Human Molecular Genetics, 2012, Vol. 21, No. 6 ***1419–1432***
doi:10.1093/hmg/ddr570
Advance Access published on November 30, 2011

Replication of loci influencing ages at menarche and menopause in Hispanic women: the Women's Health Initiative SHARe Study

Christina T. L. Chen[1,*], Lindsay Fernández-Rhodes[2], Robert G. Brzyski[4], Christopher S. Carlson[1], Zhao Chen[5], Gerardo Heiss[2], Kari E. North[2,3], Nancy F. Woods[6], Aleksandar Rajkovic[7], Charles Kooperberg[1] and Nora Franceschini[2]

[1]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA, [2]Department of Epidemiology, Gillings School of Global Public Health and [3]Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC 27514, USA, [4]Department of Obstetrics and Gynecology, University of Texas Health Science Center, San Antonio, TX 78229, USA, [5]Division of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ 85724, USA, [6]Biobehavioral Nursing and Health Systems, University of Washington, Seattle, WA 98109, USA and [7]Department of Obstetrics, Gynecology and Reproductive Science, University of Pittsburgh, Pittsburgh, PA 15213, USA

Received July 29, 2011; Revised November 4, 2011; Accepted November 28, 2011

Several genome-wide studies have identified loci associated with reproductive traits, such as ages of menarche and menopause, in women of European ancestry. In this study, we investigated the relevance of these loci in minority US Hispanic women. We utilized data from 3468 women who were genotyped as a part of the Women's Health Initiative SNP Health Association Resource. We replicated associations of eight loci (*LRP18*, *LIN28B*, *CENPW*, *INHBA*, *TMEM38B*, *ZNF483*, *NFAT5* and *OLFM2*) with age at menarche, and of two loci (*MCM8* and *BRSK1/TMEM150B*) with age at menopause. The *MCM8* locus was also associated with early menopause risk. Three loci (*CENPW*, *MCM8* and *BRSK1/TMEM150B*) were associated with the length of reproductive lifespan. We provide evidence that genetic variants influencing reproductive traits identified in European populations are also important in minority US Hispanic women.

INTRODUCTION

Menarche and menopause are two milestone events in a woman's life, marking the beginning and the end of her reproductive years. Menarche occurs at a mean age of ~13 years in individuals of European ancestry, and ~2 years after the onset of puberty (1). Menopause is defined as the irreversible cessation of menses for at least 6 to 12 months and occurs on average at ~50–52 years of age (2,3). Ages at which these events occur are associated with various health outcomes. For instance, age at menarche is correlated with adult stature and obesity (4–7). It also influences the risk for several diseases in women, including breast cancer (8), Alzheimer's disease (9), stroke (10) and type 2 diabetes (11). Age at natural menopause has been shown to be a risk factor for increased morbidity and mortality from cardiovascular diseases and osteoporosis (12–16). Menopause before the age of 40 years is referred to as premature ovarian insufficiency. Women with premature ovarian insufficiency have a 50% higher mortality rate than those reporting menopause at age 50 years or older (12,17,18). To better assist women in managing their overall health, factors influencing the timing of these events need to be elucidated.

Both environmental and genetic factors influence the ages of menarche and menopause. Body fat composition in childhood and exposures to both endogenous and exogenous estrogen and

*To whom correspondence should be address at: 1100 Fairview Ave N., M4-B402, PO Box 19024, Seattle, WA 98109, USA. Tel: +1 2066672900; Fax: +1 2066677850; Email: ctchen@fhcrc.org

© *The Author 2011. Published by Oxford University Press. All rights reserved.*
For Permissions, please email: journals.permissions@oup.com

anti-androgenic agents affect the age at menarche timing (19). Prenatal factors influencing *in utero* conditions are hypothesized to cause epigenetic changes in the child and affect the offspring's age at menarche (20,21). Socio-economic status, alcohol consumption and smoking have been significantly associated with menopausal age and premature ovarian insufficiency (3,22,23). These epidemiologic findings are supported by research showing that smoking induces apoptosis of the oocytes, accelerating ovarian aging and inducing earlier menopause (24). In addition to these epidemiologic factors, high correlations in the timing of these two reproductive events in twin and mother–daughter pairs have also suggested that genetics may play an important role (2,17,25,26). In the case of age of menarche, the estimated heritability varies from 0.49 to 0.69 in studies of non-Hispanic women of European ancestry (25,26). Several genetic variants affecting pubertal timing have been described (hypogonadism and Kallmann syndrome, reviewed in 27). Mutations and sequence variants in genes, including *NOBOX*, *GDF9*, *BMP15* and *FOXL2*, have also been associated with premature ovarian insufficiency (28–32). Recent genome-wide association studies (GWAS) identified numerous single-nucleotide polymorphisms (SNPs) associated with ages at menarche and menopause. Forty-one loci have been identified for menarche in individuals of European ancestry. These include genes implicated in reproductive hormonal regulation, and loci previously identified for obesity and body weight regulation (33). Markers on chr20p12.3 (*MCM8*), chr19q13.42 (*BRSK1/TMEM150B*), chr5q35.2 (*UIMC1*/*HK3*), chr6p24.2 (*SYCP2L*) and chr13q34 have been found to affect the timing of menopause (34,35). SNPs in four of these five clusters (*MCM8*, *BRSK1/TMEM150B*, *UIMC1/HK3*, *SYCP2L*) have also been shown to be associated with early menopause, defined as menopause before 45 years of age (36).

To date, most genetic studies on menarche and menopause have focused on women of European ancestry. Epidemiologic studies, however, suggest that the timings of these two events differ by races and ethnicity. Hispanic Americans and African Americans are two minority groups in the USA and have received relatively little attention in GWAS. To extend our understanding of the genetic variants influencing ages of menarche and menopause in minority women, we genotyped African and Hispanic American women in the Women's Health Initiative SNP Health Association Resource (WHI-SHARe). A separate study on these reproductive timings among African American women in WHI-SHARe is ongoing within the Continental Origins and Genetic Epidemiology Network (COGENT) consortium. In this current study, we sought to replicate variants identified in European ancestry populations (index and proxy SNPs) in Hispanic women in WHI-SHARe after adjusting for known environmental factors. For loci where the index and proxy SNPs were not available, we performed a comprehensive examination of variation within the region to screen for secondary alleles. We also investigated whether the same genetic variants influencing age at natural menopause were also associated with early menopause. Finally, we examined whether these variants were also associated with normal variation in reproductive lifespan.

RESULTS

Description of the Hispanic samples for age of menarche and age of menopause

Enrollment characteristics of Hispanics in the WHI-SHARe cohort are described in Supplementary Material, Table S1. For age of menarche, 3468 of 3587 self-reported US Hispanic women with genotype and phenotype data (age of menarche categories between 9 years or less and 17 years or older) were included. The average age at first menstrual cycle was 12.6 years with 95% of women starting menses between 9.5 and 15.8 years. For the age of menopause analyses, 1560 Hispanic unrelated participants were included. We excluded individuals with the history of endometrial and cervical cancer, individuals who had their last menstrual period within the past 12 months and individuals with both ovaries removed before natural menopause. The average age at menopause was 49.6 years with 95% of women undergoing menopause between 38.5 and 60.8 years. There were 20% of women with early menopause, defined as having menopause before 46 years of age, and 25% of these women with early menopause had undergone menopause prior to 40 years of age. The average reproductive lifespan defined by the difference between age of menopause and age of menarche was 37 years (95% CI: 25.7–48.4 years).

Non-genetic predictors of ages of menarche and menopause

We used linear regression to assess the association between non-genetic factors and ages of menarche and menopause. A significant birth cohort effect has been observed in previous age of menarche studies (37,38). Birth year was not available in our data. Therefore, we used age at recruitment as a proxy of birth year as it was positively associated with age at menarche (P-value $< 4.7 \times 10^{-7}$). Body mass index (BMI) at childhood was not available and hence was not evaluated as a covariate for age at menarche. Smoking exposure at the time of menopause was significantly associated with age at menopause (P-value $< 3.7 \times 10^{-6}$) and therefore included as covariate in all analyses (Supplementary Material, Table S2). On average, participants who were current smokers underwent menopause 1.9 years (95% CI: 1.5–2.3 years) earlier than women who were not. This is consistent with prior published findings (39,40). No other lifestyle choices, such as alcohol, coffee and tea use, were found to be associated with age at menopause in our study. Neither anthropometric measurements, including BMI, weights, waist to hip circumference ratio, nor use of oral contraceptives was significantly associated with age at menopause in our Hispanic cohort. Percent European ancestry was shown to be significantly associated with age at menopause (P-value $< 0.00\,246$); age of menopause was delayed by 2.1 years for every 1% increase in European ancestry (95% CI: 0.8–3.5 years). Percent European ancestry, however, was not associated with age at menarche in our sample (P-value > 0.10).

Eight loci were found to influence age of menarche in Hispanic women

We hypothesized that the loci previously been associated with age of menarche in women of European ancestry also influenced age of menarche in US Hispanic women. We used linear regression with additive genetic models adjusted for age at recruitment, recruitment center and the estimated European ancestry to account for birth cohort, site effects and admixture, respectively. We had available genotypes on published SNPs (index SNPs) or proxies ($r^2 \geq 0.8$) for 31 of the 41 published loci. Five of these SNPs were nominally associated with age at menarche ($\alpha = 0.05$): rs1079866 on chromosome 7 near *INHBA* (inhibin beta A, *P*-value < 0.037), rs12686569 on chromosome 9 near *TMEM38B* (transmembrane protein 38B, $r^2 = 0.92$ with rs2090409 in HapMap CEU, *P*-value < 0.0024), rs369065 on chromosome 6 in *LIN28B* (lin-28 homolog B, $r^2 = 0.97$ with rs7759938 in HapMap CEU, *P*-value $< 8.9 \times 10^{-5}$), rs2152876 on chromosome 6 in *CENPW* (centromere protein W, $r^2 = 0.93$ with rs4565329 in HapMap CEU, *P*-value < 0.025) and rs7245579 on chromosome 19 in *OLFM2* (olfactomedin 2, $r^2 = 0.90$ with rs1862471 in HapMap CEU, *P*-value < 0.0029) (Table 1, Fig. 1). The index SNP was available for the *INHBA* locus (rs1079866) and the allele frequency of the *G* coded allele was similar to HapMap CEU samples (0.11 in WHI-SHARe Hispanics and 0.15 in HapMap CEU) as was the direction of the effect (6.7 weeks in WHI-SHARe Hispanics and 3.9 weeks in (33)).

For 36 loci where the index SNPs failed to replicate, or where proxy SNPs could not be identified, we examined all SNPs in the same linkage disequilibrium (LD) block to identify additional significant signals within the genomic regions. The LD block for each index SNP was defined as the largest window spanning the index SNPs from previous report, bounded by a pair of polymorphisms showing r^2 of at least 0.5 with the index SNPs. Using this definition, we found three significant associations after correcting for number of tests performed per block: rs16855058-T in *LRP1B* (low density lipoprotein receptor-related protein 1B, $r^2 < 0.01$ with rs12472911 in HapMap CEU, MAF = 0.08, *P*-value $< 5.9 \times 10^{-4}$), rs7027778-G in *ZNF483* (zinc finger protein 483, $r^2 = 0.35$ with rs10980926 in HapMap CEU, MAF = 0.27, *P*-value $< 7.5 \times 10^{-3}$) and rs7190665-*T* in *WWP2* (WW domain containing E3 ubiquitin protein ligase 2, $r^2 < 0.01$ with rs1364063 in *NFAT5* in HapMap CEU, MAF = 0.02, *P*-value $< 1.3 \times 10^{-3}$). In total, we replicated associations of five previously published SNPs with age of menarche and identified possible secondary association signals in three genomic regions where the European variants failed to replicate.

MCM8 and 19q13.42 loci are associated with age at menopause in Hispanics

We compiled a list of 24 SNPs, located in 14 loci, which were previously associated with age at menopause at *P*-values $< 10^{-6}$ (34,35). Assuming each index SNP represented an independent signal in each locus, we tested the associations between each index SNP (or all available proxy SNPs if the index SNP was not available) and age at menopause. Six of the 24 SNPs were genotyped directly in WHI-SHARe and proxies were available for nine SNPs. Two of these index SNPs associations with age at menopause were replicated (Table 2, Fig. 2). Both of these SNPs, rs16991615 (*P*-value $< 1.4 \times 10^{-6}$) and rs236114 (*P*-value < 0.0024), are located within the *MCM8* (minichromosome maintenance complex component 8) gene. One copy of the major allele, *G*, at rs16991615 reduces age at menopause by 2.3 years (95% CI: 1.3–3.2 years) while one copy of major allele, *C*, at rs236114 reduces age at menopause by 0.87 years (95% CI: 0.30–1.4 years). rs16991615 is not in high LD with rs236114 ($r^2 = 0.36$ in HapMap CEU, $r^2 = 0.29$ in WHI-SHARe Hispanics), but when we accounted for the effects of both SNPs on age at menopause using conditional analyses, the association with rs16991615 still remained significant (*P*-value $< 1.6 \times 10^{-4}$); whereas the association with rs236114 diminished (*P*-value < 0.57). Thus, these SNPs appear to represent a single association. Interestingly, the effect size of rs16991615 in our cohort was larger than the previously reported value of 1.07 (34,35). We examined the LD structure among SNPs surrounding these two SNPs in our Hispanic sample and HapMap CEU population (Fig. 2, Supplementary Material, Fig. S9) and did not identify any other SNP in LD with these two SNPs. We also found that rs16991615 is associated with an increased risk for early menopause (*P*-value < 0.0052, odds ratio = 1.23–3.56), whereas rs236114 is not (*P*-value < 0.14) (Table 3). These results suggest that *MCM8* influences age at menopause among Hispanic women.

None of the proxy SNPs located in the 19q13.42 locus replicated in WHI-SHARe Hispanic women and therefore, we investigated whether any SNP on the LD block in this locus showed evidence of association with age at menopause. We identified rs17782355, an intronic SNP, as significantly associated with menopause timing after applying Bonferoni correction for number of independent SNPs tested (*P*-value < 0.0064) (Table 2, Fig. 2). One copy of the major allele, G, at rs17782355 delays menopause by 1.4 years (95% CI: 0.39–2.35 years). This SNP is likely a secondary signal from this locus as it is not in LD ($r^2 < 0.2$) with any index SNPs in either WHI-SHARe Hispanics or HapMap CEU (Fig. 2, Supplementary Material, Fig. S10). These results provide additional evidence that this locus plays a role in determining age at menopause.

Since *MCM8* and the 19q13.42 locus both influence the timing of menopause, we investigated whether there was an interaction between the two loci by including both rs16991615 and rs17782355 in a single model. We found that there was no significant interaction between the SNPs and that the magnitudes of effect on age of menopause were similar to those models with only one SNP (rs16991615: -2.20 ± 0.91 years with *P*-value $< 2.4 \times 10^{-6}$, rs17782355: 1.31 ± 0.98 years with *P*-value $< 8.9 \times 10^{-3}$). Thus, we replicated 2 of the 14 loci previously associated with age of menopause in our Hispanic cohort.

Reproductive life span is influenced by three loci

We calculated the length of reproductive lifespan for WHI-SHARe participants by subtracting their age of menarche

Table 1. Association of significant loci from previous GWAS with age at menarche in WHI-SHARe Hispanic women

Nearest gene(s)[a]	Chr	SNP	Position[b]	r^{2c}	Freq	Allele	Beta (SE) (years)	*P*-value	SNP with lowest *P*-value on the LD block				
									Block[b]	*N*	SNP(allele)	Beta (SE) (years)	*P*-value
RXRG (33)	1	rs466639 (i)	163661506	–	0.08	T	−0.016 (0.067)	0.81	163644785,163673914	16	rs157861(G)	−0.13 (0.049)	9.0×10^{-3}
SEC16B (33)	1	rs527248 (p)	176142137	1.00	0.18	G	−0.029 (0.048)	0.55	176036720, 176180142	42	rs3886720(G)[d]	0.14 (0.052)	5.8×10^{-3}
TMEM18 (33)	2	rs2947411 (i)	604168	–	0.16	A	0.045 (0.052)	0.39	590575, 643874	17	rs2867105(C)	0.077 (0.053)	0.15
CCDC85A (33)	2	rs17268785 (i)	56445587	–	0.23	G	0.018 (0.045)	0.69	56360947, 56451813	18	rs7562381(G)	0.062 (0.041)	0.14
LRP1B (33)	2	rs3936045 (p)	141944229	0.85	0.23	T	−0.025 (0.046)	0.58	141941798, 142090977	42	rs16855058(T)[d]	0.24 (0.070)	5.9×10^{-4}‡‡
NR4A2 (33)	2	rs17188434 (i)	156805022	–	–	–	–	–	156664322,156866133	17	rs1356743(G)‡	−0.070 (0.040)	0.08
PLCL1 (33)	2	rs1533535 (p)	199330953	0.81	0.42	G	−0.045 (0.038)	0.24	199320235,199364688	13	rs10202112(G)	−0.078 (0.040)	0.05
KLHDC8B (33)	3	rs7617480 (i)	49185736	–	0.17	A	5.7×10^{-3} (0.051)	0.91	48696044, 49665203	25	rs9859473(G)	−0.062 (0.054)	0.25
RBM6 (33)	3	rs17657688 (p)	50008232	1.00	0.32	C	−0.019 (0.040)	0.63	49719894, 50217391	26	rs739983(G)‡	−0.080 (0.039)	0.04
VGLL3 (33)	3	rs1437049 (p)	86929468	0.93	0.49	C	0.016 (0.038)	0.67	86828258, 87068743	14	rs17027938(T)‡	0.13 (0.051)	0.01
IGSF11 (33)	3	rs6804394 (p)	119075038	1.00	0.47	A	0.062 (0.038)	0.10	118968093, 119207083	44	rs4277680(G)	0.070 (0.040)	0.08
EEFSEC (33)	3	rs2687729 (i)	129377916	–	0.28	G	0.027 (0.042)	0.52	129214363, 129605086	36	rs2811415(G)‡	0.085 (0.054)	0.11
TMEM108, NPHP3 (33)	3	rs6439371 (i)	134093442	–	–	–	–	–	134093442, 134098154	3	rs1869151(T)	0.034 (0.038)	0.37
ECE2 (33)	3	rs3914188 (i)	185492742	–	–	–	–	–	185477230, 185546485	12	rs1881973(G)‡	0.055 (0.038)	0.15
TRA2B, ETV5 (33)	3	rs16860328 (p)	187118379	1.00	0.30	C	−0.017 (0.042)	0.69	187113427, 187162117	8	rs6802503(C)	0.080 (0.058)	0.17
PHF15 (33)	5	rs13187289 (i)	133877076	–	–	–	–	–	133877076, 133886618	3	rs10040679(T)‡	−0.068 (0.059)	0.25
KDM3B (33)	5	rs757647 (i)	137735214	–	0.29	A	−0.068 (0.042)	0.11	137698925, 137800195	10	rs7706614(G)‡	−0.29 (0.11)	8.6×10^{-3}
PRDM13, MCHR2 (33)	6	rs10485227 (p)	100314048	1.00	0.42	C	-7.1×10^{-5} (0.038)	1.00	100301327, 100316628	6	rs4240580(G)	0.054 (0.042)	0.20
LIN28B(33,53)	6	rs369065 (p)	105550751	0.97	0.29	G	0.17 (0.042)	8.9×10^{-5}‡‡, #	–	–	–	–	–
CENPW, TRMT11 (33)	6	rs4565329 (p)	126794491	0.93	0.38	G	0.090 (0.040)	0.025‡‡	–	–	–	–	–
INHBA (33)	7	rs1079866 (i)	41436618	–	0.11	G	0.13 (0.061)	0.037‡‡	–	–	–	–	–
PEX2 (33)	8	rs4735765 (p)	78260360	1.00	0.40	T	−0.024 (0.039)	0.54	78256392, 78365393	17	rs1452820(T)	0.24 (0.19)	0.21
TMEM38B (33,53)	9	rs12686569 (p)	107956532	0.92	0.31	A	−0.12 (0.041)	2.4×10^{-3}‡‡	–	–	–	–	–
ZNF483 (33)	9	rs10980926 (i)	113333455	–	–	–	–	–	113327639,113375944	5	rs7027778(G)	−0.12 (0.044)	7.5×10^{-3}‡‡
TRIM66 (33)	11	rs4929923 (i)	8595776	–	0.43	T	0.026 (0.038)	0.50	8355423, 8651406	24	rs11245788(G)	0.081 (0.039)	0.04
ARNTL (33)	11	rs7109016 (p)	13250555	1.00	0.34	T	0.049 (0.040)	0.22	13205115, 13308981	27	rs11022738(G)	0.081 (0.042)	0.05
PHF21A (33)	11	rs16938437 (i)	46009151	–	0.08	T	−0.041 (0.069)	0.56	45908887, 46107195	17	rs7950474(C)[d]	0.078 (0.038)	0.04
GAB2 (33)	11	rs10899489 (i)	77773021	–	0.28	A	0.048 (0.043)	0.27	77586662, 77939972	25	rs34956708(T)	0.25 (0.098)	0.01
BSX (33)	11	rs6589964 (i)	122375893	–	0.58	A	−0.025 (0.039)	0.53	122373415,122383305	2	rs6589964(C)	0.025 (0.039)	0.53
C13orf16, ARHGEF7 (33)	13	rs9560105 (p)	110973364	1.00	0.28	A	0.012 (0.042)	0.77	110966878, 111042012	21	rs2774440(G)	0.057 (0.038)	0.13
BEGAIN (33)	14	rs6575793 (i)	100101970	–	–	–	–	–	100098571, 100123782	3	rs10873519(T)	0.059 (0.042)	0.16
RORA (33)	15	rs3743266 (i)	58568805	–	0.29	C	−0.026 (0.042)	0.54	58474968, 58593477	19	rs16942653(G)	0.37 (0.16)	0.02
IQCH (33)	15	rs6494654 (p)	65468961	1.00	0.38	C	−0.015 (0.040)	0.70	65263372, 65687937	34	rs6494673(G)[d]	−0.15 (0.065)	0.02
FTO (33)	16	rs9939609 (i)	52378028	–	0.32	A	0.072 (0.041)	0.08	52356024, 52402988	12	rs6499646(G)[d]	0.11 (0.053)	0.04
NFAT5 (33)	16	rs12599391 (p)	68162850	0.97	0.35	C	−0.019 (0.040)	0.62	68105242, 68488959	29	rs7190665(T)	0.49 (0.15)	1.3×10^{-3}‡‡
CA10 (33)	17	rs9635759 (i)	46968784	–	–	–	–	–	46964157, 46970544	3	rs17662433(T)[d]	0.17 (0.11)	0.13
SLC14A2 (33)	18	rs2243787 (p)	41211045	1.00	0.39	A	0.029 (0.040)	0.47	41185687, 41690219	137	rs1484878(G)[d]	−0.14 (0.045)	2.4×10^{-3}
FUSSEL18 (33)	18	rs1398217 (i)	43006236	–	–	–	–	–	42839418, 43047530	13	rs8086549(T)	0.078 (0.038)	0.04
OLFM2 (33)	19	rs7245579 (p)	9861876	0.90	0.50	C	−0.11 (0.038)	2.9×10^{-3}‡‡	–	–	–	–	–
CRTC1 (33)	19	rs10423674 (i)	18678903	–	–	–	–	–	18669915, 18695124	4	rs757349(G)[d]	0.060 (0.076)	0.43
PCSK2 (33)	20	rs852069 (i)	17070593	–	–	–	–	–	16914048, 17107076	44	rs852027(G)[d]	0.11 (0.041)	8.5×10^{-3}

No known LD for block proxies rs4277680, rs1452820, rs11245788, rs11022738, rs34956708, rs10873519, rs16942653 and rs7190665 and their index SNPs. All index SNPs can be found in Supplementary Material, Table S3.

i, index SNP; p, proxy SNP; SE, standard error; *N*, number of SNP on the LD block; chr, chromosome; freq, coding allele frequency.

[a]For complete gene names, please see Supplementary Material, Table S3.

[b]Positions are in NCBI36 coordinates.

[c]LD between the index and proxy SNPs in HapMap CEU population ($r^2 \geq 0.80$).

[d]$r^2 < 0.30$ between this SNP and index SNP.

‡‡*P*-value is significant at $\alpha = 0.05$ (for index and proxy SNPs) or $\alpha = 0.05/N$ (for SNPs on the LD block).

[#]*P*-value is significant at $\alpha = 0.05/31$, for 31 index/proxy SNPs tested.

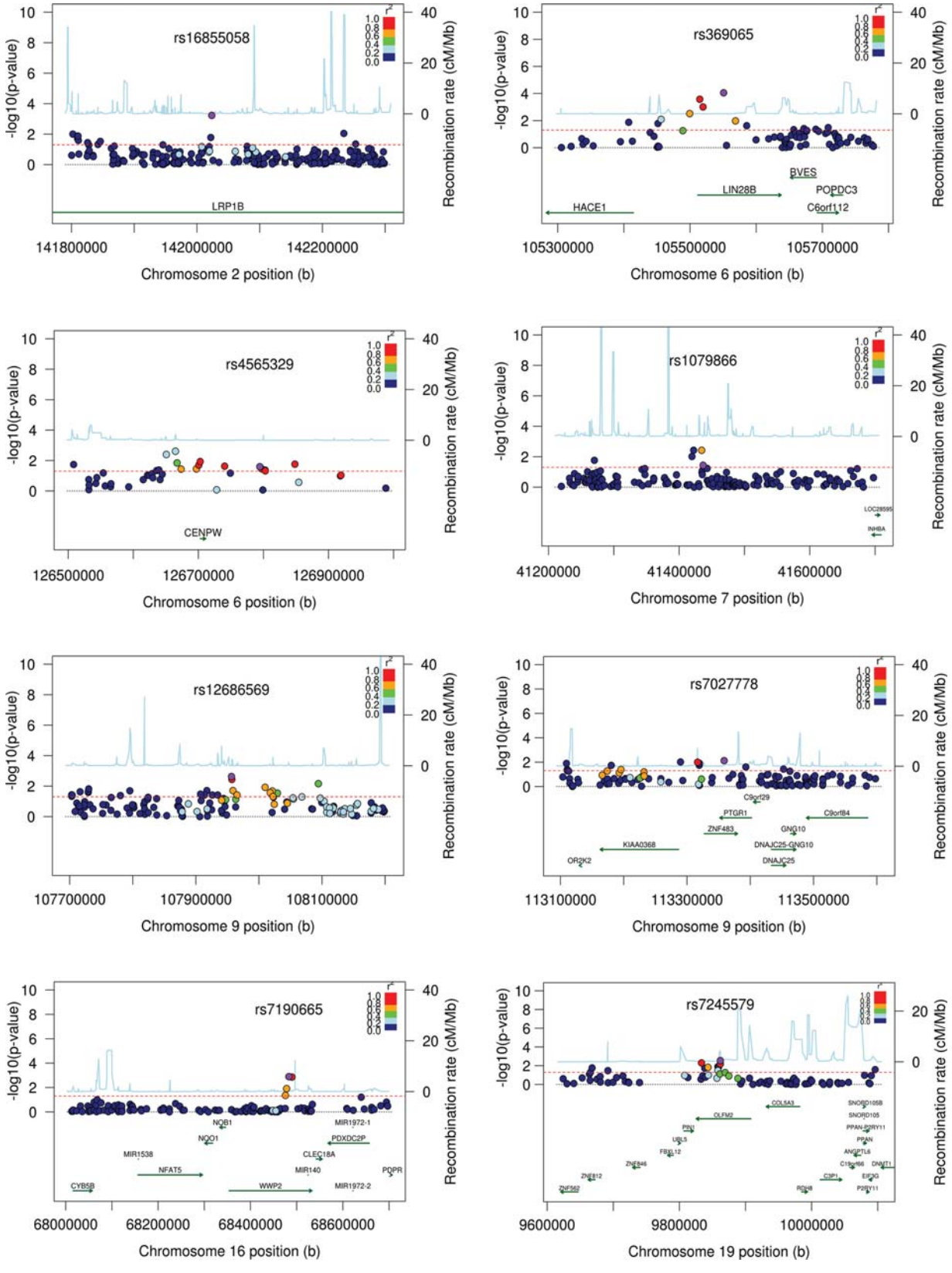

Figure 1. Regional plots for the eight regions identified to associate with age at menarche in WHI-SHARe Hispanics. Each panel shows the recombination rate in each region estimated from HapMap CEU data, pairwise LD between SNPs in the region and the SNP identified (labeled in purple) estimated from WHI-SHARe Hispanics data, *P*-values for strength of associations and genes in each region. The SNPs identified in each panel (labeled in purple) are either index/proxy SNPs or SNPs with the lowest *P*-value on the LD block. The r^2 values are color coded according to the scale on each panel.

Table 2. Association of significant loci from previous GWAS with age at menopause in WHI-SHARe Hispanic women

Nearest gene(s)[a]	Chr	SNP	Position[b]	r^{2c}	Freq	Allele	Beta (SE) (years)	*P*-value	SNP with lowest p-value on the LD block				
									Block[b]	*N*	SNP(allele)	Beta (SE) (years)	*P*-value
MMADHC (35)	2	rs11889862 (i)	150405394	–	–	–	–	–	150391583, 150407804	1	rs4132253(T)[d]	0.037 (0.26)	0.89
RPL22L1 (35)	3	rs4955755 (i)	171977103	–	–	–	–	–	171927988, 172009326	12	rs9860492(T)	−1.40 (1.06)	0.19
HK3 (34)	5	rs2278493 (i)	176247040	–	0.42	T	0.054 (0.21)	0.80	176247040, 176436169	16	rs3762971(G)[d]	−0.60 (0.34)	0.08
HK3 (34)	5	rs691141 (i)	176255904	–	0.64	G	−0.20 (0.22)	0.37					
UIMC1 (34)	5	rs601923 (p)	176349992	1.00	0.45	G	0.19 (0.21)	0.37					
SYCP2L (34)	6	rs2153157 (i)	11005474	–	0.55	G	−0.35 (0.21)	0.09	10954879, 11010318	21	rs7761983(T)	2.83 (1.01)	5.1×10^{-3}
SLC44A4 (35)	6	rs494620 (i)	31946692	–	–	–	–	–	31796497, 32001923	24	rs805293(T)	−0.49 (0.21)	0.02
PIK3CG (35)	7	rs17153527 (i)	106283045	–	–	–	–	–	106273479, 106379027	24	rs12705393(T)[d]	0.64 (0.25)	0.01
GPR124 (35)	8	rs6468442 (i)	37805907	–	–	–	–	–	37800584, 37813395	1	rs12676965(G)	−0.16 (0.31)	0.60
TLE4 (35)	9	rs1930259 (p)	81504229	0.89	0.80	G	0.27 (0.26)	0.29	81382178, 81537278	24	rs7870254(G)	1.89 (1.13)	0.10
SPATA19 (35)	11	rs4397868 (i)	133072195	–	0.82	T	−0.32 (0.32)	0.33	133067731, 133072195	5	rs713279(C)[d]	0.35 (0.23)	0.13
ARHGEF7 (34,35)	13	rs1163623 (p)	111019632	1.00	0.098	G	−0.60 (0.35)	0.09	111001762, 111038132	14	rs1756086(G)	−1.48 (0.81)	0.07
DYNC1H1 (35)	14	rs2253998 (p)	101566105	1.00	0.73	C	−0.15 (0.23)	0.51	101494082, 101633550	15	rs10142230(G)	2.07 (1.43)	0.15
BANP (35)	16	rs4843747 (i)	86548552	–	–	–	–	–	86548552, 86674516	23	rs9936008(G)	0.55 (0.23)	0.02
BRSK1 (34)	19	rs2607336 (p)	60492141	1.00	0.55	T	0.29 (0.21)	0.17	60492141, 60527897	7	rs17782355(G)[d]	1.37 (0.50)	6.4×10^{-3}‡‡
BRSK1 (34,35)	19	rs1551562 (i)	60506693	–	–	–	–	–					
BRSK1 (34,35)	19	rs4806660 (p)	60525272	0.96	0.36	G	−0.26 (0.22)	0.22					
TMEM150B (34)	19	rs734518 (p)	60524696	0.94	0.45	T	−0.026 (0.21)	0.90					
TMEM150B (34,35)	19	rs4806661 (p)	60516473	1.00	0.38	T	−0.23 (0.21)	0.28					
TMEM150B (34,35)	19	rs4806659 (p)	60516398	0.97	0.57	G	0.14 (0.21)	0.50					
MCM8 (35)	20	rs236114 (i)	5883385	–	0.16	T	0.87 (0.29)	2.4×10^{-3}‡‡,#	–	–	–	–	–
MCM8 (34)	20	rs16991615 (i)	5896227	–	0.95	G	−2.26 (0.47)	1.4×10^{-6}‡‡,#					

No known LD for block proxies rs9860492, rs7761983, rs7870254, rs1756086, rs10142230 and rs9936008 and their index SNPs. rs601923 was selected as a proxy for three *UIMC1* index SNPs: rs7718874 ($r^2 = 0.97$), rs365132 ($r^2 = 0.93$) and rs402511 ($r^2 = 1.00$). All index SNPs can be found in Supplementary Material, Table S4.

i, index SNP; p, proxy SNP; SE, standard error; *N*, number of SNP on the LD block; chr, chromosome; freq, coding allele frequency.

[a]For complete gene name, please see Supplementary Material, Table S4.

[b]Positions are in NCBI36 coordinates.

[c]LD between the index and proxy SNPs in HapMap CEU population ($r^2 \geq 0.80$).

[d]$r^2 < 0.30$ between this SNP and index SNP.

‡‡*P*-value is significant at $\alpha = 0.05$ (for index and proxy SNPs) or $\alpha = 0.05/N$ (for SNPs on the LD block).

#*P*-value is significant at $\alpha = 0.05/15$, for 15 index/proxy SNPs tested.

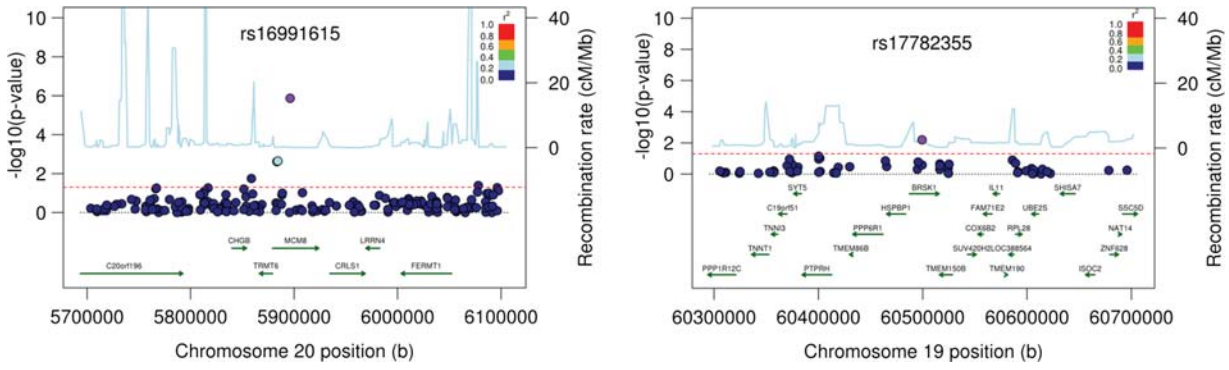

Figure 2. Regional plots for the two regions identified to associate with age at menopause in WHI-SHARe Hispanics. Each panel shows the recombination rate in each region estimated from HapMap CEU data, pairwise LD between SNPs in the region and the SNP identified (labeled in purple) estimated from WHI-SHARe Hispanics data, *P*-values for strength of associations and genes in each region. The SNPs identified in each panel (labeled in purple) are either index/proxy SNPs or SNPs with the lowest *P*-value on the LD block. The r^2 values are color coded according to the scale on each panel.

from their age of menopause and then assessed the association between reproductive lifespan and 65 index SNPs in 54 loci previously associated with either menarche or menopause. Both SNPs in *MCM8*, rs16991615 (P-value $< 8.9 \times 10^{-6}$) and rs236114 (P-value $< 4.7 \times 10^{-3}$), were found to influence the length of reproductive lifespan in Hispanic women (Table 4, Fig. 3). The *G* allele at rs16991615 reduces reproductive lifespan by 2.1 years per copy (95% CI: 1.2–3.1). The *C* allele at rs236114 reduces reproductive lifespan by 0.83 years per copy (95% CI: 0.26–1.4 years). We also found that rs17782355 in *BRSK1* (BR serine/threonine kinase 1) on chromosome 19 was also associated with this phenotype, increasing reproductive lifespan by 1.5 years per copy of the *G* allele (95% CI: 0.48–2.5, P-value $< 3.8 \times 10^{-3}$). Only one of the 41 loci previously associated with age of menarche was associated with the length of reproductive lifespan (P-value $< 2.3 \times 10^{-3}$, rs853964). The *T* allele at this locus reduces the reproductive lifespan by 2.4 years per copy in our population (95% CI: 0.84–3.9). rs853964 is likely a secondary signal in this region, as it is at low LD ($r^2 < 0.2$) with the index SNP (rs1361108) in both WHI-SHARe Hispanics and HapMap CEU populations (Fig. 3, Supplementary Material, Fig. S11). Two of the three loci found to be associated with length of reproductive span are previously associated with age at menopause, suggesting that the length of women's reproductive lifespan is more dependent on age at menopause than on when they have their first menses.

DISCUSSION

Hispanic Americans constitute 16.3% of the total population in the USA (41). Several epidemiologic studies show that the timing of menarche and menopause in Hispanic women differ from non-Hispanic Caucasian women even after adjusting for known environmental factors (42–45), but most GWAS have focused solely on women of European ancestry. Our study is one of the first to use genome-wide data to seek replication of findings from European populations in US Hispanics. Our study is limited by a modest sample size, particularly for menopause analyses, so we were not powered for GWAS discovery. Therefore, we explored replication to Hispanic women, and replicated 8 of 41 menarche loci using a sample of 3468 Hispanic women, and 2 of 14 menopause loci in 1560 Hispanic women (Supplementary Material, Fig. S14). We also replicated the risk of early menopause for one of the two loci already associated with age at menopause. In addition, we found three loci to influence the length of a woman's reproductive lifespan: one locus associated with age of menarche and two loci associated with age at menopause.

Of the eight loci we replicated (*LIN28B*, *LRP1B, NFAT5, INHBA, CENPW*, *TMEM38B*, *ZNF483* and *OLFM2*), two have experimental evidences to support their roles in affecting age at menarche. Transgenic mice expressing *LIN28A* show a delayed onset of puberty (46). The previously reported variant is located upstream of *LIN28B*, whereas the proxy variant rs369065 used in our analysis is an intronic *LIN28B* SNP that might have a direct role in affecting how this gene functions by, for example, changing the efficiency of how the protein folds. The signals we detected from *NFAT5* and *LRP1B* may be secondary signals, independent of the index SNP reported previously. The previously reported SNP, rs1364063, is upstream of *NFAT5* and its proxy, rs12599391, did not replicate in our samples. However, rs7190665 ($r^2 < 0.01$ with rs12599391 in WHI-SHARe Hispanics) located downstream of *NFAT5* did replicate (P-value < 0.001). rs7190665 is an intronic SNP located in the gene *WWP2*, a protein with ligase activity. The SNP with the lowest *P*-value in the *LRP1B* locus, rs16855058, is at low LD with both the index SNP (rs12472911) in HapMap CEU ($r^2 = 0.005$) and the proxy SNP (rs3936045) in WHI-SHARe Hispanics ($r^2 = 0.022$). All three SNPs are located in the intronic region in *LRP1B*. At *INHBA*, we replicated the previously reported SNP, rs1079866 (P-value < 0.037), located downstream of the gene. *INHBA* encodes the protein subunit inhibin beta A, which with subunit alpha, forms a hormone that is secreted by cells in the ovary. A nearby SNP, rs4141153, was also found to be significantly associated with age at menarche (P-value $<$ 0.0037). These two SNPs likely represent the same signal as rs1079866 is in strong LD with rs4141153 in both HapMap CEU ($r^2 = 0.57$) and WHI-SHARe Hispanics ($r^2 = 0.68$). We also identified that the *CENPW-T* locus not only was

Table 3. Association of significant loci from previous GWAS with early menopause in WHI-SHARe Hispanic women

Nearest gene(s)	Chr	SNP	Position[a]	r[2b]	Freq	Allele	Beta (SE) (years)	*P*-value	SNP with lowest p-value on the LD block				
									Block[a]	*N*	SNP(allele)	Beta (SE) (years)	*P*-value
MMADHC (35)	2	rs11889862 (i)	150405394	–	–	–	–	–	150391583, 150407804	1	rs4132253(T)[c]	-5.5×10^{-3} (0.12)	0.96
RPL22L1 (35)	3	rs4955755 (i)	171977103	–	–	–	–	–	171927988, 172009326	12	rs471976(T)	−0.095 (0.10)	0.36
HK3 (34)	5	rs227849 (i)	176247040	–	0.42	T	0.051 (0.094)	0.59	176247040, 176436169	16	rs251843(G)	−0.17 (0.10)	0.08
HK3 (34)	5	rs691141 (i)	176255904	–	0.64	G	0.18 (0.10)	0.080					
UIMC1 (34)	5	rs601923 (p)	176349992	1.00	0.45	G	−0.16 (0.095)	0.10					
SYCP2L (34)	6	rs2153157 (i)	11005474	–	0.55	G	0.013 (0.092)	0.89	10954879, 11010318	21	rs7761983(T)	−0.84 (0.62)	0.17
SLC44A4 (35)	6	rs494620 (i)	31946692	–	–	–	–	–	31796497, 32001923	24	rs707915(T)[c]	0.40 (0.17)	0.02
PIK3CG (35)	7	rs17153527 (i)	106283045	–	–	–	–	–	106273479, 106379027	24	rs849376(G)[c]	0.30 (0.13)	0.02
GPR124 (35)	8	rs6468442 (i)	37805907	–	–	–	–	–	37800584, 37813395	1	rs12676965(G)	0.025 (0.14)	0.86
TLE4 (35)	9	rs1930259 (p)	81504229	0.89	0.80	G	−0.17 (0.12)	0.14	81382178, 81537278	24	rs4877149(G)	0.33 (0.16)	0.03
SPATA19 (35)	11	rs4397868 (i)	133072195	–	0.82	T	0.27 (0.16)	0.084	133067731, 133072195	5	rs10894727(T)[c]	−0.20 (0.10)	0.05
ARHGEF7 (34,35)	13	rs1163623 (p)	111019632	1.00	0.098	G	0.14 (0.16)	0.36	111001762, 111038132	14	rs4771759(G)[c]	0.12 (0.099)	0.22
DYNC1H1 (35)	14	rs2253998 (p)	101566105	1.00	0.73	C	0.025 (0.11)	0.82	101494082, 101633550	15	rs11621560(C)[c]	0.16 (0.10)	0.11
BANP (35)	16	rs4843747 (i)	86548552	–	–	–	–	–	86548552, 86674516	23	Rs9936008(G)	−0.30 (0.10)	3.5×10^{-3}
BRSK1 (34)	19	rs2607336 (p)	60495632	1.00	0.55	T	−0.089 (0.094)	0.34	60492141, 60527897	7	rs17782355(G)[c]	−0.40 (0.21)	0.06
BRSK1 (34,35)	19	rs1551562 (i)	60506693	–	–	–	–	–					
BRSK1 (34,35)	19	rs4806660 (p)	60525272	0.96	0.36	G	0.053 (0.097)	0.58					
TMEM150B (34)	19	rs734518 (p)	60524696	0.94	0.45	T	0.028 (0.092)	0.76					
TMEM150B (34,35)	19	rs4806661 (p)	60516473	1.00	0.38	T	0.037 (0.095)	0.70					
TMEM150B (34,35)	19	rs4806659 (p)	60516398	0.97	0.57	G	−0.085 (0.093)	0.36					
MCM8 (35)	20	rs236114 (i)	5883385	–	0.16	T	−0.20 (0.14)	0.14	–	–	–	–	–
MCM8 (34)	20	rs16991615 (i)	5896227	–	0.95	G	0.74 (0.27)	5.2×10^{-3}‡‡					

No known LD for block proxies and their index SNPs: rs251843 (with rs227849), rs7761983 and rs9936008. rs601923 was selected as a proxy for three *UIMC1* index SNPs: rs7718874 ($r^2 = 0.97$), rs365132 ($r^2 = 0.93$) and rs402511 ($r^2 = 1.00$). All index SNPs can be found in Supplementary Material, Table S4.

i, index SNP; p, proxy SNP; SE, standard error; *N*, number of SNP on the LD block; chr, chromosome; freq, coding allele frequency.

[a]Positions are in NCBI36 coordinates.

[b]LD between the index and proxy SNPs in HapMap CEU population ($r^2 \geq 0.80$).

[c]$r^2 < 0.30$ between this SNP and index SNP.

‡‡*P*-value is significant at $\alpha = 0.05$ (for index and proxy SNPs) or $\alpha = 0.05/N$ (for SNPs on the LD block).

Table 4. Association of significant loci from previous GWAS with reproductive lifespan in WHI-SHARe Hispanic women

Nearest gene(s)[a]	Chr	SNP	Position[b]	r^{2c}	Freq	Allele	Beta (SE) (years)	*P*-value	SNP with lowest p-value on the LD block				
									Block[b]	*N*	SNP(allele)	Beta (SE) (years)	*P*-value
RXRG (33)	1	rs466639 (i)	163661506	–	0.92	G	0.43 (0.37)	0.25	163644785,163673914	16	rs157861(G)	−0.52 (0.32)	0.10
SEC16B (33)	1	rs527248 (p)	176142137	1.00	0.18	G	0.24 (0.27)	0.37	176036720, 176180142	42	rs11677045(T)	−0.42 (0.26)	0.11
TMEM18 (33)	2	rs2947411 (i)	604168	–	0.84	G	0.16 (0.28)	0.57	590575, 643874	17	rs13431365(T)	−3.20 (1.28)	0.01
CCDC85A (33)	2	rs17268785 (i)	56445587	–	0.23	G	−0.16 (0.25)	0.54	56360947, 56451813	18	rs10180182(G)[d]	0.73 (0.28)	8.2×10^{-3}
LRP1B (33)	2	rs12618990 (p)	141944795	0.85	0.23	C	0.42 (0.26)	0.10	141941798, 142090977	42	rs11677045(T)	−0.42 (0.26)	0.11
MMADHC (35)	2	rs11889862 (i)	150405394	–	–	–	–	–	150391583, 150407804	1	rs4132253(T)[d]	0.14 (0.27)	0.60
NR4A2 (33)	2	rs17188434 (i)	156805022	–	–	–	–	–	156664322,156866133	17	rs1356743(G)[d]	−0.44 (0.22)	0.05
PLCL1 (33)	2	rs1533535 (p)	199330953	0.81	0.42	G	−0.067(0.22)	0.76	199320235,199364688	13	rs10200777(G)	1.69 (0.93)	0.07
KLHDC8B (33)	3	rs7617480 (i)	49185736	–	0.17	T	0.22 (0.28)	0.43	48696044, 49665203	25	rs3870340(G)	2.41 (0.92)	8.7×10^{-3}
RBM6 (33)	3	rs17657688 (p)	50008232	1.00	0.68	T	0.15 (0.22)	0.49	49719894, 50217391	26	rs7637266(G)	−5.65 (1.83)	2.0×10^{-3}
VGLL3 (33)	3	rs7628864 (p)	86933308	0.93	0.49	T	−0.33 (0.21)	0.13	86828258, 87068743	14	rs11914380(G)[d]	−0.58 (0.32)	0.07
IGSF11 (33)	3	rs6804394 (p)	119075038	1.00	0.53	C	0.26 (0.21)	0.21	118968093,119207083	44	rs6438414(T)[d]	0.63 (0.23)	5.5×10^{-3}
EEFSEC (33)	3	rs2687729 (i)	129377916	–	0.28	G	−0.13 (0.24)	0.59	129214363,129605086	36	rs7610072(T)	1.88 (1.27)	0.14
TMEM108,NPHP3 (33)	3	rs6439371 (i)	134093442	–	–	–	–	–	134093442, 134098154	3	rs1452142(T)	0.26 (0.21)	0.21
RPL22L1 (35)	3	rs4955755 (i)	171977103	–	–	–	–	–	171927988, 172009326	12	rs9860492(T)	−1.44 (1.08)	0.18
ECE2 (33)	3	rs3914188 (i)	185492742	–	–	–	–	–	185477230, 185546485	12	rs865809(T)	0.70 (0.27)	0.01
TRA2B, ETV5 (33)	3	rs16860328 (p)	187118379	1.00	0.70	T	0.25 (0.23)	0.27	187113427, 187162117	8	rs6802503(C)	−0.52 (0.32)	0.10
PHF15 (33)	5	rs13187289 (i)	133877076	–	–	–	–	–	133877076, 133886618	3	rs10040679(T)[d]	−0.61 (0.32)	0.06
KDM3B (33)	5	rs757647 (i)	137735214	–	0.71	G	−0.25 (0.23)	0.28	137698925, 137800195	10	rs6865472(G)	−3.32 (1.95)	0.09
HK3 (34)	5	rs227849 (i)	176247040	–	0.42	T	0.077 (0.21)	0.72	176247040, 176436169	16	rs3762971(G)[d]	−0.49 (0.35)	0.16
HK3 (34)	5	rs691141 (i)	176255904	–	0.64	G	−0.17 (0.22)	0.44					
UIMC1 (34)	5	rs601923 (p)	176349992	1.00	0.45	G	0.20 (0.21)	0.36					
SYCP2L (34)	6	rs2153157 (i)	11005474	–	0.55	G	−0.31 (0.21)	0.14	10954879, 11010318	21	rs7761983(T)	2.98 (1.03)	3.9×10^{-3}
SLC44A4 (35)	6	rs494620 (i)	31946692	–	–	–	–	–	31796497, 32001923	24	rs805293(T)	−0.43 (0.21)	0.04
PRDM13,MCHR2 (33)	6	rs10485227 (p)	100314048	1.00	0.58	T	−0.0034(0.21)	0.99	100301327, 100316628	6	rs9495441(T)[d]	1.08 (0.43)	0.01
LIN28B (33,53)	6	rs369065 (p)	105550751	0.97	0.29	G	−0.12 (0.23)	0.62	105471114,105569569	7	rs6905606(G)	−1.71 (1.39)	0.22
CENPW, TRMT11 (33)	6	rs2152876 (p)	126802921	0.93	0.37	G	0.11 (0.22)	0.64	126667310, 127122393	20	rs853964(T)[d]	−2.35 (0.77)	2.3×10^{-3}‡‡
INHBA (33)	7	rs1079866 (i)	41436618	–	0.11	G	−0.12 (0.34)	0.72	41434471, 41704471	72	rs6958234(G)	−2.24 (1.04)	0.03
PIK3CG (35)	7	rs17153527 (i)	106283045	–	–	–	–	–	106273479, 106379027	24	rs12705393(T)[d]	0.060 (0.25)	0.02
GPR124 (35)	8	rs6468442 (i)	37805907	–	–	–	–	–	378000584, 37813395	1	rs12676965(G)	−0.12 (0.31)	0.69
PEX2 (33)	8	rs4735765 (p)	78260360	1.00	0.40	T	0.042 (0.21)	0.85	78256392, 78365393	17	rs2310747(G)	−0.40 (0.24)	0.09
TLE4 (35)	9	rs1930259 (p)	81504229	0.89	0.80	G	0.29 (0.27)	0.28	81382178, 81537278	24	rs7870254(G)	2.27 (1.15)	0.05
TMEM38B (33,53)	9	rs12686569 (p)	107956532	0.92	0.69	C	−0.057(0.23)	0.80	107947088, 108111896	20	rs7030412(G)[d]	1.07 (0.87)	0.22
ZNF483 (33)	9	rs10980926 (i)	113333455	–	–	–	–	–	113327639, 113375944	5	rs7027778(G)	−0.28 (0.24)	0.24
TRIM66 (33)	11	rs4929923 (i)	8595776	–	0.57	G	−0.24 (0.21)	0.25	8355423, 8651406	24	rs10743082(G)	0.51 (0.21)	0.02
ARNTL (33)	11	rs7109016 (p)	13250555	1.00	0.34	T	0.34 (0.23)	0.13	13205115, 13308981	27	rs11823291(G)	3.60 (1.53)	0.02
PHF21A (33)	11	rs16938437 (i)	46009151	–	0.82	T	−0.12 (0.37)	0.75	45908887, 46107195	17	rs11038752(G)[d]	0.50 (0.36)	0.16
GAB2 (33)	11	rs10899489 (i)	77773021	–	0.28	T	−0.11 (0.25)	0.65	77586662, 77939972	25	rs7101517(G)	−1.58 (0.85)	0.06
BSX (33)	11	rs6589964 (i)	122375893	–	0.42	C	−0.19 (0.22)	0.37	122373415, 122383305	2	rs6589964(C)	−0.19 (0.22)	0.37
SPATA19 (35)	11	rs4397868 (i)	133072195	–	0.82	T	−0.21 (0.33)	0.53	133067731, 133072195	5	rs713279(C)[d]	0.35 (0.23)	0.14
C13orf16,ARHGEF7 (33)	13	rs9560105 (p)	110973364	1.00	0.72	T	0.36 (0.23)	0.11	110966878, 111042012	21	rs7324250(T)[d]	−0.65 (0.23)	5.2×10^{-3}
ARHGEF7 (34,35)	13	rs1163623 (p)	111019632	1.00	0.098	G	−0.57 (0.36)	0.12					
BEGAIN (33)	14	rs6575793 (i)	100101970	–	–	–	–	–	100098571, 100123782	3	rs10873519(T)	0.27 (0.24)	0.26
DYNC1H1 (35)	14	rs2253998 (p)	101566105	1.00	0.73	C	−0.035 (0.24)	0.88	101494082, 101633550	15	rs12161908(G)	1.12 (0.70)	0.11
RORA (33)	15	rs3743266 (i)	58568805	–	0.29	G	0.27 (0.23)	0.23	58474968, 58593477	19	rs7183916(G)[d]	−0.62 (0.27)	0.02
IQCH (33)	15	rs6494654 (p)	65468961	1.00	0.62	T	−0.11 (0.22)	0.62	65263372, 65687937	34	rs3743350(T)	6.10 (2.23)	6.2×10^{-3}
FTO (33)	16	rs9939609 (i)	52378028	–	0.32	T	−0.13 (0.22)	0.56	52356024, 42402988	12	rs16952525(T)	2.85 (1.06)	7.0×10^{-3}
NFAT5 (33)	16	rs4783722 (p)	68139413	0.97	0.35	G	0.040(0.22)	0.86	68105242, 68488959	29	rs4608327(G)[d]	0.50 (0.30)	0.10
BANP (35)	16	rs4843747 (i)	86548552	–	–	–	–	–	86548552, 86674516	23	rs9936008(G)	0.58 (0.23)	0.01
CA10 (33)	17	rs9635759 (i)	46968784	–	–	–	–	–	46964157, 46970544	3	rs4794258(T)[d]	−0.29 (0.22)	0.20
SLC14A2 (33)	18	rs2243787 (p)	41211045	1.00	0.61	C	−0.11 (0.23)	0.62	41185687, 41690219	137	rs16978354(G)[d]	−1.85 (0.57)	1.3×10^{-3}
FUSSEL18 (33)	18	rs1398217 (i)	43006236	–	–	–	–	–	42839418, 43047530	13	rs17784795(G)[d]	−0.69 (0.34)	0.04
OLFM2 (33)	19	rs7245579 (p)	9861876	0.90	0.50	G	0.094 (0.21)	0.66	9820014, 9879802	10	rs4239546(G)	−0.30 (0.23)	0.18
CRTC1 (33)	19	rs10423674 (i)	18678903	–	–	–	–	–	18669915, 18695124	4	rs12462498(G)	−0.46 (0.23)	0.05

Continued

Table 4. *Continued*

Nearest gene(s)[a]	Chr	SNP	Position[b]	r^{2c}	Freq	Allele	Beta (SE) (years)	*P*-value	SNP with lowest p-value on the LD block				
									Block[b]	*N*	SNP(allele)	Beta (SE) (years)	*P*-value
BRSK1 (34)	19	rs2607336 (p)	60495632	1.00	0.55	T	0.38 (0.21)	0.08	60492141, 60527897	7	rs17782355(G)[d]	1.48 (0.51)	3.8×10^{-3} ‡‡
BRSK1 (34,35)	19	rs1551562 (i)	60506693	–	–	–	–	–					
BRSK1 (34,35)	19	rs4806660 (p)	60525272	0.96	0.36	G	−0.35 (0.22)	0.11					
TMEM150B (34)	19	rs734518 (p)	60524696	0.94	0.45	T	−0.095 (0.21)	0.65					
TMEM150B (34,35)	19	rs4806661 (p)	60516473	1.00	0.38	T	−0.29 (0.22)	0.17					
TMEM150B (34,35)	19	rs4806659 (p)	60516398	0.97	0.57	G	0.23 (0.21)	0.29					
MCM8 (35)	20	rs236114 (i)	5883385	–	0.16	T	0.83 (0.29)	4.7×10^{-3} ‡‡	–	–	–	–	–
MCM8 (34)	20	rs16991615 (i)	5896227	–	0.95	G	−2.13 (0.48)	8.9×10^{-6} ‡‡,#					
PCSK2 (33)	20	rs852069 (i)	17070593	–	–	–	–	–	16914048, 17107076	44	rs7264693(G)	3.04 (1.27)	0.02

No known LD for block proxies and their index SNPs: rs11677045, rs13431365, rs10200777, rs3870340, rs763726, rs7610072, rs9860492, rs6865472, rs7761983, rs6905606, rs6958234, rs7870254, rs11823291, rs7101517, rs12161908, rs3743350, rs16952525, rs9936008, rs10423674 and rs7264693. rs601923 was selected as a proxy for three *UIMC1* index SNPs: rs7718874 ($r^2 = 0.97$), rs365132 ($r^2 = 0.93$) and rs402511 ($r^2 = 1.00$). rs7324250 was selected as a proxy SNP in the LD block for rs955810 ($r^2 = 0.48$) and rs7333181 ($r^2 < 0.01$). All index SNPs can be found in Supplementary Material, Tables S3 and S4.
i, index SNP; p, proxy SNP; SE, standard error; *N*, number of SNP on the LD block; chr, chromosome; freq, coding allele frequency.
[a]For complete gene names, please see Supplementary Material, Tables S3 and S4.
[b]Positions are in NCBI36 coordinates.
[c]LD between the index and proxy SNPs in HapMap CEU population ($r^2 \geq 0.80$).
[d]$r^2 < 0.30$ between this SNP and index SNP.
‡‡*P*-value is significant at $\alpha = 0.05$ (for index and proxy SNPs) or $\alpha = 0.05/N$ (for SNPs on the LD block).
#*P*-value is significant at $\alpha = 0.05/46$, for 46 index/proxy SNPs tested.

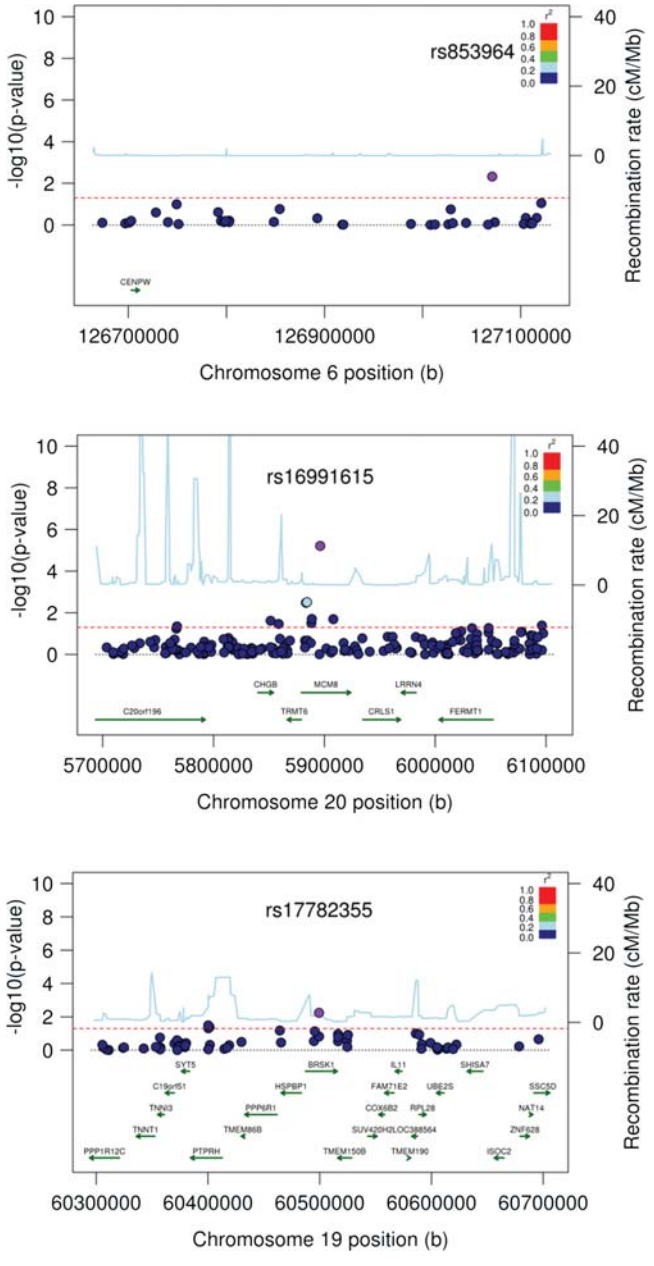

Figure 3. Regional plots for the three regions identified to associate with reproductive lifespan in WHI-SHARe Hispanics. Each panel shows the recombination rate in each region estimated from HapMap CEU data, pairwise LD between SNPs in the region and the SNP identified (labeled in purple) estimated from WHI-SHARe Hispanics data, *P*-values for strength of associations and genes in each region. The SNPs identified in each panel (labeled in purple) are either index/proxy SNPs or SNPs with the lowest *P*-value on the LD block. The r^2 values are color coded according to the scale on each panel.

associated with later menarche but also with a shorter reproductive lifespan in WHI-SHARe Hispanic women. Despite the reported association between BMI with age at menarche, none of the BMI-associated loci identified in women of European ancestry (such as *FTO*, *SEC16B*, *TRA2B*, *TMEM18*) was found to affect the age at menarche in our Hispanic samples. Our findings support the hypothesis that adiposity and Hispanic ethnicity are independently associated with earlier pubertal

development in girls (47); however, a much larger cohort of US Hispanic women is needed to study the interaction between these two factors more carefully.

BRSK1 and *MCM8* are ubiquitously expressed genes, expected to affect functions of multiple organs. In our WHI-SHARe Hispanic cohort, we successfully replicated the association between variants in these two loci and age at menopause previously published by multiple studies in European ancestry populations. rs17782355 in the *BRSK1* locus affects the age at menopause by 1.4 years on average per allele. As rs17782355 is not in high LD with any of the previously published variants in this locus, it is likely a secondary signal. We could not identify any other SNP in high LD with rs17782355 in the nearby genomic region in either WHI-SHARe Hispanics or HapMap CEU. *r*s16991615 in the *MCM8* locus decreases the age of menopause by 2.3 years per allele. rs16991615 is a non-synonymous SNP and is in low LD with rs236114 and rs236115 in both HapMap CEU ($r^2 = 0.36$) and WHI-SHARe Hispanics ($r^2 = 0.29$). rs16991615 and rs236114 have previously been identified by two different studies to be associated with age at menopause. We found that they likely represent the same signal with rs16991615 driving the effect in our sample based on our conditional analysis. rs16991615 also showed association with early menopause in Hispanic women, as has been shown in the Caucasian cohort (36). *MCM8* is thought to be involved in the initiation of eukaryotic genome replication, and likely critical in the control of cell proliferation. The association of *MCM8*, and other ubiquitously expressed genes, with early menopause, may indicate that early menopause is a marker for a more rapidly aging individual in general, rather than a specific marker for ovarian aging. This, in part, may explain the association of early menopause with adverse outcomes such as increased mortality and cardiovascular morbidity. Animal studies on these variants are lacking, and future studies are needed to ascertain the functional significance of these variants on ovarian aging and aging in general.

Disease-causing variants may have different effect sizes and frequencies in multiethnic populations with different admixture, making replication of findings from GWAS in European populations to non-European populations difficult. Collecting these large-scaled genetic data from non-European populations is therefore imperative; however, $<12\%$ of all GWAS done to-date have focused on non-European populations (48). One of the challenges in conducting genetic studies in non-European populations is the existence of complex population structure. Using genome-wide genotypes to measure relative amounts of European ancestry, we were able to adjust for population stratification using conventional methods. As more non-European GWAS data become available in the near future, we will be able to model the admixture structure in Hispanics more accurately and, thereby, hope to identify causal variants in the human genome for traits of interest.

MATERIALS AND METHODS

Study population

WHI is a long-term national health study that focuses on strategies for preventing common diseases such as heart disease, cancer and fracture in postmenopausal women. A total of 161 838 women aged 50–79 years old were recruited from 40 clinical centers in the USA between 1993 and 1998. WHI consists of an observational study, two clinical trials of postmenopausal hormone therapy (estrogen alone or estrogen plus progestin), a Calcium and Vitamin D Supplement Trial and a Dietary Modification Trial (49). Study recruitment and exclusion criteria have been described previously (50). Study protocols and consent forms were approved by the institutional review boards at all participating institutions. Medical history was updated annually (for women in the observational study) or semiannually (for women in the clinical trials) by mail and/or telephone questionnaires.

The WHI-SHARe includes 3642 self-identified Hispanic women from WHI who provided consent for DNA analysis. Approximately 1% of samples could not be genotyped. A total of 3587 participants were kept after excluding samples with call rates below 95%, in duplicates, or with genotype calls on the Y chromosome. The relatedness among participants was investigated and there were 97 sets of monozygotic twins, parent-offspring, full-siblings and half-siblings. After retaining only one individual from each set, our final sample size was 3493 Hispanic women.

Genotyping

DNA was extracted using specimens collected at time of enrollment. All samples, plus 2% blinded duplicates, were genotyped at Affymetrix Inc. (www.affymetrix.com) on the Genome-wide Human SNP Array 6.0 (909 622 SNPs). SNPs that were located on the Y chromosome or were Affymetrix quality control probes (not intended for analysis) were excluded ($n = 3280$). We also excluded SNPs that had call rates below 95% and concordance rates below 98%, leaving us 871 309 SNPs available for use in this study. The average concordance for blinded duplicate samples was 99.8%, and the average sample call rate after SNP exclusions was 99.8%.

Phenotype assessments

Self-reported age at menarche gathered from the WHI enrollment questionnaires, using the question ‘How old were you when you had your first menstrual period (menses)’. Possible responses were ‘9 or less’, ‘10’, ‘11’, ‘12’, ‘13’, ‘14’, ‘15’, ‘16’ and ‘17 or older’. For this current study, we included up to 3468 individuals for the age at menarche analysis with phenotype and genotyped data. Age at natural menopause is defined as the age when menstrual bleeding terminated naturally for at least 12 months. Questions on menopause status were asked on the enrollment questionnaires. The related questions were: ‘How old were you when you last had any menstrual bleeding? (If you are still having menstrual bleeding or periods, enter your current age.)’, ‘When was the last time you had any menstrual bleeding or spotting?’, ‘Did you ever have a hysterectomy?’ and ‘Did you ever have an operation to have one or both of your ovaries taken out?’. Eligible participants were defined as those who had at least one ovary intact before the cease of menses and who reported having last menstrual bleeding or spotting over 12 months ago at the time of the study. After removing participants with

endometrial and cervical cancer history, there were 1560 Hispanic participants. Of these 1560 participants, women who underwent menopause before they were 46 years old were defined as having early menopause. We defined the length of reproductive lifespan for each woman as the difference between her ages at menopause and at menarche. The characteristics of these participants are listed in Supplementary Material, Table S1.

Analysis

To better characterize the ancestry of these Hispanic individuals in the context of four ancestral populations (European Caucasians, Africans, Native Americans and East Asians), EIGENSTRAT was used to perform principle component analyses (PCA) (51). Specifically, we obtained principal components using 178 101 SNP markers that were common between our samples and our reference panels comprising 475 publically available samples from the YRI population, the CEU population, the Human Genome Diversity Project (HGDP) East Asian population and the HGDP Native American populations. To complement the analysis by PCA, the relative contributions of all four ancestral groups to every admixed individual were estimated using *frappe* (52).

We used linear regression models to study the association of each SNP with ages at menarche and menopause and the length of reproductive span. All analyses assume additive genetic models and include *frappe* estimates (% European Caucasian, % African and % East Asian) to account for population stratification. For age of menarche models, age at recruitment (proxy of birth year) and the recruitment centers (*West*, *South*, *Northeast*, *Midwest*) were significantly predictors for the phenotype. Therefore, these two covariates were added to the regression models. For age of menopause models, we performed univariate analyses on environmental factors which were previously associated with age of menopause. Variables on lifestyle, including smoking-related variables, uses of alcohol, coffee and tea, and medical history, including BMI, ratio of waist–hip circumferences, number of pregnancies, cancer status and use of oral contraceptives were analyzed. Smoking status at the time of menopause (current/non-current smokers) was shown to be significantly correlated with age at menopause and, therefore, was included as a covariate in age of menopause models. We defined a woman with early menopause when she underwent menopause prior to 46 years of age. We used logistic regression models with smoking status as a covariate to assess the risk of a woman having early menopause. Reproductive lifespan was defined as the number of years between a woman's ages of menopause and menarche. We included age at recruitment, recruitment centers and smoking status as covariates in the regression models used to analyze reproductive lifespan.

We first performed a genome-wide discovery, but no statistically significant associations were detected; likely due to our small sample size. To replicate previous findings, we compiled a list of 41 SNPs previously associated with age of menarche (P-value $< 10^{-8}$) and 24 SNPs previously associated with age of menopause (P-value $< 10^{-5}$) (index SNPs) (33–35,53). A different threshold for selecting index SNPs for age of menopause was used due to the small sample size in the two published discovery studies, making it possible for true variants to have diminished P-values. For index SNPs that were not available on our genotyping platform, we performed association tests for available SNPs that were at high LD in HapMap CEU population with the index SNPs (proxy SNPs). HapMap CEU population was used since these index SNPs were previously identified in European Caucasian studies. We identified proxy SNPs using the Genome Variation Server (http://gvs.gs.washington.edu/GVS) and only selected SNPs with at least $r^2 \geq 0.8$ with the index SNPs. In cases where multiple proxy SNPs were available for a given index SNP, we selected the proxy SNP with the highest LD. In cases where there were more than one proxy SNPs with the highest LD, we performed association studies for all of them to confirm that they had similar P-values. For clarity, we included only one proxy SNP for each index SNP. Lists of index SNPs can be found in Supplementary Material, Tables S3 and S4. We assessed the associations between the phenotypes and index/proxy SNPs at 5% significance level. In the event that the previously identified signals were replicated in our sample, we compared the LD structure in WHI-SHARe Hispanics to that in HapMap CEU population to compose a list of potential causal variants. We also assess the significance of these index/proxy SNPs using a conservative Bonferroni correction based on the number of index/proxy SNPs tested. To account for the possibility that the previously reported SNPs were tagging the causal SNPs in the nearby region in people with European ancestry but not in US Hispanics, we defined a LD block for each index SNP using the HapMap CEU population and tested whether any SNP genotyped on that block were associated with our phenotypes. Each LD block was defined as the largest block within 500 kb of the index SNP on which the LD between the SNPs at both ends of the block and the index SNP have $r^2 \geq 0.5$. The LD relationships among the SNPs on the same LD block were investigated and only one SNP from high LD pairs ($r^2 \geq 0.8$) was kept in the analysis. For the LD block analysis, we applied a Bonferroni correction to the P-values of association tests based on the number of SNPs in the block. This is a conservative correction as these SNPs on the LD block may have low levels of LD among them. In the events that we observe an association signal, we examined whether it represents the prior signal or is a secondary signal by comparing the LD structure in the region in both HapMap CEU and WHI-SHARe Hispanics. LD structures in WHI-SHARe Hispanics were computed using the pLINK program (http://pngu.mgh.harvard.edu/~purcell/plink/) (54). Graphs depicting the relationships between SNPs on the same LD block were created using LocusZoom (55).

Power calculations were performed using Quanto (http://hydra.usc.edu/gxe, version 1.2.4) using the gene-only model with a continuous trait phenotype and unrelated individuals. For the age at menarche, we assumed a population mean of 13 years with a standard deviation of 1.6. For the age at menopause, we assumed a population mean of 50 years with a standard deviation of 5.7. Based on a conservative assumption that the loci included in our study had similar magnitude of effects on traits of interest in WHI-SHARe Hispanics as in women with European ancestry, we were adequately powered to detect 2 of 41 loci previously associated with age at

menarche and 3 of 24 SNPs previously associated with age at menopause (Supplementary Material, Fig. S14). This estimate was a lower bound as we expected the same signals to have bigger effects in our Hispanic samples since the timing of these traits in minorities were shown to deviate from women of European ancestry.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

This manuscript was prepared in collaboration with investigators of the WHI, and has been approved by the WHI. WHI investigators are listed at http://www.whiscience.org/publications/WHI_investigators_shortlist.pdf.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Ruth L. Kirschstein National Research Service Award (5T32HL007055 to L.F.-R.); and the American Heart Association (0675001N, 1R01HL089651-01, 1U01HG004803-01 to N.F.). The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32 and 44221.

REFERENCES

1. Sulem, P., Gudbjartsson, D.F., Rafnar, T., Holm, H., Olafsdottir, E.J., Olafsdottir, G.H., Jonsson, T., Alexandersen, P., Feenstra, B., Boyd, H.A. *et al.* (2009) Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. *Nat. Genet.*, **41**, 734–738.
2. Torgerson, D.J., Thomas, R.E. and Reid, D.M. (1997) Mothers and daughters menopausal ages: is there a link? *Eur. J. Obstet.* Gynecol. Reprod. Biol., **74**, 63–66.
3. Parazzini, F. (2007) Determinants of age at menopause in women attending menopause clinics in Italy. *Maturitas*, **56**, 280–287.
4. Biro, F.M., McMahon, R.P., Striegel-Moore, R., Crawford, P.B., Obarzanek, E., Morrison, J.A., Barton, B.A. and Falkner, F. (2001) Impact of timing of pubertal maturation on growth in black and white female adolescents: The National Heart, Lung, and Blood Institute Growth and Health Study. *J. Pediatr.*, **138**, 636–643.
5. Freedman, D.S., Khan, L.K., Serdula, M.K., Dietz, W.H., Srinivasan, S.R. and Berenson, G.S. (2002) Relation of age at menarche to race, time period, and anthropometric dimensions: the Bogalusa Heart Study. *Pediatrics*, **110**, e43.
6. Freedman, D.S., Khan, L.K., Serdula, M.K., Dietz, W.H., Srinivasan, S.R. and Berenson, G.S. (2003) The relation of menarcheal age to obesity in childhood and adulthood: the Bogalusa heart study. *BMC Pediatr.*, **3**, 3.
7. Dunger, D.B., Ahmed, M.L. and Ong, K.K. (2006) Early and late weight gain and the timing of puberty. *Mol. Cell Endocrinol.*, **254–255**, 140–145.
8. Rockhill, B., Moorman, P.G. and Newman, B. (1998) Age at menarche, time to regular cycling, and breast cancer (North Carolina, United States). *Cancer Causes Control*, **9**, 447–453.
9. Cui, R., Iso, H., Toyoshima, H., Date, C., Yamamoto, A., Kikuchi, S., Kondo, T., Watanabe, Y., Koizumi, A., Inaba, Y. *et al.* (2006) Relationships of age at menarche and menopause, and reproductive year with mortality from cardiovascular disease in Japanese postmenopausal women: the JACC study. *J. Epidemiol.*, **16**, 177–184.
10. Rees, M. (1995) The age of menarche. *ORGYN*, **4**, 2–4.
11. He, C., Zhang, C., Hunter, D.J., Hankinson, S.E., Buck Louis, G.M., Hediger, M.L. and Hu, F.B. (2010) Age at menarche and risk of type 2 diabetes: results from 2 large prospective cohort studies. *Am. J. Epidemiol.*, **171**, 334–344.
12. Cooper, G.S. and Sandler, D.P. (1998) Age at natural menopause and mortality. *Ann. Epidemiol.*, **8**, 229–235.
13. Goswami, D. and Conway, G.S. (2005) Premature ovarian failure. *Hum. Reprod. Update*, **11**, 391–410.
14. Jacobsen, B.K., Nilssen, S., Heuch, I. and Kvale, G. (1997) Does age at natural menopause affect mortality from ischemic heart disease? *J. Clin. Epidemiol.*, **50**, 475–479.
15. Lindquist, O., Bengtsson, C., Hansson, T. and Roos, B. (1979) Age at menopause and its relation to osteoporosis. *Maturitas*, **1**, 175–181.
16. van Der Voort, D.J., van Der Weijer, P.H. and Barentsen, R. (2003) Early menopause: increased fracture risk at older age. *Osteoporos Int.*, **14**, 525–530.
17. Gosden, R.G., Treloar, S.A., Martin, N.G., Cherkas, L.F., Spector, T.D., Faddy, M.J. and Silber, S.J. (2007) Prevalence of premature ovarian failure in monozygotic and dizygotic twins. *Hum. Reprod.*, **22**, 610–615.
18. Coulam, C.B. (1982) Premature gonadal failure. *Fertil. Steril.*, **38**, 645–655.
19. Euling, S.Y., Selevan, S.G., Pescovitz, O.H. and Skakkebaek, N.E. (2008) Role of environmental factors in the timing of puberty. *Pediatrics*, **121**(Suppl. 3), S167–S171.
20. dos Santos Silva, I., De Stavola, B.L., Mann, V., Kuh, D., Hardy, R. and Wadsworth, M.E. (2002) Prenatal factors, childhood growth trajectories and age at menarche. *Int. J. Epidemiol.*, **31**, 405–412.
21. Gluckman, P.D. and Hanson, M.A. (2006) Evolution, development and timing of puberty. *Trends Endocrinol. Metab.*, **17**, 7–12.
22. Mikkelsen, T.F., Graff-Iversen, S., Sundby, J. and Bjertness, E. (2007) Early menopause, association with tobacco smoking, coffee consumption and other lifestyle factors: a cross-sectional study. *BMC Public Health*, **7**, 149.
23. Torgerson, D.J., Avenell, A., Russell, I.T. and Reid, D.M. (1994) Factors associated with onset of menopause in women aged 45–49. *Maturitas*, **19**, 83–92.
24. Matikainen, T., Perez, G.I., Jurisicova, A., Pru, J.K., Schlezinger, J.J., Ryu, H.Y., Laine, J., Sakai, T., Korsmeyer, S.J., Casper, R.F. *et al.* (2001) Aromatic hydrocarbon receptor-driven Bax gene expression is required for premature ovarian failure caused by biohazardous environmental chemicals. *Nat. Genet.*, **28**, 355–360.
25. Anderson, C.A., Zhu, G., Falchi, M., van den Berg, S.M., Treloar, S.A., Spector, T.D., Martin, N.G., Boomsma, D.I., Visscher, P.M. and Montgomery, G.W. (2008) A genome-wide linkage scan for age at menarche in three populations of European descent. *J. Clin. Endocrinol. Metab.*, **93**, 3965–3970.
26. Towne, B., Czerwinski, S.A., Demerath, E.W., Blangero, J., Roche, A.F. and Siervogel, R.M. (2005) Heritability of age at menarche in girls from the Fels Longitudinal Study. *Am. J. Phys. Anthropol.*, **128**, 210–219.
27. Gajdos, Z.K., Hirschhorn, J.N. and Palmert, M.R. (2009) What controls the timing of puberty? An update on progress from genetic investigation. *Curr. Opin. Endocrinol. Diabetes Obes.*, **16**, 16–24.
28. Aittomaki, K., Lucena, J.L., Pakarinen, P., Sistonen, P., Tapanainen, J., Gromoll, J., Kaskikari, R., Sankila, E.M., Lehvaslaiho, H., Engel, A.R. *et al.* (1995) Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. *Cell*, **82**, 959–968.
29. Laissue, P., Christin-Maitre, S., Touraine, P., Kuttenn, F., Ritvos, O., Aittomaki, K., Bourcigaux, N., Jacquesson, L., Bouchard, P., Frydman, R. *et al.* (2006) Mutations and sequence variants in GDF9 and BMP15 in patients with premature ovarian failure. *Eur. J. Endocrinol.*, **154**, 739–744.
30. Crisponi, L., Deiana, M., Loi, A., Chiappe, F., Uda, M., Amati, P., Bisceglia, L., Zelante, L., Nagaraja, R., Porcu, S. *et al.* (2001) The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat. Genet.*, **27**, 159–166.
31. Rizzolio, F., Bione, S., Sala, C., Goegan, M., Gentile, M., Gregato, G., Rossi, E., Pramparo, T., Zuffardi, O. and Toniolo, D. (2006) Chromosomal rearrangements in Xq and premature ovarian failure:

mapping of 25 new cases and review of the literature. *Hum. Reprod.*, **21**, 1477–1483.
32. Qin, Y., Choi, Y., Zhao, H., Simpson, J.L., Chen, Z.J. and Rajkovic, A. (2007) NOBOX homeobox mutation causes premature ovarian failure. *Am. J. Hum. Genet.*, **81**, 576–581.
33. Elks, C.E., Perry, J.R., Sulem, P., Chasman, D.I., Franceschini, N., He, C., Lunetta, K.L., Visser, J.A., Byrne, E.M., Cousminer, D.L. *et al.* (2010) Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat. Genet.*, **42**, 1077–1085.
34. He, C., Kraft, P., Chen, C., Buring, J.E., Pare, G., Hankinson, S.E., Chanock, S.J., Ridker, P.M., Hunter, D.J. and Chasman, D.I. (2009) Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat. Genet.*, **41**, 724–728.
35. Stolk, L., Zhai, G., van Meurs, J.B., Verbiest, M.M., Visser, J.A., Estrada, K., Rivadeneira, F., Williams, F.M., Cherkas, L., Deloukas, P. *et al.* (2009) Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat. Genet.*, **41**, 645–647.
36. Murray, A., Bennett, C.E., Perry, J.R., Weedon, M.N., Jacobs, P.A., Morris, D.H., Orr, N., Schoemaker, M.J., Jones, M., Ashworth, A. *et al.* (2010) Common genetic variants are significant risk factors for early menopause: results from the Breakthrough Generations Study. *Hum. Mol. Genet.*, **20**, 186–192.
37. Nichols, H.B., Trentham-Dietz, A., Hampton, J.M., Titus-Ernstoff, L., Egan, K.M., Willett, W.C. and Newcomb, P.A. (2006) From menarche to menopause: trends among US Women born from 1912 to 1969. *Am. J. Epidemiol.*, **164**, 1003–1011.
38. Malina, R.M., Pena Reyes, M.E., Tan, S.K. and Little, B.B. (2004) Secular change in age at menarche in rural Oaxaca, southern Mexico: 1968–2000. *Ann. Hum. Biol.*, **31**, 634–646.
39. Cramer, D.W., Harlow, B.L., Xu, H., Fraer, C. and Barbieri, R. (1995) Cross-sectional and case-controlled analyses of the association between smoking and early menopause. *Maturitas*, **22**, 79–87.
40. Kaufman, D.W., Slone, D., Rosenberg, L., Miettinen, O.S. and Shapiro, S. (1980) Cigarette smoking and age at natural menopause. *Am. J. Public Health*, **70**, 420–422.
41. Humes, K.R., Jones, N.A. and Ramirez, R.R. (2011) *Overview of Race and Hispanic Origin: 2010*. U.S. Census Bureau, http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf.
42. Frere, G. (1971) Mean age at menopause and menarche in South Africa. *South African J. Med. Sci.*, **36**, 21–24.
43. Kwawukume, E.Y., Ghosh, T.S. and Wilson, J.B. (1993) Menopausal age of Ghanaian women. *Int. J. Gynaecol. Obstet.*, **40**, 151–155.
44. Bromberger, J.T., Matthews, K.A., Kuller, L.H., Wing, R.R., Meilahn, E.N. and Plantinga, P. (1997) Prospective study of the determinants of age at menopause. *Am. J. Epidemiol.*, **145**, 124–133.
45. MacMahon, B. and Worcester, J. (1966) Age at menopause. United States—1960–1962. *Vital Health Stat.*, **11**, 1–20.
46. Zhu, H., Shah, S., Shyh-Chang, N., Shinoda, G., Einhorn, W.S., Viswanathan, S.R., Takeuchi, A., Grasemann, C., Rinn, J.L., Lopez, M.F. *et al.* (2010) Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. *Nat. Genet.*, **42**, 626–630.
47. Rosenfield, R.L., Lipton, R.B. and Drum, M.L. (2009) Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. *Pediatrics*, **123**, 84–88.
48. Rosenberg, N.A., Huang, L., Jewett, E.M., Szpiech, Z.A., Jankovic, I. and Boehnke, M. (2010) Genome-wide association studies in diverse populations. *Nat. Rev. Genet.*, **11**, 356–366.
49. Prentice, R., Anderson, G., Cummings, S., Freedman, L., Furberg, C., Henderson, M., Johnson, S., Kuller, L., Manson, J., Oberman, A. *et al.* (1998) Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin. Trials*, **19**, 61–109.
50. Hays, J., Hunt, J.R., Hubbell, F.A., Anderson, G.L., Limacher, M., Allen, C. and Rossouw, J.E. (2003) The Women's Health Initiative recruitment methods and results. *Ann. Epidemiol.*, **13**, S18–S77.
51. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
52. Tang, H., Peng, J., Wang, P. and Risch, N.J. (2005) Estimation of individual admixture: analytical and study design considerations. *Genet. Epidemiol.*, **28**, 289–301.
53. Perry, J.R., Stolk, L., Franceschini, N., Lunetta, K.L., Zhai, G., McArdle, P.F., Smith, A.V., Aspelund, T., Bandinelli, S., Boerwinkle, E. *et al.* (2009) Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. *Nat. Genet.*, **41**, 648–650.
54. Purcell, S., Cherny, S.S. and Sham, P.C. (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, **19**, 149–150.
55. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R. and Willer, C.J. (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26**, 2336–2337.