg* and Adam P Hill†,
Victor Chang Cardiac Research Institute,
Lowy Packer Building, 405 Liverpool Street
Darlinghurst, NSW 2010
Australia

* j.vandenberg@victorchang.edu.au
† a.hill@victorchang.edu.au

Abstract (140 words)

Cardiac electrical activity is controlled by the carefully orchestrated activity of more than a dozen different ion conductances. Yet, there is considerable variability in cardiac ion channel expression levels both within and between subjects. In this study we tested the hypothesis that variations in ion channel expression between individuals are not random but rather there are modules of co-expressed genes and that these modules make electrical signaling in the heart more robust.

Meta-analysis of 3653 public RNA-Seq datasets identified a strong correlation between expression of *CACNA1c* (L-type calcium current, I_{CaL}) and *KCNH2* (rapid delayed rectifier K^+ current, I_{Kr}), which was verified in mRNA extracted from human induced pluripotent stem cell-derived cardiomyocytes. *In silico* modeling indicates that the co-expression of *CACNA1c* and *KCNH2* limits the variability in action potential duration and reduces susceptibility to early afterdepolarizations, a surrogate marker for pro-arrhythmia.

Keywords:

Ion channels; co-expression; gene modules; meta-analysis; cardiac action potential, early afterdepolarization, arrhythmia, drug-induced long QT syndrome

Introduction

Robust electrical signaling in the heart is critical for co-ordinating the efficient pumping of blood around the body. Failure of cardiac electrical signaling, even for just a few minutes, can have fatal consequences, with sudden cardiac death accounting for up to 10% of deaths in our community (1). Despite decades of research, predicting in advance who is more or less susceptible to sudden cardiac death remains challenging (2).

The action potential (AP) of excitable cells, such as cardiac myocytes and neurons, reflects the orchestrated activity of at least a dozen distinct ion channels and electrogenic transporters (3) (4). In such complex systems, both theoretical and experimental studies have shown that there is considerable inter-individual variability in the combinations of molecular input parameters that can produce very similar integrated outputs (5) (6). This has led to a paradigm shift in computational modeling that relies not on generating idealized outputs based on mean data but rather development of populations of models that account for the observed variability in molecular inputs (7) (8). Such models are already proving useful for interrogating inter-individual variability in response to pathological stimuli (9) (10) (11). Our challenge now is to discern the underlying essence of these complex systems (12,13) so that we may then make rational interventions to treat pathology that takes into account inter-individual variation. Specifically, are there underlying principles regulating cardiac electrical activity that can provide insights into why some people are more susceptible to sudden cardiac death in response to pathological

stimuli, such as drug block of the rapid delayed rectifier potassium channel (I_{Kr}) which is the underlying basis of drug-induced long QT syndrome (di-LQTS) (14).

A common approach to discern patterns in multi-dimensional biological problems has been to look for co-expression networks (13) (15). Co-expression networks are known to encode functional information (16), with co-expression reflecting co-regulation and co-functionality (17) (18). To help identify robust co-expression modules, it is helpful to use meta-analytical approaches, as the aggregation of large numbers of individual networks across multiple independent experiments averages away noise and reinforces those correlations that reflect real signals (19) (20) (21).

Here, we have used meta-analytic co-expression analysis in large scale human gene expression data sets to identify modules of co-expressed ion channel genes which were then used to constrain population models of cardiac electrical activity. These models were then used to test the hypothesis that co-expression of repolarization and depolarization currents helps prevent irregular action potentials from emerging when human cardiac myocytes are exposed to pro-arrhythmic stimuli. We show that tight coupling of current densities for the L-type calcium current (I_{CaL}) and the rapid component of the delayed rectifier potassium current (I_{Kr}) reduced the emergence of pro-arrhythmic early afterdepolarizations (EADs) and this protection persisted in the face of highly variable expression of other ion channels, as well as in the presence of pharmacological block of I_{Kr}, a potent pro-arrhythmic stimulus (22). A very important prediction to arise out of our modelling studies is that in the context of drug block of I_{Kr} those patients with high expression of I_{CaL} and I_{Kr} experience more EADs and are therefore more likely to be susceptible to ventricular arrhythmias.

Results

The shape and duration of action potentials in cardiac myocytes are determined by the orchestrated activity of voltage-gated sodium, calcium and potassium channels, as well as a series of electrogenic transporters that regulate intracellular ion concentrations (See Supplementary data **Figure S1**). These channels, transporters and related intracellular calcium handling proteins are encoded by a few dozen genes, sometimes referred to as the rhythmonome (23) (see Supplementary data, **Table 1**).

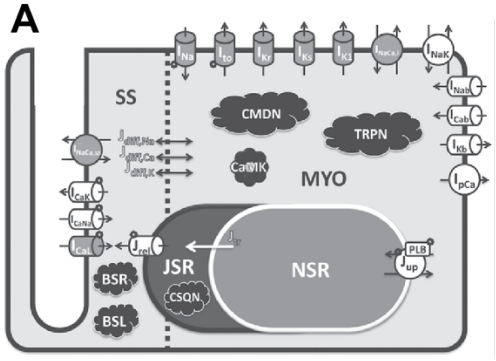
To determine whether there were any co-expression patterns among the rhythmonome genes we first undertook an untargeted screen for possible expression correlation patterns in publicly available RNA-seq data sets (see list of RNA-seq experiments in supplementary data, **Table 2**). Ranked correlation coefficients from an aggregate co-expression network that contain data from 3653 samples are illustrated in the heatmap in **Figure 1A**. High ranked correlations indicate similarity of transcriptional profiles between the genes. A clustering analysis, as shown by the dendrogram in **Figure 1A**, groups genes according to their correlation similarities, as defined by the Spearman's correlation coefficients (see

color code in **Figure 1A**). There is a large cluster of yellow/green squares in the bottom left corner indicating that there are significant levels of correlation amongst many of the genes. Furthermore, within this large cluster there are two sub-clusters. The cluster of yellow-green squares corresponding to 13 genes in the bottom left corner (red dashed box in **Figure 1A**) encode for proteins that regulate calcium fluxes as well as the transient outward K^+ current (*KCND3*), which helps to maintain the plateau potential at a level that maximizes calcium influx (24). A second sub-cluster, in the upper right quadrant of the main cluster, encompasses 10 genes (black dashed box in **Figure 1A),** including *KCNH2, SCN5a, KCNJ12* and *KCNIP2*, that encode for ion channel proteins important for regulating action potential duration (APD).

Overall, the connectivity within the set of rhythmonome genes is fairly high in comparison to their connectivity to all other genes in the co-expression network (node degree analysis, p~5.7e-14), with a central gene being *CACNA1C*, the gene encoding the alpha subunit of the L-type calcium channel (see Supplementary data, **Figure S2**). Although *CACNA1c* clusters in the group of calcium handling genes in the bottom left quadrant of the main cluster in **Figure 1A**, it also shows high levels of correlation with the cluster of ion channel genes in the top right quadrant of **Figure 1A**. In a similar fashion, the *KCNH2* gene, which encodes for I_{Kr}, is included in the ion channel cluster but also shows moderate-high correlations with a portion of the calcium handling genes in the bottom left quadrant. The highest ranked co-expression partners for *CACNA1C* and *KCNH2* are highlighted in **Figure 1B.**

The vast majority of the public RNA-Seq datasets included in our analyses were not heart specific. It is, however, noteworthy that there are no strong correlations between the expression of any of the individual ion channel or calcium handling genes and the cardiac-specific markers included in our analyses: *GATA-4*, *NKX2-5*, *MYL2* and *MYL7*. For example, the cardiac marker genes are towards the bottom of the lists in **Figure 1B**. This suggests that the correlations within the set of rhythmonome genes are not simply a reflection of cardiac specific expression but rather represent intrinsic correlations.

Figure S1: *In silico* model of ventricular Action Potential



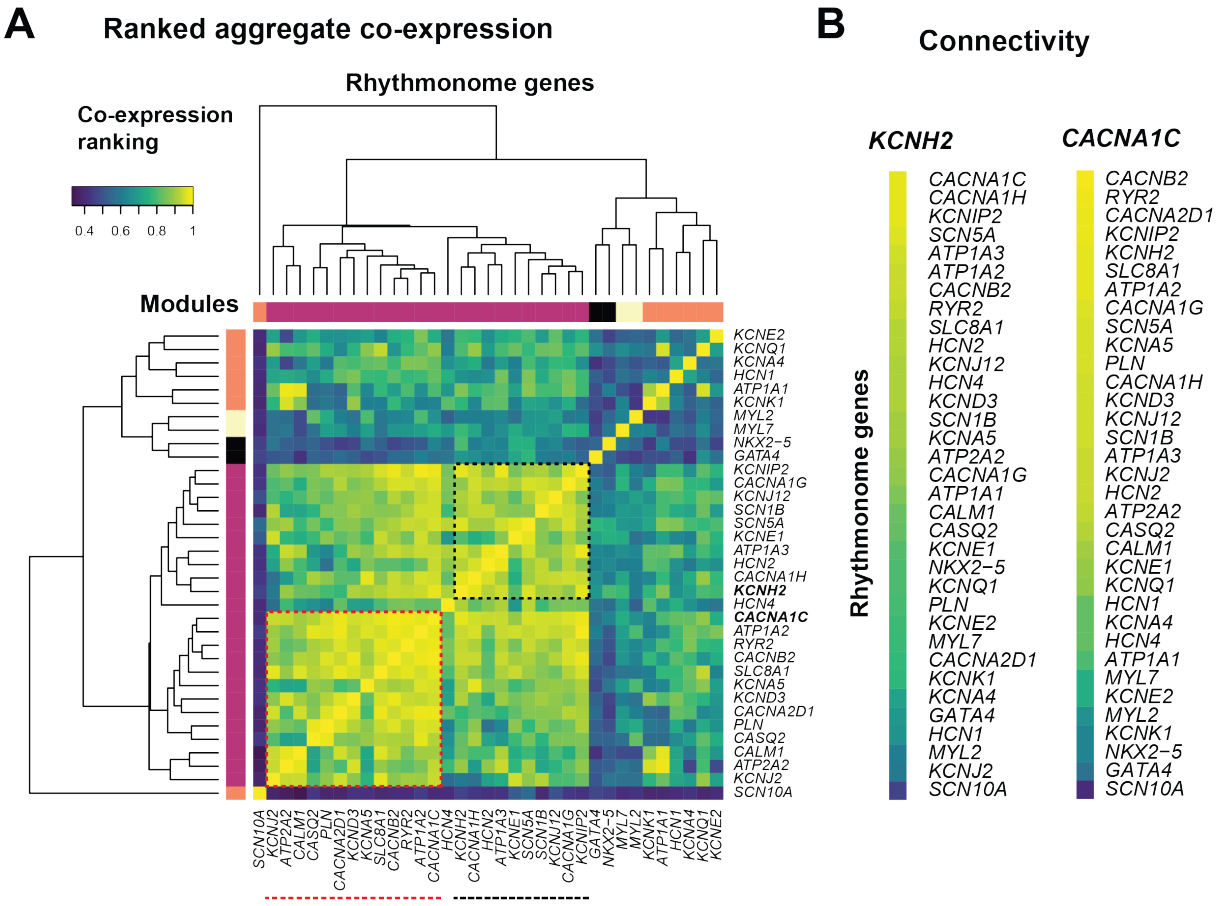
B

Conductance parameter	scalar
GKs	8.09
GKr	1.17
GCaL (PCa)	3.57
PNaCa (Gncx)	3.05
PNaK	1.91
GNaL	1.7

A. Schematic diagram of the ventricular myocyte model used in this study (reproduced from (25) under the CC-BY licence.

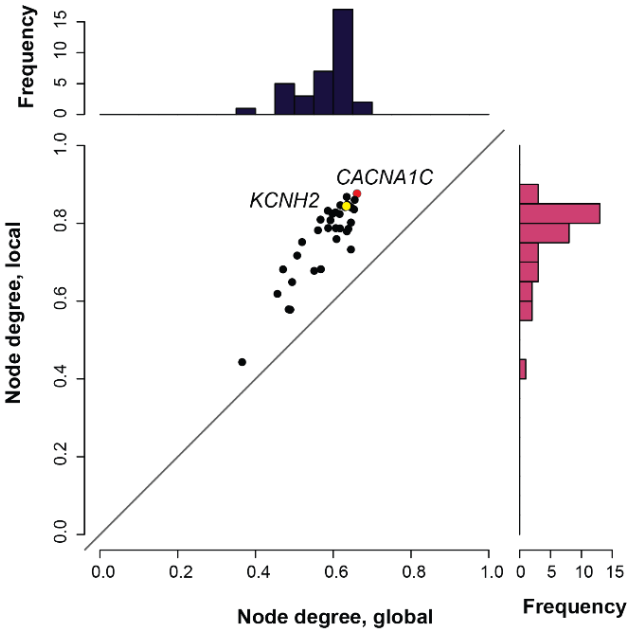
B. Scaling factors applied to the original model were taken from (26)

Figure 1: Cardiac ion channel and calcium handling protein co-expression connectivity analysis from published RNA-seq datasets



A. Sub-network of co-expression ranking across 35 genes encoding for cardiac ion channels or calcium homeostasis proteins (see Supplementary Table 1). Blue: low ranked correlations, yellow: high ranked correlations. The dendrogram illustrates the modules of genes with high levels of similarity in their transcriptional profiles. Red and black dashed boxes highlight subnetworks of correlated genes **B**. Connectivity of genes to *KCNH2* (encodes for I_{Kr}) and *CACNA1C* (encodes for I_{CaL}).

Figure S2: Network connectivity for each rhythmonome gene



The degree of network connectivity for each rhythmonome gene within the local rhythmonome network (y-axis) versus its connectivity within the whole genome (x-axis). All the rhythmonome genes are more highly connected within the local network (i.e., all the points lie above the line of identity, $p \sim 5.7e^{-14}$). *CACNA1C* is highlighted in red and *KCNH2* in yellow.

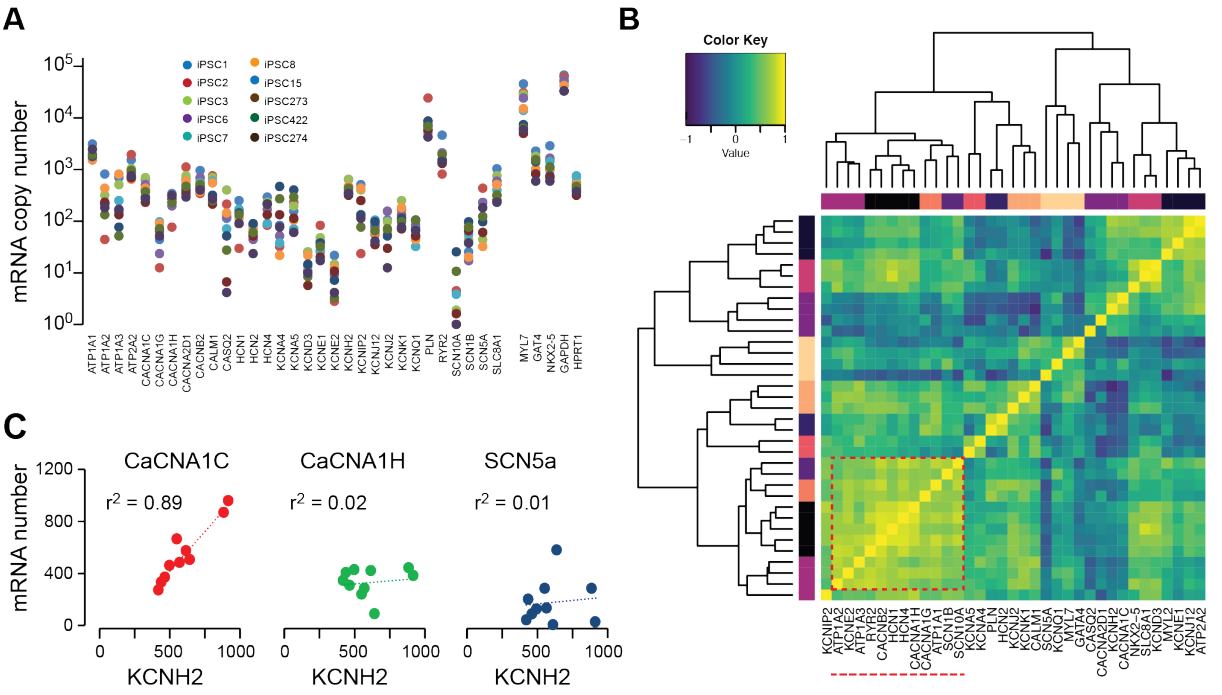
To investigate whether the meta-analytic co-expression patterns observed in **Figure 1** were also seen in human heart cells, we extracted mRNA from human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) (27) obtained from 10 patients with no known heart disease**.** All samples contained high levels of cardiac marker genes (*MYL7, GATA4* and *NKX2*.5) **(Figure 2A)**. Furthermore, the levels of expression of the housekeeper genes, *GPADH* and *HRPT1* were similar across samples (see data at right side of **Figure 2A**). The levels of expression of most rhythmonome genes showed variations between the samples that spanned approximately an order of magnitude. However, similar to the generic tissue RNA-Seq datasets, there were modules of co-expressed genes (e.g., see dashed box in bottom left quadrant of **Figure 2B**). Most of the 13 genes in the cluster in **Figure 2B** are present in the modules highlighted by the black and red dashed boxes in **Figure 1A**. Conversely, many of the ion channel genes contained in the modules highlighted in **Figure** 1 are not present in the module in **Figure 2** (e.g. *KCNJ2/KCNJ12*, which encodes for I_{K1}; *KCND3*, which encodes for I_{To}; and *SCN5a*, which encodes for I_{Na}). These genes are all known to be expressed at lower levels in embryonic hearts and so unsurprisingly they are not well expressed in the hiPSC-CM lines (28).

We next looked to see if there were any specific relationships between genes encoding depolarization and repolarization currents, within the hiPSC-CM expression profiles. In **Figure 2C,** we have plotted the expression of *KCNH2* versus the genes that encode for the depolarization currents that showed the highest levels of correlation with expression of *KCNH2* in the generic tissue datasets (see **Figure 1B**, i.e., *CACNA1c*, *CACNA1H* and *SCN5a*).

The most notable correlation that was observed in the hiPSC-CMs was that between *KCNH2* and *CACNA1c*; $r^2 = 0.89$ (**Figure 2C**).

As the only robust relationship that we observed in both the generic tissue sets and the hiPSC-CM lines was the co-expression of *KCNH2* and *CACNA1c*, we focused on this pair for our subsequent studies. To investigate whether co-expression modules of ion channel genes might influence integrated cardiac electrical function, we used an *in silico* approach. First, we simulated a population of 1000 human cardiac action potentials where random scalar values, chosen from a log normal distribution with mean 1 and standard deviation of 0.5 (**Figure 3A**, lower panel), was applied to every conductance in each iteration of the action potential model (see **Figure 3A** and Supplementary Data, **Movie S1**).

Figure 2: HiPSC-CM mRNA correlation analysis



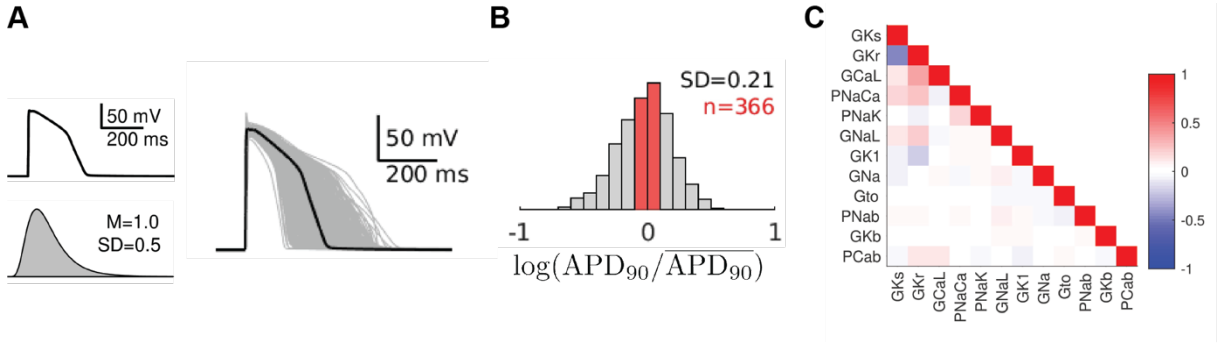
A. plot of all mRNAs (log axis) for 10 hiPSC-CM lines. Note that there are high levels of expression of cardiac markers (*MYL7* as well as *NKX2-5* and *GATA4*) in all cell lines. Also there is less variability in the levels for the housekeeping genes (*GAPDH* and *HPRT1*) compared to that seen for the ion channels and calcium handling proteins. **B.** Correlation matrix for hiPSC-CM mRNA. The red dashed box highlights the module with the highest correlations. **C.** Correlation plots for *KCNH2* versus genes that encode the major depolarizing currents (*CACNA1c*, *CACNA1H* and *SCN5a*).

The baseline action potential produced by this model (black trace in right hand panel of **Figure 3A**) has a duration at the point of 90% repolarization (APD_{90}) of 264 ms. The population of 1000 cardiac cells generated by randomly scaling the conductances exhibited APD_{90} values that ranged from ~120 ms to ~500 ms (**Figure 3A,** right panel, and **Figure 3B**). We next selected those cells with APD_{90} values that fell within 20% of the mean value

(red bars, **Figure 3B**) to determine if there were any patterns of ion channel co-expression that could contribute to keeping the APD_{90} within this narrow selected range. The correlation matrix of the conductance scaling factors for the selected cells (**Figure 3C**) reveals a positive correlation between G_{Kr} and G_{CaL} (R=0.36) as well as an inverse correlation for G_{Kr} and G_{Ks} (R=-0.46). The positive correlation between the conductance scalars G_{Kr} and G_{CaL} in the *in silico* modelling dataset (**Figure 3**) suggests that the correlation seen between *CACNA1c* and *KCNH2* mRNA expression in both the public RNA-Seq datasets (**Figure 1**) and the hiPSC-CM dataset (**Figure 2**) would contribute to reducing the population variability in APD_{90} values.

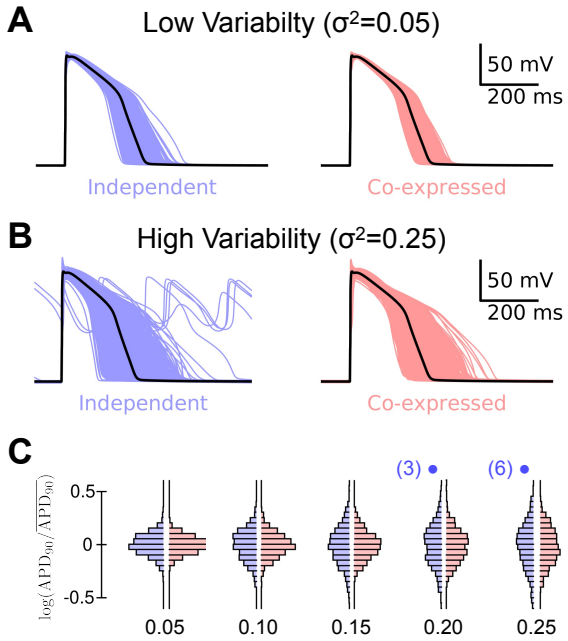
We next repeated our previous simulation of 1000 action potentials but forced the conductance scaling factors for I_{Kr} and I_{CaL} to be identical in each cell (denoted co-expression in **Figure 4,** also see Data supplement, **Movie S2**). The scaling factors for all other conductances remained independent. The distribution of APD_{90} values for both independent and co-expression cell populations becomes broader as the level of variability is increased (**Figure 4A-C**). However, the spread of APD_{90} values in the cells with identical G_{CaL}-G_{Kr} scalars is always narrower than in the cell populations with independent G_{CaL}-G_{Kr} scalar values. For example, in the case of **Figure 4B,** the variance of the APD_{90} values was 0.026 for the co-expression dataset but 0.046 for the independent dataset (see supplementary data **Figure S3-C**). Another notable feature of the data in **Figure 4C** is that early afterdepolarisations (EADs) begin to appear in the cell population with independent scalars when the scalar variability, σ^2, exceeds 0.20 (also see supplementary data **Figure S4**). The number of cells with an EAD are indicated in parentheses above each distribution in **Figure 4C.**

Figure 3: ***In silico*** **predictions of conductance scalars that give APD_{90} values in the normal range**



A. Baseline AP model and frequency histogram of scalars (log normal distribution with a mean of 1 and SD of 0.5) used to generate the population of 1000 APs with the baseline AP (all scalars set to 1.0) shown in black. **B.** Frequency histogram of APD_{90} values for the 1000 simulations with those falling within ±20% of the mean value shown in red. **C.** correlation matrix for sets of conductance scalars that gave APD_{90} values within ±20% of the mean value

Figure 4: Impact of I_{Kr}-I_{CaL} co-expression and conductance scalar variability on APD90 variability

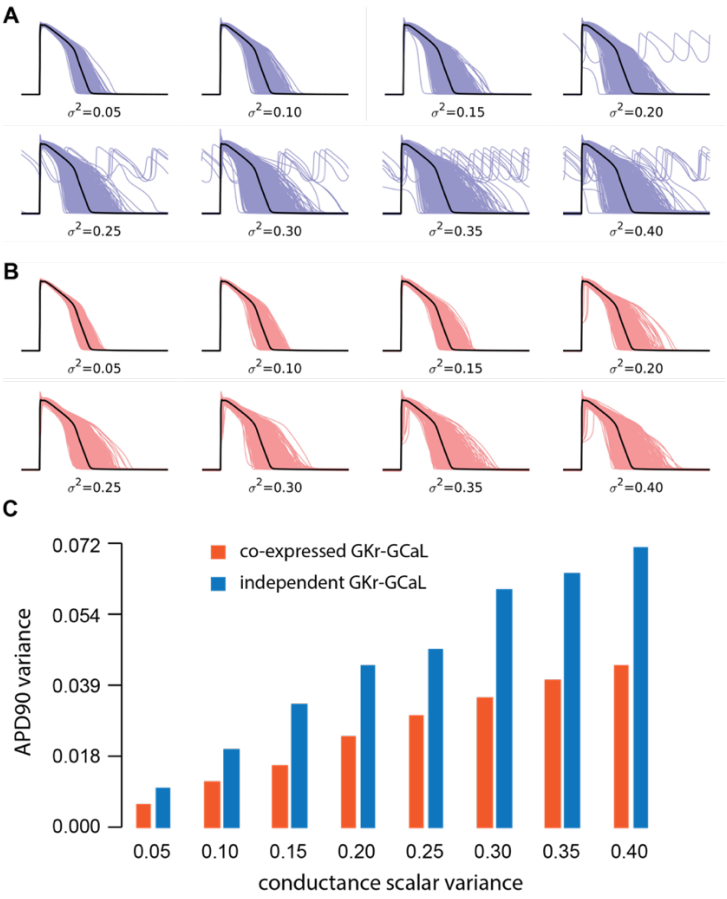


Raw APs for independent (blue) and co-expression of G_{Kr}-G_{CaL} (red) for **A.** low scalar variability (σ^2=0.05) and **B.** high scalar variability (σ^2=0.25). Note the presence of EADs in 6 of the APs in the independent group with $\sigma^2 = 0.25$. **C.** Histograms of log(APD_{90}/mean APD_{90}) distributions for co-expressed (red) or independent (blue) G_{Kr}-G_{CaL} simulations with $\sigma^2 = 0.05$, 0.1, 0.15, 0.2, 0.25. The numbers in parentheses above the 0.2 and 0.25 groups indicate the number of EADs in each independent scalars group.

Movie S1

Movie S2

Figure S3: Population models of ventricular APs with different levels of scalar variance

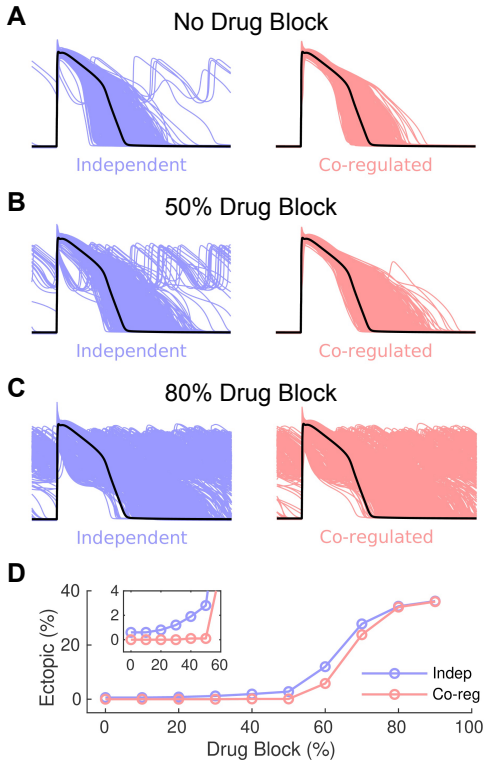


A-B. Population of models with variance systematically increased from σ2 = 0.05 to 0.40 for independent conductance scalars for G_{Kr} and G_{CaL} **(A), or** co-expression conductance scalars for G_{Kr} and G_{CaL} **(B)**. As the variance increases, the number of cells that develop EADs increases in the independent series. Conversely, for the co-expression models, there are no EADs even at variances up to 0.4. **C.** Variance of the APD_{90} distributions for each set of conductance scalar inputs for independent (blue) and co-expression (red) of I_{Kr} and I_{CaL}. Note that the variance of the APD90 distribution was always greater for the independent simulations ($p<0.008$ between groups, paired t-test)

We next investigated whether coupling of the conductance scalars for G_{Kr} and G_{CaL} influenced the generation of EADs in response to a pathological stimulus. Specifically, how cells respond to drug block of I_{Kr}, the underlying cause of di-LQTS (22). Example populations of action potentials obtained for independent and co-expression populations with I_{Kr} block of 0%, 50% and 80% are illustrated in **Figure 5.** As the extent of I_{Kr} block is increased (from A to C), the proportion of simulated action potentials producing EADs increases. It is also clear that at lower levels of I_{Kr} block, EADs were more frequent when G_{CaL} and G_{Kr} scalars were modulated independently (see **Figure 5B** and inset to **Figure 5D**). However, the proportion of simulations developing EADs in both the independent and co-expression populations becomes similar when the extent of I_{Kr} block exceeds 80% (**Figure 5D**).

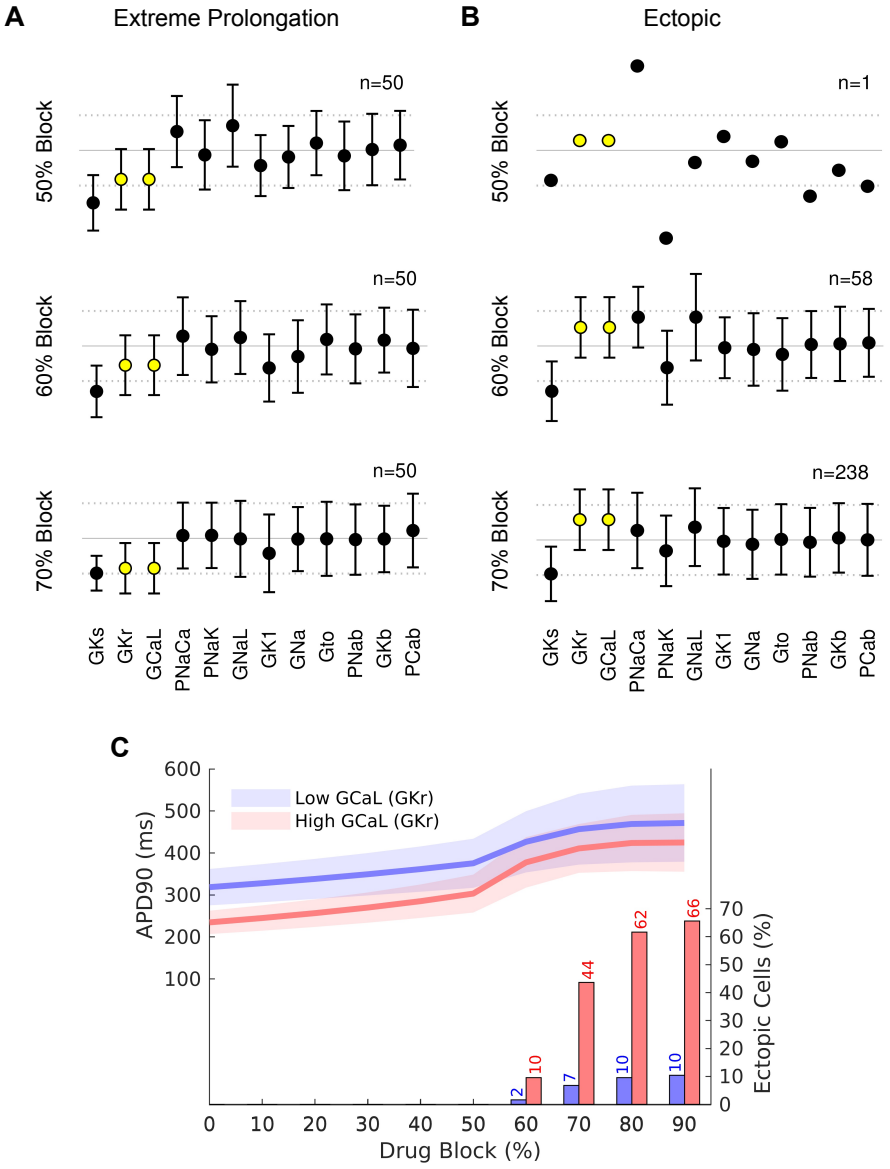
It is well established that only a subset of patients exposed to drugs that block I_{Kr} (29), or with a mutation causing 50% loss of I_{Kr} function (30), will develop life threatening arrhythmias. This is consistent with the prediction made by our simulated drug block experiments shown in **Figure 5**. We therefore asked whether the data from the co-expression datasets could tell us anything about what factors might predispose to the development of EADs in the presence of a drug that blocks I_{Kr}. Analysis of the subset of scalars within the co-expression dataset that produced the 50 longest APs without EADs, compared to the subset of scalars that resulted in APs with EADs, is illustrated in **Figure 6A and 6B** respectively. Notably, the longest APs without EADs had low G_{CaL} scalars (and hence low G_{Kr} scalars) before addition of drug block. Conversely, the APs that developed EADs had higher G_{CaL} and G_{Kr} scalars. In **Figure 6C**, we have plotted the APD_{90} values for cells in the highest (red) and lowest (blue) quartiles of G_{Cal} - G_{Kr} scalars. As expected, the low G_{CaL} group showed longer APD_{90} values on average compared to the high G_{CaL} group (see the continuous lines in **Figure 6C**). Furthermore, for the 70% I_{Kr} block scenario, 44% of the high G_{CaL} group have developed EADs whereas only 7% of the low G_{CaL} group have developed EADs (compare red and blue bars at 70% drug block in **Figure 6C**). Thus, higher G_{CaL} is associated with a greater risk of developing EADs in response to moderate levels of I_{Kr} block. A similar pattern of results was observed when G_{CaL} and G_{Kr} were allowed to vary independently, except that in this scenario the difference between the high G_{CaL} and low G_{CaL} groups was even more dramatic at lower levels of I_{Kr} block (see supplementary data **Figure S4**).

Figure 5: Impact of G_{Kr}-G_{CaL} co-expression on response to drug block of I_{Kr}



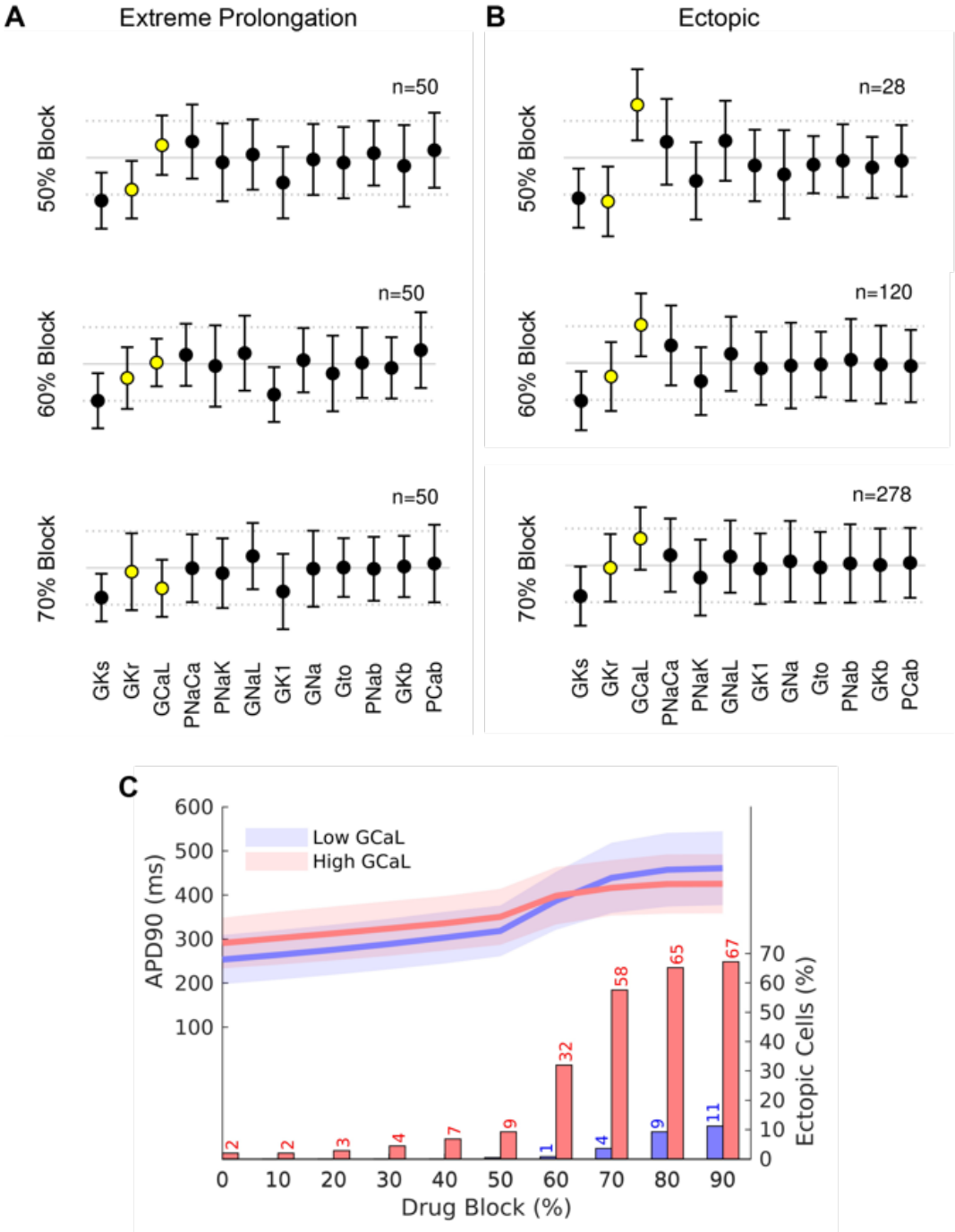
Raw APs for co-expressed G_{Kr}-G_{CaL} (red) and independent (blue) G_{Kr}-G_{CaL} populations, with scalar variance $\sigma^2 = 0.25$, for **A.** control, **B.** 50% I_{Kr} block and **C.** 80% I_{Kr} block. **D.** plot of % of AP simulations with EADs versus % I_{Kr} block (co-expression, red and independent, blue). The inset in panel D shows an expanded view of the % EADs at small levels of I_{Kr} block.

Figure 6: High G_{CaL} increases risk of EAD formation at moderate I_{Kr} block



Scalars (box and whisker plots) for the population of cells with G_{Kr}-G_{CaL} co-expression. **A.** 50 cells with the longest APD_{90} values and no EAD with increasing % I_{Kr} block. **B.** scalars (box and whisker plots) for cells with EADs with increasing % I_{Kr} block. The G_{CaL} and G_{Kr} scalars are highlighted in yellow. **C.** APD_{90} values (left axis) for the highest (red) and lowest (blue) quartile of cells according to baseline G_{CaL} scalar. The continuous lines show the mean value and shaded area shows ±1 SD). The percentage of EADs in each quartile is shown as columns (see axis on right side of graph). The corresponding plots for the population of cells with independent G_{Kr}-G_{CaL} scalars are show in the Supplementary data (**Figure S4**)

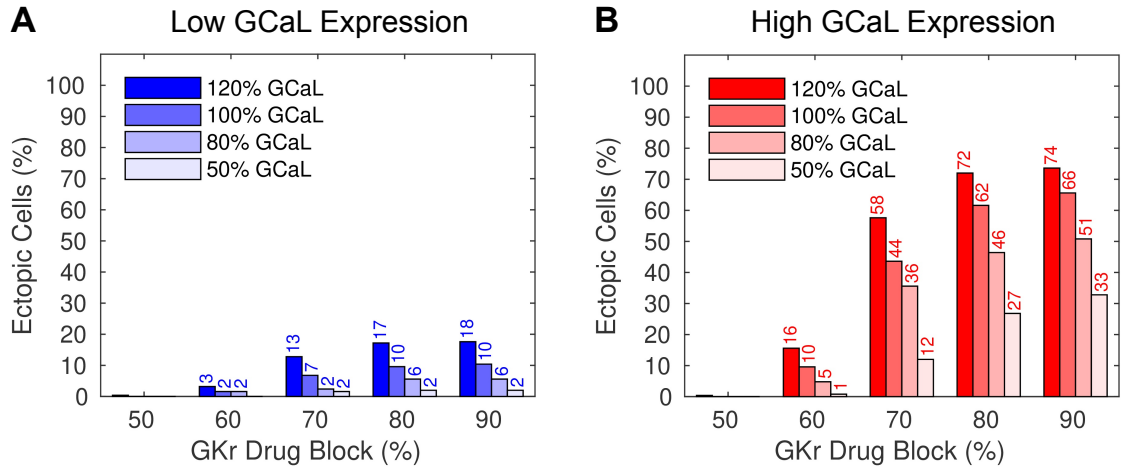
Figure S4: High G_{CaL} increases risk of EAD formation at moderate I_{Kr} block even when G_{CaL} and G_{Kr} are varied independently



Conductance scalars (box and whisker plots) that give **A.** The 50 most extreme APD values or **B.** EADs for the population of cells with independent regulation of G_{Kr} and G_{CaL}. **C.** Plot of APD_{90} values for the highest quartile of G_{CaL} (red) and lowest quartile of G_{CaL} (blue) versus % drug block of I_{Kr}. The continuous line shows mean and shaded regions show ± 1 standard deviation. The column graph shows the % of cells that develop EADs at each level of drug block.

A corollary of our prediction that patients with high G_{CaL} are more susceptible to EADs when exposed to a drug that blocks I_{Kr}, is that co-administration of a drug that blocks I_{CaL} would reduce the incidence of EADs. Magnesium, which is used in the acute management of di-LQTS (14), is a weak calcium channel blocker. Raising plasma [Mg^{2+}] from 1.5 to 2.5 mM would be expected to inhibit I_{CaL} by ~20% (31), conversely reducing [Mg^{2+}] from 1.5 to 0.5 mM would be expected to increase I_{CaL} by ~20%. When we increased I_{CaL} by 20%, the incidence of EADs in the high G_{CaL} group increased from 10% to 16% for the 60% I_{Kr} block simulation and from 44% to 58% for the 70% I_{Kr} block simulations (see **Figure 7)**. A drug that inhibited I_{CaL} by 20% caused a modest decrease in the percentage of cells with EADs and reduction of I_{CaL} by 50% had a correspondingly larger effect, for example, reducing EADs from 44% to 12% in the 70% I_{Kr} block scenario (**Figure 7**).

Figure 7: Effect of modifying I_{CaL} on incidence of EADs in response to I_{Kr} drug block.



I_{CaL} was increased 20% to mimic hypomagnesaemia, reduced 20% to mimic hypermagnesaemia (see text for details), or reduced 50% to mimic administration of a calcium channel blocker. In both the quartile of cells with the lowest G_{CaL} scalars (blue, **A**) and the quartile of cells with the highest G_{CaL} scalars (red, **B**) the incidence of EADs increased when G_{CaL} was enhanced and decreased when G_{CaL} was decreased. The observed differences were most pronounced with moderate (60-70%) I_{Kr} drug block.

Discussion

Cardiac electrical activity is regulated by the interdependent activity of a plethora of ion channels, transporters and calcium handling proteins (4). Understanding the precise details of how these conductances interact to control the rhythm of the heart has been an enduring source of fascination. In this study, we have used meta-analytical techniques to interrogate the large numbers of RNASeq datasets that have been deposited in public access databases, to look for patterns of co-expressed genes that might help decipher the control of cardiac electrical activity. The most important pair of co-expressed genes, identified both in the public RNA Seq datasets (**Figure 1**) and in heart cells derived from

hiPSC lines (**Figure 2**), was that of *CACNA1c* (I_{CaL}) and *KCNH2* (I_{Kr}). Using an *in silico* approach we demonstrated that tight co-expression of *CACNA1c* and *KCNH2*, against a background of variability of all other ion channels, helps to control the duration of repolarization (**Figure 4**). More importantly, the co-expression of *CACNA1c* and *KCNH2* helps to protect the heart from early after depolarizations when they are exposed to drugs that block I_{Kr} (**Figure 5**). Our simulations also suggest that inter-individual differences in pro-arrhythmic responses to I_{Kr} drug block can be explained by inter-individual differences in levels of *CACNA1c* expression (**Figure 6**).

Over the last few years, numerous groups have shown that there is considerable heterogeneity of ion channel expression amongst excitable cells, including neurons (32) and cardiac myocytes (7,19). Furthermore, since the pioneering work of Eva Marder and colleagues, the presence of modules of co-expressed ion channel genes has also been well appreciated(15). These previous studies, however, relied on patch clamp analysis of isolated cells (15,33) and qPCR analysis (15) of individual ion channel mRNA, which are highly laborious and so have been restricted to only a few important ion channel genes. The advent of high throughput transcriptomic analyses has greatly facilitated the identification of conserved networks of co-expressed gene modules amongst the entire set of expressed genes (21). By applying meta-analytic approaches to a large number of independent datasets one can more readily discern genuine co-expression signals from noise, as well as explore smaller gene modules (21). In our analysis of a subset of rhythmonome genes (35 members, **Table S1**) within a large library of public RNA Seq datasets, (see **Table S2**) we identified two clusters of genes: a first subset that predominantly affects calcium handling and a second subset that predominantly affects membrane potential (see **Figure 1**). The nodes within each of these clusters that are most closely connected within the rhythmonome relative to all other genes are *KCNH2* and *CACNA1c* (**Figure S2**). This is analogous to a network of networks (34), where the *KCNH2-CACNA1c* link provides an interconnection of the two networks. Independent evidence to corroborate an important link between calcium handling and regulation of cardiac action potential duration comes from the large genome-wide association studies (GWAS) of QT interval duration which identified SNPs in a number of calcium handling genes as well as *KCNH2* as being important determinants of QT interval in the population (35).

Due to the large number of associations we were testing for in our network analyses, it is possible that some would occur by chance. That we were able to confirm the presence of at least some of the co-expression modules in an independent dataset, i.e. the hiPSC-CM (see **Figure 2**) provides important corroborative evidence that these co-expression patterns are real and therefore likely to have physiological relevance. It should also be noted that many ion channels important for function in adult cardiac myocytes are not expressed at significant levels in immature cardiac myocytes, such as those derived from hiPSC (e.g. *SCN5a, KCNQ1, KCNJ12, KCND3*, (28)). It is therefore possible that we have

underestimated the number of genes within the modules of co-expressed genes in adult cardiac myocytes.

The patterns of co-expressed genes we observed show some important similarities, as well as differences to previous studies. Banyasz et al. (33) and Rees et al. (18) have both demonstrated that the expression of I_{CaL} is correlated with the sum of the major repolarizing ion currents in guinea-pig and mouse respectively. However, the molecular players involved in cardiac electrical activity in rodents are quite distinct to humans (36). In humans, the most important determinant of repolarization duration at baseline is I_{Kr} (37), whereas in guinea-pig I_{Ks} and I_{Kr} play equally important roles (33) and in mice the fast component of the transient outward current ($I_{to,f}$) and the ultra-rapid delayed rectifier (I_{Kur}) are the major repolarization currents (18). Thus, there is a common factor between our study showing co-expression of *KCNH2* and *CACNA1c* in human and the previous rodent studies, i.e., all studies show a correlation between I_{CaL} and the major repolarizing currents present in that species. Our study, however, is the first to demonstrate the important co-expressed genes in human heart tissue.

Identifying modules of co-expressed genes are the first step in seeking to understand the logic of complex systems (6). Understanding how such modules impact function in health and disease is the next challenge. In neuronal cells, Marder and colleagues have shown that modules of co-expressed ion channels play an important role in regulating action potential firing patterns (38). Rees and colleagues, have demonstrated that modules of co-expressed depolarization and repolarization currents can help to ensure normal amplitude calcium transients, a critical determinant of overall heart function (18). We have extended these studies to show that in normal heart cells, co-expression of *KCNH2* (repolarization) and *CACNA1c* (depolarization) channels help to maintain the plateau duration of the action potential, which in turn likely contributes to regulating the duration and amplitude of the calcium transient. More importantly, our studies provide the first insights into how patterns of co-expressed ion channel genes influence the hearts response to pathological stimuli.

Sudden death due to abnormalities of cardiac electrical signaling is a major cause of mortality (1). Predicting in advance who is more or less susceptible to sudden cardiac death and therefore warrants prophylactic treatment remains challenging (2). A key to being able to predict who is at greatest risk is understanding why different people respond differently to the same pro-arrhythmic stimulus. Based on the results of our *in silico* studies, we have provided two important insights into the nature of interindividual risk for developing arrhythmias in response to drugs that block I_{Kr}, the major cause of drug-induced cardiac arrhythmias (14). First, cells with low G_{CaL} (and hence low G_{Kr} at baseline) exhibited the greatest prolongation of AP duration when exposed to I_{Kr} drug block. Second, cells with high G_{CaL} (and hence high G_{Kr} at baseline) showed greater propensity for development of EADs at moderate levels of I_{Kr} drug-block (**Figure 6**). An important implication of the observation that a high G_{CaL} increases the susceptibility to EADs in

response to drug block of I_{Kr} is that the co-administration of an I_{CaL} blocker should reduce the risk of EADs (as shown in **Figure 7**). This is consistent with the observation that the administration of magnesium, which is a mild calcium channel blocker (31), is helpful in the acute management of patients with drug-induced *torsades de pointes* (14), and conversely that hypomagnesaemia, which would stimulate I_{CaL}, can exacerbate *torsades de pointes* (39). It is also consistent with the observation that drugs that block I_{CaL} and I_{Kr} (e.g., verapamil) are not associated with drug-induced arrhythmias (40) and that verapamil prevented the development of *torsades de pointes* in rabbit hearts exposed to an I_{Kr} blocker (41). However, given that calcium channel blockers are contra-indicated in some ventricular arrhythmias (42), and the likelihood that patients who have drug-induced LQTS may have other underlying cardiac conditions (14), one should be cautious about prescribing calcium channel blockers. Conversely, it would be reasonable to consider using calcium channel blockers to treat patients with LQTS type 2 (i.e. patients with an isolated loss of I_{Kr} function) who continue to have cardiac events despite treatment with ß-blockers (43).

In summary, we have demonstrated that meta-analysis of large-scale gene expression data sets is a powerful technique for discerning underlying patterns in gene expression, and that this can provide insights into disease causation at an individual level. Specifically, we have demonstrated that the co-expression of *KCNH2* (I_{Kr}) and *CACNA1c* (I_{CaL}) plays an important role in regulating cardiac repolarization both in health and in disease.

Methods

Analysis of public RNASeq datasets

An aggregate co-expression gene network was built from public data, similar to that described previously(44). Briefly, 75 human RNA-seq expression experiments (listed in Supp **Table S2**) that passed quality control and had a least 10 samples (3653 samples in total) were downloaded from the Gemma database (45). Approximately thirty thousand genes were used for the network, limited only to those with Entrez gene identifiers. A co-expression network was generated for each experiment by calculating Spearman's correlation coefficients between every gene pair and then ranking these values (44). An aggregate gene co-expression network was then generated by averaging across all the individual networks, and re-ranking the final network. This final aggregate network was then used to determine the co-expression ranking between genes that encode for the set of ion channels and calcium handling proteins that determine the shape and duration of the human ventricular AP, the so-called rhythmonome gene subset (see Supp **Table S1**). Network connectivity of the gene set was measured by comparing the weighted local node degree to the global node degree(46). Node degrees are the sum of the total connections a node (here gene) has within a network. Local node degree refers to the sum of connections (here the ranked correlation) within the rhythmonome gene set, while global node degree

is the sum of connections to that gene across all the genes in the network. Code for the analysis is available at https://github.com/sarbal/hERG-cal

Human induced pluripotent stem cell derived cardiac myoctes.

Human iPSC lines, generated from healthy patients by Stanford Cardiovascular Institute Biobank, as previously described (47), were a generous donation from Joseph Wu (Standford Cardiovascular Institute). HiPSC colonies were maintained on Matrigel® (Corning) coated plates in chemically defined medium (mTeSR1™, StemCell technologies), and passaged using Dispase (StemCell technologies). For differentiation, hiPSCs were dissociated by incubating at 37°C for 7 minutes with TrypLE™ Express (ThermoFisher) and seeded at 125000 cells/cm^2 on a Matrigel® coated 12 well plate, in mTeSR™1 medium supplemented with StemMACS™ Y27632 (Miltenyi Biotec). Once the cells reached greater than 95% confluency, differentiation was initiated using STEMdiff™ Cardiomyocyte (CM) differentiation and Maintenance Kit (StemCell technologies). At day 15, CMs were dissociated by incubation in Collagenase Type I (ThermoFisher) for 45 minutes at 37°C to break up the matrix and then incubated in 0.25% Trypsin with EDTA for 7 minutes at room temperature followed by filtering through a 40 µm cell strainer (48). The CMs were seeded on a Fibronectin coated 96-well plate (Greiner Bio-One) and maintained in CM maintenance medium for 10-15 days before they were harvested for mRNA expression analysis. Total RNA was obtained from 40,000 hiPSC-CMs lysed using QIAzol Lysis reagent (Qiagen). The RNA was purified using miRNeasy® Mini Kit (Qiagen), and all samples had RIN values >7.5, and were analysed using Agilent Bioanalyzer pico-chip. RNA samples were hybridised with probes designed to detect 35 known rhythmonome genes using nCounter (NanoString Technologies, see Supplementary **Table S1**), which was performed at the Ramaciotti Centre for Genomics (UNSW).

Computer modelling

Human cardiac APs were simulated using the endocardial configuration of the O'Hara-Rudy (ORD11) model (25) with key conductances modified as described by Krogh-Madsen et al. (26) (See Supp **Figure S1**). The original ORD11 code was adapted to run in the Brain Dynamics Toolbox for Matlab (49). To incorporate population variability in ion channel expression levels the maximum conductance for each current was multiplied by conductance scalar (G_x), that was drawn from a random log-normal distribution (37), with unit mean and variance that was systematically manipulated from 0.05 to 0.5. All models were paced at 1Hz with a stimulus of -40 mV and duration of 1 ms and allowed to equilibrate for 300 beats. We then analysed the next four beats (to allow for the possibility of development of alternans) after the equilibration stage. The peaks in those APs were identified using the Matlab *findpeaks* function. Individual beats were classified as ectopic if they had secondary peaks that were separated by more than 100 ms. The set of four beats were further classified as *alternans* if the profile of any of the APs deviated from each other by more than 1 mV at any time point. In a second set of simulations, we repeated the same method as above except that the random multipliers applied to both I_{CaL} and I_{Kr} were

identical. This case we denote co-expression whereas the former case we denote independent expression.

Acknowledgements

This work was supported by grants from the National Health and Medical Research Council (NHMRC), App1116948 (to JIV), App1074386 (to JIV), (App1164518 to APH), by the National Institutes of Health (NIH) R01LM012736 (to JAG), R01MH113005 (to JAG), and supported by the Victor Chang Cardiac Research Institute Innovation Centre, funded by the NSW Government. We also thank Terry Campbell, Dan Roden and Raj Subbiah for helpful discussions.

References

1. Zipes DP, Wellens HJ. Sudden cardiac death. Circulation. 1998;98(21):2334–51.

2. Behr E, Ensam B. New approaches to predicting the risk of sudden death. Clinical Medicine. 2016;16(3):283–3.

3. Jan LY, Jan YN. Voltage-gated potassium channels and the diversity of electrical signalling. J Physiol (Lond). 2012;590(11):2591–9.

4. Noble D, Garny A, Noble PJ. How the Hodgkin-Huxley equations inspired the Cardiac Physiome Project. J Physiol (Lond). 2012;590(11):2613–28.

5. Marder E, Goaillard J-M. Variability, compensation and homeostasis in neuron and network function. Nat Rev Neurosci. 2006;7(7):563–74.

6. Weiss JN, Karma A, MacLellan WR, Deng M, Rau CD, Rees CM, et al. "Good enough solutions" and the genetics of complex diseases. Circ Res. 2012;111(4):493–504.

7. Prinz AA, Billimoria CP, Marder E. Alternative to Hand-Tuning Conductance-Based Models: Construction and Analysis of Databases of Model Neurons. J Neurophysiol. 2003;90(6):3998–4015.

8. Britton OJ, Bueno-Orovio A, Van Ammel K, Lu HR, Towart R, Gallacher DJ, et al. Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. Proceedings of the National Academy of Sciences. 2013;110(23):E2098–105.

9. Sarkar AX, Sobie EA. Quantification of repolarization reserve to understand inter-patient variability in the response to pro-arrhythmic drugs: a computational analysis. Heart Rhythm 2011;8(11):1749–55.

10. Liberos A, Bueno-Orovio A, Rodrigo M, Ravens U, Hernandez-Romero I, Fernandez-Aviles F, et al. Balance between sodium and calcium currents underlying chronic atrial fibrillation termination_ An in silico intersubject variability study. Heart Rhythm. 2016;13(12):2358–65.

11. Passini E, Mincholé A, Coppini R, Cerbai E, Rodríguez B, Severi S, et al. Mechanisms of pro-arrhythmic abnormalities in ventricular repolarisation and anti-arrhythmic therapies in human hypertrophic cardiomyopathy. J Mol Cell Cardiol. 2016;96(C):72–81.

12. Marder E. Understanding Brains: Details, Intuition, and Big Data. PLoS Biol. 2015;13(5):e1002147–6.

13. O'Leary T, Sutton AC, Marder E. Computational models in the age of large datasets. Curr Opin Neurobiol. 2015;32:87–94.

14. Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med. 2004;350(10):1013–22.

15. Schulz DJ, Goaillard J-M, Marder EE. Quantitative expression profiling of identified neurons reveals cell-specific constraints on highly variable levels of gene expression. Proceedings of the National Academy of Sciences. 2007;104(32):13187–91.

16. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences. 1998 95(25):14863–8.

17. Gaiteri C, Ding Y, French B, Tseng GC, Sibille E. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. Genes Brain Behav. 2013;13(1):13–24.

18. Rees CM, Yang J-H, Santolini M, Lusis AJ, Weiss JN, Karma A. The Ca2+ transient as a feedback sensor controlling cardiomyocyte ionic conductances in mouse populations. eLife. 2018;7:4417.

19. Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P. Coexpression analysis of human genes across many microarray data sets. Genome Res. 2004;14(6):1085–94.

20. Gillis J, Pavlidis P. The role of indirect connections in gene networks in predicting function. Bioinformatics. 2011;27(13):1860–6.

21. Ballouz S, Verleyen W, Gillis J. Guidance for RNA-seq co-expression network construction and analysis: safety in numbers. Bioinformatics. 2015;31(13):2123–30.

22. Vandenberg JI, Walker BD, Campbell TJ. HERG K+ channels: friend and foe. Trends Pharmacol Sci. 2001;22(5):240–6.

23. Bush WS, Crawford DC, Alexander C, George AL Jr, Roden DM, Ritchie MD. Genetic variation in the rhythmonome: ethnic variation and haplotype structure in candidate genes for arrhythmias. Pharmacogenomics. 2009;10(6):1043–53.

24. Sah R, Ramirez RJ, Backx PH. Modulation of Ca 2+Release in Cardiac Myocytes by Changes in Repolarization Rate. Circ Res. 2002;90(2):165–73.

25. O'hara T, Virág L, Varró A, Rudy Y. Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation. PLoS Comput Biol. 2011;7(5):e1002061–29.

26. Krogh-Madsen T, Jacobson AF, Ortega FA, Christini DJ. Global Optimization of Ventricular Myocyte Model to Multi-Variable Objective Improves Predictions of Drug-Induced Torsades de Pointes. Frontiers in Physiology. 2017;8:e1000173–10.

27. Sayed N, Liu C, Wu JC. Translation of Human-Induced Pluripotent Stem Cells. J Am Coll Cardiol. 2016;67(18):2161–76.

28. Robertson C, Tran DD, George SC. Concise Review: Maturation Phases of Human Pluripotent Stem Cell-Derived Cardiomyocytes. Stem Cells 2013;31(5):829–37.

29. Pedersen HS, Elming H, Seibæk M, Burchardt H, Brendorp B, Torp-Pedersen C, et al. Risk Factors and Predictors of Torsade de Pointes Ventricular Tachycardia in Patients With Left Ventricular Systolic Dysfunction Receiving Dofetilide. Am J Cardiol. 2007;100(5):876–80.

30. Shimizu W, Moss AJ, Wilde AAM, Towbin JA, Ackerman MJ, January CT, et al. Genotype-phenotype aspects of type 2 long QT syndrome. J Am Coll Cardiol. 2009;54(22):2052–62.

31. Hartzell HC. Effects of magnesium on inactivation of the voltage-gated calcium current in cardiac myocytes. J Gen Physiol. 1989;94(4):745–67.

32. Marder E, Goaillard J-M. Variability, compensation and homeostasis in neuron and network function. Nat Rev Neurosci. 2006;7(7):563–74.

33. Bányász T, Horváth B, Jian Z, Izu LT, Chen-Izu Y. Sequential dissection of multiple ionic currents in single cardiac myocytes under action potential-clamp. J Mol Cell Cardiol 2011;50(3):578–81.

34. Min B, Zheng M. Correlated network of networks enhances robustness against catastrophic failures. PLoS ONE. 2018;13(4):e0195539–13.

35. The QT Interval International GWAS Consortium (QT-IGC), Lundby A, Rossin EJ, Steffensen AB, Acha MR, Newton-Cheh C, et al. Annotation of loci from genome-wide association studies using tissue-specific quantitative interaction proteomics. Nat Methods 2014;11(8):868–74.

36. Huang CL-H. Murine Electrophysiological Models of Cardiac Arrhythmogenesis. Physiol Rev. 2017;97(1):283–409.

37. Sadrieh A, Mann SA, Subbiah RN, Domanski L, Taylor JA, Vandenberg JI, et al. Quantifying the origins of population variability in cardiac electrical activity through sensitivity analysis of the electrocardiogram. J Physiol (Lond) 2013;591(17):4207–22.

38. Marder E. Variability, compensation, and modulation in neurons and circuits. Proceedings of the National Academy of Sciences. 2011;108(Supplement_3):15542–8.

39. Roden DM, Iansmith DH. Effects of low potassium or magnesium concentrations on isolated cardiac tissue. The American Journal of Medicine. 1987;82(3A):18–23.

40. Redfern WS, Carlsson L, Davis AS, Lynch WG, MacKenzie I, Palethorpe S, et al. Relationships between preclinical cardiac electrophysiology, clinical QT interval

prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. Cardiovasc Res. 2003;58(1):32–45.

41. Farkas AS, Makra P, Csík N, Orosz S, Shattock MJ, Fülöp F, et al. The role of the Na+/Ca2+ exchanger, INa and ICaL in the genesis of dofetilide-induced torsades de pointes in isolated, AV-blocked rabbit hearts. Br J Pharmacol. 2009;156(6):920–32.

42. Akhtar M, Tchou P, Jazayeri M. Use of calcium channel entry blockers in the treatment of cardiac arrhythmias. Circulation 1989;80(6 Suppl):IV31–9.

43. Komiya N, Tanaka K, Doi Y, Fukae S, Nakao K, Isomoto S, et al. A patient with LQTS in whom verapamil administration and permanent pacemaker implantation were useful for preventing torsade de pointes. Pacing and Clinical Electrophysiology. 2004;27(1):123–4.

44. Ballouz S, Verleyen W, Gillis J. Guidance for RNA-seq co-expression network construction and analysis: safety in numbers. Bioinformatics. 2015;31(13):2123–30.

45. Zoubarev A, Hamer KM, Keshav KD, McCarthy EL, Santos JRC, Van Rossum T, et al. Gemma: a resource for the reuse, sharing and meta-analysis of expression profiling data. Bioinformatics. 2012;28(17):2272–3.

46. Ballouz S, Gillis J. Strength of functional signature correlates with effect size in autism. Genome Med. 2017;9(1):64.

47. Burridge PW, Diecke S, Matsa E, Sharma A, Wu H, Wu JC. Modeling Cardiovascular Diseases with Patient-Specific Human Pluripotent Stem Cell-Derived Cardiomyocytes. Methods Mol Biol.; 2016;1353(Chapter 196):119–30.

48. Mills RJ, Titmarsh DM, Koenig X, Parker BL, Ryall JG, Quaife-Ryan GA, et al. Functional screening in human cardiac organoids reveals a metabolic mechanism for cardiomyocyte cell cycle arrest. Proceedings of the National Academy of Sciences. 2017;114(40):E8372–81.

49. Heitmann S, Aburn MJ, Breakspear M. The Brain Dynamics Toolbox for Matlab. Neurocomputing; 2018;315:82–8.