SUPPLEMENTAL MATERIAL

Supplemental Methods

Data collected from the literature. We searched for the terms “KCNH2”, “hERG”, and “KV11.1” in PubMed and identified 659 papers which reported *KCNH2* variant functional data and/or heterozygote clinical phenotypes. To these data, we added individuals heterozygous for *KCNH2* variants in the genome Aggregation Database (gnomAD) of population variation (http://gnomad.broadinstitute.org/; release 2.0)[18] and classified all as not meeting criteria for “affected” status, accepting a small classification error due to the rarity of long QT syndrome in the general population (~1 in 2,500; see also Supplemental Text).[47] From the combined literature (659 papers) and gnomAD sources, we found 4,810 individuals (2,985 from gnomAD) heterozygous for 871 unique missense or in-frame insertion/deletion (indel) *KCNH2* variants (< 0.001 MA; 538 from gnomAD). Additionally, five centers that hold cardiology clinics and conduct research gathered clinical phenotypes and genotypes for individuals heterozygous for *KCNH2* variants referred based on a suspicion of LQT2, including Unité de Rythmologie, Centre de Référence Maladies Cardiaques Héréditaires, Service de Cardiologie, Hôpital Bichat, Paris, France; the Center for Cardiac Arrhythmias of Genetic Origin Istituto Auxologico Italiano IRCCS, Milan, Italy; Shiga University of Medical Science Department of Cardiovascular and Respiratory Medicine, Shiga, Japan; National Cerebral and Cardiovascular Center, Osaka, Japan; Nagasaki University, Nagasaki, Japan. The Bichat group provided 140 unique *KCNH2* variants carried by 418 individuals, 243 of which were diagnosed with LQTS; the Istituto Auxologico Italiano IRCCS provided 141 unique KCNH2 variants carried by 467 individuals, 295 of which were diagnosed with LQTS; the combined Japanese cohort has provided 279 unique *KCNH2* variants carried by 647 individuals, 457 of which were diagnosed with LQTS. In addition to these data, we found

871 unique *KCNH2* variants (< 0.001 MAF) carried by 4,810 individuals, 1,041 with LQTS, derived from the literature and gnomAD. We have removed any potential overlapping patients between the literature and the data provided from these sites. All statistical evaluations were performed against the observed LQT2 probability from the data. The cohort dataset was withheld from all model training and used strictly for validation.

Definition of LQT2 status. Each heterozygous individual was annotated as LQT2 (affected) via the LQTS diagnosis described in the HRS/EHRA/APHRS Expert Consensus Statement on the Management of Inherited Arrhythmias;[48] however, in the absence of genotype data: an LQTS risk score > 3.5 [Schwartz et al. (1993)[49]] without secondary causes for QT prolongation or a QTc > 480 in repeated 12-lead ECG without secondary causes for QT prolongation. In cases where these data were not available in the literature, we assigned affected status according to diagnosis reported. Heterozygotes with reported QTc at baseline and history of events below this threshold were annotated as not meeting the criteria for affected status.

KCNH2 *functional data and* in silico *predictive covariates.* We collected functional data, as previously described. For variants with multiple sources of functional data, we selected data according to a hierarchy of heterozygous expression systems based on the greatest number of unique variants for each cell type. Since most variant data was collected in HEK293 cells, these were given highest preference, followed by CHO, COS-7, murine, hiPSC-CM, and Xenopus Oocytes. Functional data for each variant characterized more than once in the literature in the same expression system were averaged. Since our method has the greatest potential utility assisting in the interpretation of rare variants, we chose to calibrate to the data which will conceivably be generated at this scale. The advancement of high-throughput technologies for rapid variant reclassification is expected to further increase the quantity of data derived from heterologous expression systems. Values were permitted to range continuously, and we avoided any attempt to establish binary or categorical delineations among these data. In

addition, co-expression of *KCNH2* and the beta subunit, *KCNE2*, was relatively rare and we therefore deferred to data where only the alpha subunit was expressed. We additionally included *in silico* pathogenicity predictions from three commonly used servers: SIFT, Polyphen-2, and PROVEAN. We also included basic local alignment search tool position-specific scoring matrix (BLAST-PSSM) and point accepted mutation score (PAM). Lastly, we included the meta-predictor REVEL given favorable performance characteristics on several datasets.[16]

LQT2 probability density. Recently, an atomic-resolution structure of human K_V11.1 (*KCNH2* protein product) was determined.[31] Combined with observations of carriers of *KCNH2* variants, this structure allows for the quantitation of regions within K_V11.1 that are enriched or depleted for high-risk variants, which are the variants with higher or lower observed LQT2 probability, respectively. This allowed us to evaluate the combination of atomic resolution structure and regions in three-dimensional space enriched for disease. We applied our previously published method to calculate the average *LQT2 probability densit*y in a shell of residues surrounding a residue of interest (previously referred to as penetrance density).[11, 50]

LQT2 predictive models. We used our previous algorithm [11] to combine covariates (e.g. PROVEAN, LQT2 probability density, peak tail current, etc.) regressed to the fraction of heterozygous carriers diagnosed with LQT2. The algorithm uses a pattern mixture model to handle missing data.[11] Multiple models are fit, one for each missing data pattern. The details of this approach were described previously.[11] Briefly, we used a variation of the expectation maximization (EM) algorithm comprising two steps: 1) calculate the expected penetrance from an empirical Bayes penetrance model and 2) fit regression of our estimated penetrance on variant-specific characteristics by maximum likelihood.[11] Resulting models can be used to generate a predicted penetrance and nonparametric variance estimate based on local averaging of mean squared error (from the difference between posterior mean penetrance and prior mean penetrance). The model is then used to generate an updated prior and subsequent

posterior expected penetrance; this process is iterated until it converges (when the estimated posterior means change by less than 1.0% from the previous iteration). Using a beta-binomial model allows estimation of the prior from a predicted penetrance and its associated variance. The resulting estimates are combined with observations of individuals diagnosed and not diagnosed with LQT2 for the final estimate of LQT2 post-test probability. Variants are weighted by $1 - \frac{1}{0.01+total\ heterozygotes}$ in fitting routines to ensure variants with higher total heterozygote counts had greater influence in the model; e.g. a variant with one carrier has a weight of 0.01, a variant with two carriers has a weight of 0.5, a variant with three carriers has a weight of 0.67.

Statistical evaluation. Area under the receiver operating characteristic curves were computed using the pROC package in R. All reported correlation coefficients, Spearman ρ, Pearson's R^2, and coefficient of determination were weighted by the function $1 - \frac{1}{0.01+total\ heterozygotes}$, unless otherwise noted, to ensure variants with higher total heterozygote counts had greater influence in the resulting correlation coefficient estimate. We calculate a single weighted Spearmen correlation coefficient between the observed feature (in vitro functional data and in silico predictors) and the observed LQT2 penetrance or disease probability. The LQT2 probability/penetrance is estimated as the number of affected individuals over the total number of carriers and is the dependent variable. The Spearman correlation is then calculated using the wCorr package in R (https://rdrr.io/cran/wCorr/f/vignettes/wCorrFormulas.Rmd) which weights the ranks then calculates a weighted Pearson correlation coefficient. Our objective was to treat variants with many observations the same while down-weighting variants with very few or only one observation (see Figure S15). We suggest the weighted Spearman values be interpreted as indicating significant associations between the evaluated feature and observed LQT2 disease probability/penetrance. All scripts and data used are available at the Kroncke Lab GitHub page (https://github.com/kroncke-lab/Bayes_KCNH2_LQT2_Penetrance). Additionally, a compiled

and curated form of the data presented here are available in the *KCNH2* Variant Browser (https://oates.app.vumc.org/vancart/KCNH2/index.php; Figure S1)

Estimate of precision. We scaled the variance from the convergent EM result by a factor of `v`. At each level of `v` and for each variant, we sampled from binomial distribution with n of 100 and probability of $\frac{\text{LQT2 cases}}{\text{total carriers}}$, the observed LQT2 probability for each *KCNH2* variant, to produce the quantity "LQT2 sampled." We then calculated the resulting 95% posterior credible interval from the Beta distribution with shape parameters 1) $\text{LQT2 sampled} + \alpha_{\text{prior,EM}}$ and 2) $100 - \text{LQT2 sampled} + \beta_{\text{prior,EM}}$. We repeated this process 1000 times and calculated the rate of the posterior credible interval covering the "true" observed LQT2 probability ($\frac{\text{LQT2 cases}}{\text{total carriers}}$). As seen in Figures S10-S12, both large and small `v` values result in coverages which deviate from the expected 95%, many below and above, respectively. We selected the best `v` from the coverage plots which balances the tradeoff of over-coverage in variants with medium-low LQT2 probability and under-coverage of variants with high LQT2 probability. From this procedure, we chose a tuning parameter of 'v' = 10 (Figures S10-S12).

Inclusion of individuals from gnomAD. Individuals in gnomAD are mostly unaffected, given the rarity of LQTS; however, the data available from that resource could be contaminated with individuals presenting with LQTS, though likely at or near the rate in the general public. To test the sensitivity of our results to this type of misclassification, we performed a sensitivity analysis like what was done in our previous work on SCN5A and Brugada syndrome. We randomly switched individuals from unaffected (gnomAD) to LQTS cases for each variant, and we did this for a rate of LQTS in gnomAD at the population rate (1:2,500) and at 10x and 50x the population rate of disease. We then re-calculated the variant LQTS probability of all variants. We repeated this process 100 times and checked what is the magnitude of the change (as measured by the mean absolute change) and the expected number of variants which have any change. When we used a misassignment rate of 1x the prevalence of LQTS, we found the median rate of LQTS

probability change was 0.05% (with an interquartile range of 0.2%) and the expected number of variants with a change in LQTS probability was 3. With a misassignment rate of 10x the prevalence of LQTS, we found the median rate of LQTS probability change was 0.1% (interquartile range of 0.5%) and the expected number of variants with a change in LQTS probability was 14. With a misassignment rate of 50x the prevalence of LQTS, we found the median rate of LQTS probability change was 0.1% (interquartile range of 0.7%) and the expected number of variants with a change in LQTS probability was 51.

Supplemental Tables

Table Sl. Weighted R^2 between the fraction of heterozygotes diagnosed with LQT2 from literature and cohort for variants where heterozygously-collected peak tail current is known.

Model	Literature (n = 80)†	Cohort (n = 38)†
Heterozygous Peak Tail Current	0.34 [0.19-0.54]	0.48 [0.27-0.70]
LQT2 probability density	0.62 [0.46-0.78]	0.42 [0.17-0.65]
REVEL	0.44 [0.25-0.62]	0.50 [0.29-0.73]
Post-test LQT2 Probability Estimate	0.90 [0.82-0.95]	0.57 [0.31-0.78]

†Weighted R^2 [95% Confidence Interval], weighted by $1 - \frac{1}{0.01 + total\ heterozygotes}$

Table SII. Weighted R^2 between the fraction of homozygotes diagnosed with LQT2 from cohort and predictive features, gnomAD is not included.

Models trained with displayed subsets of features using the same subset of variants.

Model	All Variants (n = 246)†	Variants with Peak Tail Current (n = 38)†
Heterozygous Peak Tail Current	-	0.48 [0.27-0.67]
LQT2 probability density	0.13 [0.04-0.24]	0.22 [0.01-0.56]
REVEL	0.21 [0.10-0.33]	0.50 [0.31-0.72]
Post-test LQT2 Probability Estimate	0.26 [0.15-0.39]	0.51 [0.21-0.75]

†Weighted R^2 [95% Confidence Interval], weighted by $1 - \frac{1}{0.01 + total\ heterozygotes}$

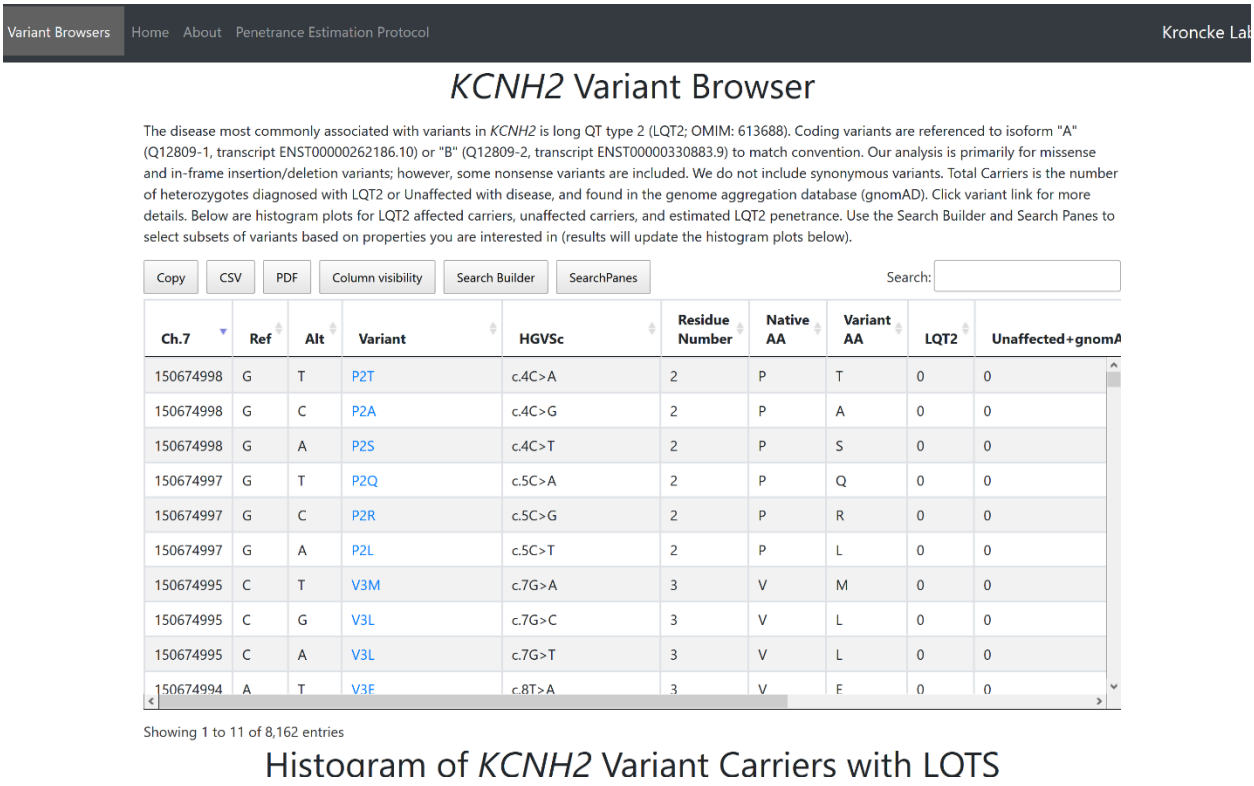
Table SIII. Weighted R^2 between the fraction of homozygotes diagnosed with LQT2 from sites within the cohort and predictive features.

LQT2 Probability Estimates	Italy (n = 84)†	Japan (n = 141)†	France (n = 77)†
LQT2 probability density	0.31 [0.13-0.52]	0.22 [0.09-0.37]	0.17 [0.02-0.39]
REVEL	0.37 [0.20-0.55]	0.17 [0.06-0.33]	0.06 [0.00-0.29]
Post-test LQT2 Probability	0.41 [0.22-0.61]	0.29 [0.15-0.45]	0.25 [0.09-0.48]

†Weighted R^2 [95% Confidence Interval] for the same subset of variants, weighted by $1 - \frac{1}{0.01+total\ heterozygotes}$, n is the number of unique *KCNH2* variants

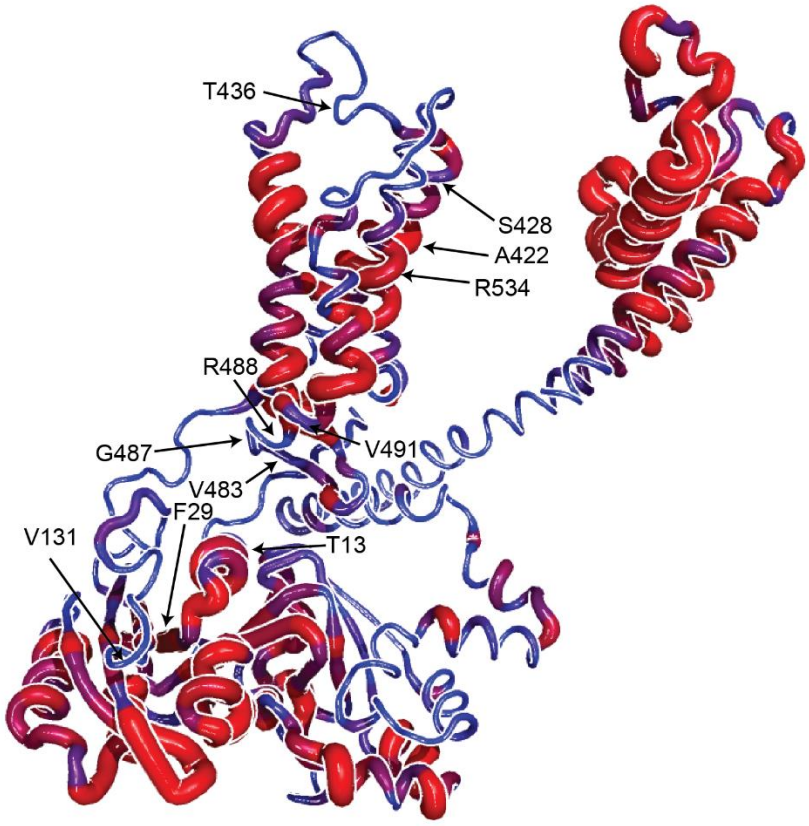
Supplemental Figures

Supplemental Figure I



Supplemental Figure I. Screen shot of the web-based searchable resource which contains the data used here.

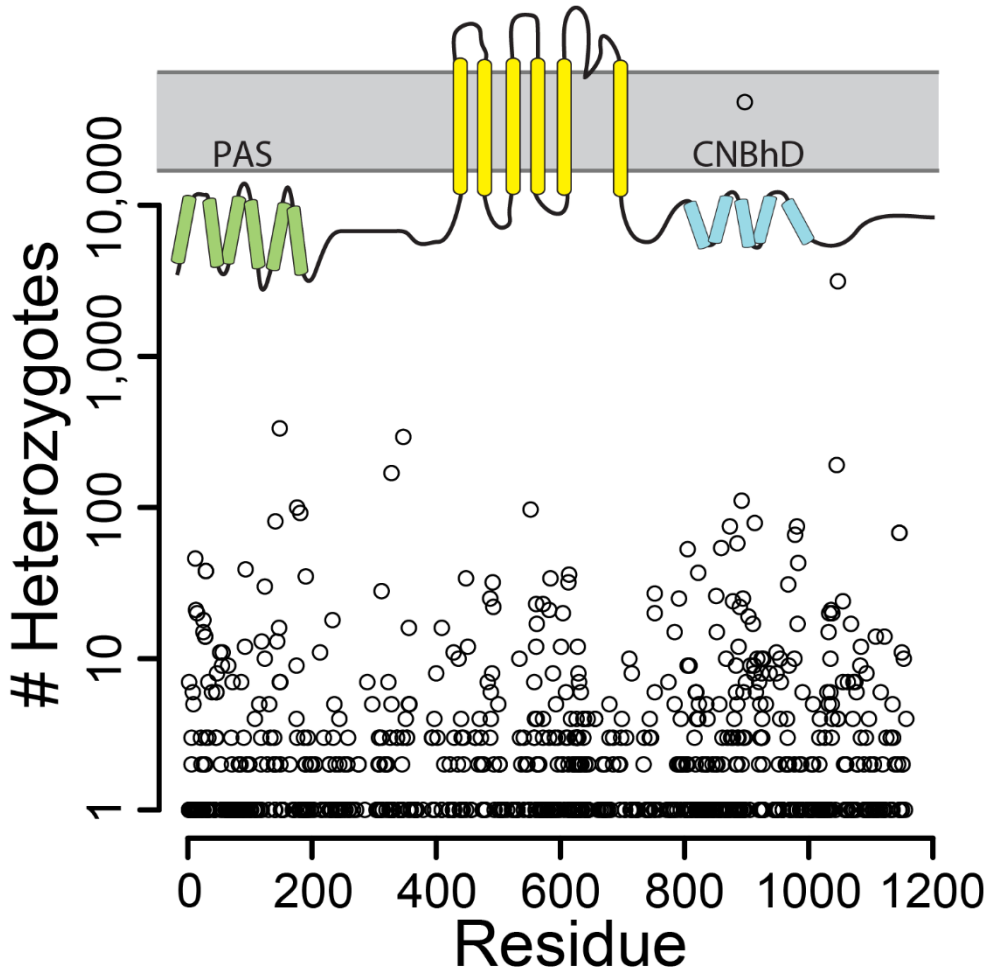
Supplemental Figure II



Supplemental Figure II. Structure of an individual subunit of the K_V11.1 channel (N-terminus through the voltage-sensing domain of one subunit of the tetrameric channel and the pore through the CNBhD of an adjacent subunit, shown this way for clarity). Larger, red segments indicate higher LQT2 penetrance density; smaller, blue segments indicate lower LQT2 penetrance density. As mentioned in the main text, the canonical domains of K_V11.1 are heterogeneous for high and low LQT2 probability. As an example, in the PAS domain, p.Thr13Asn (c.38C>A) was seen in 21 individuals in gnomAD, and 4 individuals in the France cohort (one of which met the criteria for LQT2 diagnosis), p.Val131Leu (c.391G>T) was seen in 5 individuals in gnomAD. In contrast, within the same domain, p.Phe29Leu (c.87C>A) was seen in 38 individuals, 28 met the criteria for LQT2 diagnosis; p.Phe22Tyr (c.65T>A) was seen in 5 individuals, 4 from the Italy cohort and 1 from the Japan cohort, all of which met the criteria for LQT2 diagnosis. In the voltage-sensing

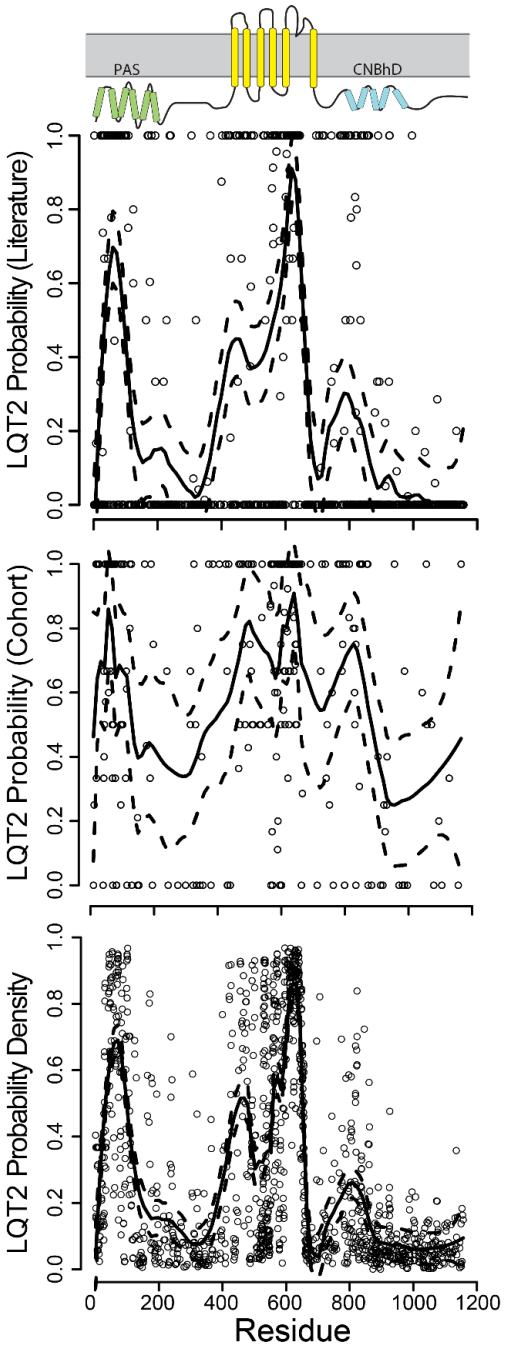
domain of K_V11.1, some residues are in areas of high LQT2 penetrance. Residues 420-426 on the S1 helix contact residues 531-534 on the S4 helix also in the voltage-sensor domain and residues 556-563 on the S5 helix in the pore. Variants within these segments observed in carriers show high penetrance: p.Ala422Thr (c.1264G>A) seen in 3 carriers, all diagnosed with LQT2, p.Pro426His (c.1277C>A) (3 heterozygotes observed, 2 with LQTS), p.Tyr420Cys (c.1259A>G, 1 heterozygote with LQTS), p.Thr421Met (c.1262C>T) (1 heterozygote with LQTS), p.Arg534Cys (c.1600C>T) on the S4 helix (22 heterozygotes, 20 with LQT2), p.Ala558Pro (c.1672G>C) on the S5 helix (7 heterozygotes, 6 with LQT2), p.Ala561Val (c.1682C>T) on the S5 helix, one of the most well-studied variants (39 heterozygotes, 35 with LQT2), and p.Ala561Thr (c.1681G>A) on the S5 helix (28 heterozygotes, 25 with LQT2). However, also in the voltage sensor domain, there are regions of variants with relatively low estimated LQT2 penetrance. For example, p.Val483Ile (c.1447G>A) (7 heterozygotes found in gnomAD), p.Gly487Ser (c.1459G>A) (24 heterozygotes in gnomAD and 1 unaffected heterozygote from literature), p.Arg488Cys (c.1461C>T) (6 heterozygotes in gnomAD), p.Val491Ile (c.1471G>A) (31 heterozygotes in gnomAD and one unaffected heterozygote from the literature). Towards the extracellular half of the voltage sensor variants are also associated with low penetrance: p.Thr436Met (c.1307C>T) 1 with Afib, 1 unaffected, 9 in gnomAD. p.Ser428Leu (c.1283C>T; 5 heterozygotes, 0 diagnosed with LQT2; 4 individuals without symptoms, one with a QTc of 470ms, others between 426 and 440, one asymptomatic individual with a QTc of 450 in the Japan cohort.

Supplemental Figure III



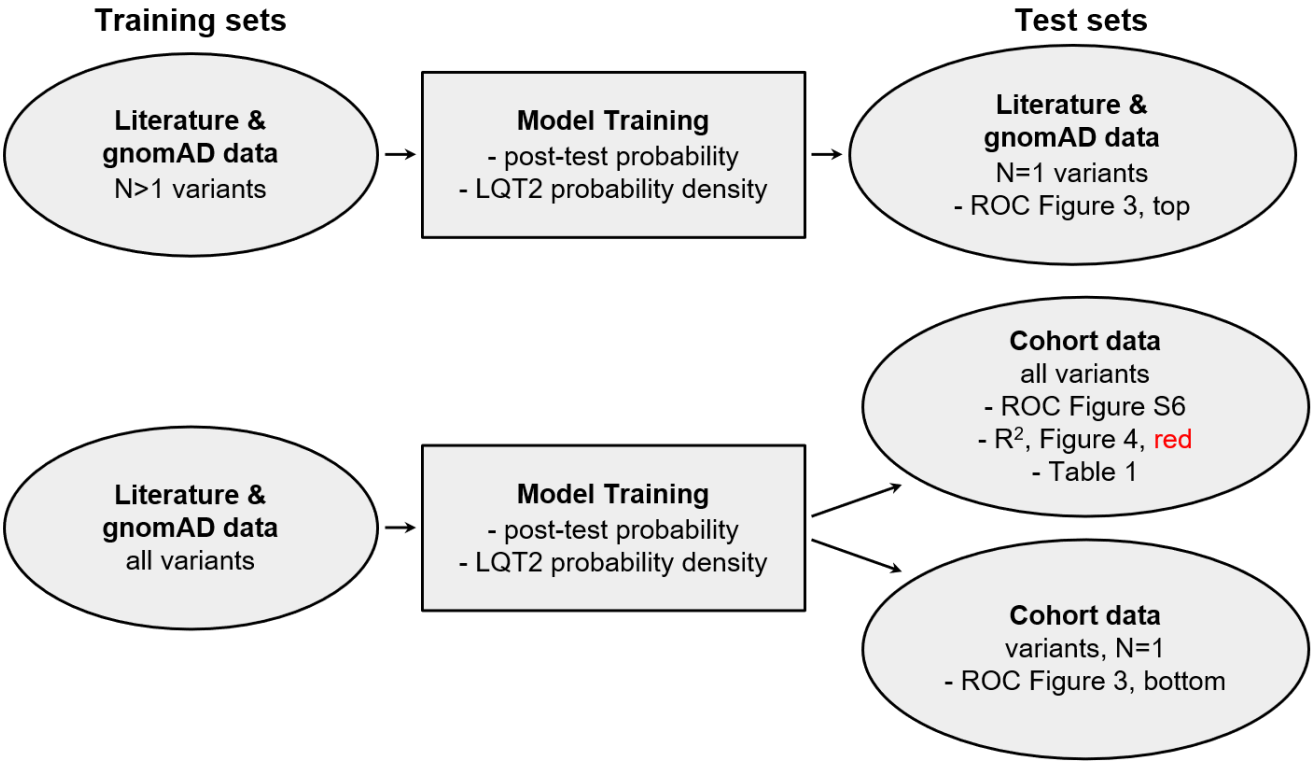
Supplemental Figure III. Heterozygote counts by residue (note y-axis log scale). Variants used to calculate LQT2 probability are uniformly distributed through *KCNH2*.

Supplemental Figure IV



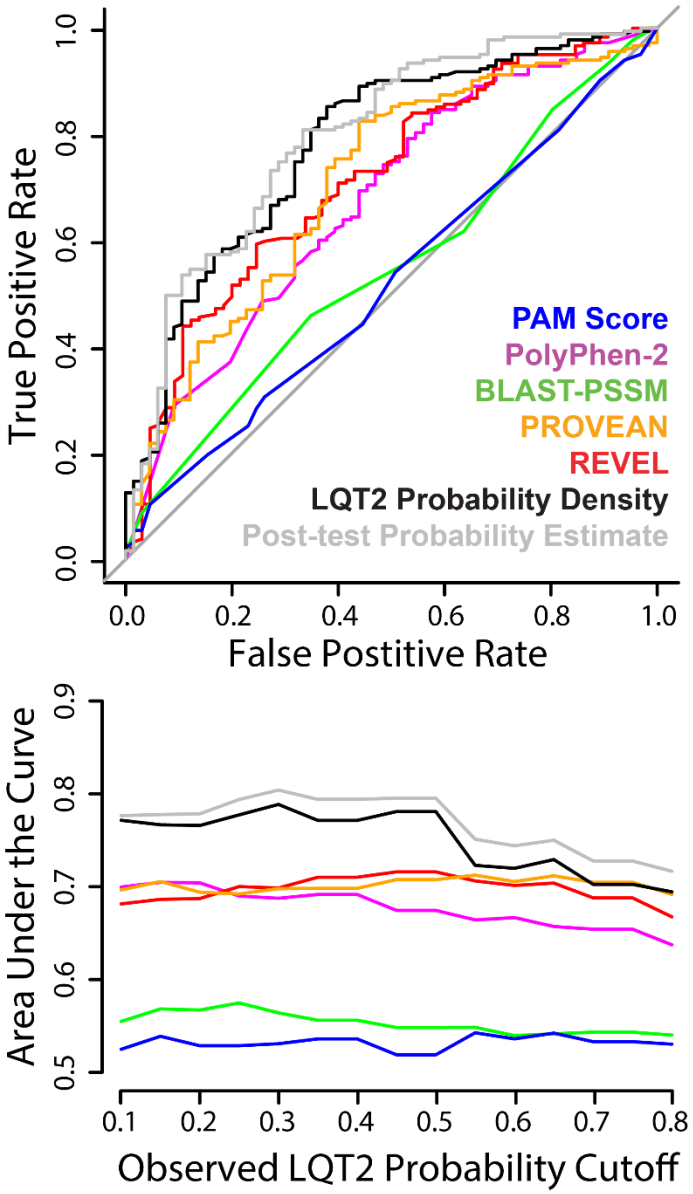
Supplemental Figure IV. LQT2 probability versus residue from the literature, cohort, and structure (density). LQT2 probability density values are calculated based on distances between residues in the K_V11.1 structure(13) and observed LQT2 probability from the literature (cohort data were not included in the calculation). The solid and dashed lines are a locally weighted average of the mean observed LQT2 probability and the 95% confidence estimates on that mean, respectively.

Supplemental Figure V



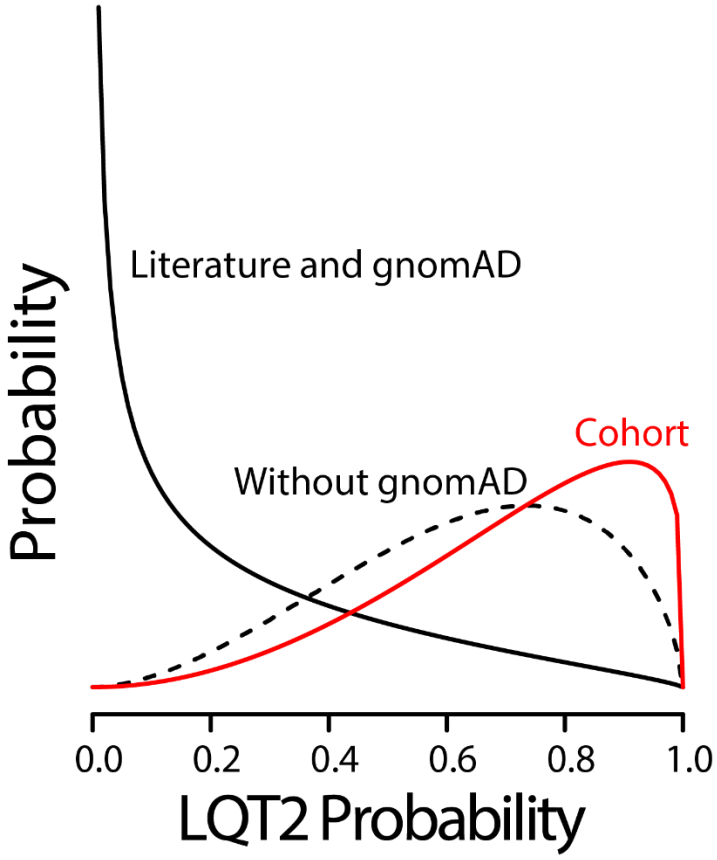
Supplemental Figure V. Graphic describing how EM and LQT2 probability density ROC curves were calculated in Figure 3 (top) and how post-test probability estimate and LQT2 probability density ROC curves and R^2 values were calculated in Figure 3 (bottom), Figure S6, and Table 1.

Supplemental Figure VI



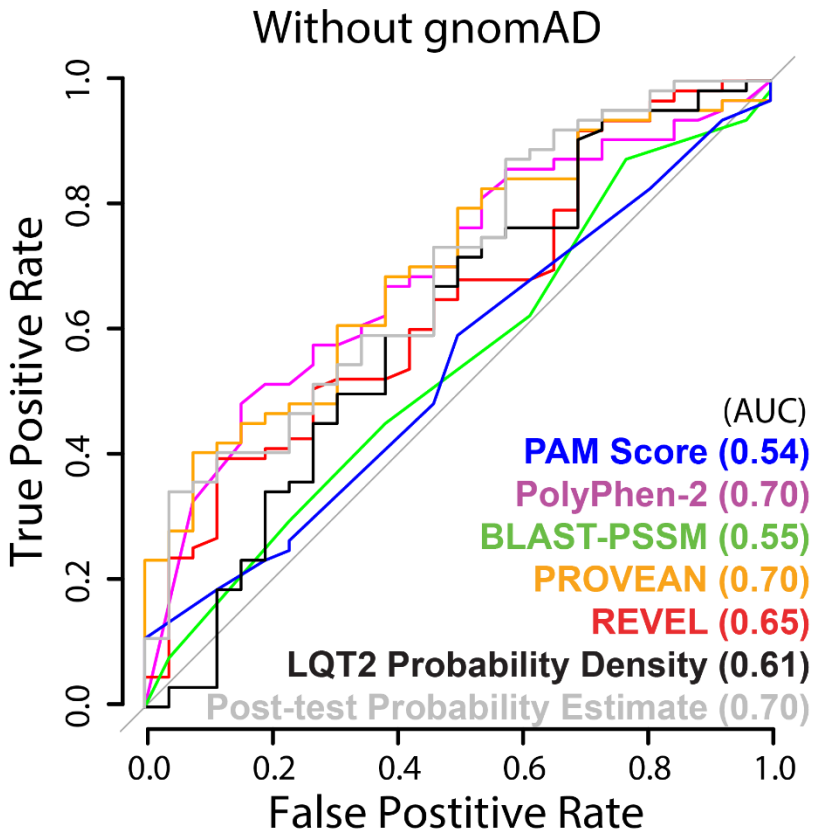
Supplemental Figure VI. Receiver operating characteristic (ROC) curves and area under the ROC curves for *KCNH2* variants in the cohort dataset by covariates and predictive models. Top: ROC curve evaluated against observed cohort probability with cutoff set to 0.2, an observed probability of 20% or greater. Bottom: AUCs for covariates and EM model evaluated at multiple cutoffs. EM prediction consistently has the highest AUC.

Supplemental Figure VII



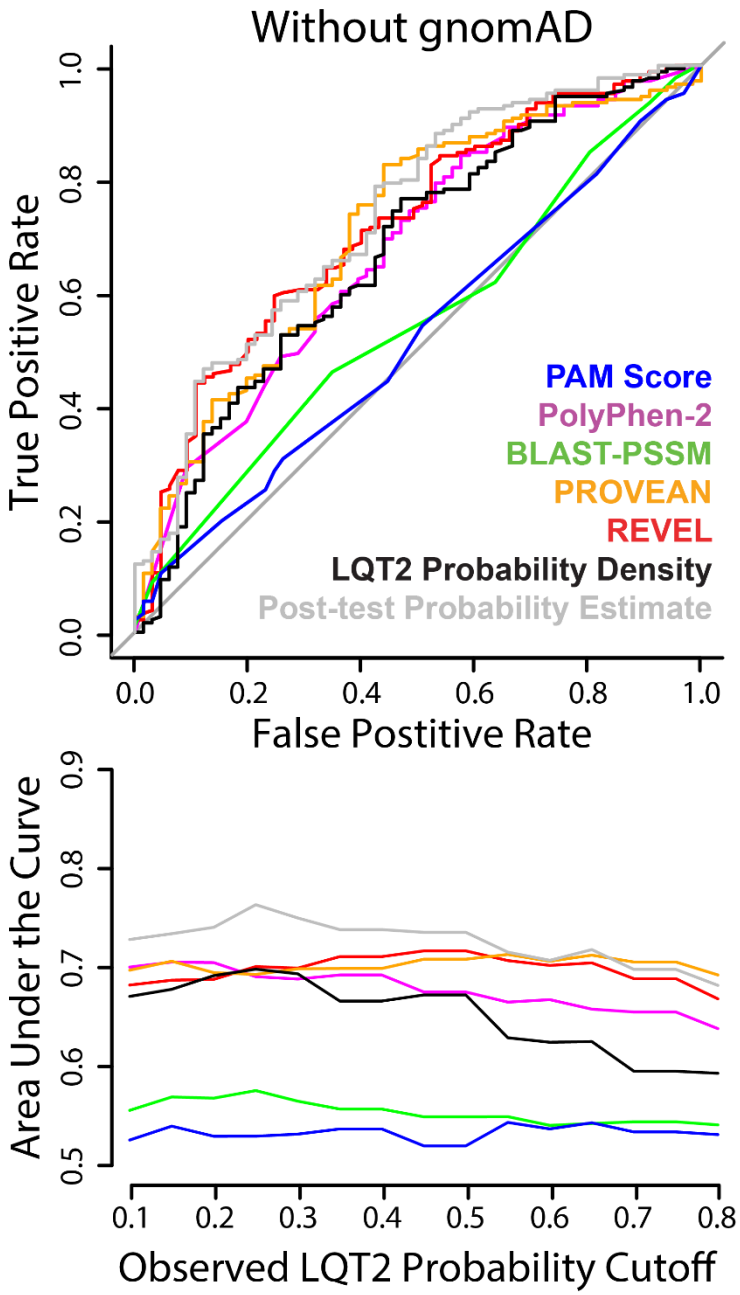
Supplemental Figure VII. Empirical priors were calculated by taking the weighted mean observed LQT2 probability as the average and mean squared error from the weighted mean to the observed value as the variance to define a beta distribution. This can be interpreted as the probability of LQT2 (given that the individual is heterozygous for an unknown missense or indel variant) if the priors were informed by the cohort data alone (red), the literature alone (black, dashed), or the literature combined with gnomAD (black, solid).

Supplemental Figure VIII



Supplemental Figure VIII. Receiver operating characteristic (ROC) curves from predictors (trained with the literature data without gnomAD) against variants in the cohort with only one observation, an individual affected with LQT2 or not. All cohort data were withheld from the post-test probability estimate and LQT2 probability density algorithms during construction. The area under the ROC curves are lower for post-test probability estimate and LQT2 probability density than when gnomAD is included.

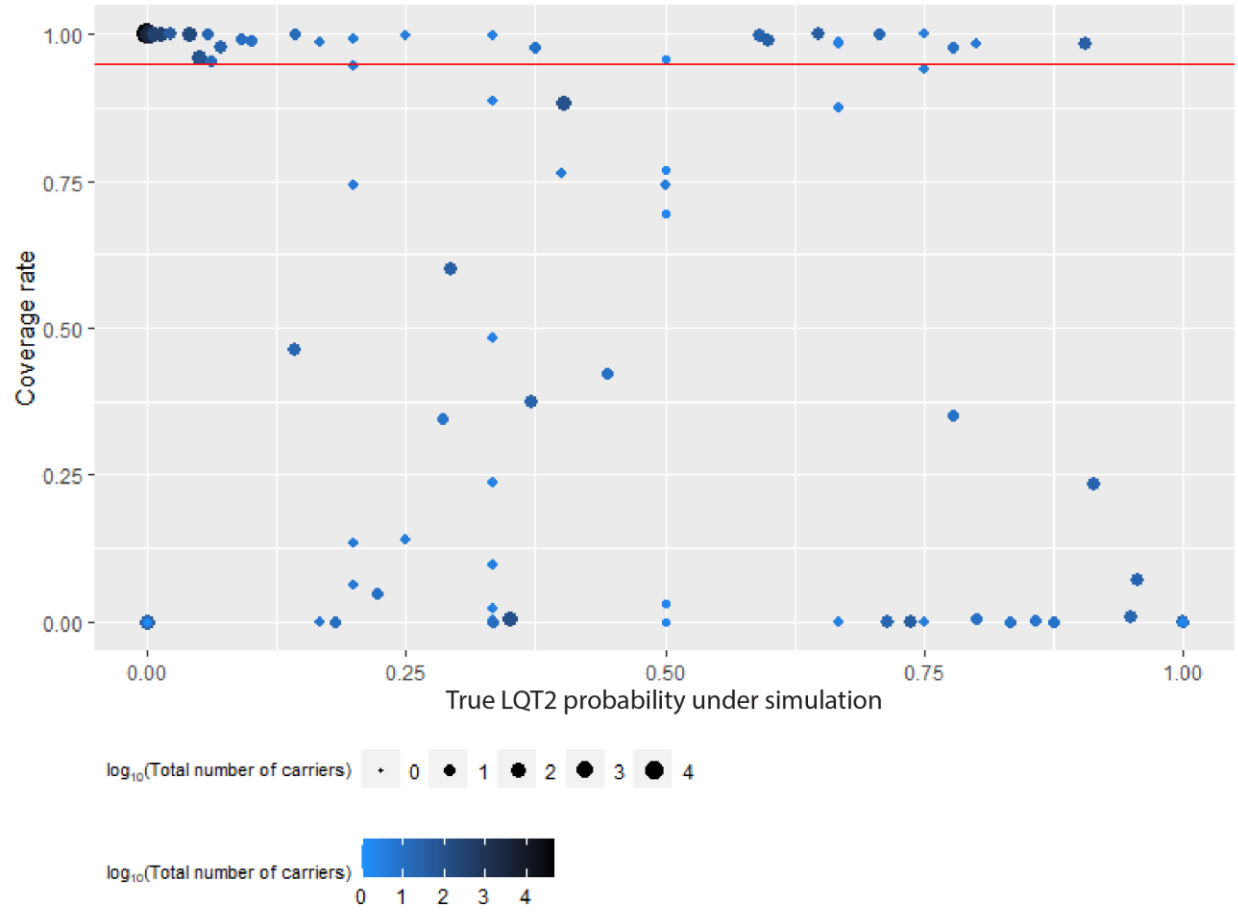
Supplemental Figure IX



Supplemental Figure IX. Receiver operating characteristic (ROC) curves from predictors (trained with the literature data without gnomAD) against all variants in the cohort dataset. All cohort data were withheld from post-test probability estimate and LQT2 probability density algorithms during construction. This plot is similar to Figure S6 except here models are trained on literature data without gnomAD.

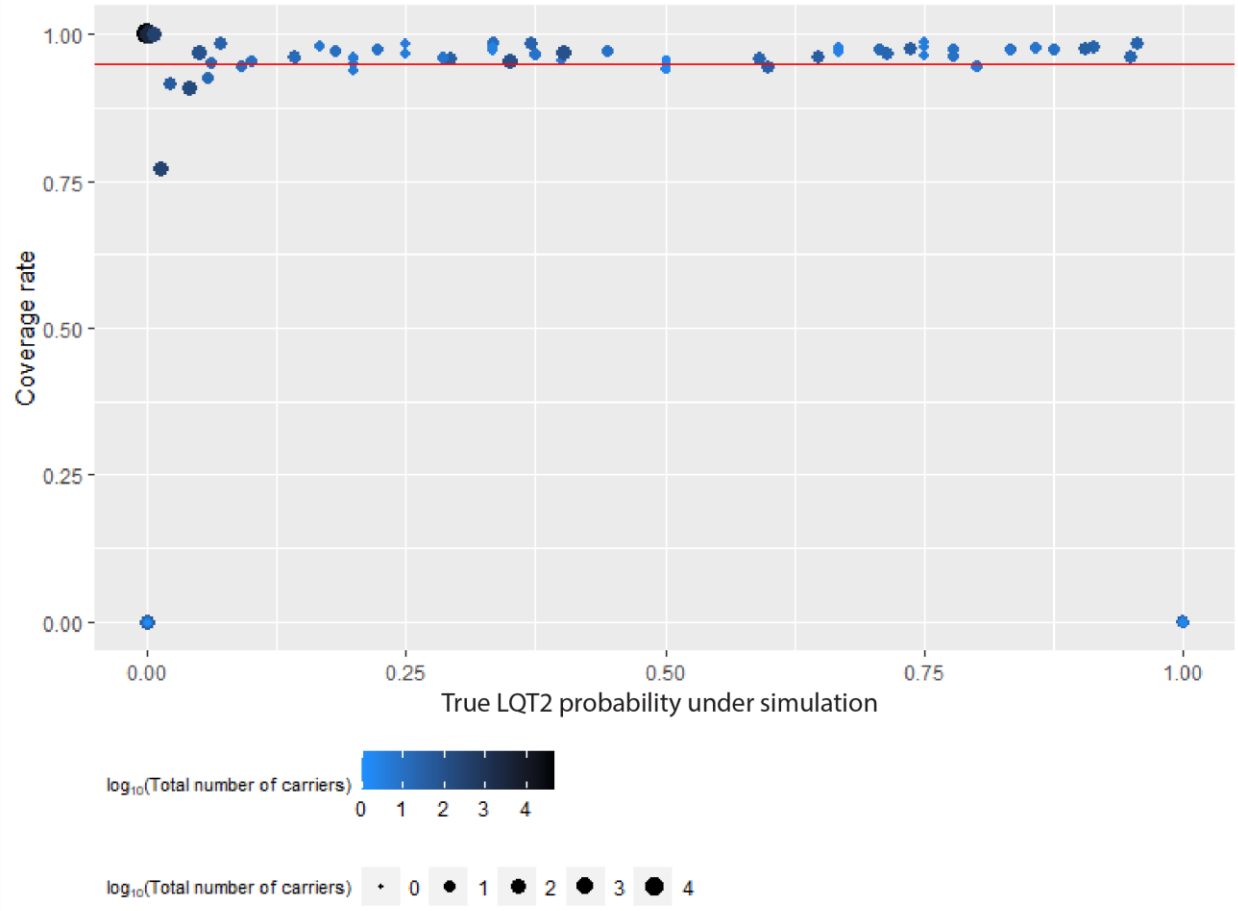
As in Figure S8, the area under the ROC curved are lower for post-test probability estimate and LQT2 probability density than when gnomAD is included.

Supplemental Figure X



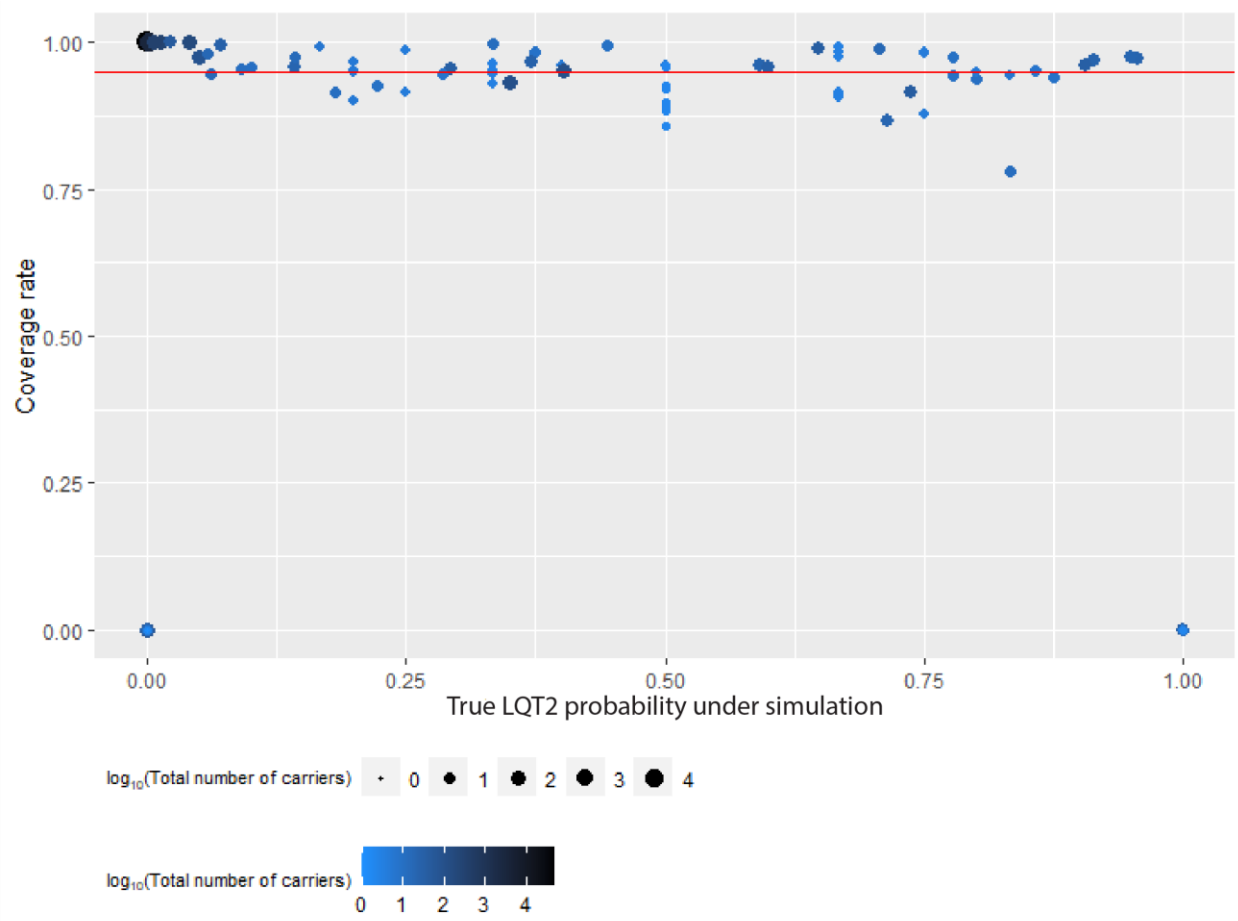
Supplemental Figure X. Estimated coverage rates for each *KCNH2* variant versus sampled "true" probability, the observed LQT2 probability. Coverage rate was calculated as defined in the supplemental text. Color and radius indicate the $\log_{10}$ of the total number of heterozygotes present in the dataset. The tuning parameter Equation 1 was set to $\nu = 100$. There is under coverage (less than 95%, the red line) for variants with high and low observed probability indicating an underestimate of the variance.

Supplemental Figure XI



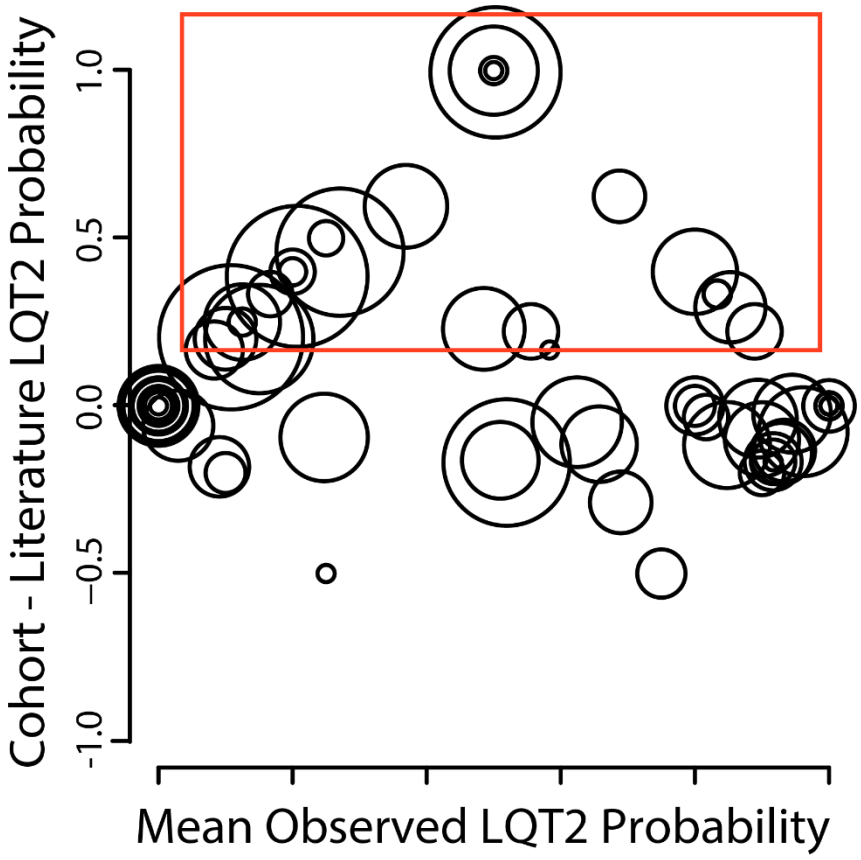
Supplemental Figure XI. Estimated coverage rates for each *KCNH2* variant versus sampled "true" probability, the observed LQT2 probability. Coverage rate was calculated as defined in the supplemental text. Color and radius indicate the $\log_{10}$ of the total number of heterozygotes present in the dataset. The tuning parameter Equation 1 was set to $\nu = 2$. There is overcoverage (greater than 95%, the red line) for variants with high and low observed LQT2 probability indicating an overestimate of the variance.

Supplemental Figure XII



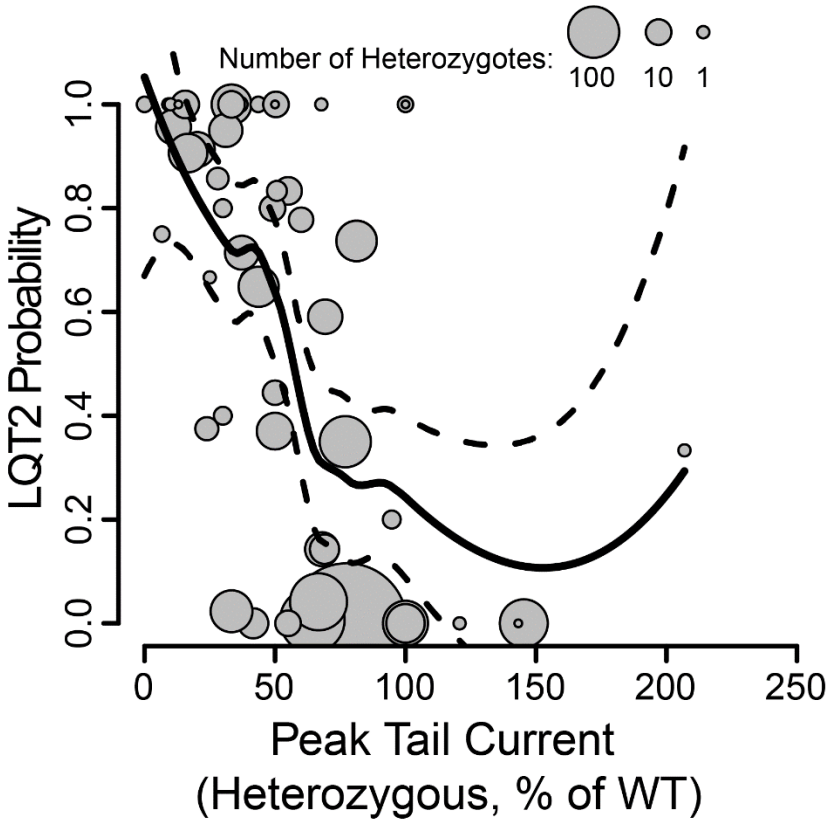
Supplemental Figure XII. Estimated coverage rates for each *KCNH2* variant versus sampled “true” probability, the observed LQT2 probability. Coverage rate was calculated as defined in the supplemental text. Color and radius indicate the $\log_{10}$ of the total number of heterozygotes present in the dataset. The tuning parameter Equation 1 was set to $\nu = 10$. The coverage is near 95%, the red line, both above and below, for variants with high and low observed probability indicating a reasonable estimate of the variance.

Supplemental Figure XIII



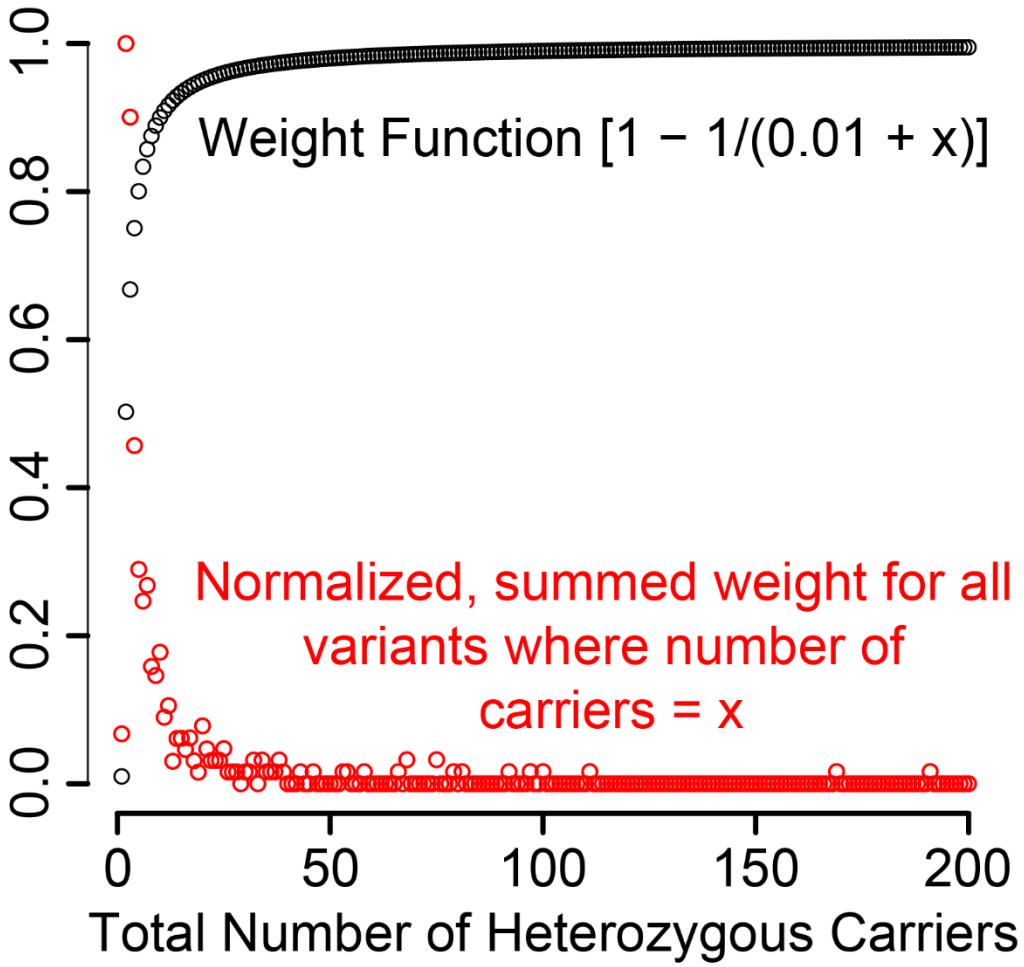
Supplemental Figure XIII. Bland-Altman plot of the mean LQT2 probability observed in the literature and cohort datasets (x-axis) versus the difference between the LQT2 probability observed (cohort – literature; y-axis). The circle radii are proportional to the log10 of the number of individuals heterozygous for these variants in the literature dataset. The red rectangle indicates the variants with higher observed LQT2 probability in the cohort than in the literature. Note the circle size for many of these is relatively large indicating many observations of individuals not meeting the criteria for affected status, assumed likely unaffected, from gnomAD.

Supplemental Figure XIV



Supplemental Figure XIV. LQT2 probability/penetrance observed in the literature and cohort combined versus heterozygous IKr peak tail current. Each grey circle is an KCNH2 variant where the radius is equal to log10(number of carriers). The solid trend line is the locally weighted average of LQT2 probability/penetrance from peak IKr tail current; the dashed lines are 95% confidence intervals for the weighted average. At a peak tail current of 100% of WT the average LQT2 probability is ~ 20% indicating biased towards higher LQT2 probability. To test the hypothesis that our LQT2 probability estimates are overly influence by functional data with an asymmetric ascertainment bias, towards higher LQT2 probability, we reran our calculations in absence of functional data and evaluated these results against the cohort group. We found an R2 of 0.31 [0.20-0.41
95%CI] and an AUC of 0.79, compared to our reported R2 of 0.30 [0.19-0.43 95%CI] and AUC 0.78 when functional data are included (Table 1 and Figure 3, respectively).

Supplemental Figure XV



Supplemental Figure XV. Weighting function referenced throughout the text (y-axis) by total number of heterozygous carriers.