SUPPLEMENTAL MATERIAL

Transethnic genome-wide association study provides insights in the genetic architecture and heritability of long QT syndrome

Najim Lahrouchi, et al.

Table of contents

SUPPLEMENTARY TABLES (in Microsoft Excel© Workbook file)

Supplementary Table I: Overview of case datasets

Supplementary Table II: Overview of control datasets

Supplementary Table III: QT interval polygenic risk score SNPs

Supplementary Table IV: Clinical characteristics of European and Japanese LQTS cases

Supplementary Table V: Multivariable Cox regression model of LAE in LQTS

Supplementary Table VI: Summary of results in the LQTS case-control GWAS for the 68 SNPs previously associated to the QT interval in the general population

Supplementary Table VII: Polygenic risk score analyses in European and Japanese LQTS patients (only non-LQTS loci included in PRS)

Supplementary Table VIII: PRS derivation from European QT GWAS and validation in European LQTS dataset

Supplementary Table IX: Overview of variant classification as per ACMG/AMP* guidelines

Supplementary Table X: Polygenic risk score analyses in European and Japanese LQTS patients (excluding patients with a variant of unknown significance)

Supplementary Table XI: Increasing PRS_{QT} quartiles in European LQTS

Supplementary Table XII: Common variant heritability on the liability scale for LQTS as implemented in GCTA (GREML) and LDAK (PCGC and GREML)

Supplementary Table XIII: Multivariable linear regression model for QTc in Europeans

SUPPLEMENTARY METHODS

1.1 Genetic testing and curation of rare variants in LQTS patients

Genetic testing was conducted at each participating centre and genetic variants were re-assessed centrally. Rare variants were classified as pathogenic, likely pathogenic or variant of unknown significance (VUS) as per the American College of Medical Genetics and Genomics and Association of Molecular Pathology (ACMG/AMP) guidelines[23] using CardioClassifier[45] and applying a channel topology driven etiological fraction based criterion for ACMG/AMP classification, as performed previously in hypertrophic cardiomyopathy genes (see below for more details).[46] For primary analyses, genotype positive and negative LQTS cases were defined as those with and without a rare variant (allele frequency $<1x10^{-4}$ gnomAD) respectively, regardless of the ACMG/AMP classification (i.e. including VUS). Sensitivity analyses were also performed after exclusion of cases with VUS.

1.2 Assessing the pathogenicity of reported LQTS variants – adapted CardioClassifier approach

The following ACMG/AMP rules were applied:

- PM2/BS1: The PM2 or BS1 rules were activated depending on whether the gnomAD exomes filtering allele frequency was below or above the calculated maximum tolerated allele frequency for LQTS. The threshold applied was $5x10^{-5}$, calculated with https://www.cardiodb.org/allelefrequencyapp/ using a disease prevalence of 1 in 2000, allelic heterogeneity of 0.02 and penetrance of 0.10.[47]
- BA1: Filtering allele frequency in gnomAD exomes >0.1%.
- PVS1: Truncating variants in *KCNQ1* and *KCNH2*, i.e. frameshift, nonsense, splice donor and splice acceptor variants.
- PS4: Variant is enriched in cases cohorts (LQTS referral cohort[48]) compared to ExAC population controls (Fisher's exact test, $p<1.79x10^{-6}$ after multiple testing correction), as applied by CardioClassifier.
- PP3/BP4: Multiple lines of computation evidence support or refute a deleterious effect, as applied by CardioClassifier.
- PM4: A protein length change as a result of an in-frame deletion or insertion in a non-repeat region.
- BP3: In-frame insertion or deletion that falls within a region annotated by repeat masker.
- PS1: Same amino acid change as a previously established pathogenic variant (multiple ClinVar submissions with no conflicting evidence).
- PM5: Novel missense change at an amino acid residue where a different missense change has previously been established as pathogenic (multiple ClinVar submissions with no conflicting evidence).
- PS1_moderate/PM5_supporting: Known disease-causing variant affecting the same residue in a paralogous protein[49], either with the same or a different amino acid substitution.
- PM1: This rule was applied based on pre-computed etiological fraction (EF) values for rare, non-truncating variants as described by Walsh et al[46]. This approach defines the prior probability, as calculated through case-control cohort analysis, that a variant in a particular gene/protein region is pathogenic. The rules applied are PM1_strong ($EF \geq 0.95$), PM1_moderate ($0.9 \leq EF < 0.95$) or PM1_supporting ($0.8 \leq EF < 0.9$). For the three LQTS genes, this analysis was performed using case data from this study and population data from gnomAD

exomes. Protein domains were defined according to Uniprot (version 207) entries P51787 (*KCNQ1*), Q12809 (*KCNH2*) and Q14524 (*SCN5A*), as well as the highly conserved *KCNQ1* C-terminal regions identified by Kapplinger et al[50]. The PM1_strong rule was applied to the *KCNQ1* N-terminus (residues 1-121), transmembrane (122-348) and C-terminus highly conserved regions (349-391,509-575,585-607) and the *KCNH2* PAS (41-70), PAC (92-144), transmembrane (404-659) and cNBD (742-842) domains. The PM1_moderate rule was applied to the *KCNH2* other N-terminus regions (1-40,71-91,145-403). The PM1_supporting rule was applied to the *KCNH2* other C-terminus regions (660-741,843-1159).

Rules based on co-segregation of variants with disease in family pedigrees (PP1/BS4), de novo inheritance with/without confirmed paternity and maternity (PS2/PM6) and well-established functional studies showing a deleterious or no effect (PS3/BS3) were not applied.

1.3. Measurement of the QT interval

Core lab measurement of ECG parameters (PR, QRS, QT, RR) was performed by an experienced electrophysiologist (YM) using ImageJ (version 1.51s).[56] The QT-interval duration was averaged from 3 consecutive beats, preferably in lead II or V5, using the tangent method (**Figure I in the Supplement**).[57] Corrected QT (QTc) was calculated using Bazett's formula ($QTc = QT/\sqrt{RR}$). ECGs recorded during atrial fibrillation as well as those with ventricular pacing or ventricular conduction abnormalities were excluded. As a validity check, the QT-interval for all cases was measured by a second physician (NL) and cases where the QTc measures differed by >50ms (e.g. due to dysmorphic T wave) were reviewed by a QT expert (PGP).

1.4 Quality control of genotypic data

We excluded ambiguous variants (A/T or C/G) and those with a missingness rate >0.05, Hardy-Weinberg equilibrium test $P<10^{-6}$, or minor allele frequency (MAF) <0.001. We removed samples with a missingness rate >0.02, inbreeding coefficient |F|>0.15, and those with a sex mismatch. We performed pairwise identity-by-descent (IBD) analyses and excluded one per pair of duplicate samples, as well as ones with cryptic relatedness (proportion IBD > 0.125) prioritizing more severely affected cases over other cases and controls. We used principal components (PC) analyses (PCA), using the 1000 Genomes populations as a reference (Phase3v5) to exclude population outliers. PCA was performed in the final European and Japanese datasets to calculate eigenvectors used as covariates in the regression analyses.

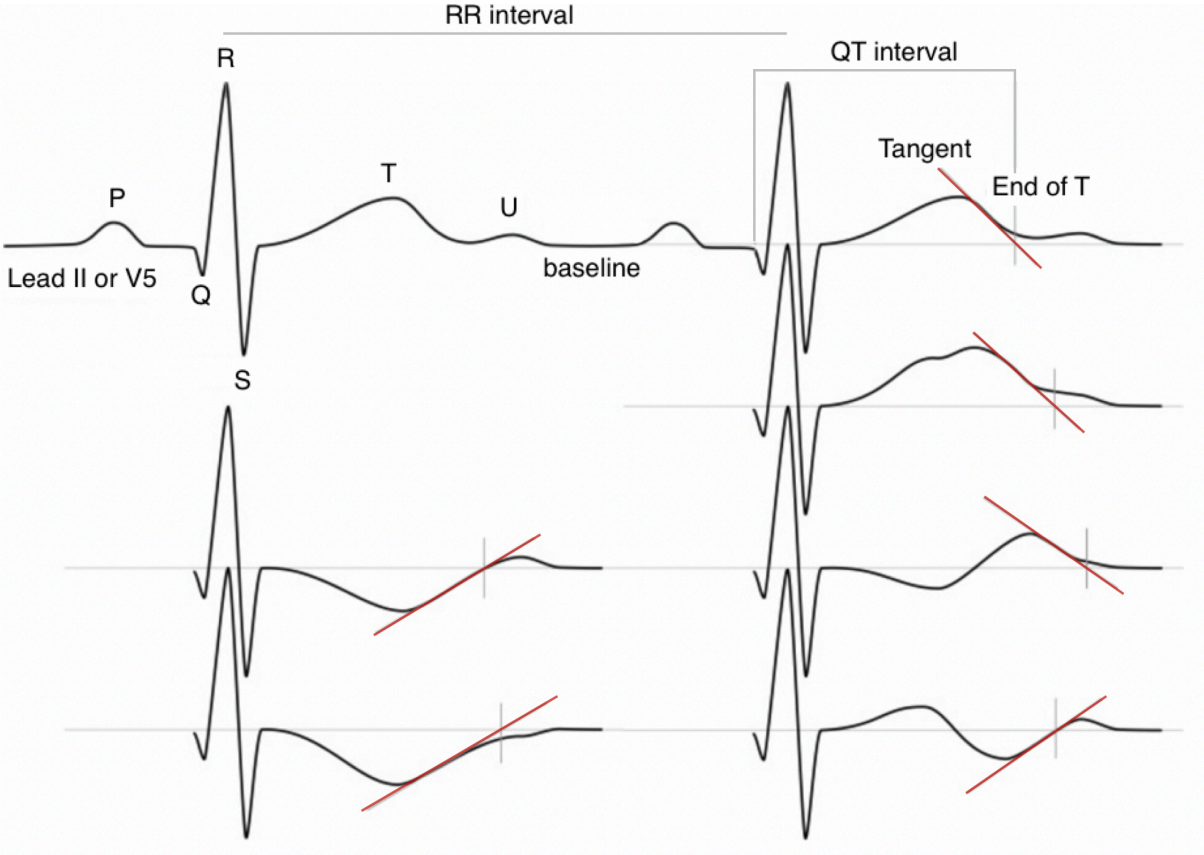
1.5 Quality control of imputed data prior to common variant heritability estimation

We used hard genotype calls after imputation (genotype probability >0.9) and restricted to variants with an imputation score >0.8. We excluded variants with genotype missingness >0.01, Hardy-Weinberg equilibrium test P<0.05, phenotype-biased missing test P<0.05, and minor allele frequency (MAF) <0.05. We then generated a genetic relationship matrix and excluded genetically related individuals (proportion IBD >0.05).

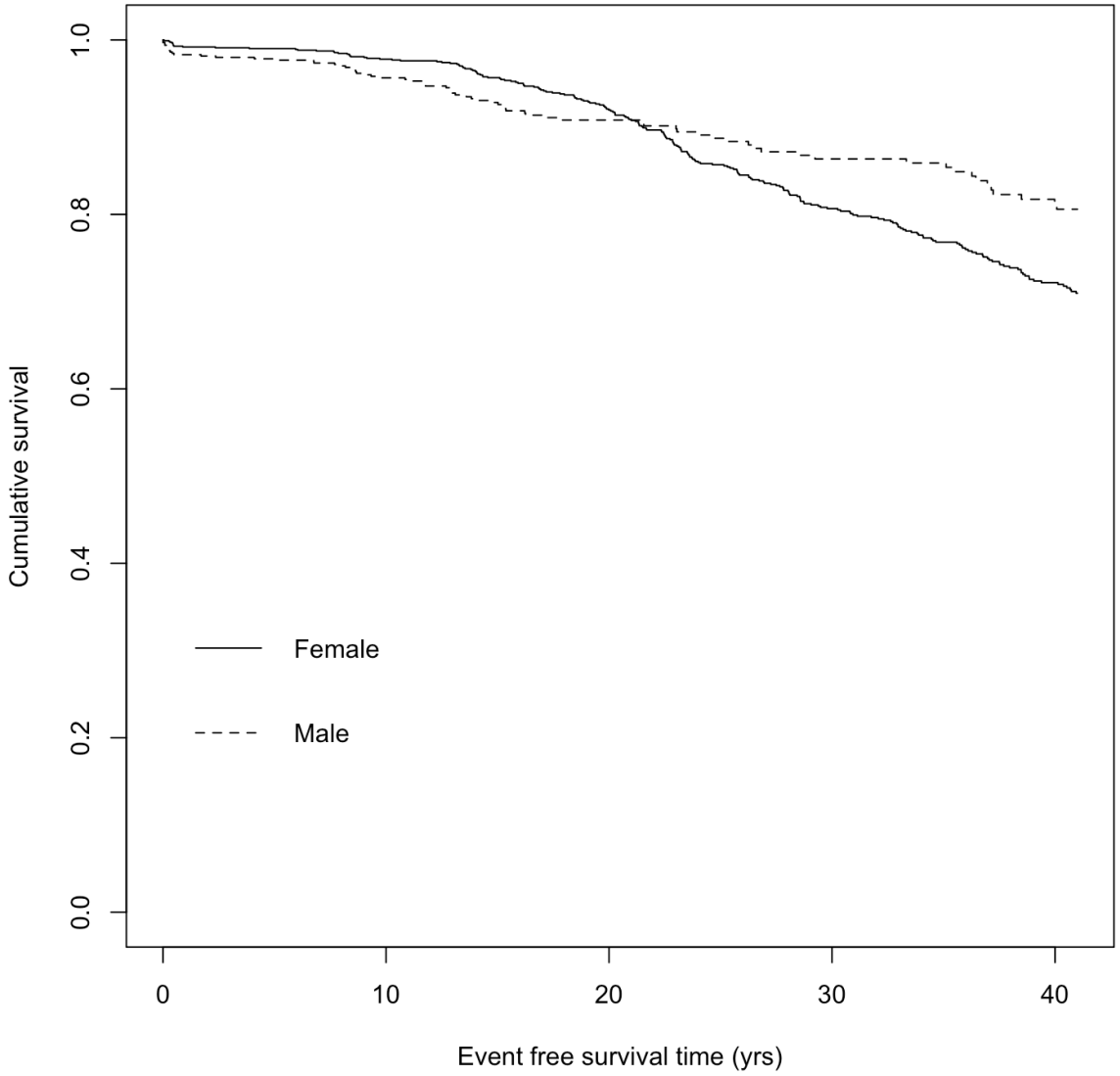
1. 6 Origin of summary statistics of cardiac electrical traits used in LD score regression

Summary statistics for cardiac electrical traits (PR[51], QRS[52], QT[30], heart rate at rest[53], heart rate in response to exercise and recovery[54], and atrial fibrillation[55]), were obtained by direct contact with the corresponding author of the study or were available online.

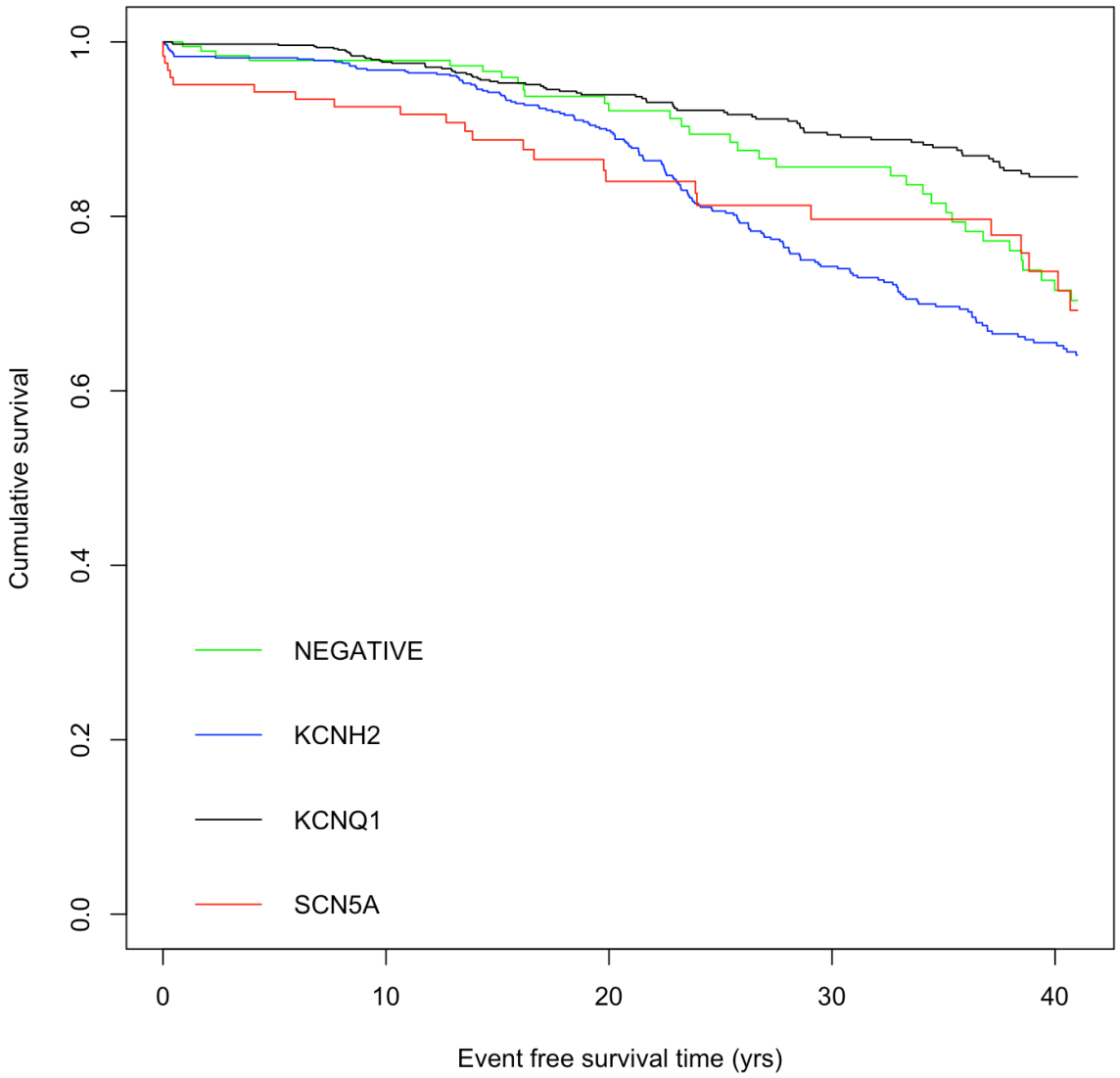
Supplementary Figure I | Schematic illustration of application of the tangent method to define the end of the T-wave in normal and abnormal TU morphologies to determine the QTc interval.[24,57]



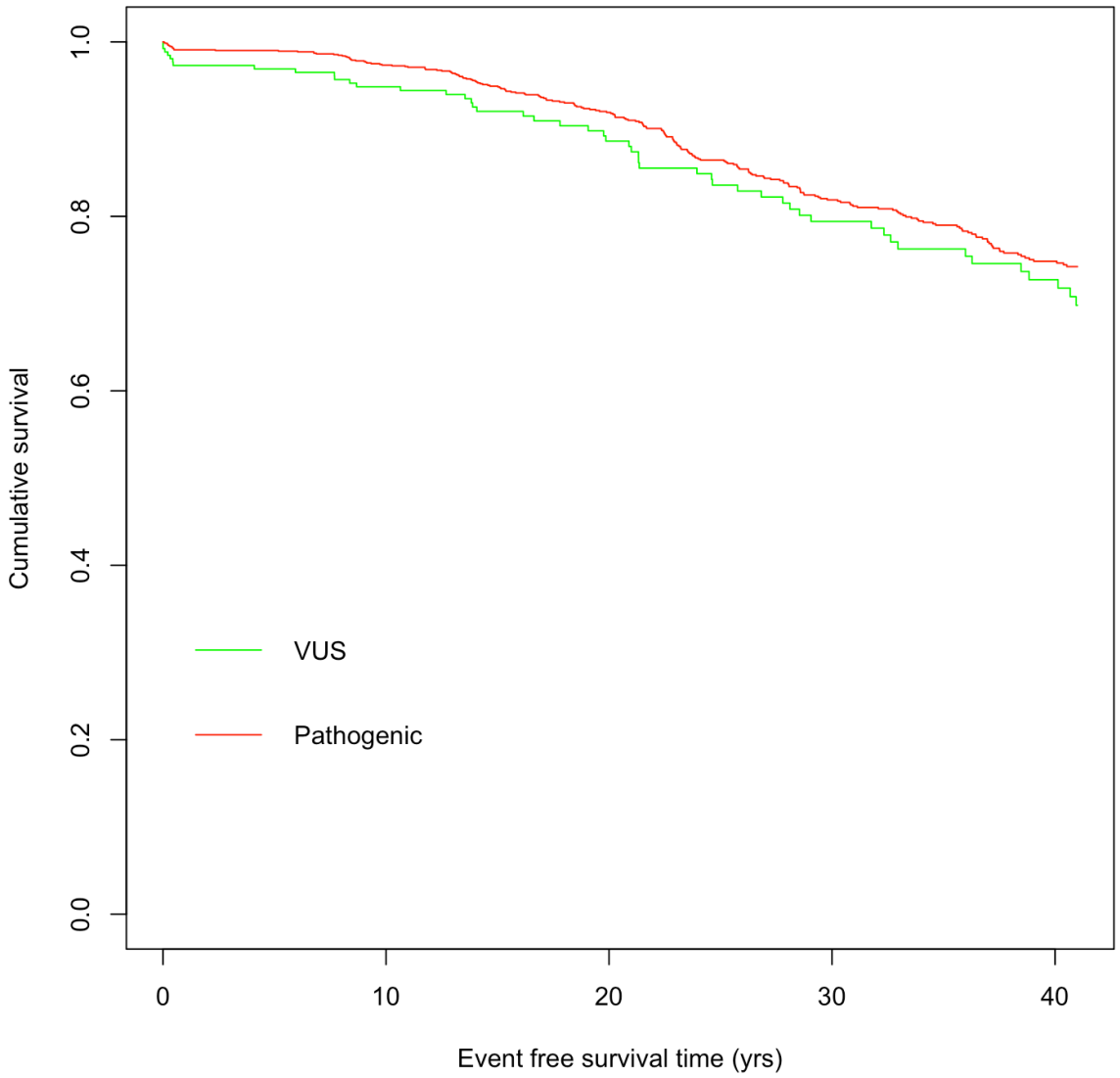
Supplementary Figure II | LAE-free Kaplan-Meier survival curve stratified by sex. Life-threatening arrhythmic events (LAE) were defined as out of hospital cardiac arrest or hemodynamically unstable ventricular tachycardia/ventricular fibrillation. Data shown include all LQTS cases, both genotype positive and negative and both European and Japanese. Yrs, years. Logrank test P=0.06.



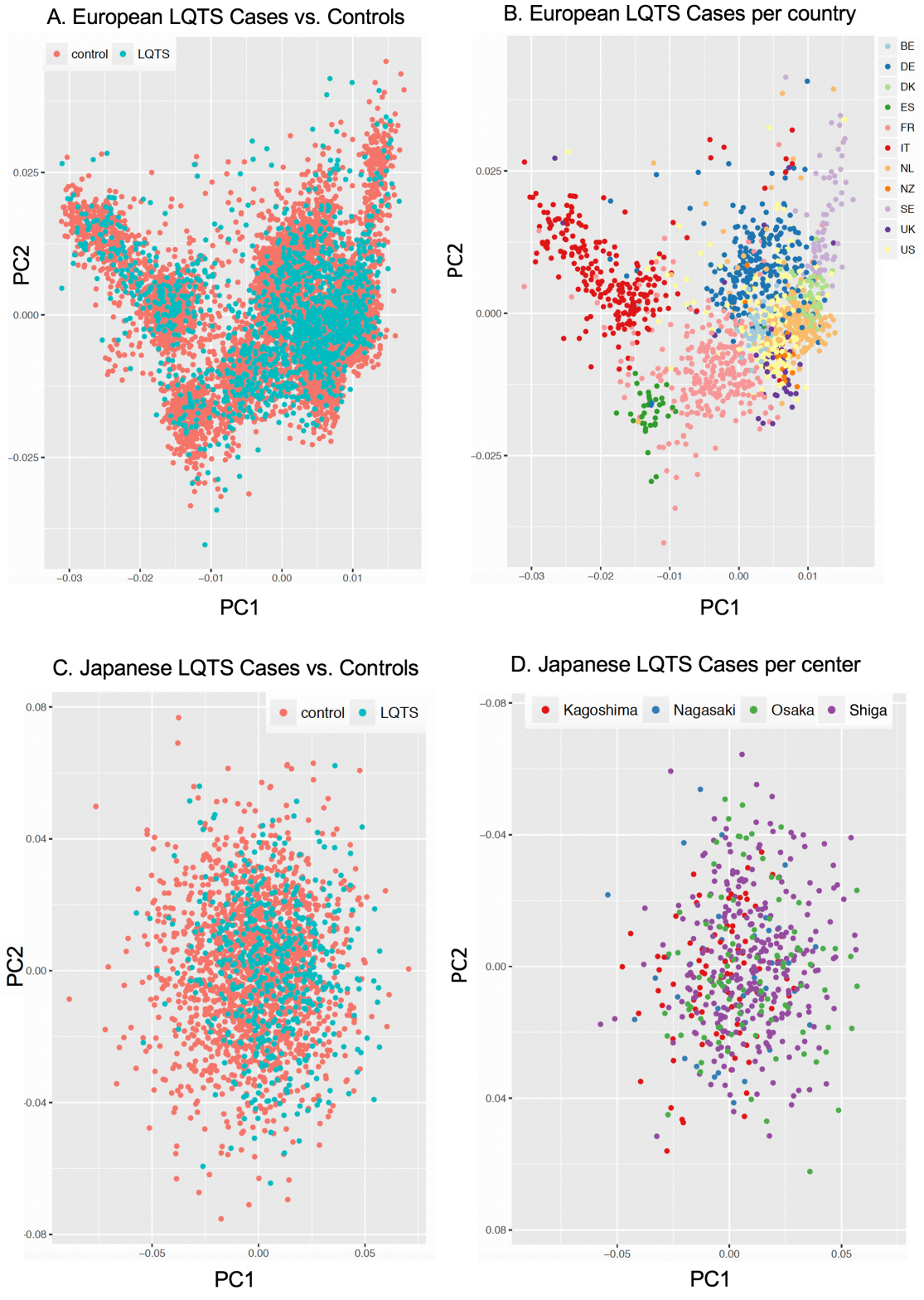
Supplementary Figure III | LAE-free Kaplan-Meier survival curve stratified by genotype. Life-threatening arrhythmic events (LAE) were defined as out of hospital cardiac arrest or hemodynamically unstable ventricular tachycardia/ventricular fibrillation. Data shown include all LQTS cases, both European and Japanese. Yrs, years. Logrank test P= $7x10^{-10}$.



Supplementary Figure IV | LAE-free Kaplan-Meier survival curve stratified by variant class (VUS vs. pathogenic or likely pathogenic). Life-threatening arrhythmic events (LAE) were defined as out of hospital cardiac arrest or hemodynamically unstable ventricular tachycardia/ventricular fibrillation. Data shown include all genotype positive cases of European and Japanese ancestry. “VUS”, rare genetic variant of unknown significance. “Pathogenic”, includes rare variants classified as pathogenic or likely pathogenic as per the American College of Medical Genetics and Genomics criteria (see **Supplementary Methods**). Logrank test P=0.2.

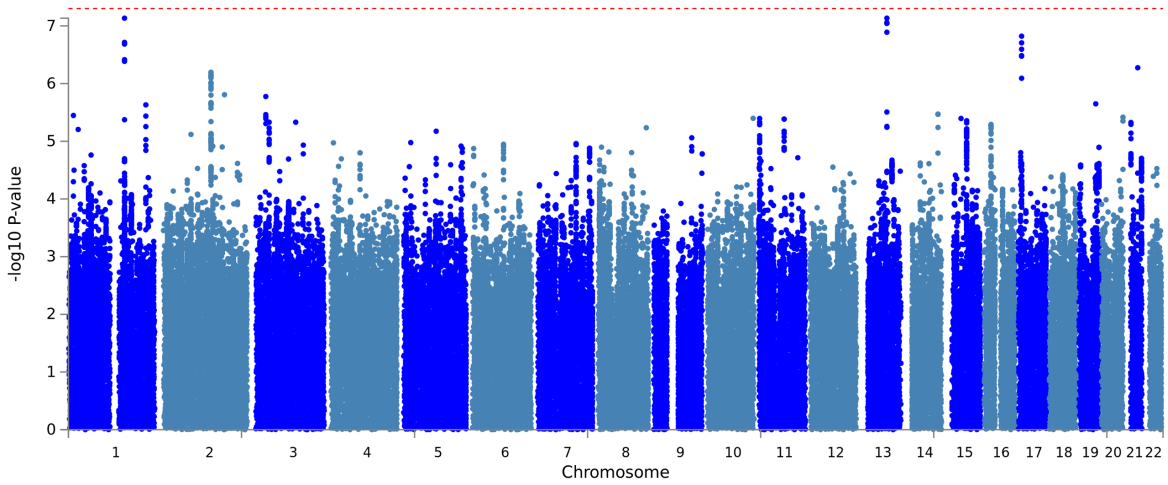


Supplementary Figure V | Final principal component (PC) analyses of European and Japanese LQTS cases and controls. The left panels (**A and C**) show the final PC1 vs. PC2 plots with overlapping of cases and controls for both the European (**A**) as well as the Japanese (**C**) dataset. The right panels (**B and D**) show the final PC1 vs. PC2 plots for European (**B**) and Japanese (**D**) LQTS cases only stratified by country or city of inclusion, respectively. BE, Belgium; DE, Germany; DK, Denmark; ES, Spain; FR, France; IT, Italy; NL, Netherlands; NZ, New Zealand; SE, Sweden; UK, United Kingdom; US, United States.

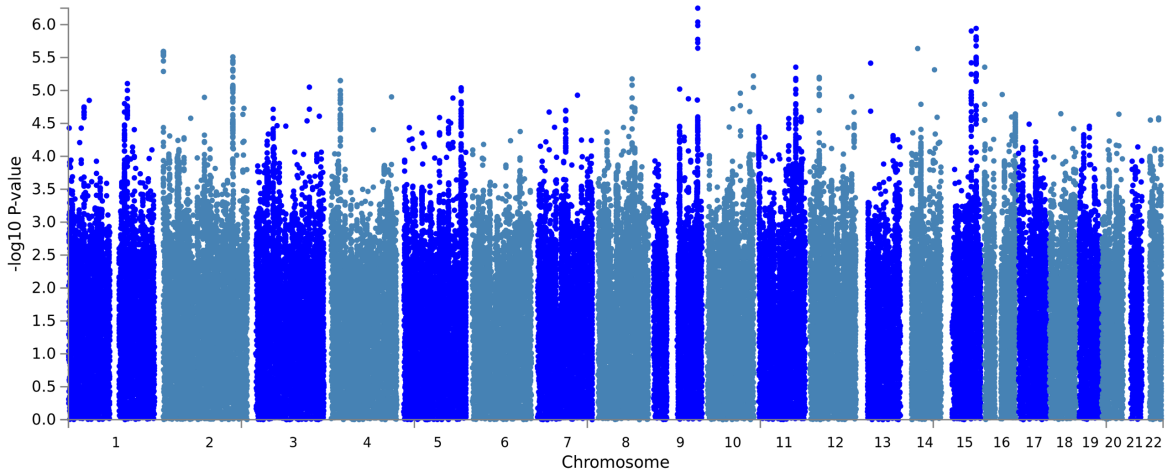


Supplementary Figure VI | Manhattan plots displaying the -log10 transformed *P* value of each SNP and LQTS risk in the all European (**A**, 1238 cases vs. 8219 controls) and Japanese (**B**, 418 cases vs. 1671 controls) case-control analyses on the *y* axis and base-pair positions along the chromosomes on the *x* axis. The red dashed line indicates genome-wide significance ($P<5x10^{-8}$). The observed genomic inflation factor (λ_{GC}) is 1.024 and 1.034 for the European and Japanese GWAS respectively.

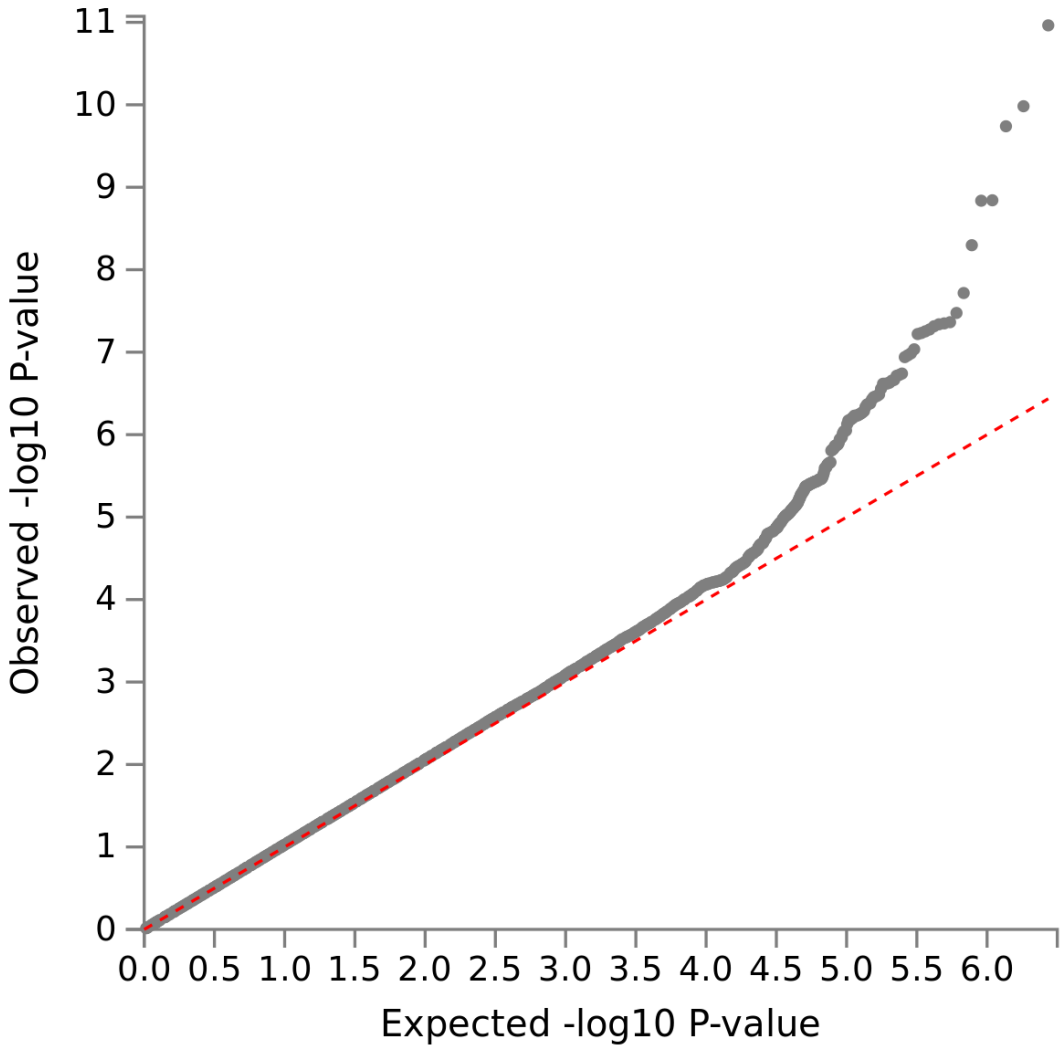
A. European case-control GWAS [n=1238]



B. Japanese case-control GWAS [n=418]

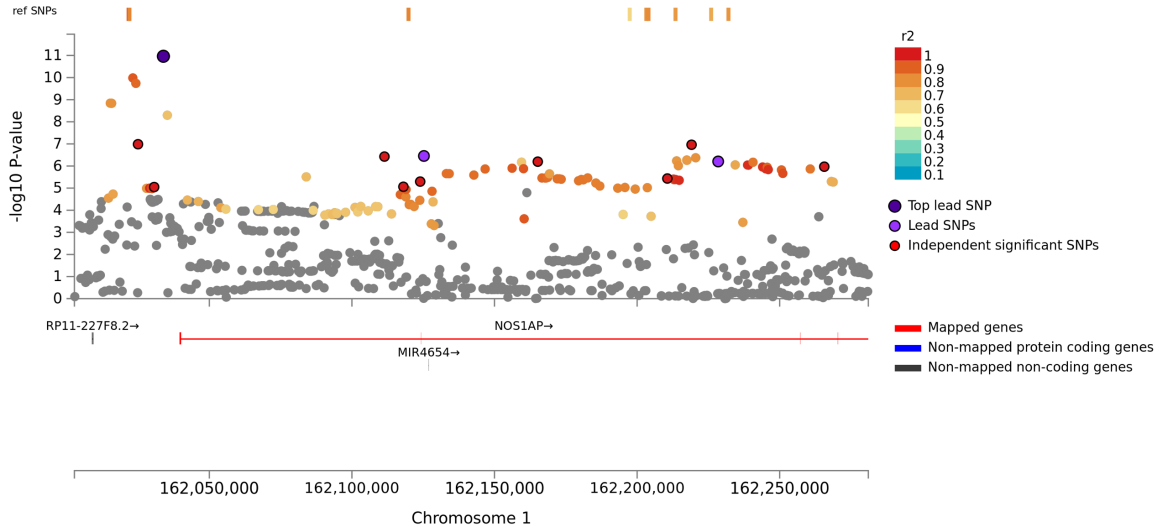


Supplementary Figure VII | Quantile-quantile (QQ) plot of the genome-wide association meta-analysis of European and Japanese LQTS GWAS (n = 1656). Observed and expected P values are on a −log10 scale (two-tailed). The red line represents the expected values under the null hypothesis. The observed genomic inflation factor (λ_{GC}) is 1.028.

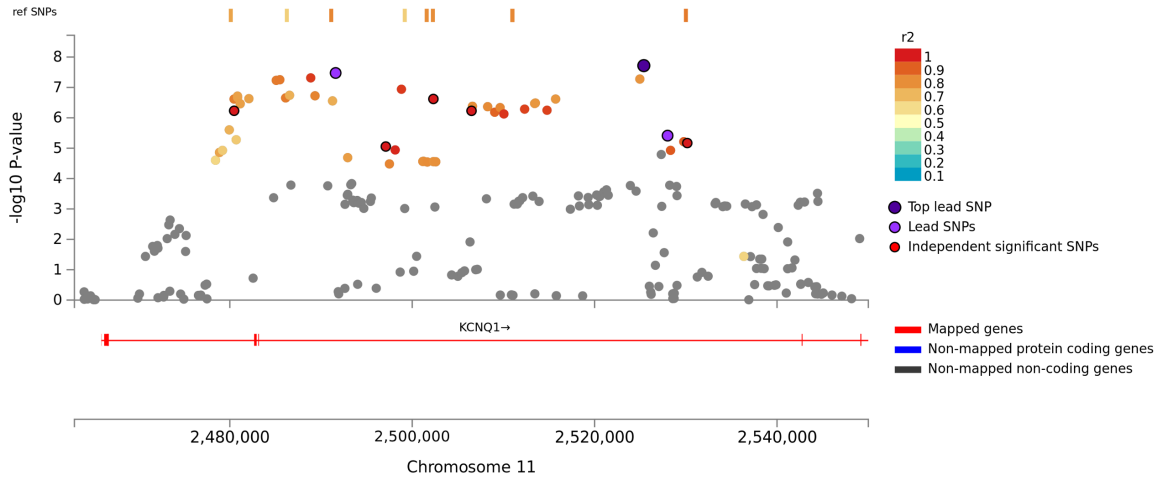


Supplementary Figure VIII | Regional association plots for four LQTS risk loci (**panels A-D**) from the European and Japanese case-control meta-analysis. The *R*2 values are from the 1000 Genomes phase3v5 EUR references samples. **GWAS P-value**: SNPs which are not in LD of any of significant independent lead SNPs in the selected region are coloured grey. The figures were created using FUMA.[28] Gene mapping was performed using positional mapping with the default parameters.

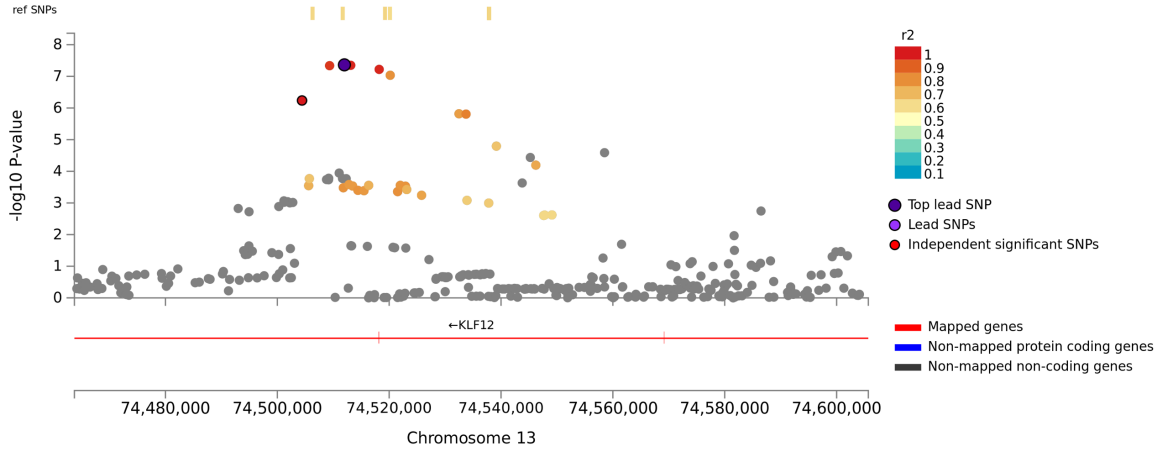
A. NOS1AP: rs2143842



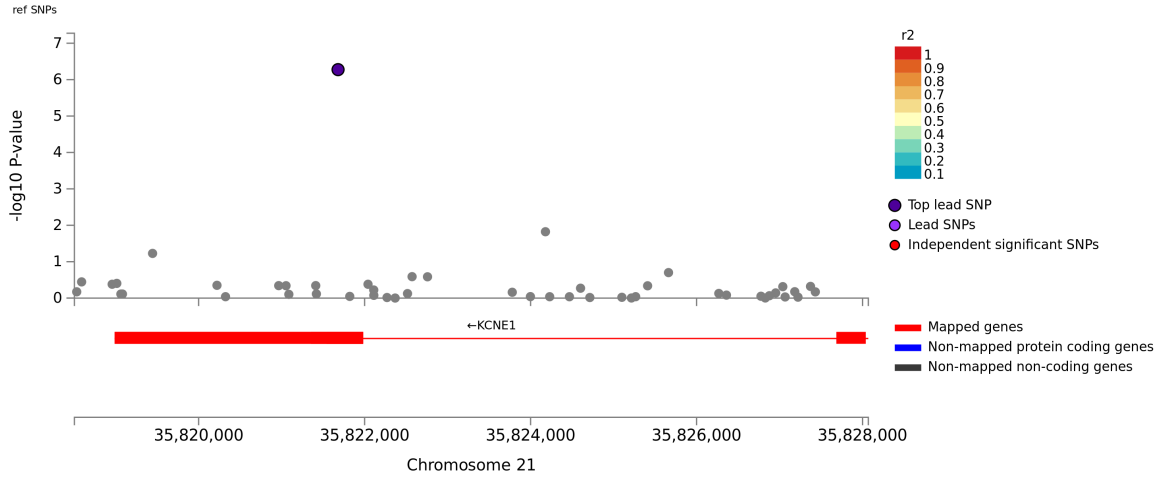
B. KCNQ1: rs179405



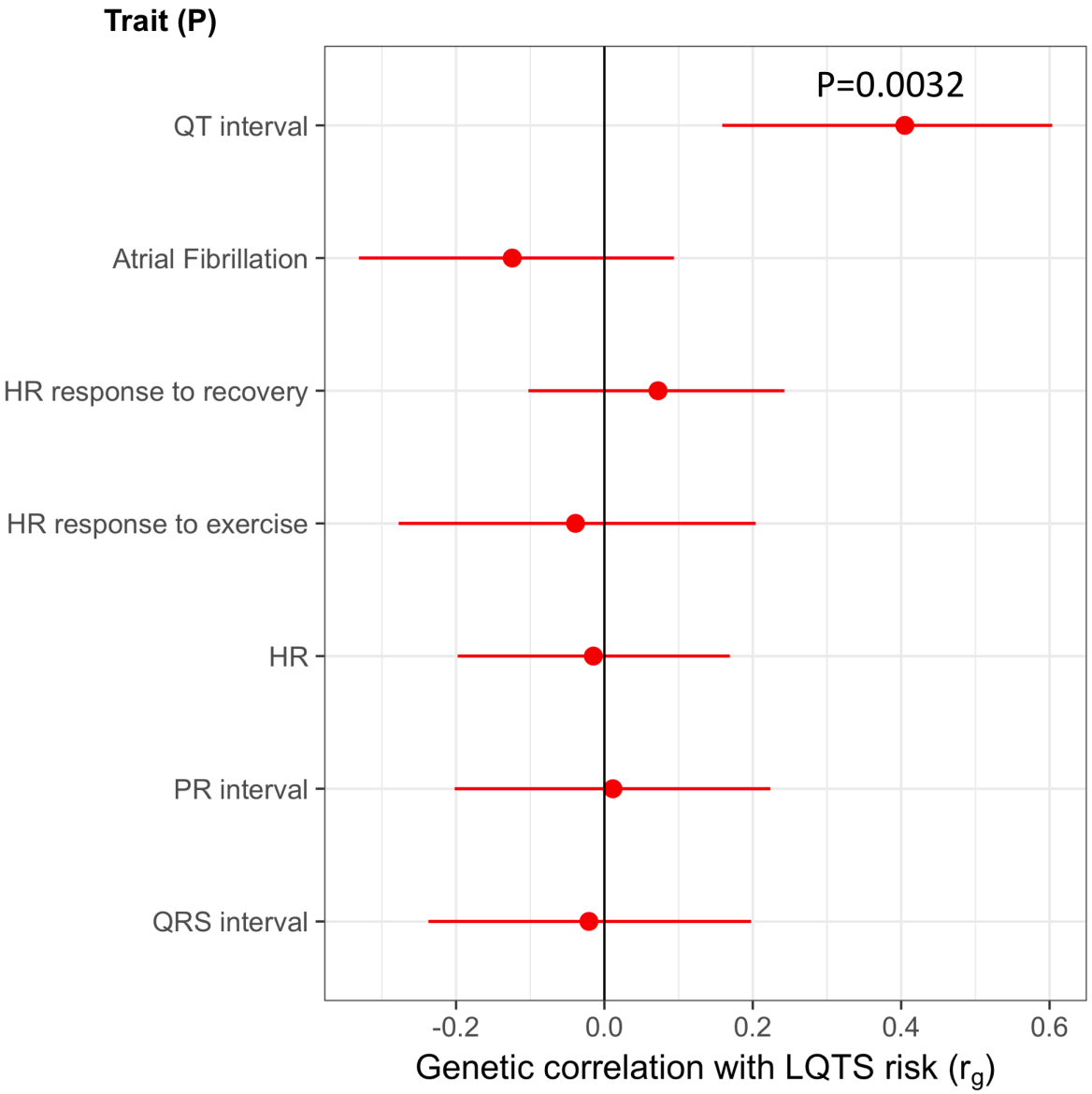
C. *KLF12*: rs17061696



D. *KCNE1*: rs1805128 (p.Asp85Asn)

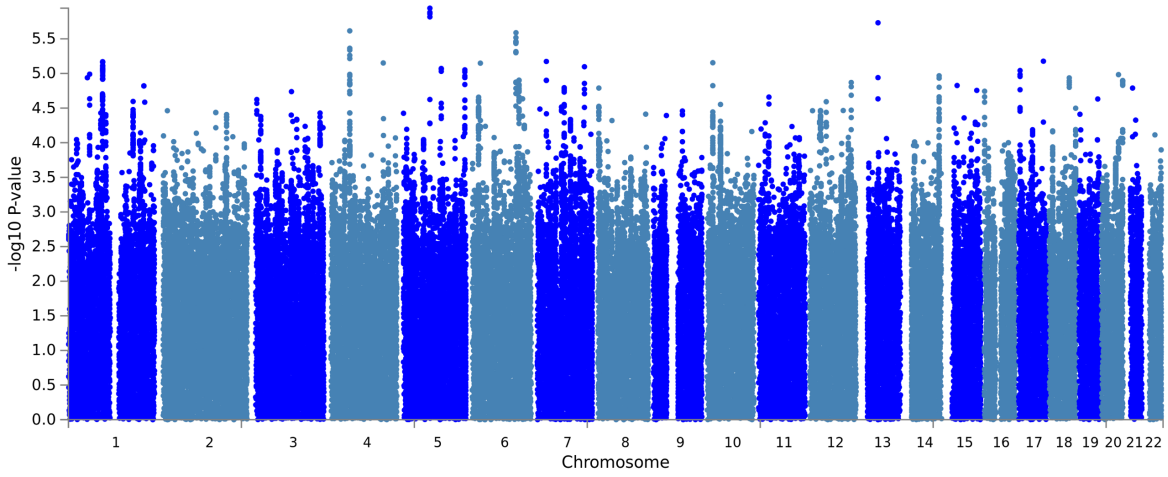


Supplementary Figure IX | Genetic correlation between LQTS risk and cardiac electrical traits. Genetic correlation (r_g) was calculated using bivariate LD score regression implemented in LDSC. The errors bars represent the 95% confidence interval. HR, heart rate.

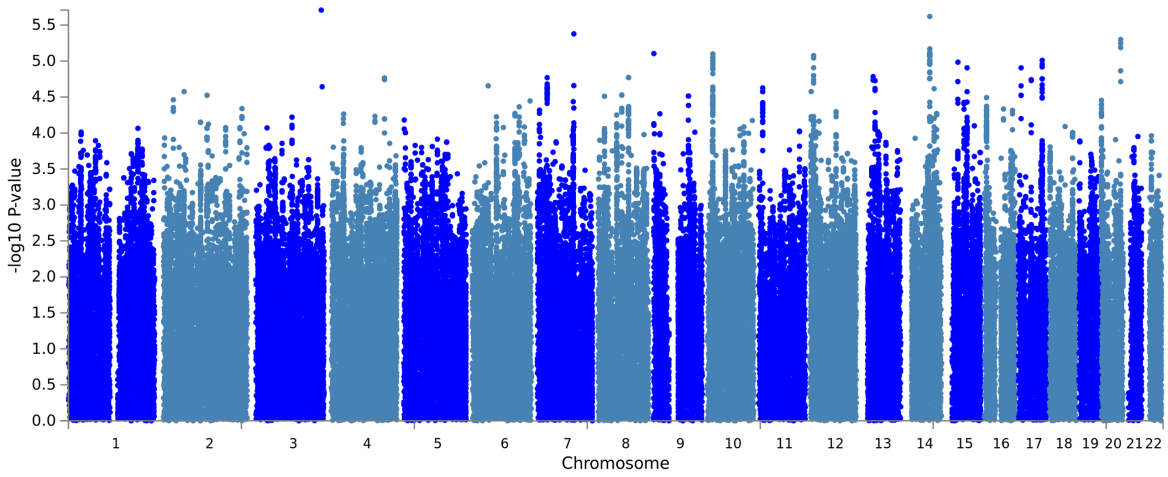


Supplementary Figure X | Manhattan plots for QTc GWAS in the European (**A**, n=1238) and Japanese (**B**, n=418) LQTS cohorts, as well as meta-analysis (**C,** n=1656) and quantile-quantile (QQ) plot of the meta-analysis (**D**). Manhattan plots (**A**,**B**,**C**) display the -log10 transformed *P* value of each SNP and QTc-interval on the *y* axis and base-pair positions along the chromosomes on the *x* axis. Quantitative trait analyses for the QTc were conducted using a linear regression model with the first 10 principal components, age at ECG recording, beta-blocker usage at ECG recording, LQTS type and sex as covariates. Meta-analysis was performed using an inverse variance weighted fixed effect meta-analysis model. QQ plot of the QTc genome-wide association meta-analysis (**D**) showing observed and expected P values, on a −log10 scale (two-tailed). The red line represents the expected values under the null hypothesis. The observed genomic inflation factor (λ_{GC}) is 0.98.

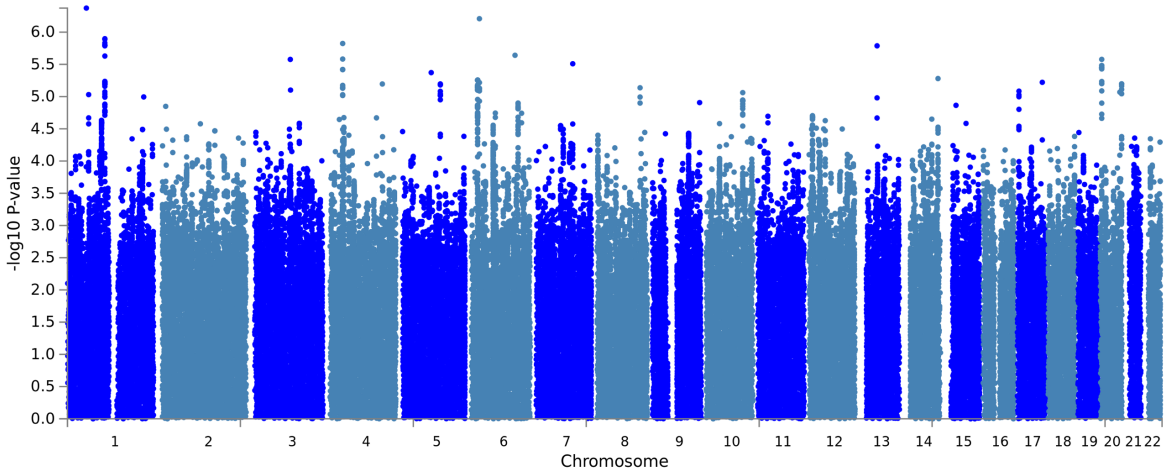
A. QTc GWAS in European LQTS [n=1238]



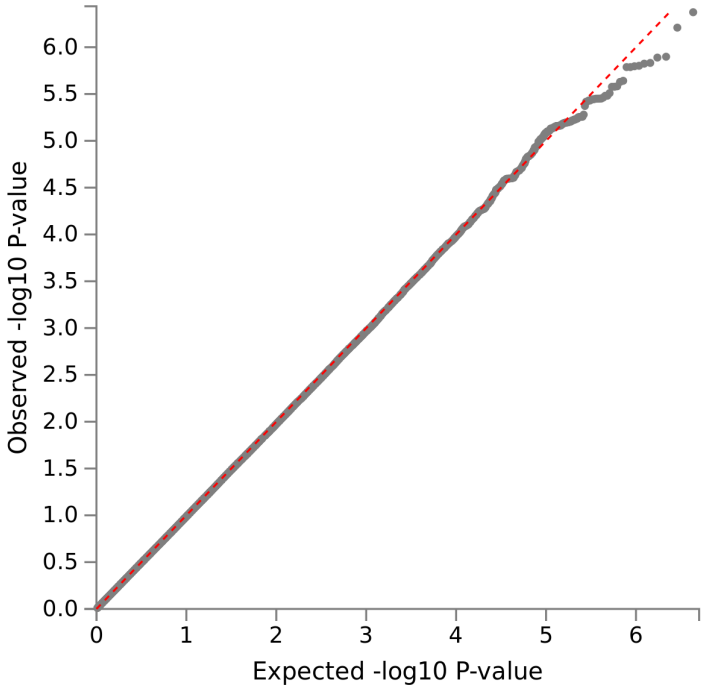
B. QTc GWAS in Japanese LQTS [n=418]



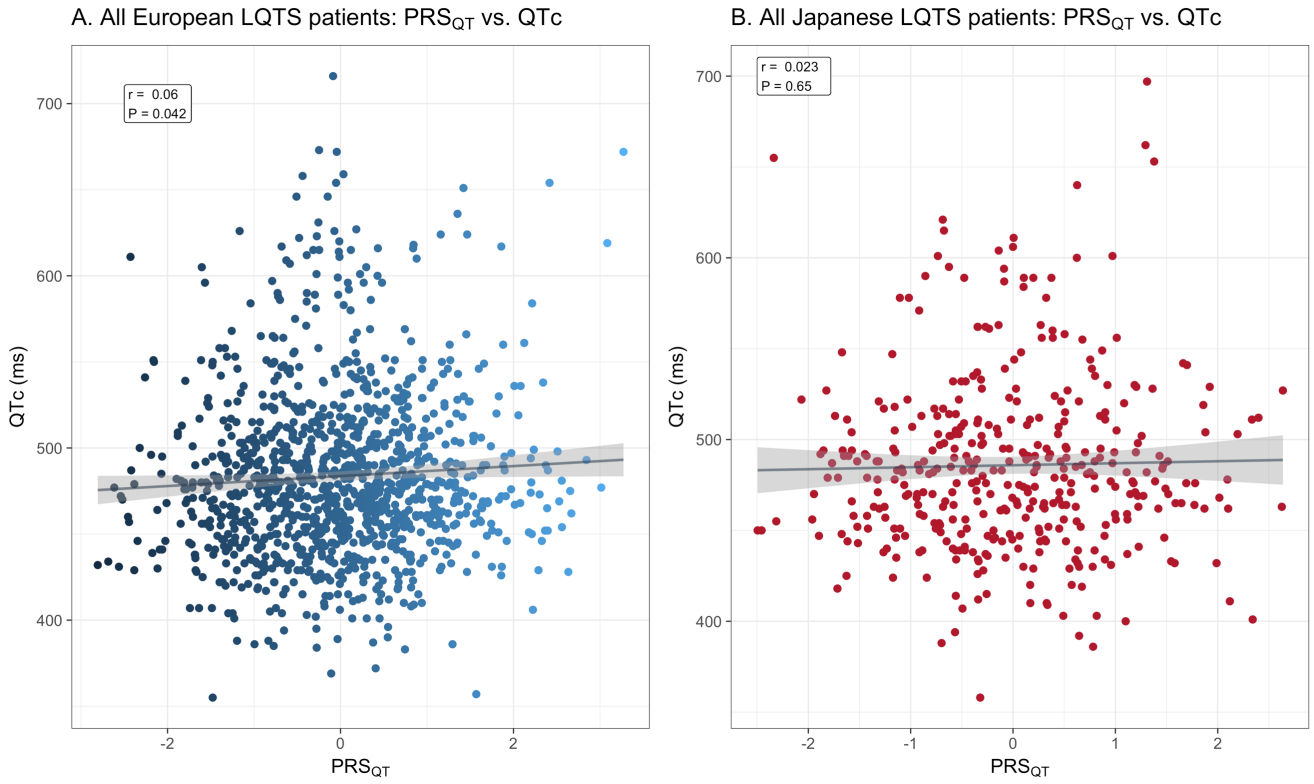
C. Meta-analysis of European and Japanese QTc GWAS [n=1656]



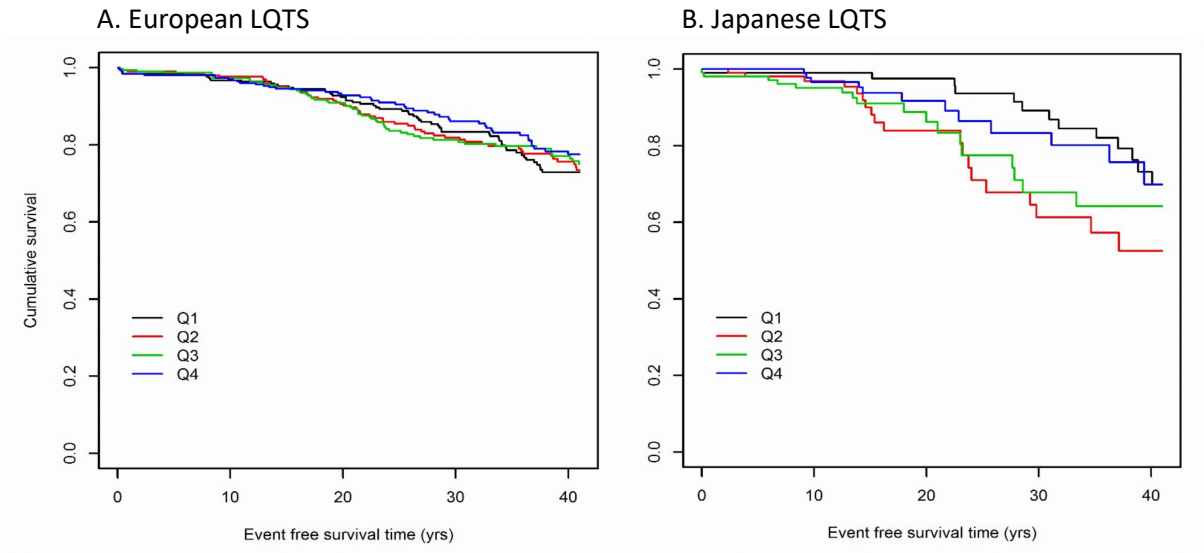
D. QQ plot of meta-analysis of European and Japanese QTc GWAS [n=1656]



Supplementary Figure XI | Correlation of QTc and the QT polygenic risk score (PRS_{QT}) in European (**A**) and Japanese (**B**) LQTS cases. Also shown are the linear regressions (red lines) and their 95 % confidence interval (gray shading) as well as the correlation coefficient (r) and correlation test P value (P) in the top left corner of each panel. PRS_{QT} were normalized to a mean of 0 and standard deviation of 1.



Supplementary Figure XII | LAE-free Kaplan-Meier survival curve stratified by PRS_{QT} quartiles (Q1-Q4) in the European (**A**) and Japanese (**B**) LQTS cohorts. Life-threatening arrhythmic events (LAE) were defined as out of hospital cardiac arrest or hemodynamically unstable ventricular tachycardia/ventricular fibrillation. Data shown include all LQTS cases, both genotype positive and negative. Logrank test P=0.73 for Europeans (**A**) and P=0.06 for Japanese (**B**).



REFERENCES

For references see main text.