Human Reproduction, Vol.28, No.7 pp. 1987–1994, 2013
Advanced Access publication on April 16, 2013 doi:10.1093/humrep/det106

human reproduction

ORIGINAL ARTICLE ***Reproductive genetics***

Shared genetic factors for age at natural menopause in Iranian and European women

Maziar Rahmani[1,2,3,†], Madalene A. Earp[1,3,†], Fahimeh Ramezani Tehrani[4], Mehran Ataee[1], Jackson Wu[1], Martin Treml[1], Ramona Nudischer[1], Sara P-Behnami[4], ReproGen Consortium, John R.B. Perry[6,7,8,9], Joanne M. Murabito[10], Fereidoun Azizi[2,*], and Angela Brooks-Wilson[1,5]

[1]Canada's Michael Smith Genome Sciences Center, BC Cancer Agency, Vancouver, BC, Canada [2]Internal Medicine and Endocrinology, Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, PO Box 19195-4763, Tehran, Iran [3]Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada [4]Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran [5]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada [6]Genetics of Complex Traits, Peninsula Medical School, University of Exeter, Exeter, UK [7]Wellcome Trust Centre for Human Genetics, Roosevelt Drive, University of Oxford, Oxford, UK [8]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK [9]Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA [10]Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA

*Correspondence address. Tel: +98-21-240-9309; Fax: +98-21-240-2463; E-mail: azizi@endocrine.ac.ir

Submitted on September 1, 2012; resubmitted on March 11, 2013; accepted on March 18, 2013

STUDY QUESTION: Do differences in heritable genetic factors explain some of the difference in age at natural menopause (ANM) among populations?

SUMMARY ANSWER: One single nucleotide polymorphism (SNP)-ANM association (rs16991615) detected in European women was replicated in Iranian women.

WHAT IS KNOWN ALREADY: Genetics plays an important role in ANM, and well-powered genome-wide association studies (GWAS) of ANM performed in European women have discovered many statistically significant SNP-ANM associations. Average ANM varies by ethnicity, and population-specific differences in ANM-associated alleles may in part explain these differences.

STUDY DESIGN, SIZE, DURATION: After quality control procedures, 97 SNPs were analyzed in genotype data of 828 Iranian women who experienced natural menopause. SNP genotyping data were used to perform linear regression analyses with ANM as a quantitative trait. Study participants were drawn from the population-based Tehran Lipid and Glucose Study based in Tehran, Iran. This study was performed between February 2009 and March 2012.

PARTICIPANTS/MATERIALS, SETTING AND METHODS: Based on an ANM-GWAS literature review, eight SNPs at four loci previously associated with ANM in European women were tested for replication in Iranian women. Linear regression analyses were performed including ($n = 828$) and excluding ($n = 783$) women who experience premature ovarian failure (ANM before 40 years of age). In addition, to search for novel population-specific ANM risk alleles, a pool-based GWAS was performed using this collection of Iranian women. Two DNA pools were constructed and compared: an 'early' ANM pool (lower 20th percentile of menopause ages, 40–45 years, $n = 165$) and a 'late' ANM pool (upper 20th percentile of menopause ages, 54–65 years, $n = 187$). Each DNA pool was assayed on four *Illumina Human1M-Duo* arrays, and allele-based tests of association were used to rank SNPs. One hundred and two highly ranked SNPs were chosen for individual genotyping by Sequenom MassARRAY and association analysis in the Iranian women.

[†] These authors contributed equally to this work.

© The Author 2013. Published by Oxford University Press on behalf of the European Society of Human Reproduction and Embryology. All rights reserved.
For Permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

MAIN RESULTS AND THE ROLE OF CHANCE: One SNP-ANM association previously detected in European women was replicated in Iranian women (rs16991615; $\beta = 1.07$, standard error (SE): 0.49, $P = 0.02$). SNPs at the previously reported 19q13.42 and 6p24.2 loci also approached statistical significance and had consistent SNP effects (magnitude and direction) in Iranian women (rs1172822; $\beta = -0.39$, SE: 0.22, $P = 0.08$; and rs2153157, $\beta = 0.41$, SE: 0.21, $P = 0.05$). We found little evidence for novel SNP-ANM associations in Iranian women; no SNP selected based on the pool-based GWAS achieved genome-wide significance.

LIMITATIONS, REASONS FOR CAUTION: Due to small sample size this study was powered to reliably detect only moderate-to-large SNP effect sizes. This limited our ability to replicate many of the previously reported SNP-ANM risk alleles and to discover novel SNP-ANM associations' specific to the Iranian population. In performing our pool-based GWAS, a reduction in power was introduced relative to a conventional GWAS.

WIDER IMPLICATIONS OF THE FINDINGS: Our results imply that European and Iranian women share ANM-associated genetic variants. Our study was underpowered but for all SNPs tested the direction of the effect was consistent with data from the European study. Therefore, we anticipate that many (if not all) of the ANM-associated SNPs discovered in European women will replicate in Iranian women upon genotyping a sufficient number of women. Our data do not support the hypothesis that population-specific SNP-ANM associations explain population-specific differences in the mean ANM.

STUDY FUNDING/COMPETING INTEREST(S): The study was supported by the Research Institute for Endocrine Sciences. The authors have no conflicts of interest to declare.

Key words: age at menopause / premature ovarian failure / genome-wide association study / Iranian population / European population

Introduction

Ovarian aging is the natural process by which a woman reaches reproductive exhaustion. Menopause is an important event in this process, and a prevailing hypothesis states that it occurs when the follicle pool in the ovary is too low to maintain regular cycles (Faddy *et al.*, 1992; Richardson, 1993; te Velde *et al.*, 1998, Voorhuis *et al.*, 2010, 2011). Menopause is a dramatic event in a woman's life history and demarks major changes in endocrine signaling, particularly a reduction in female hormone production by the ovaries (Plagnol *et al.*, 2009). It is also a risk factor for many age-related diseases. For example, early menopause is associated with an increased risk of cardiovascular disease and osteoporosis (van der Schouw *et al.*, 1996; Gallagher, 2007; Shuster *et al.*, 2010), and late menopause is associated with an increased risk of ovarian cancer (Braem *et al.*, 2010), endometrial cancer (Dossus *et al.*, 2010) and breast cancer (Titus-Ernstoff *et al.*, 1998; Shin *et al.*, 2011). Furthermore, in populations where delaying childbearing has become prevalent, age at natural menopause (ANM) is increasingly becoming a fertility issue.

ANM averages 51 years in the Caucasian population, but is roughly normally distributed between 40 and 60 years with a tail <40 years. Menopause before the age of 40 years is termed premature ovarian failure (POF) (Coulam *et al.*, 1986; Daniels *et al.*, 1998). Family- and twin-based studies have estimated the heritable component of ANM to be between 42 and 87% (Burton *et al.*, 2007; Snieder *et al.*, 1998; de Bruin *et al.*, 2001; Murabito *et al.*, 2005; Morris *et al.*, 2011); thus, the prevailing view is that genetics plays a very important role in this trait. Environmental factors, for example smoking, can influence ANM; however, collectively environmental factors explain relatively little of the variation in ANM (van Noord *et al.*, 1997; Morris *et al.*, 2011). It is not clear whether POF represents the bottom end of ANM distribution, or is an independent trait with unique genetic and environmental risk factors.

In 2009, two genome-wide association studies (GWASs) of ANM were published, collectively reporting four loci harboring genome-wide significant single nucleotide polymorphism (SNP)-ANM associations on chromosomes 20, 19, 6 and 5 (He *et al.*, 2009; Stolk *et al.*, 2009). Both GWASs were performed on European women, and encouragingly two of the loci overlapped between studies (on chromosomes 20 and 19). In 2012, a meta-analysis of 22 GWASs performed on 38 968 European women replicated these four loci and reported 13 new genome-wide significant associations ($P < 5 \times 10^{-8}$). These studies confirmed the presence of heritable genetic factors in the European population but did not address their contribution in non-European women. We aimed to identify if the most significant of these European SNP-ANM associations also explain some of the ANM variation in a non-European population.

Evidence suggests ANM varies by race/ethnicity (Gold *et al.*, 2001; Henderson *et al.*, 2008). For example, a US-based cohort study of 92 704 women from five different racial/ethnic groups found that Latina women experience menopause earlier, and Japanese women experience menopause later, than non-Latina Whites (Henderson *et al.*, 2008). Adjustment for environmental factors, including smoking, age of menarche, parity and BMI, does not change this result. These differences may be explained by environmental factors not considered, genetic differences or a combination thereof. We hypothesize that differences in heritable genetic factors can explain some of this difference, and sought to characterize population-specific differences in known and novel ANM-associated alleles. To do this we studied an Iranian cohort drawn from the Tehran Lipid and Glucose Study (TLGS) (Azizi *et al.*, 2002; Butcher *et al.*, 2005). Based on the Human Diversity Panel (Paschou *et al.*, 2010), women of Iranian descent are described as 'Middle Eastern' and are expected to be more similar to Europeans than Latinas and Japanese; however, principal component analyses show that 'Middle Eastern' peoples form a node distinguishable from that of 'Europeans' (Paschou *et al.*, 2010).

This study posed two questions: (i) Are the genetic factors identified in the 2009 ANM-GWAS of European women also associated with ANM in Iranian women? (ii) Do novel population-specific ANM risk alleles conferring moderate-to-large effects exist in Iranian

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

women? To address question 1, SNPs identified by GWASs of ANM in European women were tested for association with ANM in women from the TLGS. To address question 2, a pool-based GWAS was performed. Individual samples were physically pooled to create composite samples, and these pools were assayed on commercially available SNP arrays. Data from SNP arrays were used to estimate allele frequency in DNA pools, not to determine genotypes, hence this step is called 'allelotyping'. Pool allelotypes were then used in allele-based tests of associations to discover SNP-ANM associations. As in conventional GWAS, most GWASs using a pooling strategy are multi-stage. DNA pooling is only used in the discovery stage; subsequent replication stages use individual genotyping (IG). Pearson *et al.* (2007) empirically demonstrated that a GWAS using the DNA pooling strategy is capable of detecting associations discovered by conventional GWAS, but for a fraction of the cost. Although pooling DNA makes for a financially feasible experiment, it does reduce power; hence, we were only well powered to detect moderate-to-large SNP effects.

Materials and Methods

Study subjects

Study subjects were drawn from the TLGS (Azizi *et al.*, 2002), and included Iranian women who had experienced natural menopause. TLGS is an ongoing longitudinal study of ~15 000 individuals aged 3 years and over drawn from a geographically defined area in Tehran, Iran. Participants entered the TLGS after providing written informed consent. Approval for this ANM study was received from the Research Institute for Endocrine Sciences Clinical Research Ethics Board and from the joint Clinical Research Ethics Board of the British Columbia Cancer Agency and the University of British Columbia. Each TLGS participant completed an interview, and most provided a blood sample. Participants were questioned on reproductive history, regularity of menstrual cycles, parity, sex steroid use and menopause status (Azizi *et al.*, 2002). Menopause was defined as the absence of spontaneous menstrual bleeding for >12 months, for which no pathological cause can be determined. Women whose last menstrual period may have been induced by surgery or another obvious cause, including irradiation or hormone therapy, were excluded, as were women who reported using hormone replacement therapy during the onset of menopause. For women who reached menopause prior to entering the TLGS, the date of the last cycle based on patient recall was recorded. Only women whose four grandparents were of Iranian descent (self-reported) were included. In total, 903 women met these criteria; 828 were successfully genotyped. Of the 828 women genotyped, 45 (5%) experienced POF (ANM before 40 years). In this cohort, ANM (excluding POF) ranged from 40 to 65 years and averaged 49.8 years (SD = 4.4). The median ANM was 50 years.

Analysis of eight ANM-associated SNPs discovered by the 2009 GWAS

For each of the four genome-wide significant loci reported in the 2009 ANM-GWAS (He *et al.*, 2009; Stolk *et al.*, 2009), two SNPs (a total of eight SNPs) were selected for genotyping, including: (i) the reported SNP with the smallest *P*-value and (ii) an SNP in high linkage disequilibrium (LD) with SNP 1 (in case of assay failure) (Table I). The LD between these SNPs is given in Table I. Linear regression was used to determine the effect of the minor allele for each SNP on ANM (including and excluding POF cases). Effect sizes (beta values) were calculated in years and are per copy of the minor allele. Adjusting for birth year, BMI, smoking history oral contraceptive use and parity did not influence our results; therefore, all data are presented without correction for these variables. Data were analyzed in PLINK (v1.07).

DNA pool construction and allelotyping

To search for novel Iranian population-specific ANM risk alleles conferring moderate-to-large effects, a pool-based GWAS was performed using this cohort of women. Two DNA pools were constructed and compared: an 'early' ANM pool (lower 20% percentile of menopause ages, 40–45

Table I SNP-ANM association of eight SNPs identified in European women in Iranian women from the TLGS.

SNP information			2009 ANM-GWAS[a]			Iranian women from TLGS[b] (excluding POF, ANM ≥40)			Iranian women from TLGS[c] (including POF)	
SNP	Chr	r^2	MAF	Per allele effect (SE)	*P*-value	Tested allele frequency	Per allele effect (SE)	*P*-value	Per allele effect (SE)	*P*-value
rs7718874	5	1	G, 0.49	0.39 (0.052)	1.3E−13	0.52	0.30 (0.22)	0.17	0.48 (0.26)	0.07
rs365132	5		T, 0.49	0.39 (0.052)	8.40E−14	NA	NA	NA	NA	NA
rs2153159	6	0.87	G, NA	NA	NA	0.49	−0.20 (0.21)	0.35	−0.44 (0.26)	0.09
rs2153157	6		T, 0.49	0.29 (0.052)	5.1E−08	0.49	0.41 (0.21)	0.06	0.51 (0.26)	**0.05**
rs1172822	19	0.7	T, 0.37	−0.49 (0.054)	1.80E−19	0.38	−0.38 (0.22)	0.08	−0.27 (0.27)	0.31
rs2384687	19		C, 0.39	−0.47 (0.053)	2.40E−18	0.39	−0.36 (0.22)	0.11	−0.27 (0.27)	0.32
rs236114	20	0.32	A, 0.21	0.50 (0.077)	9.70E−11	0.18	0.45 (0.30)	0.13	0.89 (0.36)	**0.01**
rs16991615	20		A, 0.058	1.07 (0.11)	1.20E−21	0.053	1.15 (0.49)	**0.02**	1.62 (0.60)	**0.01**

SNP, single nucleotide polymorphism; ANM, age at natural menopause; GWAS, genome-wide association study; TLGS, Tehran Lipid and Glucose Study; POF, premature ovarian failure; Chr, chromosome; MAF, minor allele frequency; NA, not available; SE, standard error. rs365132 failed genotyping in our study. rs2153159 was not reported in the 2009 AMN-GWAS. Effect sizes are in years and are per copy of the minor allele. LD information is based on 1000 genome data (CEU population). Bold SNPs have an unadjusted *P*-value <0.05

[a]2009 ANM-GWAS data are from He *et al.* (2009) 'joint analysis', with the exception of rs236114, which is from Stolk *et al.* (2009) 'overall' meta-analysis.

[b]Results are based on genotyping of 783 Iranian women who experienced natural menopause ≥40 years. The tested allele was set to be the minor allele as reported in the 2009 ANM-GWAS. With the exception of rs7718874, the tested allele is also the minor allele in the Iranian women.

[c]Results are based on genotyping of 828 Iranian women who experienced natural menopause, including POF cases. The tested allele was set to be the minor allele as reported in the 2009 ANM-GWAS.

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

years, $n = 165$ samples) and a 'late' ANM pool (upper 20% percentile of menopause ages, 54–65 years, $n = 187$ samples). Although the range of ANM in the late ANM pool is large, 70% of women in this group reached menopause at age 54 years. By selecting women from the ends of the ANM distribution we aimed to retain much of the power that would be afforded by surveying the entire ANM distribution (Sham et al., 2002; Eshraghi et al., 2007). DNA was extracted according to the established methods at the Research Institute for Endocrine Sciences (Eshraghi et al., 2007; Kahrizi et al., 2009). DNA samples were quantified in duplicate by fluorometry using PicoGreen™ (Molecular Probes, Eugene, OR, USA). Pools were constructed by combining 20 ng of each DNA sample by manual pipetting, and brought to a concentration of 100 ng/μl. Each DNA pool was assayed on four *Illumina Human1M-Duo* (1M-Duo) arrays, carried out at the Center for Applied Genomics at the Hospital for Sick Children, Toronto. Replicate arrays are used to reduce the error in allele frequency estimation (Earp et al., 2011).

Analysis of pool-based GWAS data

Allele-based tests of association were used to rank SNPs assayed by the 1M-Duo arrays, and 102 highly ranked SNPs were chosen for IG in our collection of Iranian women. The 1M-Duo data were analyzed using the publicly available program GenePool (http://genepool.tgen.org, source GenePool 0.9.1); in particular, the 'SINGLEMARKER' test statistic (a modified *t*-test) (Pearson et al., 2007). Two categories of SNPs were excluded from this analysis prior to data collection: (i) SNPs without a minor allele frequency (MAF) estimate or an MAF estimate <10% (HapMap CEU, Release 27) and (ii) Y chromosome SNPs, mitochondrial genome (mtDNA) SNPs and copy-number variant (CNV) probes. Our study is underpowered to detect SNP-ANM associations in low MAF SNPs, prompting their removal to avoid detecting associations that are most likely to be spurious. Y chromosome SNPs are not relevant in an all-female cohort, mtDNA SNPs were excluded due to concerns regarding homology (>98%) between nuclear and mtDNA sequence confounding results, and probes designed to assay CNV non-polymorphic sites do not have meaningful two-colour fluorescence intensity data (needed to estimate allele frequency). After data collection, 759 SNPs that were missing from one or more array, and those 5% of SNPs with the greatest variability in allele frequency estimation on replicate arrays, were removed from pool-based analyses. In total 694 326 SNPs were tested for association with ANM.

We prioritized highly ranked SNPs (ranked by ascending SINGLEMARKER test statistic *P*-value) by taking into consideration proxy SNP association results. To do this we determined all SNPs on the 1M-Duo array in LD ($r^2 > 0.8$, based on HapMap CEU R27 data) with a given high-ranked SNP, and calculated the median SINGLEMARKER rank of these SNPs (we called this Cluster analysis). We performed Cluster analysis on the 1000 top-ranked SNPs (we call these 'primary' SNPs) and re-ranked the primary SNPs by ascending median SINGLEMARKER rank for each LD-based cluster. Many of the 1000 top-ranked SNPs investigated fell within overlapping clusters, in which case only one SNP was chosen to represent a cluster, the SNP with the smallest SINGLEMARKER *P*-value. We chose 95 SNPs for IG using this approach. Seven SNPs were chosen based solely on their SINGLEMARKER test statistic *P*-value (i.e. highly ranked SNPs without a proxy SNP).

Analysis of 102 ANM-associated SNPs from pool-based GWAS

Based on the pool-based GWAS array data, 102 SNPs were chosen for further IG and association analysis in Iranian women. Linear regression was used to determine the effect of the minor allele for each SNP on ANM (including and excluding POF cases). Effect sizes (beta values) were calculated in years and are per copy of the minor allele. Adjusting for birth year, BMI, smoking history, oral contraceptive use and parity did not influence our results; therefore, all data are presented without correction for these variables. Data were analyzed in PLINK (v1.07).

IG and quality control

In total, 110 SNPs (8 + 102) were genotyped in 903 Iranian women. Genotyping was performed using the Sequenom iPlex Gold® assay (Sequenom, Inc., San Diego, CA, USA) and carried out at the McGill University and Génome Québec Innovation Center. With respect to SNP quality control (QC), 10 SNPs had a call rate of 0% and were removed (i.e. failed to genotype), 4 SNPs had a call rate of 91–95% and were retained in analyses (rs1172822, rs10787495, rs4787423 and rs13325331). No remaining SNPs had a call rate <90%. Thus, 100 SNPs were available for analysis. Of these, three were found to deviate from Hardy–Weinberg equilibrium and were removed (rs13325331, rs7586884 and rs6597754), leaving 97 SNPs for analysis. With respect to sample QC, 55 samples had a 0% genotype rate and were removed (insufficient DNA) and 20 samples had a call rate below 90% and were removed. Thus, 828 individuals were available for analysis. Based on 40 duplicate samples (4.4% of 903 initial samples), the concordance rate was 99% (one sample was discordant for one SNP).

Results

SNPs associated with ANM in Europeans also influence ANM in Iranians

Treating ANM as a quantitative trait, the MAF, direction and effect size of SNPs associated with menopause in European women were consistent in Iranian women (Table I). For most SNPs tested, the effect sizes were smaller than those reported in the 2009 ANM-GWAS. This is anticipated due to 'winners curse', the tendency of a discovery study to overestimate the true effect size (Zollner and Pritchard, 2007). However, rs16991615 (chromosome 20) had a larger effect size in our collection of Iranian women, increasing ANM by ~14 months (1.15 years) per allele instead of ~13 months (1.07 years) in European women. Notably, this effect size estimate has a large standard error because of the low MAF of the SNP and our study's sample size. rs2153157 (chromosome 6) also had a larger effect size in Iranian women, increasing ANM by ~5 months (0.41 years) per allele instead of ~3.5 months (0.29 years in European women). To date, the most significant SNP-ANM association reported is that of rs16991615; this SNP had a *P*-value of <0.05 in our data. rs1172822 (chromosome 19) and rs2153157 (chromosome 6) narrowly missed this level of significance.

There were 45 women with POF in the TLGS cohort for whom DNA was available, and we performed a second linear regression analysis including these samples (828 women in total). With the exception of the chromosome 19 SNPs, the SNPs tested had a larger estimated effects size and smaller *P*-value in this analysis (Table I). These data are consistent with the SNPs on chromosomes 5, 6 and 20 also influencing menopause age in women who experience POF.

Test for novel SNP-ANM associations conferring moderate-to-large effects in Iranian women

Because 12 SNPs failed QC; a total of 90 (of 102) SNPs were analyzed. Treating ANM as a quantitative trait, 10 SNPs selected for IG

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

based on our pool-based GWAS had a *P*-value of <0.05 in our Iranian women (783 women, excluding POF) (Table II). The most significant of these was rs10140275 (P-value $= 4.0 \times 10^{-4}$), and increased ANM by ~13 months (1.09 years) per allele. No SNPs were significant at the genome-wide level; however, rs10140275 was significant based on Bonferroni correction for 90 tests (α-level $= 5.6 \times 10^{-4}$). Of the 783 women in the analysis, 210 were included in the DNA pools and 573 were not. When women participating in the pooling stage were excluded from the regression analysis (i.e. independent replication), 2 SNPs (rs4304553 and rs4663953) had a *P*-value of <0.05 in the women tested (573 individuals, POF cases excluded) (Table II). Notably, the ANM association with rs10140275 was much weaker in this analysis ($\beta = 0.32$, $P = 0.185$). A second linear regression analysis including POF samples (828 women in total) was performed. Only 2 SNPs reported in Table II (rs10140275 and rs4413314) had a *P*-value of <0.05 and a consistent direction of SNP effect in this analysis; no other SNPs (of 90 tested) had a *P*-value of <0.05. Both rs10140275 and rs4413314 have a larger SNP effect size and a smaller *P*-value when POF samples were included. With POF samples included, rs10140275 increased ANM by ~17 months (1.48 years) ($P = 1.0 \times 10^{-4}$) and rs4413314 increased ANM by ~8 months (0.70 years) per allele ($P = 0.013$).

Ten SNPs (Table II) were investigated for SNP-ANM association in the ReproGen consortium's GWAS meta-analysis data, which included 38 968 women of European descent experiencing ANM between 40 and 60 years of age (Stolk *et al.*, 2012). Only one SNP, rs10840211, had a *P*-value of <0.05 in the data (Table II), and eight SNPs in high LD ($r^2 > 0.8$) with rs10840211 had similar *P*-values in the ReproGen data (data not shown). In Iranian women rs10840211 decreased ANM by ~6 months (-0.48 years, $P = 0.044$) per allele and in European women it decreased ANM by ~48 days (-0.11 years, $P = 0.0013$). The smaller SNP effect in European women is consistent with the winner's curse.

Discussion

Principal findings

One SNP-ANM association detected in European women on chromosome 20 was replicated in Iranian women (rs16991615; $\beta = 1.15$, standard error (SE) $= 0.49$, $P = 0.02$). SNPs at the previously reported 19q13.42 and 6p24.2 loci also approached statistical significance and had consistent SNP effects (magnitude and direction) in Iranian women (rs1172822; $\beta = -0.39$, SE: 0.22, $P = 0.08$; and rs2153157, $\beta = 0.41$, SE: 0.21, $P = 0.05$). We were underpowered to detect many of the SNP-ANM associations reported by the 2009 ANM-GWAS; however, for all SNPs tested the direction of the effect was consistent with that observed in European women. Hence, we anticipate that many (if not all) of the ANM-associated SNPs discovered in European women will replicate in Iranian women upon genotyping a sufficient number of women. Our results imply that European and Iranian women share ANM-associated genetic variants. Chen *et al.* (2011) recently demonstrated that genetic variants influencing ANM in European women also influenced ANM in Hispanic women. Like our study, they replicate the ANM association with rs16991615, and they find this SNPs effect size to be larger than the previously reported (European) value of 1.07. They also find that rs16991615 is associated with an increased risk of early menopause, similar to our analysis including POF samples. These results imply that human populations that differ by race/ethnicity, including European, Hispanic and Iranian populations, share ANM-associated genetic variants; however, the magnitude of the SNP effects in each population may differ.

We tested for novel ANM risk alleles conferring moderate-to-large effects in Iranian women using a pool-based GWAS; however, we found little evidence for such SNPs. The most significant SNP detected, rs10140275, had a P-value $= 4.0 \times 10^{-4}$ when analyzed in 783 Iranian women who experienced menopause, and a P-value $= 1.0 \times 10^{-4}$ when analyzed in 828 Iranian women who experienced menopause or POF: in both analyses, the effect size was large ($\beta = 1.09$ and $\beta = 1.48$ years, respectively). However, this SNP did not have a *P*-value of <0.05 when samples used in the pooling stage and POF samples were excluded, i.e. when independent replication in Iranian women was performed. Given a sample size of 573 Iranian women, and that these 573 samples were primarily from the middle of the ANM distribution (ANM between 46 and 53 years), this replication analysis may have been underpowered to detect this association. Notably, when 45 samples from women with POF were included in the replication (pooling samples still excluded), rs10140275 had a *P*-value of <0.05 ($\beta = 0.86$, SE: 0.39, $P = 0.027$). Additional genotyping in Iranian women (who experience the full range of menopause ages) is needed to confirm this association. rs10140275 was not associated with ANM in European women (Table II); therefore, this could represent an SNP-ANM association unique to Iranian women.

One SNP (of 90 tested), rs10840211, had a *P*-value of <0.05 when analyzed in Iranian women ($\beta = -0.48$, $P = 0.044$), and a *P*-value of <0.05 in European women ($\beta = -0.11$, $P = 0.0013$). This SNP did not have a *P*-value of <0.05 in Iranian women when samples used in the pooling stage were excluded (Table II); a possible reason for this is as discussed above. Inclusion of samples from women with POF in the replication did not change this result (rs10840211, $\beta = -0.35$, SE: 0.29, $P = 0.22$). Additional genotyping in Iranian women is needed to confirm this association. This could represent an SNP-ANM association shared between Iranian and European women, but having a larger effect size in Iranian women.

This study's greatest limitation was power. Due to small sample size this study was powered to reliably detect only moderate-to-large SNP effect sizes. This limited our ability to replicate many of the previously reported SNP-ANM associations and to discover novel SNP-ANM associations' specific to Iranian women. Further, in performing our pool-based GWAS we accepted a loss of power relative to a conventional GWAS, stemming from the fact that SNP allele frequencies must be estimated from pools rather than directly calculated from individual genotypes. This introduces error into the calculation of any test of association and consequently reduces power. To reduce the size of these errors, replicate arrays are used to assay the same DNA pool, and this strategy was used here (using four replicate arrays per pool). Another factor limiting power is that the pool-based GWAS approach necessitates allele-based tests of association, and these are not as powerful as genotype-based tests. In practice, very few pool-based GWAS focus only on SNPs that achieve genome-wide significance in the discovery pooling stage. Rather, SNPs are ranked according to the strength of their association, and practical

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

Table II SNP-ANM association of 10 SNPs identified in a pool-based GWAS of Iranian women (excluding POF, ANM ≥40 years).

SNP	Chr	Iranian samples used in the pooling stage		Iranian replication samples (excluding POF, ANM ≥40 years)		Iranian women (excluding POF, ANM ≥40 years)				European women (ReproGen consortium)			
		β_1	*P*-value	β_1	*P*-value	Alleles[a]	AF[a]	β_1	*P*-value	Alleles[b]	AF[b]	β	*P*-value
rs10140275	14	3.59	**3×10^{-4}**	0.32	0.185	C/T	0.14	1.09	**4.0E−4**	A/G	0.81	−0.03	0.41
rs10144724	14	−2.32	**0.01**	−0.19	0.45	C/T	0.14	−0.82	**0.01**	T/C	0.84	−0.03	0.51
rs7766409	6	−0.31	0.10	−1.21	0.09	C/T	0.37	−0.56	**0.02**	A/G	0.70	0.01	0.70
rs4413314	3	0.35	0.05	1.10	0.13	T/C	0.40	0.54	**0.02**	T/C	0.50	−0.02	0.52
rs4304553	1	−0.24	0.29	−1.85	**0.05**	A/C	0.17	−0.64	**0.03**	A/C	0.14	−0.06	0.16
rs6762123	3	−0.42	0.11	−1.52	0.13	T/C	0.12	−0.71	**0.03**	A/G	0.12	−0.03	0.45
rs10840211	11	−1.33	0.07	−0.19	0.32	T/C	0.32	−0.48	**0.04**	A/G	0.25	−0.11	**1.0E−3**
rs4902619	14	2.81	**0.02**	0.14	0.58	G/T	0.15	0.66	**0.04**	A/C	0.87	0.04	0.32
rs2211313	13	−2.58	**0.01**	0.09	0.74	A/C	0.14	−0.65	**0.05**	T/G	0.13	0.01	0.86
rs4663953	2	−0.16	0.55	−1.99	**0.05**	C/T	0.13	−0.65	**0.05**	A/G	0.85	−0.05	0.26

Table is ordered by smallest *P*-value in the 'Iranian women' regression analysis (783 women who had experienced menopause, including 573 replication samples and 210 samples used in the pooling stage). Bold SNPs have an unadjusted *P*-value <0.05. Data for European women is from the Reprogen consortium's ANM-GWAS meta-analysis (Stolk *et al.*, 2012).

[a]Major allele/minor allele; AF, allele frequency for minor allele.

[b]Coded allele/non-coded allele; AF, allele frequency for coded allele. The coded allele was tested for association.

considerations dictate how many top-ranked SNPs are chosen for subsequent IG and replication (where stringent significance thresholds are then imposed). This is the approach used in this study. Finally, detecting, excluding or adjusting for individuals whose ancestry differs from that of other samples is not possible with pooled DNA, nor is removing individuals showing cryptic relatedness.

The greatest strength of this study was the population-based, geographically focused and homogenous collection of Iranian women with recently collected epidemiological data relating to fertility and available DNA samples. Although this study was limited to detecting larger SNP effects due to sample size, if variants conferring these effect sizes existed, we were well positioned to discover them. Furthermore, 45 women in our collection experienced POF, which allowed for a preliminary assessment of ANM-associated SNPs in this phenotype. One concern with this data set was that ANM was ascertained by self-reporting, which is subject to recall bias. This is most challenging for women who entered the TLGS post-menopause. The mean ANM for women in this study ($\mu = 49.8$, SD = 4.4) was very similar to that reported in another collection of Iranian women ($\mu = 50.4$, SD = 4.3), leading us to conclude that ANM was, on average, accurately reported by our participants.

A weakness of our pool-based GWAS was the inability to fully validate SNPs of interest before replication. Validation is typically performed in pool-based GWAS to remove SNPs that are associated with the trait of interest due to technical error during pooling; this helps to reduce the number of tests performed during replication. Insufficient quantities of DNA for some samples prevented us from performing IG on all of the samples included in the pooling stage.

Conclusion

We find evidence for shared ANM-associated genetic variants between different racial/ethnic populations. Our data do not support the hypothesis that novel population-specific SNP-ANM associations explain population-specific differences in mean ANM. Population-specific differences in the mean ANM could be caused by environmental factors, or genetic variants which have yet to be discovered, owing to small effect size or low MAF.

Acknowledgements

We thank the TLGS study and the ReproGen Consortium participants and investigators. We also thank the staff of the Centre for Applied Genomics at the Hospital for Sick Children in Toronto and the staff of the McGill University Genome Quebec Innovation Centre for expert genotyping.

Authors' roles

M.R., F.R.T. and F.A. handled the recruitment of patients, sample collection and supervised the clinical aspects of the work. M.R., M.A.E., M.A., J.W., M.T., R.N. and S.P.-B. undertook all the laboratory work. M.R., M.A.E., F.R.T., J.R.B.P., J.M.M., F.A. and A.B.-W. contributed to data analysis. M.R., M.A.E., F.R.T., F.A. and A.B.-W. designed the overall study. All authors contributed to writing the report.

Funding

TLGS is funded by the Research Institute for Endocrine Sciences (RIES). The ANM study was funded by the RIES (grants to M.R., F.R.T. and F.A.). M.R. received Postdoctoral Research Awards from the Heart and Stroke Foundation of Canada/Pfizer (HSFC) and the Michael Smith Foundation for Health Research (MSFHR). M.A.E. was the recipient of an NSERC Doctoral Postgraduate Scholarships (PGSD) and a Four Year Doctoral Fellowship from the University of British Columbia. A.B.W. is a Senior Scholar of the MSFHR.

Conflict of interest

None declared.

References

Azizi F, Rahmani M, Emami H, Mirmiran P, Hajipour R, Madjid M, Ghanbili J, Ghanbarian A, Mehrabi Y, Saadat N *et al.* Cardiovascular risk factors in an Iranian urban population: Tehran lipid and glucose study (phase 1). *Sozial Praventivmed* 2002;**47**:408–426.

Braem MG, Onland-Moret NC, van den Brandt PA, Goldbohm RA, Peeters PH, Kruitwagen RF, Schouten LJ. Reproductive and hormonal factors in association with ovarian cancer in the Netherlands cohort study. *Am J Epidemiol* 2010;**172**:1181–1189.

Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ *et al.* Genome-wide association study of 14 000 cases of seven common diseases and 3 000 shared controls. *Nature* 2007; **447**:661–678.

Butcher LM, Meaburn E, Knight J, Sham PC, Schalkwyk LC, Craig IW, Plomin R. SNPs, microarrays and pooled DNA: identification of four loci associated with mild mental impairment in a sample of 6000 children. *Hum Mol Genet* 2005;**14**:1315–1325.

Chen CT, Fernández-Rhodes L, Brzyski RG, Carlson CS, Chen Z, Heiss G, North KE, Woods NF, Rajkovic A, Kooperberg C *et al.* Replication of loci influencing ages at menarche and menopause in Hispanic women: the Women's Health Initiative SHARe Study. *Hum Mol Genet* 2011; **21**:1419–1432.

Coulam CB, Adamson SC, Annegers JF. Incidence of premature ovarian failure. *Obstet Gynecol* 1986;**67**:604–606.

de Bruin JP, Bovenhuis H, van Noord PA, Pearson PL, van Arendonk JA, te Velde ER, Kuurman WW, Dorland M. The role of genetic factors in age at natural menopause. *Hum Reprod* 2001;**16**:2014–2018.

Dossus L, Allen N, Kaaks R, Bakken K, Lund E, Tjonneland A, Olsen A, Overvad K, Clavel-Chapelon F, Fournier A *et al.* Reproductive risk factors and endometrial cancer: the European Prospective Investigation into Cancer and Nutrition. *Int J Cancer* 2010;**127**:442–451.

Earp MA, Rahmani M, Chew K, Brooks-Wilson A. Estimates of array and pool-construction variance for planning efficient DNA-pooling genome wide association studies. *BMC Med Genom* 2011;**4**:81.

Eshraghi P, Hedayati M, Daneshpour MS, Mirmiran P, Azizi F. Association of body mass index and Trp64Arg polymorphism of the beta3-adrenoreceptor gene and leptin level in Tehran Lipid and Glucose Study. *Br J Biomed Sci* 2007;**64**:117–120.

Faddy MJ, Gosden RG, Gougeon A, Richardson SJ, Nelson JF. Accelerated disappearance of ovarian follicles in mid-life: implications for forecasting menopause. *Hum Reprod* 1992;**7**:1342–1346.

Gallagher JC. Effect of early menopause on bone mineral density and fractures. *Menopause* 2007;**14**:567–571.

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024

Gold EB, Bromberger J, Crawford S, Samuels S, Greendale GA, Harlow SD, Skurnick J. Factors associated with age at natural menopause in a multiethnic sample of midlife women. *Am J Epidemiol* 2001;**153**:865–874.

He C, Kraft P, Chen C, Buring JE, Paré G, Hankinson SE, Chanock SJ, Ridker PM, Hunter DJ, Chasman DI. Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat Genet* 2009;**41**:724–728.

Henderson KD, Bernstein L, Henderson B, Kolonel L, Pike MC. Predictors of the timing of natural menopause in the Multiethnic Cohort Study. *Am J Epidemiol* 2008;**167**:1287–1294.

Kahrizi K, Mohseni M, Nishimura C, Bazazzadegan N, Fischer SM, Dehghani A, Sayfati M, Taghdiri M, Jamali P, Smith RJ *et al.* Identification of SLC26A4 gene mutations in Iranian families with hereditary hearing impairment. *Eur J Pediatr* 2009;**168**:651–653.

Morris DH, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Familial concordance for age at natural menopause: results from the Breakthrough Generations Study. *Menopause* 2011;**18**:956–961.

Murabito JM, Yang Q, Fox C, Wilson PW, Cupples LA. Heritability of age at natural menopause in the Framingham Heart Study. *J Clin Endocrinol Metab* 2005;**90**:3427–3430.

Paschou P, Lewis J, Javed A, Drineas P. Ancestry informative markers for fine-scale individual assignment to worldwide populations. *J Med Genet* 2010;**47**:835–847.

Pearson JV, Huentelman MJ, Halperin RF, Tembe WD, Melquist S, Homer N, Brun M, Szelinger S, Coon KD, Zismann VL *et al.* Identification of the genetic basis for complex disorders by use of pooling-based genomewide single-nucleotide-polymorphism association studies. *Am J Hum Genet* 2007;**80**:126–139.

Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. *Biostatistics* 2009;**10**:327–334.

Richardson SJ. The biological basis of the menopause. *Baillieres Clin Endocrinol Metab* 1993;**7**:1–16.

Sham P, Bader JS, Craig I, O'Donovan M, Owen M. DNA pooling: a tool for large-scale association studies. *Nat Rev Genet* 2002;**3**:862–871.

Shin A, Song YM, Yoo KY, Sung J. Menstrual factors and cancer risk among Korean women. *Int J Epidemiol* 2011;**40**:1261–1268.

Shuster LT, Rhodes DJ, Gostout BS, Grossardt BR, Rocca WA. Premature menopause or early menopause: long-term health consequences. *Maturitas* 2010;**65**:161–166.

Snieder H, MacGregor AJ, Spector TD. Genes control the cessation of a woman's reproductive life: a twin study of hysterectomy and age at menopause. *J Clin Endocrinol Metab* 1998;**83**:1875–1880.

Stolk L, Zhai G, van Meurs JB, Verbiest MM, Visser JA, Estrada K, Rivadeneira F, Williams FM, Cherkas L, Deloukas P *et al.* Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat Genet* 2009;**41**:645–647.

Stolk L, Perry JR, Chasman DI, He C, Mangino M, Sulem P, Barbalic M, Broer L, Byrne EM, Ernst F *et al.* Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways. *Nat Genet* 2012;**44**:260–268.

te Velde ER, Scheffer GJ, Dorland M, Broekmans FJ, Fauser BC. Developmental and endocrine aspects of normal ovarian aging. *Mol Cell Endocrinol* 1998;**145**:67–73.

Titus-Ernstoff L, Longnecker MP, Newcomb PA, Dain B, Greenberg ER, Mittendorf R, Stampfer M, Willett W. Menstrual factors in relation to breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 1998; **7**:783–789.

van der Schouw YT, van der Graaf Y, Steyerberg EW, Eijkemans JC, Banga JD. Age at menopause as a risk factor for cardiovascular mortality. *Lancet* 1996;**347**:714–718.

van Noord PA, Dubas JS, Dorland M, Boersma H, te Velde E. Age at natural menopause in a population-based screening cohort: the role of menarche, fecundity, and lifestyle factors. *Fertil Steril* 1997; **68**:95–102.

Voorhuis M, Onland-Moret NC, van der Schouw YT, Fauser BC, Broekmans FJ. Human studies on genetics of the age at natural menopause: a systematic review. *Hum Reprod Update* 2010;**16**:364–377.

Voorhuis M, Broekmans FJ, Fauser BC, Onland-Moret NC, van der Schouw YT. Genes involved in initial follicle recruitment may be associated with age at menopause. *J Clin Endocrinol Metab* 2011; **96**:E473–479.

Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case–control data. *Am J Hum Genet* 2007;**80**;605–615.

Downloaded from https://academic.oup.com/humrep/article/28/7/1987/612495 by guest on 24 April 2024