Predicting the Functional Impact of KCNQ1 Variants with Artificial Neural Networks

Saksham Phul[1,3], Georg Kuenze[1,3,6], Carlos G. Vanoye[4], Charles R. Sanders[1,2], Alfred L. George, Jr.[4], Jens Meiler[1,3,4,6,*]

[1]Center for Structural Biology, Vanderbilt University, Nashville, TN 37240, USA

[2]Department of Biochemistry, Vanderbilt University, Nashville, TN 37232, USA

[3]Department of Chemistry, Vanderbilt University, Nashville, TN 37235, USA

[4]Department of Pharmacology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA

[5]Department of Pharmacology, Vanderbilt University, Nashville, TN 37240, USA

[6]Institute for Drug Discovery, Leipzig University, Leipzig, SAC 04103, Germany

*Correspondence to: Jens Meiler, PhD, jens@meilerlab.org

Abstract:

Predicting the functional impact of KCNQ1 variants of uncertain significance (VUS) can assist physicians in taking appropriate treatment decision for patients with genetic heart rhythm disorder. This work presents three artificial neural network (ANN)-based predictive models that classify four key functional parameters of KCNQ1 variants as normal or dysfunctional using evolutionary and/or biophysical descriptors. Recent advances in predicting protein structure and variant

properties with artificial intelligence (AI) rely heavily on the availability of evolutionary features and thus fail to directly assess the biophysical underpinnings of a change in structure and/or function. The central goal of this work was to develop an ANN model based on structure and physiochemical properties of KCNQ1, that performs better or comparable with algorithms only on evolutionary features. These biophysical features highlight the structure-function relationships that govern protein stability, function, and regulation. The input sensitivity algorithm incorporates the roles of hydrophobicity, polarizability, and functional densities on key functional parameters of the KCNQ1 channel. Inclusion of the biophysical features outperforms exclusive use of evolutionary features in predicting variant activation voltage and deactivation time. As AI is increasing applied to problems in biology, biophysical understanding will be critical with respect to 'explainable AI', i.e., understanding the relation of sequence, structure, and function of proteins. Our model is available at www.kcnq1predict.org.

Keywords: KCNQ1, Neural Network, Phenotype, Biophysics, VUS

Author summary:

Heartbeat is maintained by electrical impulses generated by ion-conducting channel proteins in the heart such as the KCNQ1 channel. Heritable mutations in KCNQ1 can lead to channel loss-of-function and predisposition to fatal irregularities of the heartbeat. Machine learning methods that can predict the outcome of a mutation on KCNQ1 structure and function would be of great value in helping to assess the risk that a mutation could lead to a heart rhythm disorder. Recently, machine learning has made great progress in predicting the structures of proteins from their sequences. However, there are limited studies that link the effect of a mutation and change in protein structure with its function. This work presents the development of neural network models designed to predict mutation-induced changes in KCNQ1 functional parameters such as peak

current density and voltage of activation. We compare the predictive value of features extracted from sequence, structure, and physicochemical properties of KCNQ1. Moreover, input sensitivity analysis connects biophysical features with specific functional parameters that provides insight into underlying molecular mechanisms for KCNQ1 ion channels. The best performing neural network model is publicly available as a webserver, called Q1VarPredBio, that delivers predictions about the functional phenotype of KCNQ1 variants to researchers and physicians.

1. Introduction

Congenital long QT syndrome (LQTS) is a genetic disorder of heart rhythm caused by mutations in cardiac ion channel genes (1,2) that predisposes to potentially life-threatening cardiac arrhythmia. The most prevalent subtype, LQT1, is associated with genetic variants in the *KCNQ1* gene (3,4) which encodes the pore forming subunit of the voltage-gated K^+ channel $K_V7.1$ (referred to as KCNQ1) (5). In the heart, KCNQ1 forms a channel complex with KCNE1 to generate the slow delayed rectifier current, I_{Ks}, which is an essential driver of myocardial repolarization during the cardiac action potential (6,7). Variants that cause KCNQ1 loss-of-function (LOF) lead to diminished I_{Ks} and impaired repolarization that is manifest by prolongation of the QT interval on surface electrocardiograms (8).

LQT1 is among the most common inherited disorders (9). More than 1000 genetic variants of KCNQ1 have been identified (10,11), but for many variants there are insufficient data to classify each as either pathogenic or benign. Correlating these variants of uncertain significance (VUS) to their clinical outcomes and determining the risk of LQTS remain major challenges (12,13). Large-scale functional characterization of KCNQ1 variants has been made feasible by using automatic patch-clamp recording (14) and this strategy helped reclassify variants with previously conflicting

or unknown interpretations according to the ClinVar database (10). Moreover, the mechanistic basis underlying mutation-induced KCNQ1 dysfunction has been investigated (15–19). For instance, Huang et al. (18) studied the impact of mutations in the KCNQ1 voltage-sensing domain (VSD) on protein cell surface expression, trafficking, protein folding, and structure. More than half of LOF mutations examined were found to destabilize the VSD structure resulting in impaired trafficking and lower cell surface expression. This observation underscores the growing notion that mutation-induced destabilization and mistrafficking of membrane proteins are common disease mechanisms in LQT1. However, this study also identified LOF variants that did not exhibit trafficking and folding defects, indicating heterogeneity in the molecular mechanisms responsible for KCNQ1 LOF that cause LQT1. The molecular function of many variants in other regions of the KCNQ1 channel have yet to be characterized, and it is expected that these investigations will reveal additional pathogenic mechanisms (20).

Despite this progress in functional characterization of KCNQ1 variants, experimental assays remain labor-intensive, and this limits their applicability in a clinical setting. Computational approaches can support experimental testing and have the potential to help elucidate the molecular function of KCNQ1 variants as well as predicting associated clinical outcomes (21–24). Computational methods trained on information from genome-wide genetic variation data are commonly used for protein variant effect prediction (25–29), but these tools have limited applicability for KCNQ1. Prediction accuracy of genome-wide methods is low and varies between targets. This reflects that development of these methods was based on heterogenous datasets including a wide range of proteins with diverse functions and associated diseases. Furthermore, these methods fail to establish the precise effect of a variant on KCNQ1 function parameters (22). To overcome these difficulties, specific machine learning models tailored to predict the functional

effects of KCNQ1 variants were developed (22,23). Similar approaches were also applied to other cardiac ion channels (30,31).

Q1VarPred (22) is a KCNQ1-specific channel function predictor. Criteria for dysfunction were calibrated by examining experimentally determined electrophysiology parameters for KCNQ1 (i.e., peak current density, voltage of half-maximal activation) that were then used to train a neural network with input features derived from protein sequence. Q1VarPred achieved greater accuracy than genome-wide tools, which perform poorly for membrane proteins (32). Additional predictive power may be gained by analyzing the spatial clustering of variants in the 3-dimensional structure of KCNQ1 (30). Functionally critical channel regions, such as the ion selectivity filter and cytosolic gate in the pore domain (PD) and the S4 helix in the VSD, are "hotspots" for variants causing the greatest perturbations in peak current density and voltage of activation, respectively. This suggests that protein structure features can aid variant prediction.

In this study, we used machine learning to develop an advanced and innovative KCNQ1 variant prediction tool called Q1VarPredBio (www.kcnq1predict.org). The functional classification categories of Q1VarPred were expanded by a scheme that predicts variant-specific changes in four electrophysiological KCNQ1 parameters: peak current density, voltage of half-maximal activation ($V_{1/2}$), and activation and deactivation time constants (τ_{act}, τ_{deact}). We evaluated the performance of artificial neural networks (ANNs) trained on evolutionary and biophysical features for KCNQ1 and observed that a combination of both features produced a model with optimal predictive accuracy. Our machine learning approach can be useful to obtain insights into basic sequence-structure-function relationships for the KCNQ1 channel. Moreover, Q1VarPredBio may help differentiate between potential pathogenic dysfunctional KCNQ1 variants from those with normal channel function.

2. Results

We developed three types of ANN models: one trained with only evolutionary features, one trained with biophysical features, and a third one with both evolutionary and biophysical features. All three models were trained to predict the mutation-induced change in four functional parameters (peak current density, voltage of half-maximal activation, and activation and deactivation time constants) of the KCNQ1 channel as either normal (label 0) or dysfunctional (label 1) phenotype. All ANN models had an input layer, two hidden layers, and an output layer with four neurons. A schematic representation of our model development workflow is shown in Figure 1. These multitask ANN models were trained to improve accuracy of predicting all parameters combined. All models predicted four outputs between 0 and 1. A decision boundary was identified at the best possible accuracy and presented our threshold to classify a variant parameter as being either normal or dysfunctional. Accurate prediction based on this decision boundary was a criterion for determining the model performance. To measure the performance of the model, we adopted a 25-fold cross validation technique wherein performance was evaluated on a variant test set that was not included in model training and monitoring. Further details on ANN architecture, training, and model performance are described in the methods and materials section.

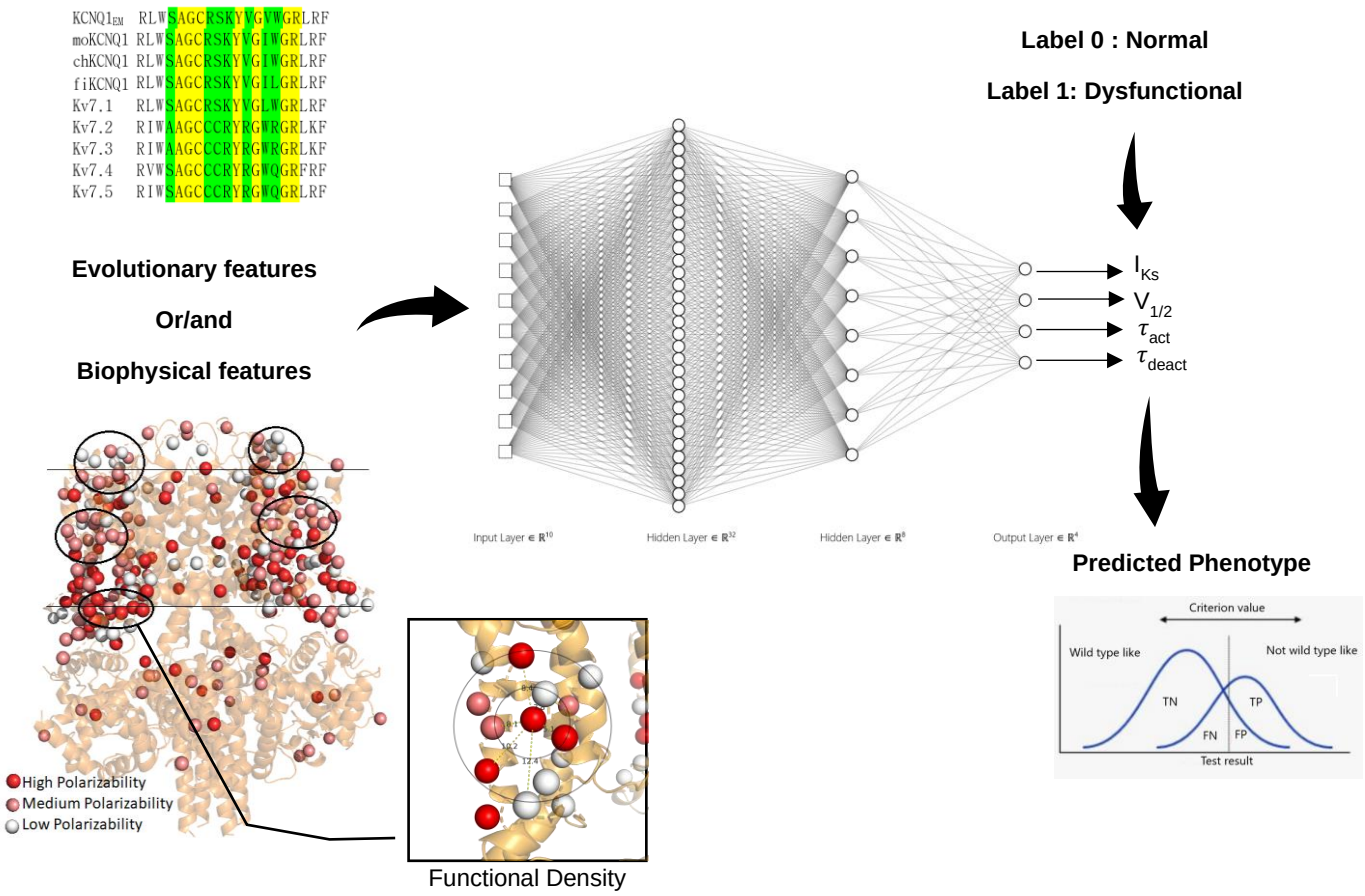

Figure 1: A Schematic flowchart of the ANN Models

Dataset and criteria

The dataset for this work contained electrophysiological data for 125 KCNQ1 variants that were generated, tested, and functionally analyzed using the same approach (14,33). The KCNQ1 variants were tested in the homozygous state transiently co-expressed with wild type KCNE1 in CHO cells. The electrophysiological data measured for each variant consists of four biophysical parameters: peak current density (designated as I_{Ks}), voltage of half-maximal channel activation ($V_{1/2}$), activation time constant (τ_{act}), and deactivation time constant (τ_{deact}). In order to compare the functional properties of variants tested across many months, the values for each parameter were normalized to the values obtained from cells expressing the wild type channel that were transfected and tested in parallel. Normalized values equal to 1 (or 100% WT) were considered wild-type-

like. A parameter phenotype was classified as dysfunctional if it satisfies the criteria in Table 1. These thresholds were derived collectively from values defined in Li et al(22), disease-causing variants in literature and by evaluating model performance at different thresholds.

For training and testing of our model, both gain-of-function and loss-of-function (according to Table 1) were classified as dysfunctional. Biophysical parameters that could not be determined for some variants (e.g., voltage-insensitive [no $V_{1/2}$] variants that do not deactivate [no τ_{deact}]), were defined as dysfunctional and for variants with peak current density ≤ 17% WT, all four biophysical parameters were considered dysfunctional.

Due to the scarcity of the functional data from certain protein regions, we introduced 345 'silent' variants, one for each of the 345 amino acids included in the KCNQ1 structural model (34) thereby increasing the size of our dataset to 470. All four functional parameters for the silent variants were considered WT. These silent variants create a 'baseline' for the protein region where data was limited allowing the ANNs to train on a greater number of instances.

Table 1: Criteria for classifying variants as dysfunctional used in this work

Loss of function (%WT)	Phenotype	Gain of function (%WT)
< 55%	$\mathbf{I_{ks}}$	> 115%
> 130 %	$\mathbf{V_{1/2}}$	< 80%
> 170%	$\boldsymbol{\tau_{act}}$	< 70%
< 75%	$\boldsymbol{\tau_{deact}}$	> 125%

Identification of biophysical features

In total, we used 14 structural and physicochemical properties of KCNQ1 to develop an ANN model based solely on biophysical features. These features were extracted by importing the KCNQ1 structure into Biochemical Library (BCL) software. The KCNQ1 cryo-EM structure model utilized for this work, had bound calmodulin (CAM), no PIP_2 and represented a decoupled state with activated voltage sensor domain and a closed pore domain (34). We explored different

biophysical features based on existing understanding of molecular mechanisms underlying KCNQ1 function and the location of critically significant regions in the KCNQ1 structure important for protein stability and channel gating. These biophysical features inform about the amino acid local environment, exposure to solvent, burial in the membrane, change in amino acid physicochemical properties at the mutation site, steric hindrances near to the α-carbon atom, and the mutation-induced change in water-membrane transfer free energy. Distance from the KCNQ1 channel pore axis was used to help the model distinguish variants in the channel pore domain from those in the voltage sensor domain (see Figure S1). Furthermore, it is more likely that a variant residue buried in the membrane will negatively impact protein function. Thus, the degree of burial of a variant in the membrane was assessed by using a three-layered membrane model and calculating a membrane-depth dependent weight calculated with a distribution function (Supplemental Equation 1). A significant weight was given to variants embedded inside the membrane (see Figure S2). Steric hindrances near the α-carbon atom were examined using a steric parameter (see Supplemental Table S1), which is a graph shape index that encodes complexity, branching and symmetry of amino acid side chain (35).

Transfer free energy of an amino acid between a hydrophilic and hydrophobic environment plays a crucial role for protein folding and stability. Thus, we used hydrophobicity of native and variant amino acids to investigate the mutation-induced changes in the free energy for transfer into the membrane (see Figure S3). The mutation-induced change in water-membrane transfer free energy at an amino acid site was examined using the hydrophobicity scale reported by Koehler *et al* (36). Highly hydrophobic or hydrophilic amino acids are usually surrounded by similarly hydrophobic or hydrophilic amino acids. Thus, we calculated the polarizability and hydrophobicity of amino acids at and around the variant site by functional density (30). Functional density is based on k-

nearest neighbors' algorithm, wherein average physiochemical property around the site of variant is weighted by inverse of their distance from the site of mutation (see Figure S4). These features examined hydrophobicity and polarizability of the neighborhood around the site of variants. We found that there exists a correlation wherein high polarizability regions in the protein are correlated with dysfunctional peak current density (see Figure S5). More details on the implementation of these biophysical features are described in the supplementary material.

We also introduced changes in the physicochemical properties [number of hydrogen bond donor sites, number of hydrogen bond acceptor sites, and van der Waals volume (35)] of amino acid at the site of mutation, in order to help the ANN model, learn whether an amino acid substitution represents a missense or silent mutation. These properties also improved the prediction for missense mutations.

Exposure of variant sites to solvent was quantified using the neighbor vector method (37) (see Figure S6). Neighbor vector (37) improved the predictions especially for peak current density whereas neighbor count (37) did not. Backbone conformation (Φ, Ψ, Ω angles) for the native amino acid and other descriptors like the location of a mutation on a helix, mutation-caused change in amino acid polarizability as well as change in hydrophobicity and solvent accessible surface area for amino acids did not improve prediction accuracy.

Evolutionary feature: PSSM-based amino acid substitution score

We used PSI-BLAST search to calculate a position-specific scoring matrix (PSSM) which measures the likelihood of amino acid substitution at a mutation site (38). PSSM was created by searching UniRef50 (39) and the NCBI non-redundant sequence databases (40). The difference of PSSM scores between variant and WT amino acid from these two databases was utilized as evolutionary features. We found that these evolutionary features were solely sufficient in

predicting the functional properties of silent mutations. More details can be found in the supplementary material.

Biophysical features outperform evolutionary features in predicting activation $V_{1/2}$ and τ_{deact}.

Model accuracy was evaluated using Matthew's correlation coefficient (MCC) and receiver operating characteristic (ROC) plots by testing a variant set that was omitted from model training and monitoring. Decision boundary was identified at the best possible accuracy (MCC value) whereas ROC plots were independent of decision boundary. These MCC values and ROC plots for different ANN models are reported in Figure 2 and 3 respectively. More details on MCC and ROC are described in Methods and Material section.

We were able to model $V_{1/2}$ and τ_{deact} using evolutionary features, biophysical features, and both features combined. MCCs for these feature sets are reported in Figure 2. Although evolutionary features achieved satisfactory performance in predicting $V_{1/2}$ with a MCC of 0.56, biophysical features perform better by attaining MCCs greater than 0.62. Area under the curve (AUC) for $V_{1/2}$ determined for biophysical features was greater than that for evolutionary features (Figure 3) . For $\tau_{deact,}$ evolutionary features achieved satisfactory performance in predicting τ_{deact} with MCC 0.50, biophysical features perform better by attaining MCCs greater than 0.56. Area under the curve (AUC) for τ_{deact} determined for biophysical features was greater than that for evolutionary features (Figure 3). Biophysical features clearly dominate in predicting the activation $V_{1/2}$ and τ_{deact} suggesting that the ANN model could determine and distinct structure-activity relationships underlying the voltage-dependence of KCNQ1 activation and deactivation of KCNQ1 kinetics.

Biophysical features can predict current density but cannot outperform evolutionary features

We were able to model peak current density (I_{Ks}) using evolutionary and biophysical features. For peak current density (I_{Ks}), biophysical features attained MCC close to 0.38 whereas evolutionary

features achieved MCC ≥ 0.43, suggesting that evolutionary features perform better in predicting peak current density. Moreover, AUC for biophysical features is 0.68 whereas for evolutionary features is 0.69 (see Figure 3). This suggests that biophysical features are comparable with evolutionary features in distinguishing between normal and dysfunctional variants. BLAST based

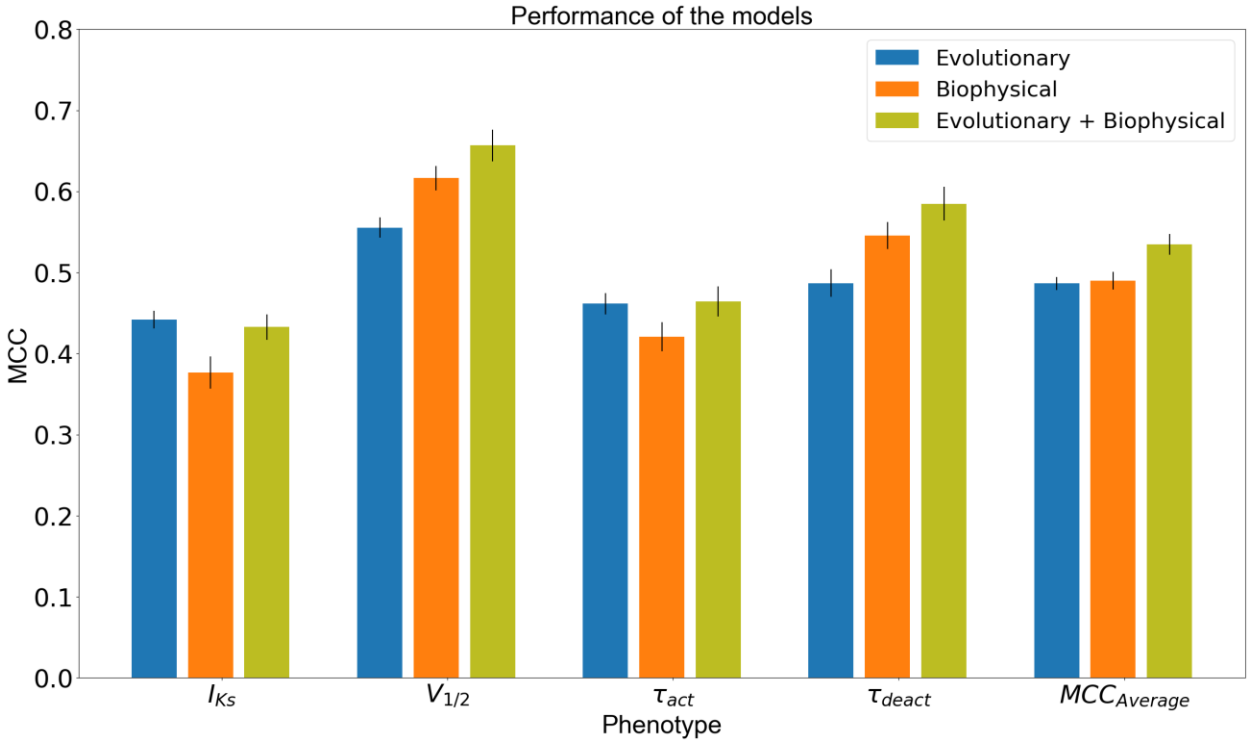

Figure 2: MCC values reported for three different ANN models. The error bar represents 1x standard derivation

PSSM-derived amino acid substitution scores, which is our evolutionary feature, has a significant association with peak current density (I_{Ks}) as previously reported by Kroncke *et al* (30).

Evolutionary features predict τ_{act} better than biophysical features

We modeled τ_{act} using both biophysical and evolutionary features. Evolutionary features yielded a MCC of 0.44 whereas biophysical features yielded a MCC of 0.40, suggesting evolutionary features perform better than biophysical features. Similarly, AUC for evolutionary features was close to 0.75 better than that for biophysical features (AUC:0.73). We observed that the ANN

model makes better predictions for peak current when activation time labels are simultaneously present in the training dataset.

Performance of biophysical features is comparable with evolutionary features

To compare models using a common performance metric, the MCCs calculated for four functional parameters were averaged ($MCC_{average}$). The $MCC_{average}$ for biophysical features was 0.49 wheras $MCC_{average}$ for evolutionary features was 0.48 (Figure 2). Thus, the two models are comparable in combined performance, irrespective of their performance for the individual functional parameters. The average AUC for biophysical features was 0.77 whereas evolutionary features was 0.75, suggesting biophysical features are comparable with evolutionary features.

Most accurate predictions are achieved by combining biophysical and evolutionary features.

For training the ANN model with both biophysical and evolutionary features, we used eleven biophysical features and the difference of PSSM score determined with NCBI non-redundant sequence database (Δ PSSM(NR)). We found that three out of 14 features used for biophysical model failed to improve the predictions on the unseen dataset, therefore were excluded. This could

be due to similar information carried by these three features [Neighbor Vector, Mutant steric parameter, and Native steric parameter] and evolutionary feature. The eleven biophysical features

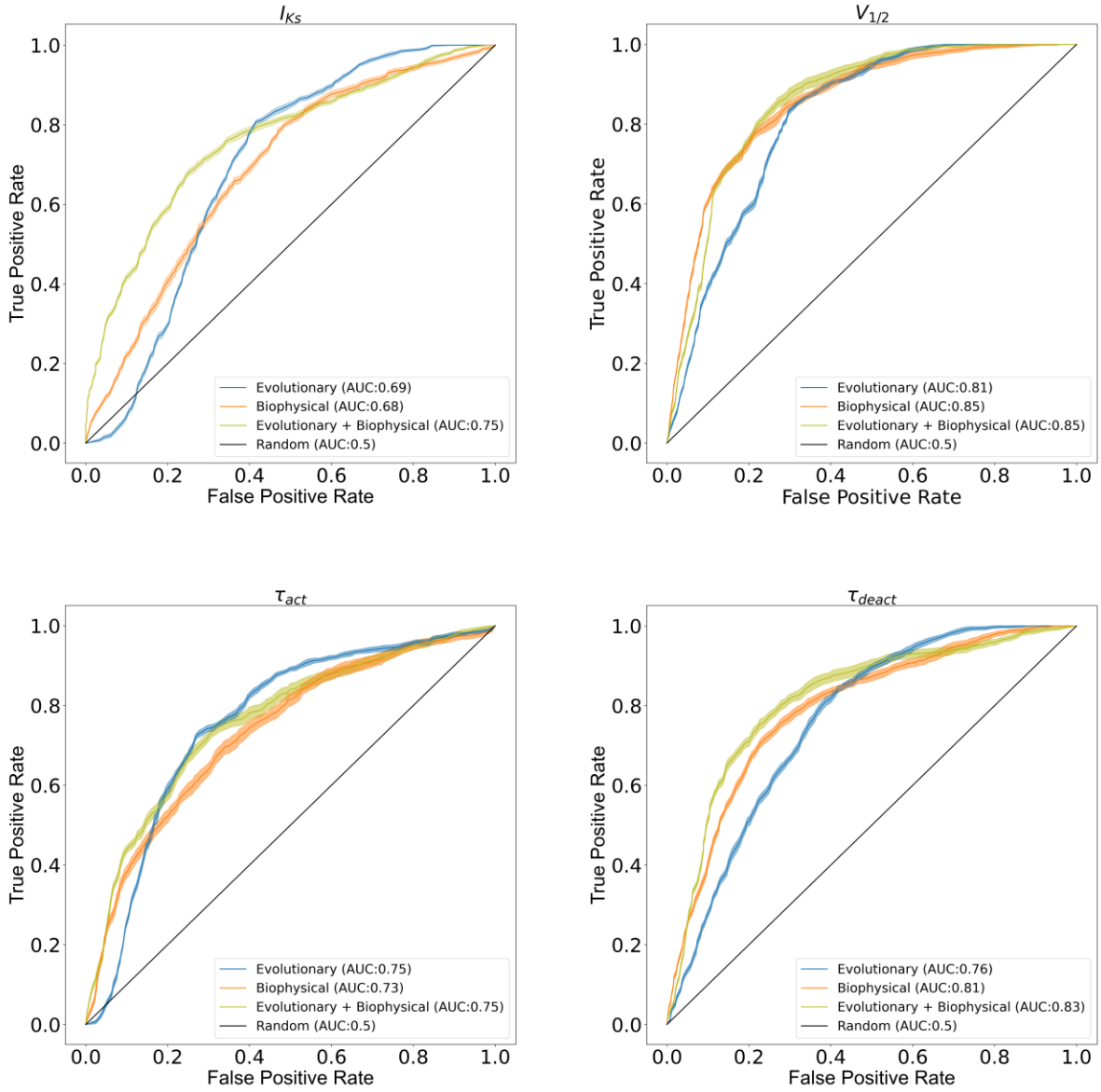

Figure 3: ROC Plots with 99% confidence interval (shaped region) for all-functional parameters of KCNQ1 channel considered in this study. Different color scheme is used to depict performance of the three feature sets.

included were: hydrophobicity of mutant amino acid, polarizability of mutant amino acid, functional density of amino acid polarizability with neighborhood sizes of 6.5 Å and 12 Å, functional density of amino acid hydrophobicity with neighborhood sizes of 1 Å and 6.5 Å, change

in number of hydrogen donor sites, change in number of hydrogen acceptor sites, change in Van der Waals volume, distance from the pore axis, and depth of the mutation site in the membrane. MCC > 0.44 and an AUC > 0.75 for peak current density suggests that although evolutionary and biophysical features combined do not improve prediction accuracy (MCC) for peak current density but improve the spread (AUC) between normal and dysfunctional variants predicted by the model. Moreover, significant improvement in performance was observed for $V_{1/2}$ and τ_{deact} by combining evolutionary and biophysical features. MCC increases by 15% for $V_{1/2}$ and by 10% for τ_{deact} when compared with the single-feature best model, i.e., biophysical model. The results also suggest that biophysical features have a significant association with τ_{deact} and $V_{1/2}$ making the performance of biophysical features for these parameters better than that of evolutionary features. For τ_{act}, we observed no improvement in MCC or AUC (Figures 2 and 3) for the combined model when compared with evolutionary features, indicating redundant nature of evolutionary and biophysical features in capturing sequence-structure-function relationships about KCNQ1 $\tau_{act.}$ Overall, in the combined features ANN model, $MCC_{average}$ increases to 0.54, which corresponds to a 12% improvement in $MCC_{average}$ from the best single-feature model.

Wildtype like variants were slow to activate

We observed that variants H105N, T118S, V129I, and E146G had peak current density, $V_{1/2}$, τ_{deact} parameters close to WT values but significantly larger activation time constants. Additionally, it was difficult to model τ_{act} with any of the ANN models using threshold 80% - 120 % as normal. Based on those observations, we adjusted our threshold for normal τ_{act} to 70% - 170% of wild type. This change improves MCC for evolutionary features from 0.22 to 0.44 and yielded 10% improvement in average MCCs across all models.

Performance of ANN models on silent variants

The MCC and AUC reported in figures 2 and 3 were evaluated on 125 experimentally validated variants. The correct classification of silent mutations signifies that ANNs can distinguish no change in protein structure results into a benign variant. This is an important feature as the wild-type sequence is not directly input into the ANN but encoded as changes in evolutionary and biophysical parameters. Thus, these silent mutations ensure that a change of zero in all parameters are understood as wild-type. All models were able to classify roughly 98% of the 345 silent variants accurately as benign. Features like change in number of hydrogen donor sites, change in number of hydrogen acceptor sites, and change in amino acid volume was required by ANN model trained with only biophysical features to help model learn silent mutations. When both features were combined, evolutionary features were found to be sufficient in correctly predicting silent mutations.

Input sensitivity highlights feature importance and their association with KCNQ1 functional parameters

To study the contribution of different features in our ANN model, we examined the input sensitivity of input features on output labels. Since considering the magnitude of input sensitivity for feature importance can be misleading due to the issue in rescaling the input features (41), we considered sign of the input sensitivity with respect to output label. Input sensitivity is defined as zero when half instance of the variants predicts a positive change with respect to the result label and the other half predict a negative change with respect to the result label. Similarly, an input sensitivity close to one signifies that the input feature strongly correlates with the output label. More details on input sensitivity analysis can be found in the materials and methods section.

The input sensitivities for our best predictive model (eleven biophysical features and one evolutionary feature) are averaged across 320 models with each model having 25 instances of input sensitivity using 25 different monitoring data subsets. These averaged input sensitivities are reported in Figure 4.

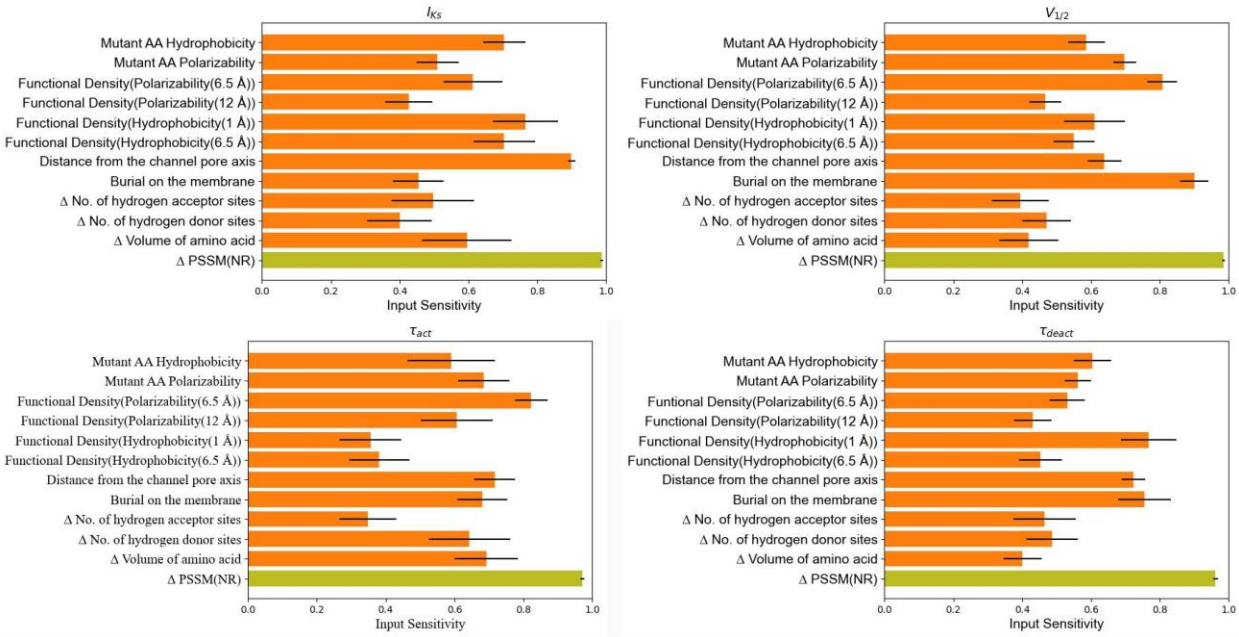

Figure 4: Average input sensitivities for combined evolutionary and biophysical features. Biophysical features are shown in orange and evolutionary features in olive color.

PSSM-based amino acid substitution score (Δ PSSM(NR)) was found to be most the sensitive feature for all the functional parameters, suggesting that sequence-based features are of high quality. For peak current density (I_{Ks}), four biophysical features were highly sensitive: hydrophobicity of mutant amino acid, hydrophobicity around mutation site within radius of 1 Å and 6.5 Å, and distance of mutation site from channel pore axis. For predicting $V_{1/2}$, the most sensitive biophysical features were burial of mutation site in the membrane, mutant polarizability, and neighborhood polarizability around mutation site with 6.5 Å radius. For τ_{act}, the most sensitive biophysical properties were neighborhood polarizability around mutation site with 6.5 Å radius, distance from the channel axis and change in volume of the amino acid. And for predicting τ_{deact},

the most sensitive features were burial of mutation site on the membrane, distance of mutation site from channel pore axis, and neighborhood hydrophobicity around mutation site within 1 Å radius.

We also report the input sensitivity for an ANN model trained solely with 14 biophysical features. These input sensitivities are also averaged across 320 models with each model having 25 instances of input sensitivity using 25 different monitoring data subsets. The averaged input sensitivities for the biophysical features only model is reported in Figure 5. We observed that all biophysical features are sensitive to functional parameters, but some features were more sensitive to specific functional parameter than other biophysical features. Distance from the channel pore axis, change in number of hydrogen donor sites and hydrophobicity at neighborhood size of 1 Å were the top highly sensitive biophysical features for I_{Ks}. Mutant amino acid hydrophobicity, neighbor vector

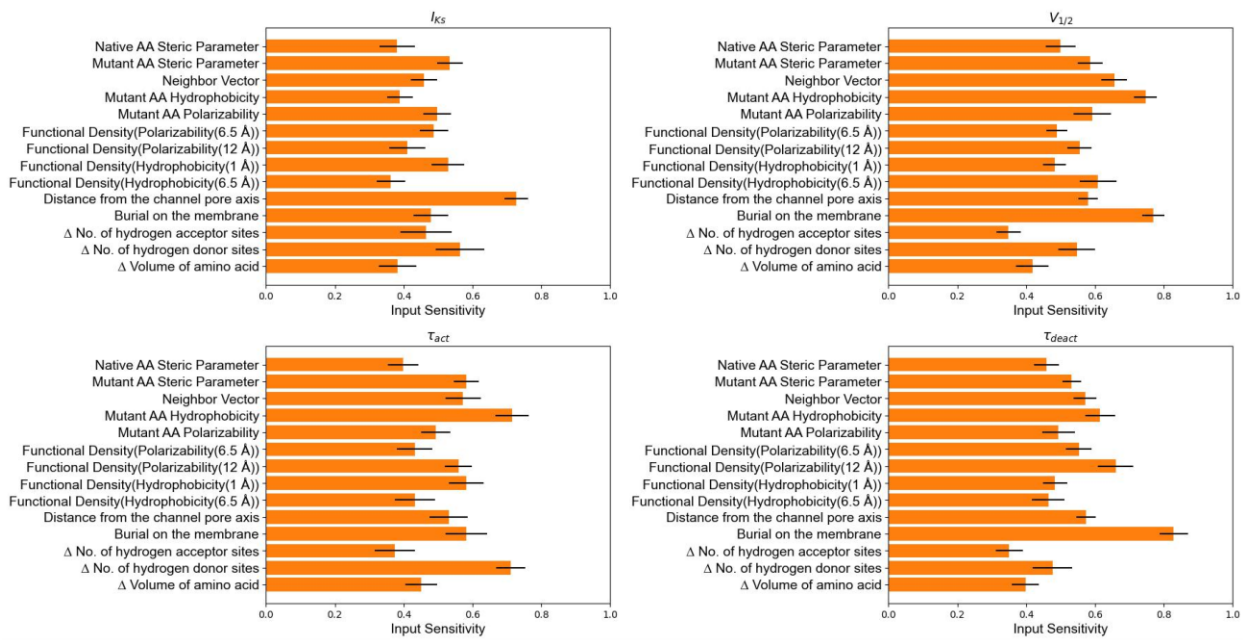

Figure 5: Average input sensitivities for ANN model with only biophysical features

and burial on the membrane for $V_{1/2}$. Change in number of hydrogen donor sites and mutant amino acid hydrophobicity for τ_{act} and burial on the membrane and neighborhood polarizability within 12 Å radius for τ_{deact}.

3. Discussion

Biophysical features perform as well as evolutionary features in predicting KCNQ1 variant function

Biophysical features performed as well as evolutionary features in predicting the functional outcomes of KCNQ1 variants. This suggests that ANN with only biophysical features recognized relationships between KCNQ1 structure, function, and mutation-induced dysfunction. However, biophysical features outperformed evolutionary features in predicting the $V_{1/2}$ and τ_{deact}. This could be linked to the prevalence of variants located in the VSD amongst all variants in the training dataset, thus, allowing the network to effectively learn about this channel domain. We found that presence of $V_{1/2}$ labels in the training dataset improves the τ_{deact} predictions and that biophysical features that were highly sensitive to $V_{1/2}$, were also found to be sensitive to τ_{deact} (Figure 4, 5). These observations could indicate that similar molecular determinants are important for voltage-dependent channel activation as well as kinetics of channel deactivation. Biophysical features like burial on the membrane and distance from channel pore axis were most important for the ANN to learn the phenotype of variants in the VSD. Other features like polarizability of mutant amino acid and functional density of amino acid polarizability within a radius of 6.5Å and 12Å around the mutation site were also sensitive to $V_{1/2}$ and τ_{act} (see Figure 4). This could indicate the fact that high polarizability of amino acids in the VSD confers sensitivity to transmembrane voltage that is required for KCNQ1 activation mechanisms. However, not all the sites in VSD actively participate in activation; some exist to stabilize the protein fold of the VSD(18). Thus, variant hydrophobicity and neighbor vector have high sensitivity suggesting that protein stability and exposure to the membrane are important features for predicting functional phenotype of $V_{1/2}$ and τ_{deact}.

We were able to model peak current density and τ_{act} using only biophysical features, however, this model performance was not as good as the one obtained with only evolutionary features. In summary, it was challenging to find biophysical features that significantly improve the peak current density and τ_{act} predictions over evolutionary features. Interestingly, we observed some interdependency of peak current density and τ_{act} predictions, similar as for $V_{1/2}$ and τ_{deact}. Due to the multi-task classification scheme, the network model can benefit and learn from other predicted labels, thus, increasing prediction accuracy compared to a single-task classification network. As per our dataset, it is more probable for $V_{1/2}$, τ_{act} and τ_{deact} to be dysfunctional if peak current is also dysfunctional. This is the case for 42 variants in the training set, for which I_{Ks} is less than 17% WT giving rise to dysfunctional $V_{1/2}$, τ_{act} and τ_{deact}. In summary, biophysical model performs well in predicting variant phenotype when all four parameters are dysfunctional or normal, but it does not perform well when only one or two parameters are impaired.

Combination of evolutionary and biophysical features improve functional phenotype predictions

When combining both types of features, the ANN model predicts functional parameters with better average accuracy than models trained with only biophysical or evolutionary features. For the voltage-dependence of activation $V_{1/2}$, prediction accuracy markedly improves by combining both feature sets. On examination of the input sensitivities in Figure 4, functional density of polarizability within 6.5Å radius, burial on the membrane, and mutant polarizability are most the biophysical features that have highest sensitivity on $V_{1/2}$ prediction. The movement of helices in VSD under the influence of electric field is due to the sites with high polarizability residues and neighborhoods with high polarizability. Therefore, mutant polarizability and functional density of polarizability are biophysical features with high sensitivity and can be linked to KCNQ1 activation

mechanism. Burial on the membrane indicates higher prominence of mutations that are inside the membrane and in close proximity to VSD region.

Prediction accuracy for τ_{deact} increases when biophysical and evolutionary features are combined, MCC improves to 0.60 and AUC increases to 0.83. Based on input sensitivity (see Figure 4) burial on the membrane, distance from the channel pore, mutant hydrophobicity, and functional density of hydrophobicity within radius 1 Å were highly sensitive biophysical features for τ_{deact} prediction. The high sensitivity for hydrophobicity-based features indicates that stability at the site of the mutation affects the phenotype of τ_{deact}. The improvement in MCC and AUC scores for τ_{deact} may be due to the improved ability of the model to identify the VSD region by using biophysical features like burial on the membrane and distance from the channel pore axis. We also observed that τ_{deact} predictions improved when $V_{1/2}$ labels were present and established, highlighting to a relation between $V_{1/2}$ and τ_{deact}.

For peak current density, there was a small improvement in MCC, suggesting that for this parameter, both feature types carry similar information content. Based on our input sensitivity analysis for peak current density shown in Figure 4, we deduce that evolutionary features (i.e., PSSM score) carry valuable information about protein folding stability and even though hydrophobicity-based features are highly sensitive features for peak current density, adding them to the network training process did not improve performance. This means that PSSM score, and hydrophobicity feature carry redundant information for predicting peak current density. The fact that hydrophobicity features like mutant hydrophobicity and functional density with hydrophobicity are amongst the most highly sensitive features highlights the relationship of peak current density with protein structure stability. It is possible that due to protein structure instability, KCNQ1 protein tends to misfolding resulting into dysfunction. The work from Huang *et al.*

observed that many variants in the KCNQ1 VSD negatively impact protein folding stability, leading to trafficking defect and consequently low peak current density (18). Considering that VSD region variants predominate in our database and the high input sensitivity for hydrophobicity-based features, we also conclude that protein stability is a major effect of mutations in VSD.

Combination of biophysical and evolutionary features did not improve the MCC for τ_{act} whereas AUC increased from 0.81 to 0.83. Among the different biophysical features used for predicting τ_{act}, we found that the change in number of hydrogen donor sites due to amino acid substitution significantly improved the predictions for τ_{act}. Likewise, the change in number of hydrogen bond donor sites was also found to be sensitive to τ_{act}. There were 43 variants in our dataset that experienced loss of donor sites due to amino acid substitution. Possible explanations for the number of hydrogen bond donor sites and τ_{act} association are that τ_{act} is dependent on the hydrogen bonds assisting in activation mechanism of VSD region of the protein, or these hydrogen bonds interact with PIP2 complex in sending a signal to the pore domain. Sun *et al*. found that residues near the S4 helix and S4-S5 linker helix interface interact with PIP2(34). PIP_2 has charge -3, -4 or -5 depending on pH of the surroundings, and the presence of the negative charge makes PIP2 a proton acceptor. Interestingly, the majority of the residues such as Arginine, Lysine, Tyrosine that are located within 4 Å distance from the PIP_2 interacting sites at the S4 helix and S4-S5 helix are proton donors. Thus, mutations in this protein region could lead to impairment in the function of the KCNQ1 protein due to the loss of hydrogen donor sites, impacting regulation by PIP_2. The loss of hydrogen donor sites and consequently hydrogen bonds, can also be linked to stability at the mutation site. The high sensitivity of τ_{act} for mutant hydrophobicity (see Figure 4 & 5) indicate protein destabilization as a likely cause for impairment. Furthermore, the high sensitivity to change in amino acid side chain volume highlights the structural impact on local environment near the site

of mutation. τ_{act} was also found to be sensitive to polarizability-based features highlighting the role that this feature places as driving force (response by amino acids under electric field) in ion channel activation.

Significance of this study

This work demonstrates the capability of ANN models and biophysical features to predict the phenotype of four functional significant KCNQ1 parameters. Recently, explainable AI has been an important consideration for usage of AI in medicine (42,43). We argue that ANNs trained on biophysical rather than evolutionary features enable understanding of determinants of function in a more transparent way. Moving forward with AI in biology, this understanding be critical with respect to explainable AI, i.e., understanding the relation of sequence, structure, and function of proteins. Moreover, input sensitivity analysis link biophysical features with functional parameters providing insights on underlying molecular mechanisms.

Limitations and future directions

The primary limitation of this study is the size of the dataset and over representation of variants from the VSD. There exists a substantial amount of functional data available in literature, however, those results were generated and analyzed using different expression and testing systems which may complicate model training. Our models were trained on data obtained with the same experimental approach and measurement protocols. Although we were able to train the models with 125 variants, there is a need for more functional data especially for variants in the pore domain (PD). Moreover, the incomplete KCNQ1 structure limits the ability of our models to make predictions for only half of amino acid sites (< 350 sites), mostly localized to the S1-S6 region.

An interesting question is whether analyzing multiple, possibly functional conformations of the KCNQ1 protein, modeled by Monte-Carlo or molecular dynamics simulations, can provide an

orthogonal set of information which can then be used to further improve the ANN model. There is experimental evidence that the KCNQ1 VSD can exist in one of three states (resting, intermediate, activated), which are coupled to the pore domain and influence opening. By incorporating structural and biophysical data about those states, our ANN model could learn molecular properties that underlie ion channel gating and how these properties are changed by variants. This extra information could allow the model to predict KCNQ1 gating parameters with further improved accuracy than ANNs trained on a single, static structure and address the effect side chain packing on protein stability. Furthermore, it will be valuable to investigate whether the dynamical properties of KCNQ1 protein determined in molecular dynamics simulations, can be used for interpreting the effects of variants.

4. Conclusion

We developed a model using biophysical features that can predict functional consequence of KCNQ1 VUS with comparable accuracy as a model using evolutionary features. We found that combining evolutionary and biophysical features together created optimal model performance. We used biophysical features derived from three-dimensional structure of KCNQ1 that demonstrated these features have the ability to predict function, highlighting vital structure-function relationships. Moreover, Q1VarPredBio will be a helpful tool to evaluate variants of uncertain significance and improve the accuracy of genetic diagnoses for LQTS. Q1VarPredBio is publicly available as a webserver at www.kcnq1predict.org.

5. Method and material

Neural Network Architecture and Training

A fully connected multitask feed forward ANN with a leaky rectifier linear unit was utilized for all the models. The number of nodes in the input layers was equal to the number of predictive features i.e., 14 for biophysical features, two for evolutionary features, and 12 for biophysical and evolutionary features together. The output layer of each network had four neurons, one for each phenotype prediction for functional parameters of KCNQ1. The first hidden layer with 32 neurons was found optimal for all three models with a dropout rate of 33%. For the networks trained with evolutionary features and biophysical plus evolutionary features, twelve neurons in the second hidden layer with a dropout of 33% were found optimal whereas for the model trained with biophysical features, eight neurons in the second hidden layer without the dropout rate performed better. Additionally, the biophysical model had a 20% dropout rate in the input layer whereas a 5% dropout rate was sufficient to prevent overfitting for evolutionary model and evolutionary plus biophysical model. These three networks were trained on binary labels (1 for dysfunctional and 0 for normal) for each phenotype with backpropagation of errors. Based on these errors, weights were updated for 1200 iterations with a learning rate set to 0.001 and momentum equal to 0.5. We utilized accuracy at 0.5 as our objective function during training. All neurons utilized a leaky rectifier transfer function

$$f(x) = \begin{cases} x & x > 0 \\ 0.05 & x \leq 0 \end{cases} \qquad \dots\dots\dots (1)$$

where x is the total input to a neuron. We observed that introduction of second hidden layer improved the performance of the models on silent mutations with no effect on the prediction accuracy for functional variants.

For better generalizability and balancing the different classes within the training, monitoring, and independent subsets, we adopted 25-fold cross-validations, wherein 23 subsets were utilized for training, 1 subset for monitoring, and 1 subset for prediction. For further balancing the different classes and removing any effect of biased subset used for prediction and monitoring, we randomly shuffled our data before dividing them into 25 subsets. It was observed that decreasing the size of these subsets hampers the performance suggesting an incomplete dataset for training the model. Moreover, training data was over-sampled with experimental data (=125) at a ratio of 3:1 for training the model, restricting the model to overtrain on silent mutations (=345).

Performance Metrics

The presence of size-imbalanced classes in the dataset led us to adopt Matthew's Correlation Coefficient (MCC) which is proven as the best performance metric available especially when classes in data are imbalanced (44). It considers all four parameters of the confusion matrix: numbers of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) as shown in equation (1). MCC value of 1 signifies perfect classification, the value of 0 indicates random classification, and the value of -1 means opposite classification. MCC measures the correlation of predicted value with observed value at a specific threshold. MCC values were computed by using Biochemical Library (BCL) wherein the threshold was adjusted individually for I_{Ks}, $V_{1/2}$, τ_{act} and τ_{deact} from 0 to 1 to achieve the best MCC for each phenotype.

$$\mathrm{MCC} = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TN}+\mathrm{FN})(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})}} \quad \ldots\ldots\ldots.(2)$$

Moreover, to measure the robustness of our predictive models without being dependent on this threshold value, we utilized a receiver operating characteristic (ROC) curve that summarizes the performance of different feature sets on the positive class. In ROC plots, the x-axis indicates a

false positive rate (FPR), and the y-axis indicates a true positive rate (TPR). The area under the curve (AUC) in a ROC plot quantifies the performance of the model which can be utilized to compare different models. Higher the AUC, the better the model is in distinguishing between normal (negative class) and dysfunctional (positive class) phenotypes. AUC value more than 0.5 signify that the classifier is better than a random classifier (AUC = 0.5) in distinguishing dysfunctional variants from normal variants. Similarly, AUC value less than or equal to 0.5 indicates that the classifier is unable to distinguish between positive and negative classes.

$$TPR = \frac{TP}{TP + FN} \qquad \ldots\ldots\ldots (3)$$

$$FPR = \frac{FP}{FP + TN} \qquad \ldots\ldots\ldots (4)$$

We evaluated the ability of evolutionary features, biophysical, and the combination of both the features to correctly classify the dysfunctional versus normal variants by plotting ROC curves to compare feature sets by their binary classification capability. The ROC curves for 25-fold cross validated models are shown in Figure 3 with shaped region depicting 99% confidence interval.

Input Sensitivity Analysis

The prediction of phenotypes for functional parameters are average predictions of 25 independent subsets using 25 monitoring subsets. We can analyze the average effect of input features on these 25 prediction datasets using the concept of input sensitivity. However, we acknowledge from Brown *et al.* (41) that calculating the magnitude of input sensitivity for feature importance cannot be meaningfully used due to the issue in rescaling the input features. Thus, we recall the consistency method adopted by Brown *et al.* (41) to evaluate the consistency of feature perturbation on our four result labels across the cross-validation models. Here, we iterate across

all input features of the training set and change feature value by a small amount to record the movement of the result label. For each feature with a corresponding label, we count the number of models that will improve the prediction by a change in the descriptor. The net consistency is defined as zero when half of the variants predict a positive change with respect to the result label and the other half predicts a negative change with respect to the result label. This sensitivity result reported in the result section is averaged across all the 320 models simulated by shuffling dataset 320 times, with each model individually averaged for 25 cross validation models, and instances in the training dataset for each individual feature for desired result label.

6. Acknowledgements

We acknowledge constructive discussions with Bian Li.

7. Funding source

This work is supported by NIH R01 HL122010, NIH S10 OD016216, and NIH S10 OD020154. This work is further supported by NIH R01 GM080403, NIH R01 DA046138, and NIH R01 GM129261.

8. References

1. Peter J. Schwartz, MD; Lia Crotti, MD, PhD; Roberto Insolia P. Long-QT Syndrome From Genetics to Management. Arrhythmogenic Disorders of Genetic Origin. 2012;Volume 5(Issue 4):868–77.
2. Goldenberg I, Moss AJ. Long QT Syndrome. Vol. 51, Journal of the American College of Cardiology. 2008. p. 2291–300.
3. Schwartz PJ, Stramba-Badiale M, Crotti L, Pedrazzini M, Besana A, Bosi G, et al. Prevalence of the congenital long-qt syndrome. Circulation. 2009;120(18).
4. Kapplinger JD, Tester DJ, Salisbury BA, Carr JL, Harris-Kerr C, Pollevick GD, et al. Spectrum and prevalence of mutations from the first 2,500 consecutive unrelated

patients referred for the FAMILION® long QT syndrome genetic test. Heart Rhythm. 2009 Sep 1;6(9):1297–303.

5. Wang Q, Curran ME, Splawski I, Burn TC, Millholland JM, Vanraay5 TJ, et al. Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias Refined genetic and physical localization of LQT1 [Internet]. 1996. Available from: http://www.nature.com/naturegenetics

6. M. C. Sanguinetti, M. E. Curran, A. Zou, J. Shen, P. S. Spector DLA& MTK. Coassembly of KvLQT1 and minK (lsK) proteins to form cardiac fKs potassium channel. 1996.

7. Barhanin J, Lesage F, Guillemare E, Fink M, Lazdunski M, Romey G. K(V)LQT1 and IsK (minK) proteins associate to form the I(Ks) cardiac potassium current. Nature. 1996;384(6604):78–80.

8. Wu J, Ding W-G, Horie M. Molecular pathogenesis of long QT syndrome type 1. Available from: http://dx.doi.org/10.1016/j.joa.2015.12.006

9. Apgar TL, Sanders CR. Compendium of causative genes and their encoded proteins for common monogenic disorders. Protein science : a publication of the Protein Society [Internet]. 2021 [cited 2021 Nov 23]; Available from: https://pubmed.ncbi.nlm.nih.gov/34515378/

10. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Research [Internet]. 2016;44. Available from: https://www.

11. Stenson PD, Mort M, Ball · Edward v, Evans K, Hayden M, Heywood S, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet [Internet]. 2017;136:665–77. Available from: https://www.

12. Ackerman MJ. Genetic purgatory and the cardiac channelopathies: Exposing the variants of uncertain/unknown significance issue [Internet]. Heart Rhythm. 2015 [cited 2021 Sep 21]. p. 2325–31. Available from: https://reader.elsevier.com/reader/sd/pii/S154752711500884X?token=267BCA7D23D522C8DCEC1CF5E75FF4BF10F0E67DE8E2C5F813DC501D2674500B8D2CBF5ADD74593C8E3BB5878D9BBF6A&originRegion=us-east-1&originCreation=20210922195928

13. Giudicessi JR, Ackerman MJ. Genetic testing in heritable cardiac arrhythmia syndromes: differentiating pathogenic mutations from background genetic noise. 2012;

14. Vanoye CG, Desai RR, Fabre KL, Gallagher SL, Potet F, DeKeyser J-M, et al. High-Throughput Functional Evaluation of KCNQ1 Decrypts Variants of Unknown

Significance. Circulation: Genomic and Precision Medicine [Internet]. 2018 Nov 1 [cited 2021 Sep 21];11(11):e002345. Available from: https://www.ahajournals.org/doi/abs/10.1161/CIRCGEN.118.002345

15. Yang T, Chung S-K, Zhang W, Mullins JGL, Mcculley CH, Crawford J, et al. Biophysical Properties of 9 KCNQ1 Mutations Associated With Long-QT Syndrome. 2009; Available from: http://www.cgl.ucsf.edu/chimera/

16. Restier L, Cheng L, Sanguinetti MC. Mechanisms by which atrial fibrillation-associated mutations in the S1 domain of KCNQ1 slow deactivation of I Ks channels. J Physiol. 2008;586:4179–91.

17. Eldstrom J, Wang Z, Werry D, Wong N, Fedida D. Microscopic mechanisms for long QT syndrome type 1 revealed by single-channel analysis of IKs with S3 domain mutations in KCNQ1. Heart Rhythm [Internet]. 2015 Feb 1 [cited 2021 Sep 21];12(2):386–94. Available from: http://www.heartrhythmjournal.com/article/S1547527114012466/fulltext

18. Huang H, Kuenze G, Smith JA, Taylor KC, Duran AM, Hadziselimovic A, et al. Mechanisms of KCNQ1 channel dysfunction in long QT syndrome involving voltage sensor domain mutations. Science Advances. 2018 Mar 7;4(3).

19. Aromolaran AS, Subramanyam P, Chang DD, Kobertz WR, Colecraft HM. LQT1 mutations in KCNQ1 C-terminus assembly domain suppress I Ks using different mechanisms. Available from: https://academic.oup.com/cardiovascres/article/104/3/501/2930990

20. Huang H, Chamness LM, Vanoye CG, Kuenze G, Meiler J, George AL, et al. Disease-linked supertrafficking of a potassium channel. The Journal of biological chemistry [Internet]. 2021 Jan 1 [cited 2021 Nov 24];296. Available from: https://pubmed.ncbi.nlm.nih.gov/33600800/

21. Bhuiyan ZA. Silent mutation in long QT syndrome: Pathogenicity prediction by computer simulation. Heart Rhythm [Internet]. 2012 Feb 1 [cited 2021 Nov 24];9(2):283–4. Available from: http://www.heartrhythmjournal.com/article/S1547527111012148/fulltext

22. Li B, Mendenhall JL, Kroncke BM, Taylor KC, Huang H, Smith DK, et al. Predicting the Functional Impact of KCNQ1 Variants of Unknown Significance. Circulation: Cardiovascular Genetics [Internet]. 2017 Oct 1 [cited 2021 Nov 24];10(5). Available from: https://www.ahajournals.org/doi/abs/10.1161/CIRCGENETICS.117.001754

23. Kernik DC, Yang P-C, Kurokawa J, Wu JC, Clancy CE. A computational model of induced pluripotent stem-cell derived cardiomyocytes for high throughput risk stratification of KCNQ1 genetic variants. PLOS Computational Biology [Internet].

2020 Aug 1 [cited 2021 Sep 21];16(8):e1008109. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008109

24. Giudicessi JR. Machine Learning and Rare Variant Adjudication in Type 1 Long QT Syndrome. Circulation: Cardiovascular Genetics [Internet]. 2017 Oct 1 [cited 2021 Sep 21];10(5). Available from: https://www.ahajournals.org/doi/abs/10.1161/circgenetics.117.001944

25. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature Methods. 2010 Apr;7(4):248–9.

26. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature Protocols 2009 4:7 [Internet]. 2009 Jun 25 [cited 2021 Sep 21];4(7):1073–81. Available from: https://www.nature.com/articles/nprot.2009.86

27. Davydov E v., Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS computational biology [Internet]. 2010 Dec [cited 2021 Nov 24];6(12). Available from: https://pubmed.ncbi.nlm.nih.gov/21152010/

28. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research [Internet]. 2019 Jan 8 [cited 2021 Sep 21];47(Database issue):D886. Available from: /pmc/articles/PMC6323892/

29. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. The American Journal of Human Genetics. 2016 Oct 6;99(4):877–85.

30. Kroncke BM, Mendenhall J, Smith DK, Sanders CR, Capra JA, George AL, et al. Protein structure aids predicting functional perturbation of missense variants in SCN5A and KCNQ1. Computational and Structural Biotechnology Journal. 2019 Jan 1;17:206–14.

31. Sallah SR, Sergouniotis PI, Barton S, Ramsden S, Taylor RL, Safadi A, et al. Using an integrative machine learning approach utilising homology modelling to clinically interpret genetic variants: CACNA1F as an exemplar. European Journal of Human Genetics 2020 28:9 [Internet]. 2020 Apr 20 [cited 2021 Sep 21];28(9):1274–82. Available from: https://www.nature.com/articles/s41431-020-0623-y

32. Kroncke BM, Duran AM, Mendenhall JL, Meiler J, Blume JD, Sanders CR. Documentation of an Imperative To Improve Methods for Predicting Membrane Protein Stability. Biochemistry [Internet]. 2016 Sep 13 [cited 2021 Sep

21];55(36):5002–9. Available from: https://pubs.acs.org/doi/full/10.1021/acs.biochem.6b00537

33. Vanoye CG, Thompson CH, Desai RR, DeKeyser JM, Chen L, Rasmussen-Torvik LJ, et al. Functional evaluation of human ion channel variants using automated electrophysiology. Methods in enzymology [Internet]. 2021 Jan 1 [cited 2021 Nov 24];654:383–405. Available from: https://pubmed.ncbi.nlm.nih.gov/34120723/

34. Sun J, MacKinnon R. Structural basis of human KCNQ1 modulation and gating. Cell [Internet]. 2020 Jan 23 [cited 2021 Sep 26];180(2):340. Available from: /pmc/articles/PMC7083075/

35. FAUCHÈRE J -L, CHARTON M, KIER LB, VERLOOP A, PLISKA V. Amino acid side chain parameters for correlation studies in biology and pharmacology. International journal of peptide and protein research [Internet]. 1988 [cited 2021 Nov 24];32(4):269–78. Available from: https://pubmed.ncbi.nlm.nih.gov/3209351/

36. Koehler J, Woetzel N, Staritzbichler R, Sanders CR, Meiler J. A Unified Hydrophobicity Scale for Multi-Span Membrane Proteins. Proteins [Internet]. 2009 Jul [cited 2021 Sep 26];76(1):13. Available from: /pmc/articles/PMC2761718/

37. Durham E, Dorr B, Woetzel N, Staritzbichler R, Meiler J. Solvent accessible surface area approximations for rapid and accurate protein structure prediction. Journal of Molecular Modeling [Internet]. 2009 [cited 2021 Sep 26];15(9):1093. Available from: /pmc/articles/PMC2712621/

38. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research [Internet]. 1997 Sep 1 [cited 2021 Sep 26];25(17):3389–402. Available from: https://academic.oup.com/nar/article/25/17/3389/1061651

39. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH. UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics [Internet]. 2007 May 15 [cited 2021 Sep 26];23(10):1282–8. Available from: https://academic.oup.com/bioinformatics/article/23/10/1282/197795

40. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Research [Internet]. 2007 Jan [cited 2021 Nov 24];35(Database issue):D61. Available from: /pmc/articles/PMC1716718/

41. Brown BP, Mendenhall J, Geanes AR, Meiler J. General Purpose Structure-Based Drug Discovery Neural Network Score Functions with Human-Interpretable Pharmacophore Maps. Journal of Chemical Information and Modeling [Internet]. 2021 Feb 22 [cited 2021 Sep 26];61(2):603–20. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jcim.0c01001

42. Kundu S. AI in medicine must be explainable. Nature medicine [Internet]. 2021 Aug 1 [cited 2021 Dec 1];27(8):1328. Available from: https://pubmed.ncbi.nlm.nih.gov/34326551/

43. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley interdisciplinary reviews Data mining and knowledge discovery [Internet]. 2019 Jul 1 [cited 2021 Dec 1];9(4). Available from: https://pubmed.ncbi.nlm.nih.gov/32089788/

44. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 2019 21:1 [Internet]. 2020 Jan 2 [cited 2021 Sep 26];21(1):1–13. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-019-6413-7