ORIGINAL RESEARCH ARTICLE

Genome-wide association study of PR interval in Hispanics/Latinos identifies novel locus at *ID2*

Amanda A Seyerle,[1,2] Henry J Lin,[3,4] Stephanie M Gogarten,[5] Adrienne Stilp,[5] Raul Méndez Giráldez,[1] Elsayed Soliman,[6,7] Antoine Baldassari,[1] Mariaelisa Graff,[1] Susan Heckbert,[8,9] Kathleen F Kerr,[5] Charles Kooperberg,[10] Carlos Rodriguez,[7,11] Xiuqing Guo,[3,4] Jie Yao,[3,4] Nona Sotoodehnia,[9,12] Kent D Taylor,[13] Eric A Whitsel,[1,14] Jerome I Rotter,[3,4] Cathy C Laurie,[5] Christy L Avery[15,16]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/heartjnl-2017-312045).

For numbered affiliations see end of article.

Correspondence to
Dr Amanda A Seyerle, Division of Epidemiology and Community Health, University of Minnesota, 1300 S 2nd Street, Suite 300, Minneapolis, MN 55454, USA; aseyerle@umn.edu

Received 28 June 2017
Revised 10 October 2017
Accepted 16 October 2017
Published Online First 10 November 2017

ABSTRACT

Objective PR interval (PR) is a heritable electrocardiographic measure of atrial and atrioventricular nodal conduction. Changes in PR duration may be associated with atrial fibrillation, heart failure and all-cause mortality. Hispanic/Latino populations have high burdens of cardiovascular morbidity and mortality, are highly admixed and represent exceptional opportunities for novel locus identification. However, they remain chronically understudied. We present the first genome-wide association study (GWAS) of PR in 14 756 participants of Hispanic/Latino ancestry from three studies.

Methods Study-specific summary results of the association between 1000 Genomes Phase 1 imputed single-nucleotide polymorphisms (SNPs) and PR assumed an additive genetic model and were adjusted for global ancestry, study centre/region and clinical covariates. Results were combined using fixed-effects, inverse variance weighted meta-analysis. Sequential conditional analyses were used to identify independent signals. Replication of novel loci was performed in populations of Asian, African and European descent. ENCODE and RoadMap data were used to annotate results.

Results We identified a novel genome-wide association ($P<5\times10^{-8}$) with PR at *ID2* (rs6730558), which replicated in Asian and European populations ($P<0.017$). Additionally, we generalised 10 previously identified PR loci to Hispanics/Latinos. Bioinformatics annotation provided evidence for regulatory function in cardiac tissue. Further, for six loci that generalised, the Hispanic/Latino index SNP was genome-wide significant and identical to (or in high linkage disequilibrium with) the previously identified GWAS lead SNP.

Conclusions Our results suggest that genetic determinants of PR are consistent across race/ethnicity, but extending studies to admixed populations can identify novel associations, underscoring the importance of conducting genetic studies in diverse populations.

Hispanic/Latino populations are chronically understudied in medical research, despite a growing population that will represent 30% of the US population by 2060.[1] Hispanic/Latino populations shoulder a higher burden of cardiovascular diseases (CVDs), including myocardial infarction, stroke and death from CVD, compared with European American populations.[2–4] Hispanic/Latino populations are highly admixed, tracing their origins to Europe, Africa and the Americas.[5] Thus, Hispanic/Latino populations are uniquely positioned to inform on the genetic architecture of not just populations from the Americas but also on populations of European, Asian and African descent. Hispanic/Latino populations represent a unique genetic architecture in which to study the genetics of CVD.[6]

The PR interval (PR), a heritable measure of atrial depolarisation and atrioventricular nodal conduction,[7] is a promising yet unexplored candidate for study in Hispanic/Latino populations. PR is associated with atrial fibrillation, a risk factor for stroke, pacemaker implantation, heart failure and all-cause mortality.[8–10] Previous genome-wide association studies (GWAS) in European, African American and Asian populations have identified more than 15 genetic loci associated with PR, including *SCN5A*, *SCN10A*, *CAV1–CAV2*, *TBX5–TBX3* and *SOX5*.[11–19]

Hispanic/Latino populations pose an interesting paradox. The prevalence of prolonged PR is higher in Hispanics/Latinos, whereas their risk for atrial fibrillation is lower compared with European-descent populations.[20] As researchers begin to develop genetic risk scores for conditions such as atrial fibrillation, it is important to acknowledge evidence supporting population-specific variants underlying common complex traits such as atrial fibrillation.[21] Population-specific variants make it critical that genetic risk scores used in Hispanic/Latino populations incorporate results from genetic studies in Hispanic/Latino populations rather than extending genetic risk scores developed in European-descent populations. We present the first GWAS of PR in Hispanic/Latino populations, bringing together three large cohorts of Hispanics/Latinos, including the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), the Women's Health Initiative (WHI) and the Multi-Ethnic Study of Atherosclerosis (MESA).

MATERIALS AND METHODS

Study populations

Three cohorts participated in this meta-analysis (n=14 756): the HCHS/SOL (n=11 703), the MESA (n=1480) and the WHI (n=1573; table 1

To cite: Seyerle AA, Lin HJ, Gogarten SM, *et al*. *Heart* 2018;**104**:904–911.

Table 1 Descriptive characteristics of participants in three cohorts in Hispanic/Latino genome-wide association study

Population	N	Mean age, years (SD)	Mean PR, ms (SD)	% Women	Mean height, cm (SD)	Mean BMI, kg/m^2 (SD)	Mean SBP, mm Hg (SD)
HCHS/SOL	11 703	46 (14)	157 (21)	59	162 (9)	30 (6)	122 (18)
MESA	1480	61 (10)	162 (22)	52	162 (9)	29 (5)	127 (22)
WHI	1573	60 (6)	156 (21)	100	157 (6)	29 (5)	127 (17)

BMI, body mass index; HCHS/SOL, Hispanic Community Health Study/Study of Latinos; MESA, Multi-Ethnic Study of Atherosclerosis; PR, PR interval; SBP, systolic blood pressure; WHI, Women's Health Initiative.

and online supplemental text). Each study was approved by the institutional review board at the respective sites, and all participants provided written consent.

PR interval measurement

Resting, supine or semirecumbent ECGs were digitally recorded in each study at baseline by certified technicians using standard 12-lead ECGs using either Marquette MAC12 or MAC PC machines (GE Healthcare, Milwaukee, Wisconsin, USA; online supplementary table 1). Comparable procedures were used for preparing participants, placing electrodes, recording, transmitting, processing and controlling the quality of the ECGs. The PR interval was measured electronically using the Marquette 12SL algorithm.

Genotyping

Participants were genotyped on the Affymetrix Genome-Wide Human SNP Array V.6.0 (MESA, WHI) or an Illumina custom array. Genotypes were imputed using the 1000 Genomes Phase 1 reference panel (additional details in online supplemental text and online supplementary table 1).

Statistical analysis

Genome-wide analyses were performed by each cohort independently across approximately 25–39 million single-nucleotide polymorphisms (SNPs) assuming an additive genetic model, using linear regression (MESA, WHI) or linear mixed models (HCHS/SOL, online supplemental text). All analyses were adjusted for age, sex, study centre or region, height, body mass index (BMI), systolic blood pressure, heart rate, beta-blocker use and study principal components of ancestry to maintain consistency with prior PR GWAS.[11–19] Study-specific test statistics were corrected for genomic inflation (λ, online supplementary figure 1) and combined with inverse-variance weighted meta-analysis. For all significant index SNPs, we performed conditional analyses by adjusting for the index SNPs and any subsequently significant SNPs until no remaining genome-wide significant SNPs were identified.

Generalisation of previously reported PR SNPs

For SNPs previously associated with PR in published GWAS reports,[11 16] we examined evidence for generalisation using the approach described by Sofer *et al* (online supplemental text).[22] The method assigns an r value to every index SNP. The generalisation null hypothesis testing generalisation of the PR index SNPs to Hispanics/Latinos was rejected when the r value was less than 0.05, controlling the false discovery rate of the generalisation null hypothesis.

Replication of novel associations

For novel associations, we attempted to replicate our results in populations of East Asian,[14] European and African (WHI study, online supplemental text) descent using exclusion criteria and analytical procedures that overlapped with previously described approaches.

Linkage disequilibrium analysis

Linkage disequilibrium (LD) was calculated with the r^2 correlation statistic using 1000 Genomes metapopulations as reference (AMR, EUR, AFR, ASN). RFMix and ASAFE were used to estimated local ancestry and ancestry-specific allele frequencies (online supplemental text).

Bioinformatic annotation

For loci associated with PR at a genome-wide significance level in our Hispanic/Latino population, we used epigenetic data from the ENCODE and RoadMap projects to examine evidence of functional annotation. Results were restricted to annotation found in available heart tissue (fetal heart, right atrium, and right or left ventricle). Additionally, publically available data from genetic association studies of clinical phenotypes associated with PR were queried for associations with lead Hispanic/Latino SNPs (online supplemental text).

RESULTS

Study population characteristics

A total of 14 756 participants from three cohorts (MESA, HCHS/SOL, WHI) contributed to this study (table 1). The majority of participants were from HCHS/SOL (n=11 703), who were, on average, 14 years younger than MESA and WHI participants. In addition, participants were predominantly women (52%–100%) and obese (average BMI=29–30 kg/m^2). After study-specific quality control and filtering by effective sample size, studies contributed between 8 217 098 (MESA) and 17 322 742 (HCHS/SOL) imputed SNPs (online supplementary table 1), which together represented 18 828 993 unique SNPs.

Novel association at *ID2*

A total of 454 SNPs at seven loci were genome-wide significant ($P<5\times10^{-8}$, figure 1, table 2), with no evidence of genomic inflation (study-specific λ range: 0.975 to 1.02, meta-analysis λ=1.004; online supplementary figure 1). One of these seven loci located upstream of *ID2* (inhibitor of DNA binding 2, HLH protein), a gene encoding a transcriptional regulator, was novel. The Hispanic/Latino lead SNP, rs6730558 ($P=3\times10^{-8}$; table 2, figure 1), was common in our population (coded allele frequency (CAF)=0.49) and had homogeneous results ($P_{het}=0.7$). Rs6730558 also was common in 1000 Genomes metapopulations (African (AFR): CAF=0.67, European (EUR): CAF=0.39, East Asian (ASN): CAF=0.57; table 3). Furthermore, when we calculated the local ancestry at each index SNP in the HCHS/SOL population and then calculated the CAF by ancestral haplotype, the T allele of rs6730558 was common across all three ancestral populations (European descent: 0.39, African descent: 0.70, Native American: 0.56; table 3). Examination of LD patterns from four 1000 Genomes metapopulations (admixed

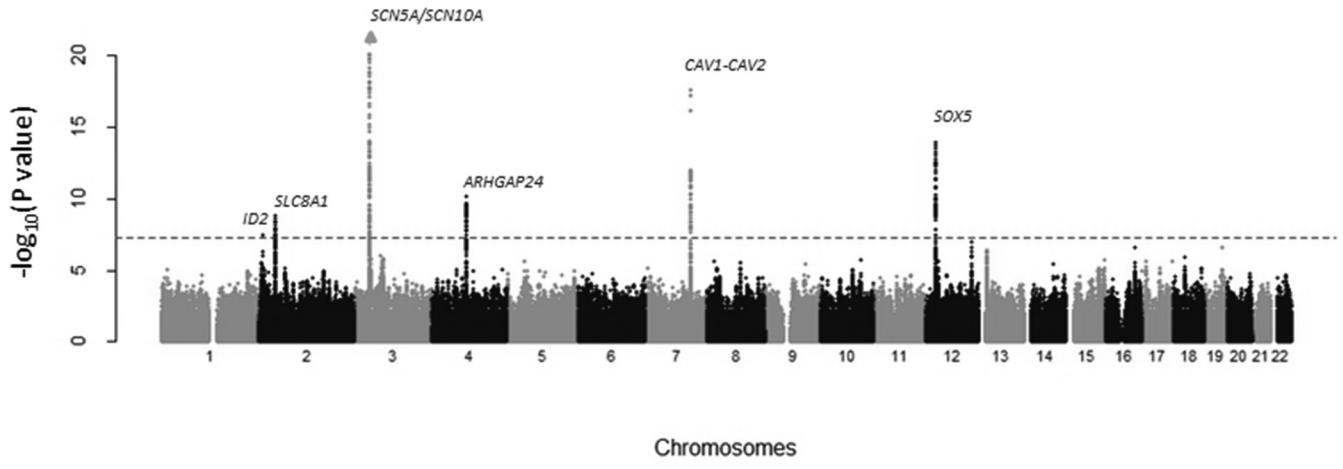

Figure 1 Manhattan plot of genome-wide association meta-analysis results in Hispanic/Latino population (n=14 756). Loci meeting genome-wide significance ($P<5\times10^{-8}$, denoted with dashed line) are annotated with nearest gene name.

American (AMR), AFR, EUR and ASN) showed that rs6730558 resides in a very narrow LD block, with particularly weak LD in AFR populations (figure 2). Conditional analyses did not identify any independent SNPs at this locus.

We attempted to replicate the association of rs6730558 with PR in independent European descent (n=4296), African American (n=3763) and Asian (n=6805) populations given the admixed nature of Hispanic/Latino populations (table 4). A Bonferroni-corrected significance level of 0.017 (ie, 0.05/3 tests) was used to determine significance. For replication in Asian populations, we identified rs3856447 as a proxy in high LD with rs6730558 (r^2 >0.9), which was associated with PR at $P=7\times10^{-3}$. Rs6730558 was also associated with PR in the European descent replication population (P=0.002) but failed to replicate within the African American replication population (P=0.05). However, the direction of effect and magnitude of the effect size was consistent with the Hispanic/Latino lead SNP across all three populations (range: 1.07–1.60 ms).

The *ID2* locus showed no associations among previously published genetic associations with atrial fibrillation (online supplementary table 2).

Replication of previously identified PR loci

Six loci previously associated with PR in other race/ethnicities were significantly associated with PR in Hispanics/Latinos (table 2, figure 1). Of the lead SNPs at these six loci in Hispanics/Latinos, three (rs3922844 at *SCN5A*, rs6801957 at *SCN10A* and rs3807989 at *CAV1–CAV2*) were identical to previously reported index SNPs (table 5). Furthermore, at *ARGHAP24*, the Hispanic/Latino lead SNP was in high LD with an index SNP previously identified in only European populations and a signal previously identified in only African American populations ($r^2=0.93$ and 0.82, respectively).

The remaining two loci have only been previously described in a single race/ethnicity. The Hispanic/Latino index SNP was in high LD with the previously described index signals at both the *SOX5* (European populations) and *SLC8A1* (Asian populations) loci ($r^2=0.93$ and 0.97, respectively). The lead SNP at the *SLC8A1* locus was rare in the EUR population (CAF=0.03) and also resided in a narrow LD block in the AFR population, although LD was similar across AMR, EUR and ASN populations (table 3, figure 2). Rs17026148 at *SLC8A1* was common in

Table 2 Common genetic variants (coded allele frequency >0.05) associated with PR interval in Hispanic/Latino population (n=14 756) at genome-wide significance ($P<5\times10^{-8}$)

Locus	Position	Index SNP	CA	CAF	P value	Effect estimate (SE)	I^2	P_{het}
*ID2**	2p25.1	rs6730558	T	0.49	3×10^{-8}	1.36 (0.25)	0	0.7
SLC8A1	2p22.1	rs17026148	A	0.17	2×10^{-9}	1.98 (0.33)	0	0.4
SCN5A	3p21.0	rs3922844	T	0.37	4×10^{-41}	−3.39 (0.25)	58	0.09
		rs7374004†	A	0.65	3×10^{-19}	2.66 (0.30)	2	0.3
		rs455675334‡	A	0.87	1×10^{-11}	−2.49 (0.37)	0	0.4
SCN10A	3p22.0	rs6801957	T	0.38	1×10^{-55}	3.90 (0.25)	0	0.9
ARHGAP24	4q22.1	rs13105921	A	0.41	6×10^{-11}	1.76 (0.27)	64	0.06
CAV1–CAV2	7q31.1	rs3807989	A	0.41	2×10^{-18}	2.18 (0.25)	0	0.7
SOX5	12p12.1	rs146974314	A	0.14	9×10^{-15}	−2.66 (0.35)	46	0.2

*Novel association.
†Conditioned on rs3922844 and rs6801957.
‡Conditioned on rs3922844, rs6801957 and rs7374004.
CA, coded allele; CAF, coded allele frequency in our study population meta-analysis; I^2, percentage of total variation across cohorts due to heterogeneity; P_{het}, P value for heterogeneity calculated using Cochran's Q statistic; SNP, single-nucleotide polymorphism.

Table 3 Comparison of coded allele frequencies in Hispanic/Latino population and in 1000 Genomes metapopulations at signals associated with PR interval in Hispanic/Latino population

			CAF by ancestral haplotype*			1000 Genomes population CAF		
Locus	Index SNP (coded allele)	Hispanic/Latino CAF	EU	AF	NA	EUR	AFR	ASN
ID2	rs6730558 (T)	0.49	0.39	0.70	0.56	0.39	0.67	0.57
SLC8A1	rs17026148 (A)	0.17	0.03	0.14	0.70	0.03	0.16	0.43
SCN5A	rs3922844 (T)	0.37	0.32	0.69	0.30	0.29	0.67	0.14
SCN10A	rs6801957 (T)	0.38	0.44	0.11	0.40	0.42	0.14	0.22
ARHGAP24	rs13105921 (A)	0.41	0.67	0.12	0.02	0.65	0.14	0.07
CAV1–CAV2	rs3807989 (A)	0.41	0.42	0.72	0.23	0.40	0.73	0.31
SOX5	rs146974314 (A)	0.14	0.17	0.001	0.17	0.15	0.01	0.13

*CAF was calculated in the Hispanic Community Health Study/Study of Latinos after calculating local ancestry (EU, AF or NA). From there, the CAF was calculated for each ancestral population at each separate chromosome region.
AF, African ancestry; AFR, 1000 Genomes African metapopulation; ASN, 1000 Genomes East Asian metapopulation; CAF, coded allele frequency; EU, European ancestry; EUR, 1000 Genomes European metapopulation; NA, Native American ancestry; SNP, single-nucleotide polymorphism.

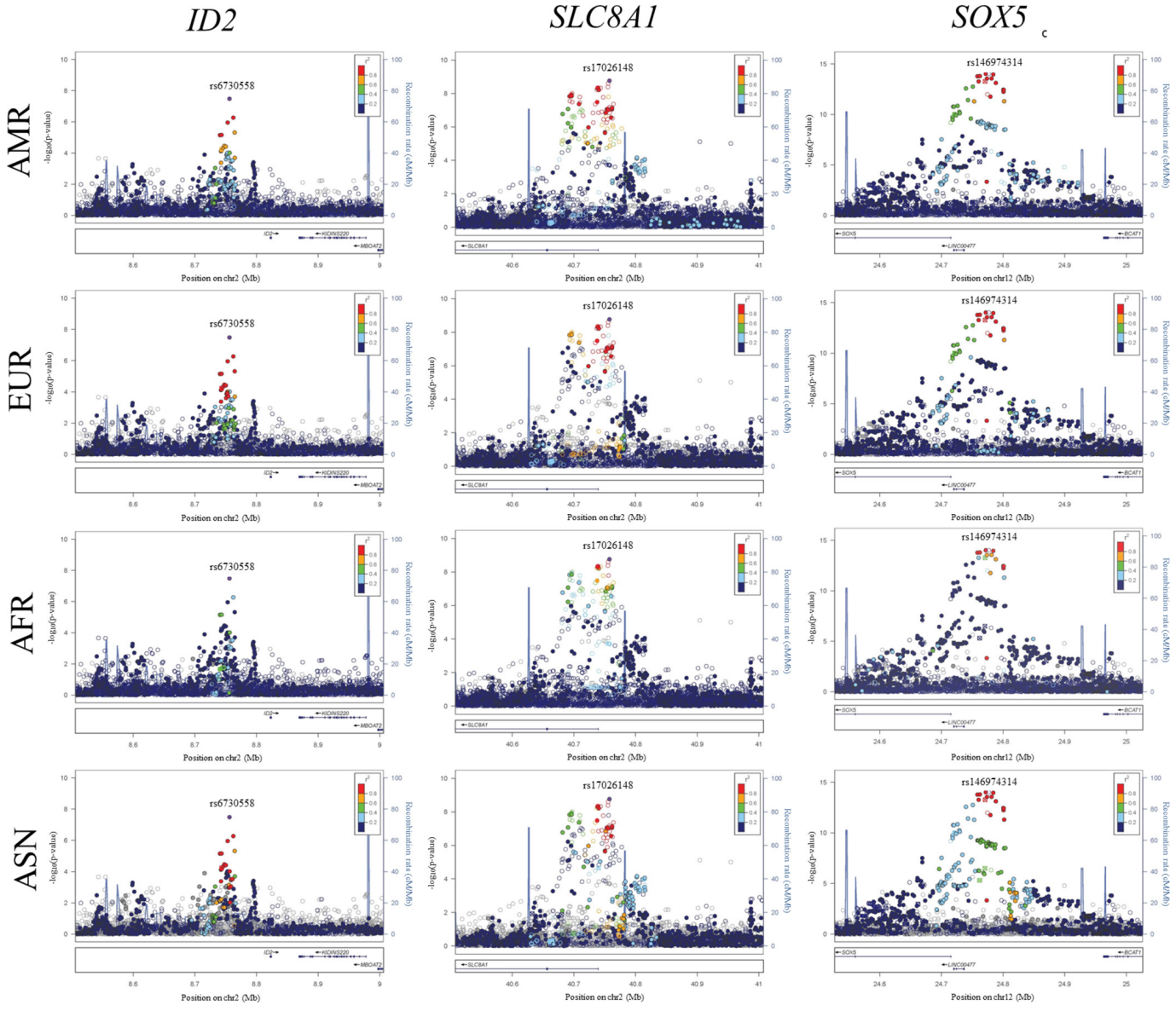

Figure 2 LocusZoom plots of common (coded allele frequency >0.05) single-nucleotide polymorphisms at three select regions (*ID2*, *SLC8A1*, *SOX5*) associated with PR interval in Hispanic/Latino populations. Results from the genome-wide association study meta-analysis in Hispanics/Latinos are plotted on linkage disequilibrium patterns calculated using 1000 Genomes metapopulations representing admixed American (AMR), European (EUR), African (AFR) and East Asian (ASN) populations.

Table 4 Replication of *ID2* Hispanics/Latinos signal in Asian (n=6805), African American (n=3763) and European American (n=4296) populations

Race/ethnic group	Replicated SNP	r^2 between replicated SNP and Hispanics/Latino index SNP	CA	CAF	P value	Effect estimate (SE)
Asian*	rs3856447	0.9	A	0.44†	0.007	1.07 (0.40)
African American‡	rs6730558	NA	T	0.61	0.05	1.16 (0.60)
European‡	rs6730558	NA	T	0.38†	0.002	1.60 (0.52)

*Asian results obtained from Hong *et al*.[14] The lead SNP in Hispanics/Latinos, rs6730558, was not present in the dataset. A proxy SNP in Asian populations was identified using the SNAP database and the r^2 value reported in the table represents the linkage disequilibrium between rs3856447 and rs6730558 in the combined Han Chinese (CHB) and Japanese (JPT) 1000 Genomes populations.
†CAF not available in study dataset. Reported CAF obtained from 1000 Genomes metapopulations, EAS (East Asian) and EUR (European).
‡African American and European results obtained using Women's Health Initiative populations. African American population from the SHARe study. European populations from the WHIMS study.
CA, coded allele; CAF, coded allele frequency; SNP, single-nucleotide polymorphism.

both the ASN population as well as in HCHS/SOL participants with Native American local ancestry at the rs17026148 chromosomal position (CAF=0.43 and 0.70, respectively; table 3). Similarly, the Hispanic/Latino association at *SOX5* was rare in the AFR population (CAF=0.01) and the Hispanic/Latino lead SNP had less LD in the ASN population than the EUR population, which showed similar LD patterns to the AMR population.

Conditional analyses identified independent signals at the *SCN5A/SCN10A* locus (rs7374004 and rs45567533) but did not identify independent signals at any other loci (table 2). Rs6801957 was also identified as a secondary signal in African Americans[11] and was in high LD with previously published index SNPs rs6800541 (r^2=0.93),[12 14 16] rs6795970 (r^2=0.93)[12 13] and rs6798015 (r^2=0.83)[19] in HapMap CEU populations and rs6798015 (r^2=0.87) in CHB+JPT populations (online supplementary table 3). Secondary signal rs45567533 was not in high LD with any previously published index SNP (r^2<0.8). There was little LD at this locus in the HapMap YRI or 1000 Genomes AFR populations (online supplementary figure 2).

Table 5 Comparison of genetic loci associated with PR interval in Hispanic/Latino populations with results from previous genome-wide association studies in European-descent,[13 16] African-descent[11 19] and Asian-descent[14 17 18] populations

Locus	Index SNP in Hispanics/Latinos	Previously identified index SNPs		
		Index SNP	Population	r^2
SLC8A1	rs17026148	rs17026114	AS	0.93
		rs17026156	AS	0.88
SCN5A	rs3922844	rs3922844	EU, AA	–
		rs7638909	AS	0.008
		rs11708996	EU	0.06
		rs6599222	AA	0.0003
		rs6763048	AA	0.05
SCN10A	rs6801957	rs6801957	EU, AA, AS	–
		rs6795970	EU, AS	0.79
		rs6800541	EU, AS	0.77
		rs6798015	AA	0.80
		rs6599257	AS	0.57
ARHGAP24	rs13105921	rs7692808	EU	0.93
		rs7660702	EU	0.74
		rs11732231	AA	0.82
CAV1-CAV2	rs3807989	rs3807989	EU	–
		rs11773845	EU, AA, AS	0.99
SOX5	rs146974314	rs11047543	EU	0.97

AA, African-descent population; AS, Asian-descent population; EU, European-descent population; r^2, correlation between previously identified index SNP and index SNP identified in Hispanic/Latino population using Hispanic/Latino linkage disequilibrium pattern; SNP, single-nucleotide polymorphism.

Generalisation

Additionally, we evaluated evidence of generalisation for 15 index SNPs reported by previous GWAS in Europeans (n=9) and African Americans (n=6).[11 16] Eight of the nine SNPs previously identified in the European population and all six SNPs previously identified in the African American population generalised to Hispanics/Latinos (r value <0.05) (online supplementary table 4 and supplementary figure 3). Notably, SNPs that generalised to Hispanics/Latinos included four SNPs at the *MEIS1*, *ITGA9*, *TBX5/TBX3* and *NKX2–5* loci in addition to six index SNPs at loci identified above as genome-wide significant in Hispanics/Latinos.

Bioinformatic annotation of associated variants

Finally, we performed bioinformatic annotation of the lead SNPs associated with PR in our Hispanic/Latino population (online supplementary table 5). The *ID2* locus (rs6730558) showed evidence of enhancer activity and H3K4me1 and H3K27ac modification, also associated with enhancer activity, in heart tissues (left ventricle, right ventricle, right atrium and fetal heart). The lead SNPs at *SCN5A*, *SCN10A* and *CAV1–CAV2* also showed evidence of cardiac enhancer activity. Our lead SNP at *ARHGAP24* (rs13105921) showed evidence of promoter and transcriptional activity, including H3K4me3 and H3K9ac modification. Signals at *SLC8A1* and *SOX5* had no evidence of functional annotation.

DISCUSSION

In this study, the first PR GWAS in Hispanics/Latinos, we identified a novel association with PR at *ID2* and generalised 10 loci, including 6 at genome-wide significant levels, to Hispanics/Latinos. Additionally, we identified one population-specific secondary signal at *SCN5A* and generalised another secondary signal at *SCN5A*. Finally, the results at *SLC8A1*, a locus previously only identified in East Asian populations,[14 17] demonstrated that evaluation of an admixed population can aid in improving characterisation of loci that had not previously extended across populations. These results emphasise the need to conduct genetic studies in admixed populations.

We identified a novel association at rs6730558, located 54 kb downstream of *ID2*, a gene encoding a DNA binding inhibitor protein. *ID2* has been implicated in the development of the cardiac conduction system of mice[23] but has yet to be associated with electrophysiology phenotypes in humans. However, other

DNA regulatory proteins, such as *TBX5/TBX3*, have been implicated in the regulation of PR in both GWAS and animal models. Associations with regulatory genes indicate that transcriptional control can be involved in PR duration. The evidence of enhancer activity at the *ID2* Hispanic/Latino lead SNP (rs6730558) also suggests that the underlying association may be affecting PR through the increased transcription of *ID2*, which in turn would affect PR duration.

There are several reasons *ID2* may not have been identified in previous GWAS of PR. First, rs6730558 may be a population-specific signal. However, it is more likely that the *ID2* signal was poorly captured in previous analyses. While rs6730558 is common across genetic ancestries, it lies on a narrow LD block, suggesting that previous genotyping and imputation platforms may not have adequately captured the *ID2* signal. In particular, African populations demonstrate extremely low LD with the lead SNP identified in this analysis, making it more difficult to characterise the region in African-descent populations.

The admixed genomes in Hispanic/Latino populations enabled us to detect two previously population-specific associations. While the *SLC8A1* and *SOX5* loci had previously not been replicated in any other populations, we were able to successfully replicate the same signal at both of these two loci in our highly admixed Hispanic/Latino population. There are several possible explanations for the lack of findings at the *SLC8A1* and *SOX5* loci. Similar to *ID2*, the signals could have been inadequately genotyped or imputed in previous studies, particularly those using HapMap imputation panels in non-EU populations. Imputation accuracy is an existing question in admixed populations such as African Americans, which have large proportions of African ancestry and shorter LD blocks.[24] The narrow LD blocks found in our study also suggest that poor imputation anywhere in the narrow LD block would make it difficult to adequately capture the signal at that locus. Use of a broader range of populations along with the increasing use of whole genome sequencing (WGS) data, for both genetic association studies and imputation, will aid future genetic research in better capturing disease association signals.

However, another possible explanation is that the *SLC8A1*, *SOX5* and *ID2* signals lie on a European or Asian/Native American haplotypes. There are several lines of evidence that could support the hypothesis that these signals lie on population-specific haplotypes. For example, although our index signal at *ID2* (rs6730558) is common across ancestral populations (EUR, AFR, AMR and ASN), there is almost no LD at the *ID2* locus in AFR populations. Conversely, the *SLC8A1* and *SOX5* loci both demonstrate more LD across all four ancestral populations. However, the index SNP at *SLC8A1*, rs17026148, is rare in European populations, and the index SNP at *SOX5*, rs146974314, is rare in African populations. Prior work on other disease states have found population-specific alleles and haplotypes associated with diseases such as schizophrenia,[25] glaucoma[26] and asthma.[27]

Furthermore, rare and population-specific variants are now being identified through WGS, which continues to lead to the discovery of new population-specific signals. For example, sequencing of a Sardinian population found ~76 000 variants common among Sardinians but rare in the 1000 Genomes panel, including four novel loci associated with lipid levels and inflammatory markers.[28] Additionally, whole exome sequencing results have found low-frequency (1%–5%) variants and rare variants (<5%) associated with asthma unique to either Hispanics/Latinos or African Americans.[29] Results from sequencing studies suggest that rare and low-frequency variants can be ethnic specific.

Despite this knowledge, genomics research remains dominated by studies in European ancestry populations. As of 2016, minority populations account for 20% of all populations in GWAS.[30] However, much of this is focused in populations of Asian and some African ancestry, leaving little research in populations of Hispanic ancestry. Minority populations are better represented in WGS efforts. However, even in the National Heart, Lung, and Blood Institute's Trans-Omics for Precision Medicine (TOPMed), African American ancestry participants account for 30% of the TOPMed population while Hispanics/Latinos remain chronically undersampled, accounting for just 10% of the TOPMed population.[30] Our study, along with those discussed herein, emphasises the need to include these minority populations in health research, but work still needs to be done to adequately include these populations in future research.

Furthermore, clinical applications of genetic research will require diverse ancestral populations. Work discussed herein, as well as the results of this study, demonstrates that the genetic underpinnings of disease-related traits can vary by population. As precision medicine advances, it is imperative that genetic research expands to include diverse populations, such as Hispanics/Latinos. Current advances in precision medicine disproportionately benefit European-descent populations, which can exacerbate health disparities, but work such as that presented herein ensures that advances in the knowledge of disease aetiology and any subsequent clinical applications are applicable to relevant populations.

This study had several notable limitations. First, our sample size of Hispanics/Latinos was only half that of the largest study in EU populations.[16] However, our sample was comparable to previous work in African American ancestry populations[11] and twice that of previous Asian ancestry GWAS of PR.[14 17] Furthermore, even with our reduced sample size, we were able to identify a novel association at *ID2* and replicate all but one previous association, suggesting that our diverse population allowed us to better capture areas of the genome such as the *ID2* locus. Our ability to examine differences among Hispanic/Latino ancestries was limited also by the composition of our cohorts. While our

Key messages

What is already known on this subject?
The PR interval (PR) is a heritable measure of atrial depolarisation and atrioventricular nodal conduction. Previous genome-wide association studies (GWAS) have identified more than 15 genetic loci associated with PR, but much of the heritability remains unexplained. Hispanic/Latino populations are particularly understudied in genetic analyses, including those of PR.

What might this study add?
We have conducted the first GWAS of PR in Hispanic/Latino populations, identifying one novel locus associated with PR and generalising a further six loci (which had previously been associated with PR in other race/ethnicities) to Hispanics/Latinos.

How might this impact on clinical practice?
Hispanic/Latino populations pose an interesting paradox, with a higher prevalence of prolonged PR but a lower risk of atrial fibrillation compared with European descent populations. Better understanding the genetic architecture of PR in Hispanics/Latinos can help better identify individuals at risk of negative health outcomes due to prolonged PR.

study included a broad range of genetic ancestries (Mexican, Cuban, Dominican, Puerto Rican, South American, Central American), Mexican ancestry accounted for more than double the proportion of our population than the next best represented ancestry. Finally, despite our broad range of ancestries, we were unable to further refine our populations of South or Central American ancestries. However, our study represents the first study of the genetics of PR in Hispanics/Latinos and includes a broad representation of Hispanic/Latino populations, improving on previous genetic studies of Hispanics/Latinos, which were previously focused on Mexican Americans.

In conclusion, our study identified a novel association with PR near *ID2* and generalised 10 previously identified loci to our Hispanic/Latino population. Our results support consistency in multiple genetic determinants of PR across race/ethnicities and underscore the importance of conducting genetic studies in diverse populations. Genetic research seeking to elucidate the genetic architecture of PR as well as research aiming to eliminate health disparities must make greater efforts to include diverse populations.

Author affiliations
[1]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[2]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota, USA
[3]Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA
[4]Department of Pediatrics, Division of Medical Genetics, Harbor-UCLA Medical Center, Torrance, California, USA
[5]Department of Biostatistics, University of Washington, Seattle, Washington, USA
[6]Division of Public Health Sciences, Wake Forest School of Medicine, Epidemiology Cardiology Research Center (EPICARE), Winston-Salem, North Carolina, USA
[7]Department of Medicine, Section of Cardiology, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA
[8]Department of Epidemiology, University of Washington, Seattle, Washington, USA
[9]Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, USA
[10]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA
[11]Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA
[12]Division of Cardiology, University of Washington, Seattle, Washington, USA
[13]Division of Genomic Outcomes and Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA
[14]Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[15]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[16]Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Correction notice Since this paper was first published online the middle initial L has been added to the author name Christy Avery.

Acknowledgements Hispanic Community Health Study/Study of Latinos (HCHS/SOL): We thank the participants and staff of the HCHS/SOL study for their contributions to this study. The baseline examination of HCHS/SOL was carried out as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236) and San Diego State University (N01-HC65237).

Contributors The following individuals were involved in the planning stages of the manuscript, including project development and analysis design: AAS, HJL, SMG, AS, RMG, ES, SH, KFK, CK, CR, NS, KDT, EAW, JIR, CCL and CA. The following individuals were involved in data collection and cleaning: HJL, ES, SH, CK, CR, XG, NS, EAW, JIR and CCL. The following individuals were involved in data analysis: AAS, SMG, AS, AB, MG, XG, JY and CCL. The following individuals were involved in manuscript preparation, editing and revision: AAS, HJL, AS, RMG, SH, KFK, NS, KDT, EAW, JIR, CCL and CA. Furthermore, all authors reviewed the final version of the manuscript and provided critiques.

Funding The following Institutes/Centers/Offices contributed to the first phase of HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, and NIH Institution–Office of Dietary Supplements. The Genetic Analysis Center at University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03). Genotyping efforts were supported by NHLBI HSN 26220/20054C, NCATS CTSI grant UL1TR000124 and NIDDK Diabetes Research Center (DRC) grant DK063491. AAS was also supported by training grants T32HL7055 and T32HL07779. NS was supported by R01HL116747 and R01 HL111089. Multi-Ethnic Study of Atherosclerosis (MESA): This research was supported by the Multi-Ethnic Study of Atherosclerosis (MESA) contracts HHSN2682015000031, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and by grants UL1-TR-000040, UL1-TR-001079 and UL1-RR-025005 from NCRR. Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. We also thank the other investigators, the staff and the participants of MESA for their valuable contributions. A full list of participating MESA investigators and institutions can be found online (http://www.mesa-nhlbi.org).

Competing interests None declared.

Ethics approval North Carolina.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Complete summary level results from this analysis will be made available on dbGap.

REFERENCES

1 Colby SL, Ortman JM, U.S. Census Bureau. Projections of the size and composition of the U.S. population: 2014 to 2060. *Current Population Reports* 2015:25–1143.
2 Goff DC, Nichaman MZ, Chan W, *et al*. Greater incidence of hospitalized myocardial infarction among Mexican Americans than non-Hispanic whites. The Corpus Christi Heart Project, 1988–1992. *Circulation* 1997;95:1433–40.
3 Morgenstern LB, Smith MA, Lisabeth LD, *et al*. Excess stroke in Mexican Americans compared with non-Hispanic Whites: the Brain Attack Surveillance in Corpus Christi Project. *Am J Epidemiol* 2004;160:376–83.
4 Pandey DK, Labarthe DR, Goff DC, *et al*. Community-wide coronary heart disease mortality in Mexican Americans equals or exceeds that in non-Hispanic whites: the Corpus Christi Heart Project. *Am J Med* 2001;110:81–7.
5 Manichaikul A, Palmas W, Rodriguez CJ, *et al*. Population structure of Hispanics in the United States: the multi-ethnic study of atherosclerosis. *PLoS Genet* 2012;8:e1002640.
6 Seldin MF, Pasaniuc B, Price AL. New approaches to disease mapping in admixed populations. *Nat Rev Genet* 2011;12:523–8.
7 Hanson B, Tuna N, Bouchard T, *et al*. Genetic factors in the electrocardiogram and heart rate of twins reared apart and together. *Am J Cardiol* 1989;63:606–9.
8 Cheng S, Keyes MJ, Larson MG, *et al*. Long-term outcomes in individuals with prolonged PR interval or first-degree atrioventricular block. *JAMA* 2009;301:2571–7.
9 Roy D, Talajic M, Dubuc M, *et al*. Atrial fibrillation and congestive heart failure. *Curr Opin Cardiol* 2009;24:29–34.
10 Soliman EZ, Prineas RJ, Case LD, *et al*. Ethnic distribution of ECG predictors of atrial fibrillation and its impact on understanding the ethnic distribution of ischemic stroke in the Atherosclerosis Risk in Communities (ARIC) study. *Stroke* 2009;40:1204–11.
11 Butler AM, Yin X, Evans DS, *et al*. Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts. *Circ Cardiovasc Genet* 2012;5:639–46.
12 Chambers JC, Zhao J, Terracciano CM, *et al*. Genetic variation in SCN10A influences cardiac conduction. *Nat Genet* 2010;42:149–52.
13 Holm H, Gudbjartsson DF, Arnar DO, *et al*. Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* 2010;42:117–22.
14 Hong KW, Lim JE, Kim JW, *et al*. Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. *Hum Mol Genet* 2014;23:6659–67.
15 Newton-Cheh C, Guo CY, Wang TJ, *et al*. Genome-wide association study of electrocardiographic and heart rate variability traits: the Framingham Heart Study. *BMC Med Genet* 2007;8 Suppl 1:S7.
16 Pfeufer A, van Noord C, Marciante KD, *et al*. Genome-wide association study of PR interval. *Nat Genet* 2010;42:153–9.
17 Sano M, Kamitsuji S, Kamatani N, *et al*. Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and

confirms previously reported associations. *Hum Mol Genet* 2014;23:6668–76.
18 Smith JG, Lowe JK, Kovvali S, *et al*. Genome-wide association study of electrocardiographic conduction measures in an isolated founder population: Kosrae. *Heart Rhythm* 2009;6:634–41.
19 Smith JG, Magnani JW, Palmer C, *et al*. Genome-wide association studies of the PR interval in African Americans. *PLoS Genet* 2011;7:e1001304.
20 Rane S, Patton KK. Impact of sex and ethnicity on arrhythmic risk. *Curr Cardiol Rep* 2015;17:50.
21 Avery CL, Wassel CL, Richard MA, *et al*. Fine mapping of QT interval regions in global populations refines previously identified QT interval loci and identifies signals unique to African and Hispanic descent populations. *Heart Rhythm* 2017;14:572–80.
22 Sofer T, Heller R, Bogomolov M, *et al*. A powerful statistical framework for generalization testing in GWAS, with application to the HCHS/SOL. *Genet Epidemiol* 2017;41:251–8.
23 Moskowitz IP, Kim JB, Moore ML, *et al*. A molecular pathway including Id2, Tbx5, and Nkx2-5 required for cardiac conduction system development. *Cell* 2007;129:1365–76.
24 Chanda P, Yuhki N, Li M, *et al*. Comprehensive evaluation of imputation performance in African Americans. *J Hum Genet* 2012;57:411–21.
25 Liu J, Li M, Su B. GWAS-identified schizophrenia risk SNPs at TSPAN18 are highly diverged between Europeans and East Asians. *Am J Med Genet B Neuropsychiatr Genet* 2016;171:1032–40.
26 Nakano M, Ikeda Y, Tokuda Y, *et al*. Novel common variants and susceptible haplotype for exfoliation glaucoma specific to Asian population. *Sci Rep* 2014;4:5340.
27 Torgerson D, Galanter JM, Roth LA, *et al*. Admixture mapping from existing genome-wide association data identifies SMAD2 as a population-specific risk factor for asthma in Latinos. *C66 THE INTERACTION BETWEEN GENES, PROTEINS AND ENVIRONMENT*: Am Thoracic Soc, 2013:A6102–A.
28 Sidore C, Busonero F, Maschio A, *et al*. Genome sequencing elucidates Sardinian genetic architecture and augments association analyses for lipid and blood inflammatory markers. *Nat Genet* 2015;47:1272–81.
29 Igartua C, Myers RA, Mathias RA, *et al*. Ethnic-specific associations of rare and low-frequency DNA sequence variants with asthma. *Nat Commun* 2015;6:5965.
30 Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 2016;538:161–4.