GWAS of QT interval in Hispanics/Latinos generalizes 13 loci and identifies population-specific signals

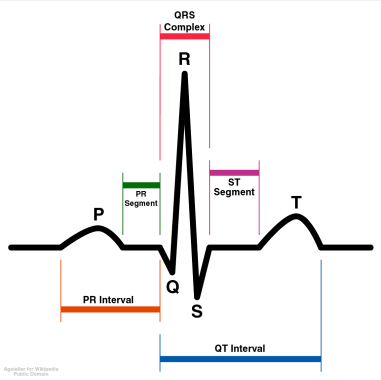
Raúl Méndez Giráldez

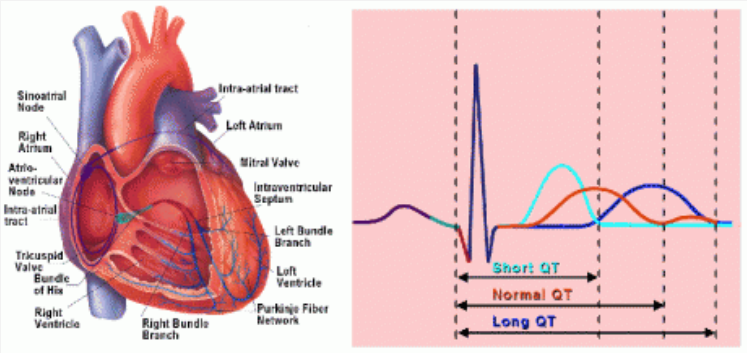
ENCODE Research Applications and Users Meeting

Stanford, June 8 - 10, 2016

UNC
GILLINGS SCHOOL OF
GLOBAL PUBLIC HEALTH

BACKGROUND: QT Interval





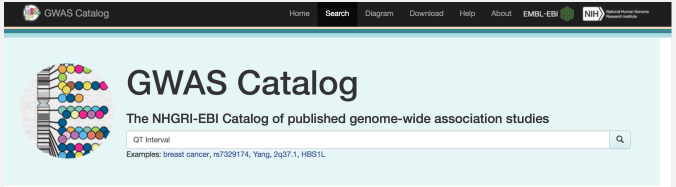
- QT is an ECG measure of ventricular depolarization and repolarization.
- Long QT interval, can cause ventricular tachyarrhythmias (*torsade de pointes*)
- Short QT can also yield ventricular fibrillations.
- Both LQT & SQT increase risk of Sudden Cardiac Death

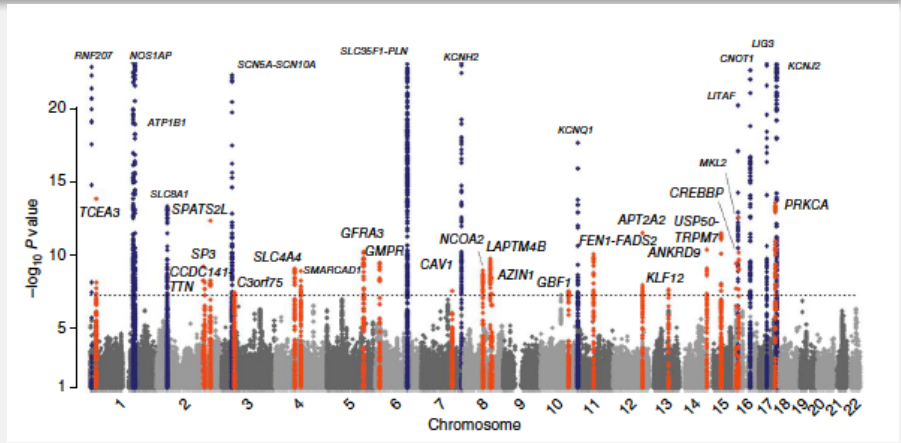
MASONIC MEDICAL RESEARCH LABORATORY

Genome Wide Association Studies (GWAS)

$$QT(ms) \sim \beta_0 + \beta_1 \times (SNP_i) + \sum_j^k \beta_j \times Cov_j + \varepsilon_i$$

$$\frac{\beta_1}{\sqrt{Var}} \sim \mathrm{N}(0,1) \rightarrow P_{val}$$



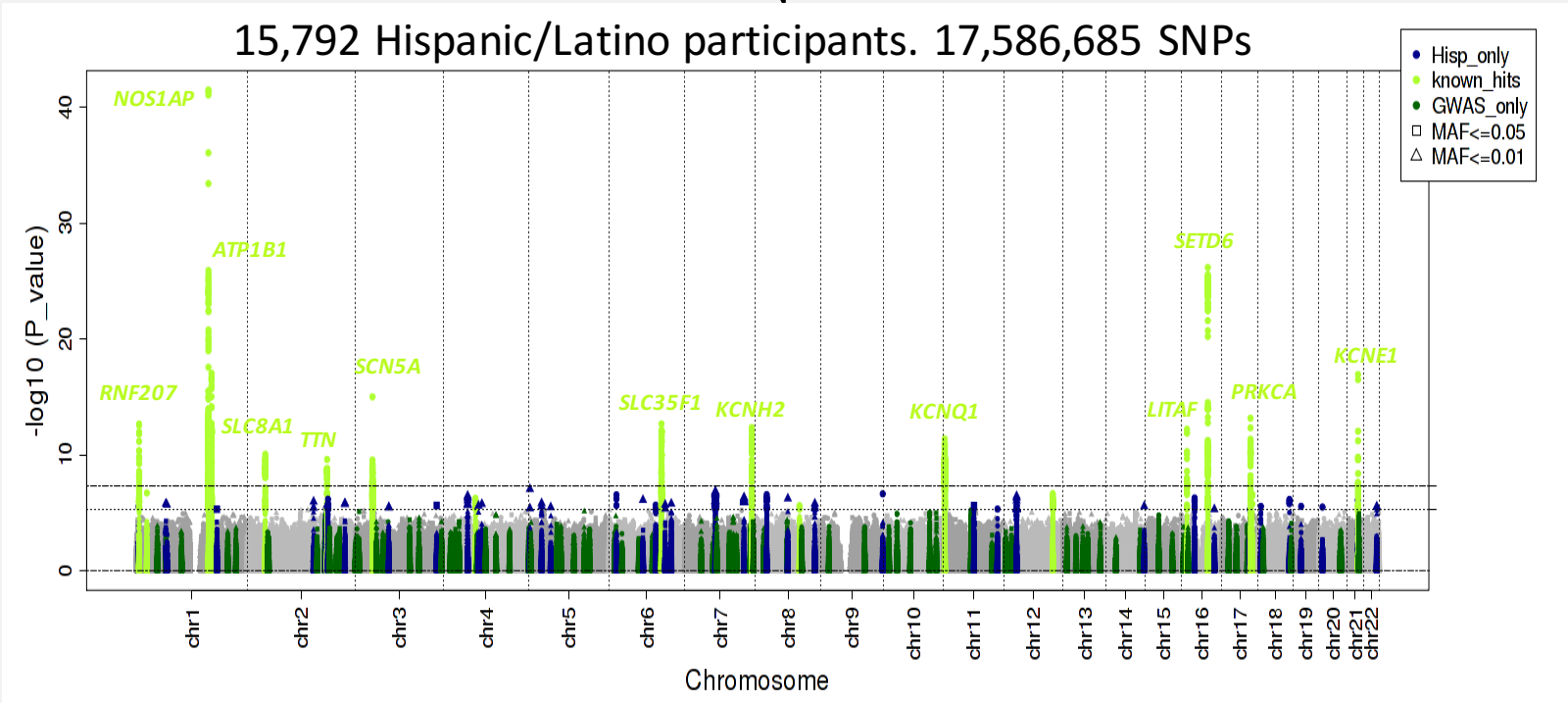


Arking DE *et al.* Nat Genet. 2014 Aug;46(8):826-36

- Association tests for ~ $10^{6\text{-}7}$ SNPs
- 170 SNPs associated to QT interval (~ 104 different genes)
- 11 GWAS (8 in EUR, 2 AA, 1 ASN)
- Need to analyze genetic structure in Hispanic/Latinos

Results: fixed effects meta-analysis

$$w_i = \frac{1}{se_i^2}; \beta = \frac{\sum_{i=1}^{N} \beta_i w_i}{\sum_{i=1}^{N} w_i}; SE = \sqrt{\frac{1}{\sum_{i=1}^{N} w_i}} \rightarrow \frac{\beta}{SE} \sim N(0,1) \rightarrow P_{val}$$



UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

Results: secondary signals & population specificity

Conditional analysis

Nearest Gene	SNP	Chr	Position (hg19)	Coded	Non-coded	CAF	β (ms)	SE (ms)	P-val
NOS1AP	rs3934467	1	162182677	T	C	0.28	2.46	0.23	9.44e-26
	rs73017364	1	162184746	T	C	0.87	2.23	0.31	5.29e-13
ATP1B1	rs1320977	1	169073388	A	G	0.15	-2.09	0.29	6.05e-13
	rs1138486	1	169101935	T	C	0.14	-1.85	0.31	1.78e-09
SCN5A	rs6762565	3	38582191	T	C	0.19	-1.69	0.29	9.42e-09
KCNQ1	rs78695585	11	2644544	A	G	0.04	3.69	0.59	3.18e-10

$$QT(ms) \sim \beta_0 + \beta_1 \times (SNP_i) + \sum_j^k \beta_j \times Cov_j + \beta_{j+1} \times SNP_1 + \beta_{j+2} \times SNP_2 + \cdots + \varepsilon_i$$

$$SNP_1 \equiv dosage\ for\ SNP\ conditoned\ on\ \equiv 2 \times P_{AA} + P_{AB}$$

Locus	Chr	GWAS SNP	Position (GRCh38)	Discovery Population (GWAS)	Hispanic/Latino Lead SNP	Position (GRCh38)	Discovery Population (Hisp)	r2 EUR	r2 AMR	r2 AFR	r2 ASN
SCN5A	3	rs11129795	38547672	EUR	rs3922844	38582762	NA	<0.05	<0.05	<0.05	<0.05
	3	rs12053903	38551902	EUR	rs3922844	38582762		<0.05	<0.05	<0.05	<0.05
	3	rs6793245	38557546	EUR	rs3922844	38582762		<0.05	<0.05	<0.05	<0.05
	3	rs11708996	38592432	EUR	rs3922844	38582762		0.06	0.09	<0.05	<0.05
	3	rs11710077	38616408	EUR	rs3922844	38582762		0.16	0.16	<0.05	0.31
	3	rs6599234	38673809	EUR	rs3922844	38582762		<0.05	<0.05	<0.05	<0.05
	3	rs6801957	38725824	EUR	rs3922844	38582762		<0.05	<0.05	<0.05	<0.05
KCNE1	21	rs1805128	34449382	EUR	rs12626657	34455875	NA	<0.05	<0.05	<0.05	<0.05

Population structure $r^2 = \left[\frac{P_{AB} - P_A \times P_B}{\sqrt{P_A \times (1-P_A) \times P_B \times (1-P_B)}}\right]^2$

ATP1B1 locus

↓

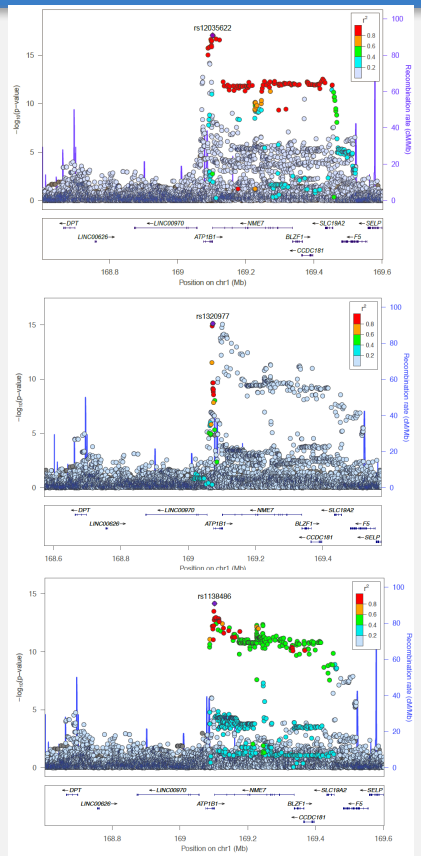
No condition

↓

Cond. rs12035622

↓

Cond. rs12035622+ rs1320977

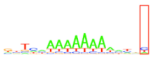
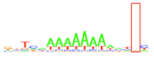


UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

A look into ENCODE/RoadMap data

RegulomeDB results for lead SNP rs12035622

Motifs Filter:

Method	Location	Motif	Cell Type	PWM	Reference
PWM	chr1:169102323..169102340	Srf			19443739
PWM	chr1:169102323..169102340	Elf3			19443739
PWM	chr1:169102324..169102341	Srf			19443739
PWM	chr1:169102324..169102341	Elf3			19443739

Single nucleotides Filter:

Method	Location	Affected Gene	Cell Type	Additional Info	Reference
eQTL	chr1:169102339..169102340	ATP1B1	Pons	cis	20485568

Chromatin structure Filter:

Method	Location	Cell Type	Additional Info	Reference
DNase-seq	chr1:169102300..169102450	Th17		ENCODE
DNase-seq	chr1:169102300..169102450	Th1		ENCODE
DNase-seq	chr1:169102300..169102450	Th2		ENCODE

Histone modifications Filter: Heart

Method	Location	Chromatin State	Tissue Group	Tissue	Reference
ChromHMM	chr1:169099400..169111200	Weak transcription	Heart	Aorta	REMC
ChromHMM	chr1:169100800..169116200	Quiescent/Low	Heart	Right Atrium	REMC
ChromHMM	chr1:169098200..169105000	Weak transcription	Heart	Right Ventricle	REMC
ChromHMM	chr1:169098800..169102400	Strong transcription	Heart	Left Ventricle	REMC
ChromHMM	chr1:169101200..169102800	Weak transcription	Heart	Fetal Heart	REMC

- TF binding site from PWM that overlap marginally with our lead SNP
- eQTL in cis with gene ATP1B1 itself in Pons mid brain tissue.
- DNase I hypersensitivity data in T-helper cells, and the results of ChromHMM in heart tissue.

CONCLUSIONS

- The same loci influence QT across populations, although there is population-specific variation.
- Do this variation involve differential gene regulation in Hispanic/Latinos? Need of population specific data in ENCODE.
- QT genetic risk scores need to be tailored to the target population (i.e. Hispanic/Latinos)

ACKNOWLEGMENTS:

Christy Avery

Eric Whitsel

Amanda Seyerle

Rahul Gondalia

SOL STUDY OF LATINOS
Hispanic Commun
Hispanic Community Health Study

WOMEN'S HEALTH INITIATIVE

MESA

Members of the Hispanic/Latino ECG Genomics Working Group