Human Reproduction, Vol.27, No.2 pp. 594–608, 2012
Advanced Access publication on November 24, 2011 doi:10.1093/humrep/der391

human reproduction

ORIGINAL ARTICLE *Reproductive genetics*

Genetic variants and environmental factors associated with hormonal markers of ovarian reserve in Caucasian and African American women

Sonya M. Schuh-Huerta[1,2,*], Nicholas A. Johnson[3], Mitchell P. Rosen[4], Barbara Sternfeld[5], Marcelle I. Cedars[4], and Renee A. Reijo Pera[1,2]

[1]Institute for Stem Cell Biology and Regenerative Medicine, Center for Human Embryonic Stem Cell Research and Education, School of Medicine, Stanford University, 265 Campus Drive, Stanford, CA 94305, USA [2]Department of Obstetrics and Gynecology, School of Medicine, Stanford University, Stanford, CA 94305, USA [3]Department of Statistics, Stanford University, Stanford, CA 94305, USA [4]Department of Obstetrics, Gynecology and Reproductive Sciences, Center for Reproductive Health, Women's Health Center, University of California San Francisco, 2356 Sutter Street, San Francisco, CA 94115, USA [5]Division of Research, Kaiser Permanente, Oakland, CA 94612, USA

*Correspondence address. E-mail: sonyas1@stanford.edu

Submitted on May 31, 2011; resubmitted on October 13, 2011; accepted on October 24, 2011

BACKGROUND: The ovarian reserve (number and quality of oocytes) is correlated with reproductive potential as well as somatic health, and is likely to have multiple genetic and environmental determinants. Several reproductive hormones are closely linked with the oocyte pool and thus can serve as surrogate markers of ovarian reserve. However, we know little about the underlying genes or genetic variants.

METHODS: We analyzed genetic variants across the genome associated with two hormonal markers of ovarian reserve, FSH and anti-Mullerian hormone, in a reproductively normal population of Caucasian ($n = 232$) and African American ($n = 200$) women, aged 25–45 years. We also examined the effects of environmental or lifestyle factors on ovarian reserve phenotypes.

RESULTS: We identified one variant approaching genome-wide significance (rs6543833; $P = 8.07 \times 10^{-8}$) and several nominal variants nearby and within the *myeloid-associated differentiation marker-like* (*MYADML*) gene, that were associated with FSH levels in African American women; these were validated in Caucasian women. We also discovered effects of smoking and oral contraceptive use on ovarian reserve phenotypes, with alterations in several reproductive hormones.

CONCLUSIONS: This work is the largest study on ovarian reserve in women of reproductive age and is the only genome-wide study on ovarian reserve markers. The genes containing or near the identified variants have no known roles in ovarian biology and represent interesting candidate genes for future investigations. The discovery of genetic markers may lead to better long-range predictions of declining ovarian function, with implications for reproductive and somatic health.

Key words: FSH / anti-Mullerian hormone / antral follicle count / single nucleotide polymorphism

Introduction

Infertility is a significant problem in society, affecting 15% of reproductive-aged couples (Hull *et al.*, 1985; Menken and Larsen, 1994; Reijo *et al.*, 1995). Many cases of infertility appear to be the result of defects in germ cell development that lead to reduced numbers or quality of gametes. The quantity and quality of oocytes within the ovaries, or ovarian reserve, are closely correlated with reproductive potential and are likely to have multiple genetic and environmental determinants. Further, ovarian reserve is affected by aging and disease, and is closely tied to both reproductive and somatic health. Many diseases including osteoporosis, cardiovascular disease and various cancers are associated with loss of oocytes/follicles, ovarian dysfunction and/or the length of the female reproductive lifespan (Cooper and Sandler, 1998; Hartge, 2009). Hence, better prediction of both normal and aberrant oocyte depletion would have significant clinical implications.

Several reproductive hormones are closely correlated with the ovarian reserve and the inevitable process of follicle depletion. Current methods for measuring ovarian reserve and predicting age at fertility loss and menopause include assessment of serum levels of FSH, inhibin B and anti-Mullerian hormone (AMH), which decreases ~10 years prior to menopause (van Rooij *et al.*, 2004). AMH and FSH are direct and indirect reflections of follicle number, respectively. AMH is a member of the transforming growth factor-β superfamily; it suppresses the cyclical recruitment of primordial follicles into the pool of growing follicles and is specifically secreted by the granulosa cells of pre-antral and antral follicles in the ovary (Durlinger *et al.*, 2002). Thus, AMH is present in proportion to the number of antral follicles. FSH on the other hand, is an indirect reflection of the follicle pool; it is a gonadotrophin secreted by the anterior pituitary in response to hypothalamic release of GnRH. It stimulates follicular growth and is a key component of the brain-pituitary-ovary axis (Gromoll and Simoni, 2005). FSH is sensitive to the inhibitory influence of inhibin B. Inhibin B is secreted by the granulosa cells of developing and early antral follicles. So as the number of follicles declines, so do inhibin levels and FSH begins to rise. Thus, there is an association between high FSH and low follicle number with ovarian aging.

Antral follicle count (AFC), transvaginal ultrasound-guided assessment of follicles, is routinely performed during fertility assessments. We have recently analyzed the correlations of several hormonal markers and AFC with age in Caucasian women (unpublished data), finding that only AFC and AMH parallel the decline of follicles with age noted in histological ovarian specimens (Hansen *et al.*, 2008). In contrast, FSH shows an inverse relationship with AFC and follicle number *in vivo*; as follicle number and AMH decline, FSH rises. It is noteworthy that both AMH and FSH levels are highly variable among women, even measured at the same point in the menstrual cycle.

The age of menopause has a strong heritability (Snieder *et al.*, 1998; de Bruin *et al.*, 2001; Murabito *et al.*, 2005). Ovarian aging (OVA), which leads to menopause with the exhaustion of the follicular pool, also seems to have a genetic component (Rosen *et al.*, 2010a). Genetic markers associated with ovarian reserve may have the advantage of providing better long-range predictions of risk for declining ovarian reserve. Thus, determining the genetic associations of rising FSH and declining AMH may allow for improved counseling regarding reproductive health and aging.

Most previous work has focused on measuring the reproductive events of menarche and menopause, or comparing ovarian reserve markers and their ability to predict reproductive outcomes in the context of infertile populations or small observational studies. Moreover, few studies have investigated the genes or variants associated with specific reproductive hormones. Therefore, we sought to identify genes and variants across the genome associated with hormonal markers of ovarian reserve in a multi-ethnic cohort recruited from the general population. We also sought to examine the effects of environmental or lifestyle factors including smoking and oral contraceptive use, on ovarian reserve phenotypes.

Materials and Methods

Study population

The final study population included 232 Caucasian and 200 African American women, aged 25–45 years, who were prospectively enrolled in a community-based cohort of the OVA Study. These women were not seeking treatment for infertility or other medical problems. The OVA Study comprises a multi-ethnic cohort and is focused on elucidating clinical and genetic markers of OVA. Subjects were recruited from potentially eligible female members of Kaiser Permanente (KP) Northern California, who resided within proximity to the Women's Health Clinical Research Center at the University of California San Francisco (UCSF). The socio-demographic and health-related characteristics of the KP patient population are representative of the population of Northern California (Gordon, 2006), and the final OVA participants were similar to the KP population. Subjects were included if they were ovulatory with normal predictable menstrual cycles of 22–35 days. Women with oligo- or anovulation, surgically diagnosed endometriosis, ovarian failure (loss of all follicles before the age of 40 years) or a history of uterine or ovarian surgery were excluded. Subjects were also excluded if they had taken oral contraceptives or medications containing estrogen or progestin that alter the menstrual cycle within the 3 months prior to enrollment. This work was approved by the Institutional Review Boards at KP, UCSF and Stanford University, and informed consent was obtained for all subjects.

Phenotypes and covariates of ovarian reserve

Blood samples were collected by venipuncture on Day 2–4 of the menstrual cycle and used for all serum hormonal assays and genotyping. FSH and LH levels were measured with standardized two-site chemiluminescence immunoassays. The intra-assay coefficients of variation (CVs) were 1.9–2.1%, whereas the inter-assay CVs were 5.2–6.8%. AMH was measured using a commercially available ELISA, which uses a direct competitive immunoassay (Beckman Coulter, Marsee, France). The AMH CVs for intra- and inter-assays were <5.6 and $<15.3\%$, respectively. Estrone (E_1) and inhibin B were measured using a commercially available ELISA (Diagnostic Systems Laboratories, Webster City, TX, USA). The intra-assay CVs for E_1 and inhibin B were 4.8–9.9 and 3.3–7.2%, respectively, whereas the inter-assay CVs were 7.7–13.4 and 7.8–17%, respectively. Estradiol (E_2) and testosterone were measured with an automated chemiluminescent assay using the Bayer Diagnostics ACS:180 (Bayer Diagnostics Corp., Tarrytown, NY, USA). The intra- and inter-assay CVs for E_2 were 6.5–6.9 and 13.6–16.1%, respectively. For total testosterone, the intra-assay CVs were 2.9–4.0% and the inter-assay CVs were 6.3–6.6%.

AFCs were performed as previously described (Rosen *et al.*, 2010b). Briefly, all subjects underwent transvaginal ultrasound assessment of ovarian volumes and AFC on Day 2–4 of their menstrual cycle. All follicles with a mean diameter (of two dimensions) of 2–10 mm were counted; all echo-free structures were regarded as follicles (Supplementary data, Fig. S1). To eliminate experimental variability all women were examined at the same time in their menstrual cycle, by one of two physicians, using the same equipment, and unclear ultrasound results were excluded. Inter-examiner concordance exceeds 90%.

All subjects completed a comprehensive questionnaire to obtain detailed health, reproductive, lifestyle and demographic information, as previously described (Rosen *et al.*, 2010b). The age of each woman was rounded off to the nearest 10th of a year. Body measurements including weight, height and waist and hip circumference, were also obtained. Parity or the number of children born to each woman was also assessed. Cigarette smoke exposure was quantified from responses to questionnaire data and categorized as 'never, past or current' smoker and pack-years of smoking. Previous history of taking estrogen- and/or progestin-containing oral contraceptive pills (OCPs) was also quantified from questionnaire data and categorized as 'never or past' and total months of OCP use.

Genotyping and quality control

Genotypic data for the Caucasian and African American women was obtained using the Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA). This microarray contains more than 900 000 single nucleotide polymorphism (SNP) markers and 940 000 copy number variant (CNV) markers; approximately one marker every 700 bases across the genome (Affymetrix, 2007). Genomic DNA was extracted and purified from white blood cells using the QIAamp DNA Blood Maxi Kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions. The DNA ligation, hybridization and microarray scanning steps were completed according to the manufacturer's protocols (Affymetrix).

Stringent quality control (QC), concordance checks and SNP filtering were employed using Genotyping Console (GTC) v.4.0 software (http://www.affymetrix.com) to eliminate poor samples and poorly performing or rare SNPs. The QC measures used were: QC call rate or percentage of a specific set of 3022 control SNPs genotyped or 'called' in the samples using the dynamic algorithm (filtered for QC call rate $\geq$86%); and contrast QC, which is an estimate of the ability to resolve SNP signals into three genotype clusters (filtered for contrast QC $\geq$0.4). Contrast QC, the most important QC measure, uses a static set of 10 000 randomly chosen SNPs and quantifies the difference between the highest and lowest contrast values of the homozygote and heterozygote genotypes. Laboratory technical QC criteria also included SNP fingerprints for sample tracking and early detection of sample misidentification, the use of HapMap controls to check genotype quality, gender confirmation based on mean intensities of SNP probes on the X and Y chromosomes, and tracking of reagent and instrument performance. In the Caucasian group, of the final 232 samples that passed QC and filter criteria, the mean contrast QC was 1.87 $\pm$ 0.046, the QC call rate was 93.9 $\pm$ 0.2% and the overall sample call rate was 98.0 $\pm$ 0.08%. In the African American group, of the final 200 samples, the mean contrast QC was 2.23 $\pm$ 0.036, the QC call rate was 97.0 $\pm$ 0.1% and the overall sample call rate was 99.4 $\pm$ 0.08%. SNPs were genotyped using GTC v.4.0, using the Birdseed v.2.0 algorithm (Korn *et al.*, 2008), which is an extension of the Bayesian Robust Linear Model with Mahalanobis Distance using perfect-match probes. Clustering analysis and examination of relative sample intensity, as well as filtering for high sample call rates ($\geq$95%) ensured high sample quality prior to downstream association analysis. Individual SNPs across the genome were filtered for SNP call rates >95%, missing call rates <5%, minor allele frequency (MAF) >1% and lack of significant deviation from Hardy–Weinberg equilibrium ($P > 0.007$). Of 909 622 SNP probes on the array, 677 261 and 738 185 in the Caucasian and African American cohorts, respectively, passed QC and SNP filter criteria and were independently tested for association with FSH and AMH. Using the 940 000 CNV probes on the array, future investigations will examine CNVs across the genome and their association with hormonal and clinical markers of ovarian reserve. All SNP locations are referenced to National Center for Biotechnology Information (NCBI) human genome build 37 (February 2009) and dbSNP build 131 (April 2010) (http://www.ncbi.nlm.nih.gov/projects/SNP). Gene names provided in all tables and text represent the HUGO Gene Nomenclature Committee (HGNC) gene symbols of the gene containing, or closest to, the identified SNP (http://www.genenames.org/).

Determination of population structure

For ethnicity validation and tests of population structure a principal component analysis (PCA) using singular value decomposition (svd) was performed. Subjects of the HapMap CEU (Utah residents with N. and W. European ancestry from the CEPH collection) and YRI (Yoruban in Ibadan, Nigeria) cohorts were genotyped using the same genotyping platform (http://hapmap.ncbi.nlm.nih.gov/) (International HapMap Consortium, 2005). The OVA cohorts were combined with the HapMap CEU and YRI cohorts, and missing genotypes were imputed using the program BEAGLE v.3.0.2 (http://www.stat.aukland.ac.nz/~bbrowning/beagle/beagle) (Browning, 2006; Browning and Browning, 2007). Genotypes at 10 000 SNPs, selected to have low levels of linkage disequilibrium (LD) and MAF >0.01, distributed across several chromosomes were coded as 0, 1 or 2 (AA, AB and BB) and were mean-centered. PCA was performed using the 'svd' function of the statistical program R v.2.11.1 (http://www.r-project.org/) (Gentleman and Ihaka, 2008) to analyze allele composition and population homogeneity (Supplementary data, Fig. S2).

Statistical analysis

Phenotypic and covariate analysis

Descriptive statistics were calculated for all hormonal, environmental and demographic variables in the study cohorts. The relationship between hormonal variables, follicle number and age was analyzed in linear regression analyses. For binary variables including smoking and OCP use, linear regressions of the hormone concentration versus age were compared between women that currently or previously smoked or used OCPs, and those that did not. Statistical differences between the slopes and intercepts of the regressions were analyzed using the *F*-test for variable inclusion with subject age mean-centered. $P < 0.05$ was considered statistically significant. Three models with different effects of smoking (or OCP use) were tested against the simple null model of an association between the hormone alone (hormone $\sim$ Age). The three models tested were: (1) smoking changes the intercept (hormone $\sim$ Age + Smoking); (2) the slope (hormone $\sim$ Age + Smoking $\times$ Age) or (3) both slope and intercept (hormone $\sim$ Age + Smoking + Smoking $\times$ Age). For these analyses, women that previously smoked were grouped with smoking women, as we reasoned if a woman ever smoked, there may have been irreversible effects on her ovarian reserve. Similar analyses were performed on the associations between follicle number and smoking or OCP use. Pack-years of smoking and total months of OCP use were also analyzed for their association with serum hormone levels, with control for age. Outliers were removed based on sensitivity analyses. Mean total serum hormone levels or AFCs were also compared between smoking (or OCP-using) and non-smoking (or non-OCP-using) women, using the Student's *t*-test or Welch 2-Sample *t*-test within R v.2.11.1 (Gentleman and Ihaka, 2008). Covariates, including smoking, OCP use and follicle number, that were significantly associated with FSH or AMH were also controlled for in downstream genome-wide association (GWA) analyses.

SNP association analysis

GWA analyses were performed in the context of a robust regression of FSH or AMH versus age using the application quantreg v.4.44 (Koenker, 2009) within R.A. Fisher's exact test for single-marker (SNP) and multiple-marker (haplotype) association using the allelic test was performed to determine associations with serum hormone levels using BEAGLE v.3.0.2 and quantreg v.4.44; both programs were integrated in R v.2.11.1. The effect sizes or the difference in the mean hormone level from the overall regression fit, by allele or genotype were calculated for all SNPs. In parallel validation analyses, allelic and genotypic tests using the Kruskal–Wallace rank-sum test determined SNPs significantly associated with FSH or AMH, as quantitative variables. For each SNP allele and genotype, a linear regression fit of FSH or AMH versus age was obtained, the residual was computed for each woman, and the women were ranked into groups based on the allele/genotype to determine whether one allele/genotype had higher or lower residuals (effect sizes) than expected by random chance alone. Associations determined by the

Fisher's exact test were also significant as determined by the Kruskal–Wallace test, for all top hits.

LD of variants, the number of SNPs involved in the association signals, the genes and the recombination rates of the regions surrounding the top hits were analyzed within BEAGLE, Haploview v.4.2 (http://www.broadinstitute.org/mpg/haploview) (Barrett *et al.*, 2005) and LocusZoom v.1.1 (http://csg.sph.umich.edu/locuszoom/) (Pruim *et al.*, 2010). LD tagging was performed with the Tagger program (de Bakker *et al.*, 2005) within Haploview to capture SNPs with squared correlation coefficients of $R^2 \geq 0.8$ and LD of flanking SNPs was similarly analyzed within LocusZoom.

To verify associations and determine final significance levels, permutation testing (1000 tests) was performed and $P < 0.05$ were selected for at the chromosome-wide level. Bonferonni correction for multiple testing and family-wise errors rates were computed to verify significant associations at the genome-wide level. *P*-values were computed and analyzed at both the chromosome- and genome-wide-levels, owing to our sample size/power constraints. Further, as several chromosomes have loci associated with ovarian reserve and function, variants were also analyzed specifically at the chromosome level. Although accepted genome-wide significance thresholds are typically $\sim 5 \times 10^{-8}$, because of our marginal sample size (which is large for a clinical/biological study but moderate for a GWA study), assuming a slightly higher genome-wide threshold of $P < 1 \times 10^{-7}$, we note that the top hit identified in this study is at a marginal genome-wide significance level. Custom-made programs and algorithms were used to mediate streamlined analyses and visualization of results within R.

Results

Genetic association analysis

As shown in Table I, in 232 Caucasian and 200 African American women, the average FSH levels were 7.1 ± 0.2 and 7.4 ± 0.3 IU/l, respectively. The mean AMH level was 30.1 ± 1.5 pM in Caucasians and 22.8 ± 1.7 pM in African Americans, with a significant difference between the two racial groups ($P < 0.001$). Both hormones were correlated with female age (Fig. 1) and follicle number ($P < 0.05$). In Caucasian women, FSH had an extrapolated rate of increase of ~0.27 IU/l per year of female age (Fig. 1A). In African American women, similar relationships were observed, with a rise of ~0.16 IU/l of FSH per year of female age (Fig. 1A). AMH was most closely correlated with both female age (Caucasians: $R^2 = 0.22$; African Americans: $R^2 = 0.043$) and follicle number (Caucasians: $R^2 = 0.67$; African Americans: $R^2 = 0.36$; $P < 0.05$); AMH decreased with increasing age and declining AFCs. In Caucasian women, AMH decreased ~2 pM per year, while in African American women it decreased ~0.84 pM per year (Fig. 1B). Both FSH and AMH levels were less strongly correlated with age in the African American population, compared with the Caucasian population (FSH: $R^2 = 0.06$ versus 0.15; AMH: $R^2 = 0.04$ versus 0.22, respectively). FSH and AMH levels were more variable among African American women and higher follicle numbers were not always predictive of higher AMH levels.

As two good hormonal markers of ovarian reserve, FSH and AMH were independently analyzed for associations across the genome. Figure 2 shows the GWA results for FSH in African American (Fig. 2A and B) and Caucasian (Fig. 2C and D) women. Figure 3 depicts the genome-wide results for AMH. In the African American cohort, the top three genetic loci associated with FSH levels localized to chromosomes 2 and 10 (Fig. 2B and Table II). Notably, one marginally genome-wide significant SNP (rs6543833) was associated with FSH in the African American population ($P = 8.07 \times 10^{-8}$). This variant, at 2p22.3, is located just upstream of the *myeloid-associated differentiation marker-like* (*MYADML*) gene and was associated with a mean increase of 1.0 ± 0.2 IU/l and a decrease of 0.6 ± 0.3 IU/l FSH for the CC and GG genotypes, respectively (MAF = 0.27; Table II and Fig. 4A). The next most significant SNP, rs12465811, at the same loci on chromosome 2, was also associated with FSH ($P = 4.28 \times 10^{-7}$; MAF = 0.32). This variant was associated with an increase of 1.0 ± 0.2 IU/l and a decrease of 0.4 ± 0.2 IU/l FSH for the GG and TT genotypes, respectively (Table II and Fig. 4A). Regional association analyses of the 800 kb surrounding these loci indicated that these variants were in LD and several nearby SNPs were involved in the association signal (Fig. 2E). The third nominally associated variant rs11255291, localized to 10p14, upstream of the *inter-alpha (globulin) inhibitor H2* (*ITIH2*) gene and was associated with a decrease of 1.3 ± 0.2 IU/l and an increase of 1.2 ± 0.4 IU/l FSH for the TT and CC genotypes, respectively (Table II).

In the Caucasian cohort, the top three genetic loci associated with FSH localized to chromosomes 12, 5 and 11 (Fig. 2C and D). As summarized in Table II, the two most significant SNPs in the Caucasian women, rs6488619 and rs10061804, were nominally associated with FSH ($P = 9.43 \times 10^{-7}$ and 1.47×10^{-6}; MAFs = 0.25 and 0.28) and localized to 12p13.1 and 5p13.3, respectively.

Table I Variables of the study population.

	Caucasians		African Americans	
	Mean ± SEM	*n*	Mean ± SEM	*n*
Age (years)	35.3 ± 0.3	232	35.6 ± 0.4	200
FSH (IU/l)	7.1 ± 0.2	232	7.4 ± 0.3	200
AMH (pM)	30.1 ± 1.5*	232	22.8 ± 1.7*	200
LH (IU/l)	5.8 ± 0.2*	232	4.5 ± 0.2*	200
Inhibin B (pg/ml)	54.9 ± 1.9*	232	44.4 ± 2.4*	200
Estradiol (pg/ml)	48.3 ± 3.6	232	45.7 ± 1.7	200
Estrone (pg/ml)	61.7 ± 1.6	232	63.3 ± 1.6	200
Total testosterone (ng/dl)	52.9 ± 1.0*	232	45.3 ± 1.0*	200
Past/current smokers[a]	45.5% (15.9 ± 1.0)	106	23.3% (14.3 ± 1.4)	47
Non-smokers[a]	54.5% (14.9 ± 0.7)	127	76.6% (16.0 ± 0.9)	154
OCP users[a]	79% (14.9 ± 0.6)	195	72% (14.3 ± 0.8)	145
Non-OCP users[a]	21% (17.2 ± 1.3)	48	28% (18.9 ± 1.6)	56

Values are shown as mean ± SEM or percentage of the cohort.
AMH, anti-Mullerian hormone; OCP, oral contraceptive pill.
[a]Parentheses denote AFC values (mean ± SEM).
*Significant differences between Caucasian and African American women ($P < 0.001$).

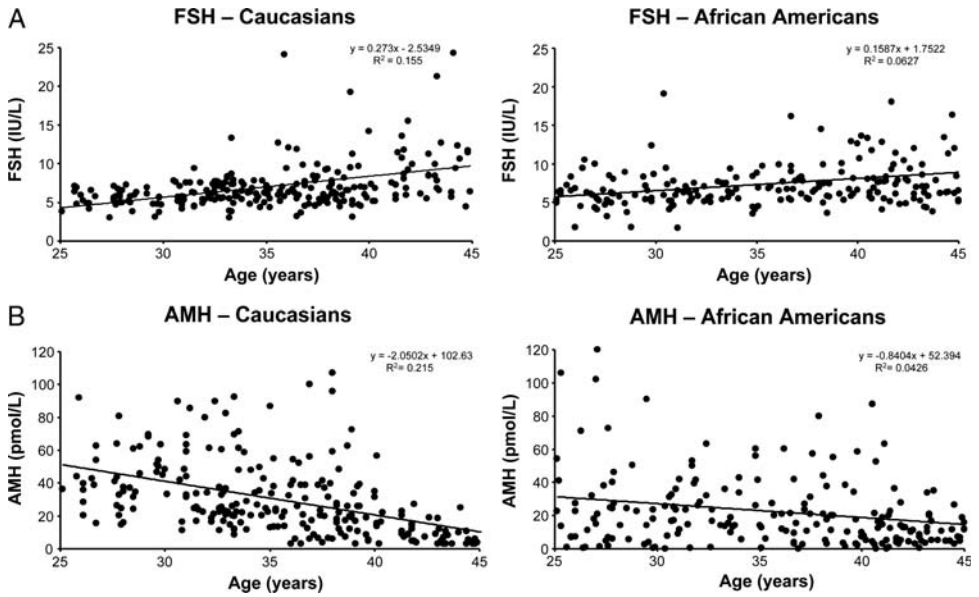

Figure 1 Measurements of FSH and AMH versus age in women of the entire study population. Total serum concentrations of FSH and female age in Caucasian and African American women indicate that FSH increases with age in both ethnic groups (**A**). Total serum concentrations of AMH versus age indicate that AMH decreases with age but is highly variable between women and is more variable among African American women (**B**). The corresponding correlation coefficients (R^2) and linear equations are shown ($n = 200–232$ women).

Variant rs6488619, is located within an intron of the *glutamate receptor, ionotropic, N-methyl D-aspartate 2B* (*GRIN2B*) gene and led to a mean decrease of 1.1 ± 0.4 IU/l and an increase of 1.2 ± 0.3 IU/l FSH for the CC and TT genotypes, respectively. Analysis of the surrounding region indicated that several variants at this loci were associated with FSH and a few of these were in LD (Supplementary data, Fig. S3A). Variant rs10061804 on chromosome 5 is located within an intron of the *natriuretic peptide receptor C/guanylate cyclase C* (*NPR3*) gene and was associated with increases of 2.1 ± 0.8 and 1.1 ± 0.4 IU/l FSH for the TT and CT genotypes, respectively. There were several SNPs in close proximity at the 5p13.3 region that approached chromosome-level significance (see Fig. 2B). The last SNP, rs4944804 at 11q13.4, weakly associated with FSH, is located immediately upstream of the *ATG autophagy related 16-like 2* (*Saccharomyces cerevisiae*) (*ATG16*) gene and was associated with an increase of 1.2 ± 0.4 IU/l FSH for the GG genotype.

There were fewer genetic variants associated with AMH compared with FSH, in both Caucasian and African American women. In the African American cohort, weakly associated variants localized to 3p25.1 and 11p15.1 near or within the *tetra-peptide repeat homeobox-like* (*TPRXL*) and *transmembrane protein 86A* (*TMEM86A*) genes, respectively (Table III and Fig. 3A and B). The top genetic loci in the Caucasian cohort were located on chromosome 6, with only very weakly associated variants on chromosomes 18, 4 and 13 (Fig. 3C and D). The most significant SNP rs12213875, which was nominally associated with AMH ($P = 2.96 \times 10^{-7}$; MAF = 0.37) localized to 6p23 and is upstream of the *Jumonji, AT rich interactive domain 2* (*JARID2*) gene. This variant was associated with a mean increase of 13.8 ± 4.2 pM, an increase of 9.9 ± 2.1 pM and a decrease of 1.3 ± 1.4 pM AMH for the AA, AG and GG genotypes, respectively (Table III and Fig. 4B). Several variants adjacent to the top SNP associated with AMH in both ethnic groups showed some degree of LD and many variants in these regions were associated with the signals (Supplementary data, Fig. S4). Many other SNPs were only weakly associated with AMH levels and are located within or near the *thioredoxin-related transmembrane protein 3* (*TMX3*), *inositol polyphosphate-4-phosphatase, type II* (*INPP4B*) and *WAS protein family, member 3* (*WASF3*) genes. Interestingly, several variants associated with FSH or AMH (this study), or AFC (previous unpublished work), localized to the same genomic regions of 12p13.1–13.2 and 13q12.13.

In replication studies, the top variants associated with FSH were examined in the Caucasian population as an independent replication cohort. The top SNP rs6543833, had a *P*-value of 0.012 with an MAF of 0.32, while the next most significant SNP rs12465811, had a *P*-value of 0.071 with an MAF of 0.33 (Supplementary data, Table S1). When only controlling for these two tests, the top SNP was significant in the replication cohort at a $P < 0.05$ level. As shown in Supplementary data, Table S1, the 300 SNPs flanking this variant were also examined and it was discovered that nine adjacent SNPs had relatively low *P*-values of 1.5×10^{-3} to 7.2×10^{-3}. The regional plots of these variants and their association signals are also indicated in Supplementary data, Fig. S3B. Comparing the expected versus the observed numbers of SNPs in this region with these low *P*-values indicates that this cluster of SNPs was unlikely the result of random chance. There were three times as many SNPs with low *P*-values than expected based on observed *P*-value distributions. These findings provide confirmation of the top SNP as well as other associated loci in the surrounding region. Interestingly, all of these variants occurred within or upstream of the *MYADML* gene, which may represent an

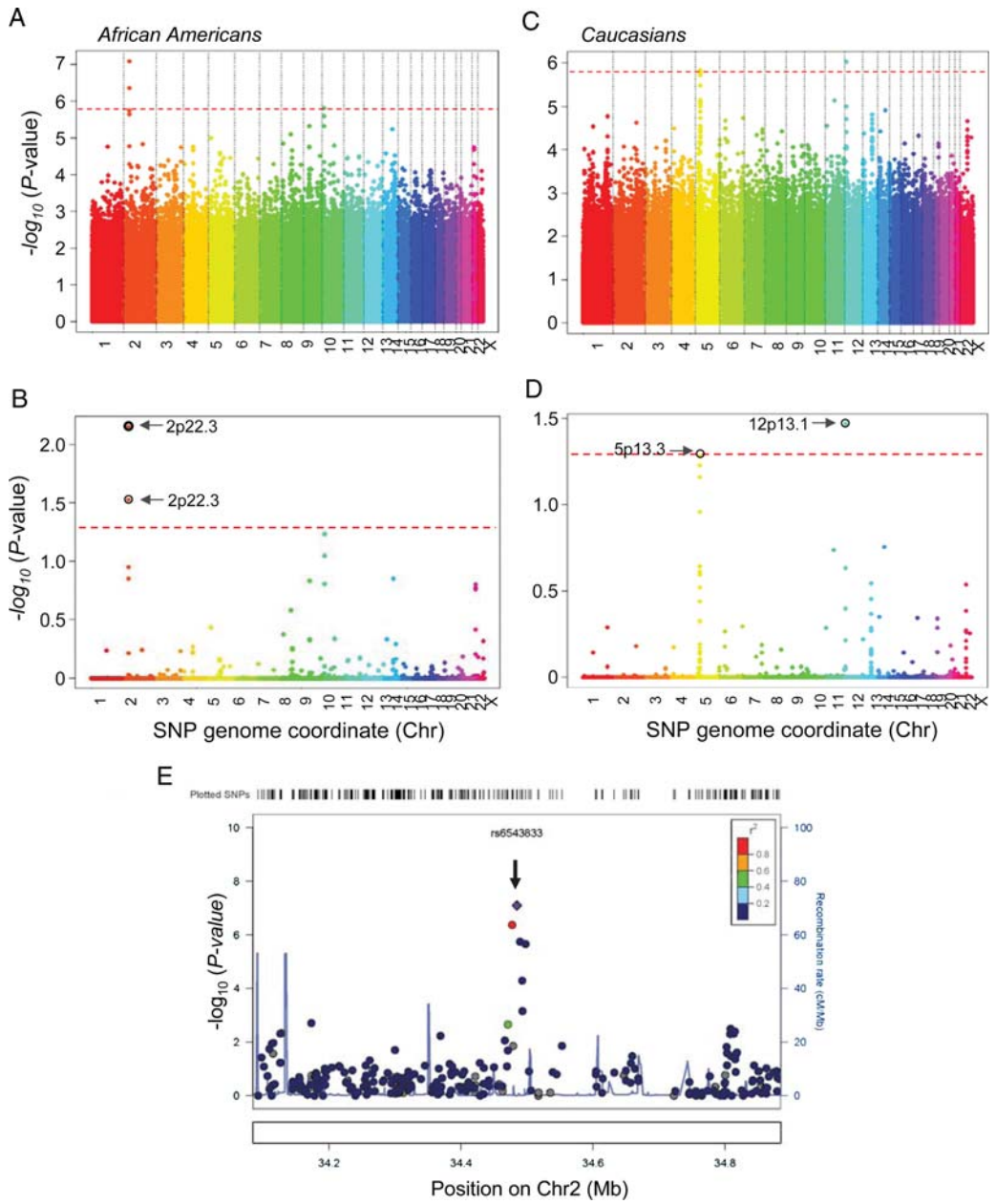

Figure 2 Summary of GWA results for FSH by chromosome showing Manhattan plots of the uncorrected (**A** and **C**) and corrected (**B** and **D**) *P*-values of the Fisher's exact test in the African American (A and B) and Caucasian (C and D) populations. Each point represents a SNP from the single SNPs and haplotype tests of association remaining after QC and SNP filtering (677 261 in the Caucasian cohort and 738 185 in the African American cohort). The red dashed lines indicate *P*-values of 1.5×10^{-6} (uncorrected) and 0.05 (corrected), chromosome-wide significance level. (**E**) Regional association analyses of 800 kb (272 SNPs) flanking the top variant associated with FSH in the African American cohort (rs6543833), including the number of SNPs involved in the signal, their *P*-values, the LD among these variants, genes in the region and the recombination rates (hg build 18, HapMap Phase II CEU reference, legend showing LD).

excellent candidate gene for future studies in women of both European and African ancestry. Replication of these findings in independent cohorts of reproductive-aged women and meta-analyses in larger combined GWA studies are needed.

In an attempt to further examine the genetic variants associated with FSH and AMH, we tested the top hits for association with another well-known marker of ovarian reserve—AFC. A total of four and five SNPs associated with FSH or AMH levels in Caucasian or African American women, respectively, were tested for association with AFC. Final *P*-values were corrected for the number of tests and considered significant at a level of $P < 0.05$. Effect sizes in terms of direction (higher or lower FSH/AMH levels or AFCs) and magnitude (total differences in the mean hormone concentration or mean number of follicles from the overall regression fit) for the effect

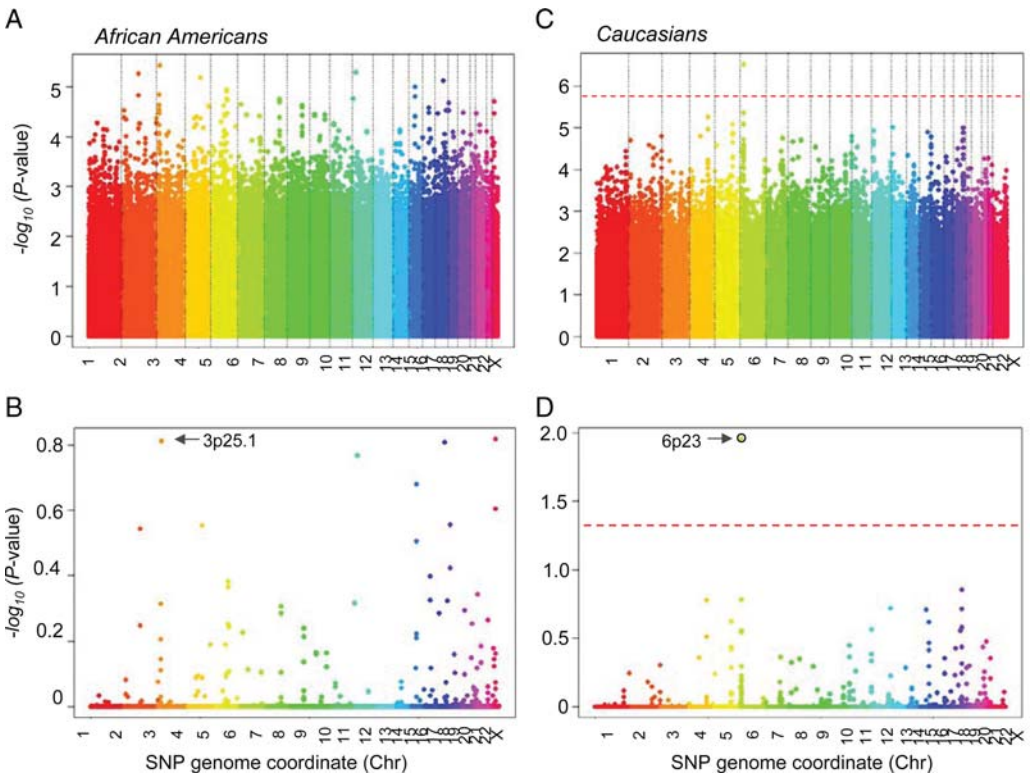

Figure 3 Summary of GWA results for AMH by chromosome showing Manhattan plots of the uncorrected (**A** and **C**) and corrected (**B** and **D**) *P*-values of the Fisher's exact test in the African American (A and B) and Caucasian (C and D) populations. Each point represents a SNP from the single SNPs and haplotype tests of association remaining after QC and SNP filtering (677 261 in the Caucasian cohort and 738 185 in the African American cohort). The red dashed lines indicate *P*-values of 1.5×10^{-6} (uncorrected) and 0.05 (corrected), chromosome-wide significance level.

allele/genotype were compared for all variants tested. Interestingly, as shown in Table IV several variants associated with FSH or AMH were also associated with the number of antral follicles. The most significant variant, rs6543833, was associated with a decrease of 0.62 ± 0.34 IU/l FSH and an increase of 1.09 ± 0.32 antral follicles for the GG genotype ($P = 9.25 \times 10^{-3}$). Variant rs9875589 associated with AMH in African Americans, had the strongest association with follicle count; the effect allele, C, led to an increase of 14.02 ± 3.72 pM AMH and a corresponding increase of 4.88 ± 0.63 follicles ($P = 1.84 \times 10^{-4}$). Notably, in almost all cases the magnitudes and directions of effects were as expected, with variants associated with lower FSH levels correlated with higher follicle counts and variants associated with higher AMH levels correlated with higher follicle counts. Controlling for follicle number in the GWA of FSH/AMH did not affect the top associations. Further, there were both shared and independent associations between the hormonal markers and follicle number, indicating that both of these traits are independently associated with distinct genetic variants and are significant markers of ovarian reserve.

Phenotypes and covariates of ovarian reserve

Serum levels of various circulating hormones potentially related to ovarian function and OVA were measured as shown in Table I. The means and ranges of hormone levels were within the spectrum of normal variation. There were significant differences between the hormone profiles of the two racial groups, with pronounced differences in AMH ($P = 0.00074$), LH ($P = 2.28 \times 10^{-5}$), inhibin B ($P = 0.00049$) and testosterone ($P = 0.00041$). All of these hormones were significantly lower in the African American women. The associations between environmental variables including smoking and OCP use, and ovarian reserve markers were also analyzed. Among the Caucasian population, 45.5% were either current or previous smokers (14.6 current and 30.9% previous smokers) with an overall average of 9.6 ± 0.6 years of smoking and a range of 1–25 years. Fewer African Americans were current or previous smokers with only 23.3% being current (11.4%) or previous (11.9%) smokers, with an overall average of 9.8 ± 1.2 years of smoking and a range of 1–27 years. There was no significant difference between the mean AFCs of women that were current or previous smokers versus that of non-smokers in the Caucasian cohort ($P = 0.36$) nor in the African American cohort ($P = 0.34$). In the Caucasian population, when controlling for age there was no correlation between the total pack-years of smoking and follicle number ($R^2 = -0.057$, $P > 0.05$). However, there was a trend for decreased follicle numbers among the smoking versus non-smoking African American women (14.3 ± 1.4 versus 16.0 ± 0.9, respectively). Although, owing to the small number of African American smokers further comparisons were not feasible.

As shown in Fig. 5, the circulating serum levels of several reproductive and sex steroid hormones were significantly different between

Table II Variants associated with FSH levels.

SNP	Chr	Gene	Position	Cytoband	Alleles[a]	MAF	AA effect[b]	AB effect[b]	BB effect[b]	*P*-value[c]	Corrected *P*-value[d]
African Americans											
rs6543833	2	*MYADML*	34631032	P22.3	G/C	0.27	−0.62 ± 0.34	0.46 ± 0.17	0.96 ± 0.21	8.07×10^{-8}	*0.0070 (0.059)*
rs12465811	2	*MYADML*	34623796	P22.3	T/G	0.32	−0.38 ± 0.24	0.58 ± 0.18	0.92 ± 0.20	4.28×10^{-7}	0.030 (0.32)
rs11255291	10	*ITIH2*	7736904	P14	T/C	0.25	−1.34 ± 0.22	0.17 ± 0.40	1.21 ± 0.38	1.53×10^{-6}	0.059 (1)
Caucasians											
rs6488619	12	*GRIN2B*	13927896	P13.1	C/T	0.25	−1.10 ± 0.35	0.15 ± 0.24	1.16 ± 0.33	9.43×10^{-7}	0.034 (0.65)
rs10061804	5	*NPR3*	32748637	P13.3	T/C	0.28	2.05 ± 0.80	1.08 ± 0.44	0.091 ± 0.23	1.47×10^{-6}	0.052 (1)

Genes: *MYADML*, myeloid-associated differentiation marker line; *ITIH2*, inter-alpha (globulin) inhibitor H2; *GRIN2B*, glutamate receptor, ionotropic, N-methyl D-aspartate 2B; *NPR3*, natriuretic peptide receptor C/guanylate cyclase C; SNP, single nucleotide polymorphism; Chr, chromosome; MAF, minor allele frequency.
[a]Alleles are shown as minor/major allele.
[b]The difference in mean (±SEM) serum FSH levels (IU/l) for a given genotype. 'A' denotes the minor allele and 'B' denotes the major allele.
[c]*P*-values are based on the Fisher's exact test for SNP association with FSH.
[d]*P*-values corrected at the chromosome-level (and genome-wide level) by permutation testing and Bonferroni correction.

smoking and non-smoking Caucasian women. Testosterone levels were 56 ± 1.6 versus 50.3 ± 1.2 ng/dl ($P = 0.0035$), AMH levels were 33.5 ± 2.4 versus 27.4 ± 1.7 pM ($P = 0.038$) and E_1 levels were 65.2 ± 2.8 versus 58.8 ± 1.7 pg/ml ($P = 0.043$), among smoking and non-smoking women, respectively (Fig. 5G). The regression fit of these and other hormones with female age was also significantly different between smoking and non-smoking women, with the most pronounced differences in E_2 and testosterone levels ($P = 0.0045$ and 0.0053, respectively; Fig. 5E and F). The regression of AFC versus age was also significantly different between smoking and non-smoking women, with a greater extrapolated rate of follicle decline among smoking women ($P = 0.039$; Fig. 5A).

Seventy-nine percent of the Caucasian women and 72% of the African American women had previously used OCPs. Total months of use were 63.0 ± 4.2 (range = 1–216 months) in the Caucasian population and 56.0 ± 5.4 (range = 1–264 months) in the African American population. In the Caucasian women, there were no statistically significant differences between the mean serum hormone levels or AFCs of women who previously took OCPs versus those who did not ($P > 0.05$). There was also no dose–response or association between the total number of months of OCP use and the total number of follicles ($R^2 = 0.014$ and 1×10^{-6} in Caucasians and African Americans, respectively). In contrast, in the African American population there was a significant increase in the number of follicles in women who had never taken OCPs ($P = 0.0046$; Fig. 5H). In both ethnic groups, there were trends for increased follicle numbers in women who had never taken OCPs compared with those who had (Caucasians: 17.2 ± 1.3 versus 14.8 ± 0.6, respectively, $P = 0.058$; African Americans: 18.9 ± 1.6 versus 14.3 ± 0.8, $P = 0.0046$; Fig. 5H). There were similar trends with several reproductive hormones and AFC regressed on age, between previous OCP-users and non-OCP-users (Supplementary data, Fig. S5). Despite the effects of smoking and OCP use on ovarian reserve phenotypes, adjusting for them in linear regression analyses did not affect the genetic associations. This indicates that the effect of each variant likely occurs directly on ovarian reserve/hormone levels, rather than through one of these covariates.

Discussion

Genetic variants associated with FSH and AMH

As FSH and AMH are two established clinical markers of ovarian reserve and are closely associated with the number of antral follicles within the ovaries, we aimed to determine whether they might be associated with underlying genetic loci. Identification of these loci could provide markers and long-range predictors of poor ovarian reserve and infertility. Here, we used a genome-wide approach to analyze genetic variants associated with these reproductive hormones, as surrogate measures of ovarian reserve in a multi-ethnic fertile population of women. We identified 1 SNP, rs6543833, approaching genome-wide significance at $P = 8.07 \times 10^{-8}$ and several nominal SNPs associated with FSH levels in African American and Caucasian women, respectively, whereas, only nominally significant variants were associated with AMH levels. In validation studies, the top hit associated with FSH was significant in the

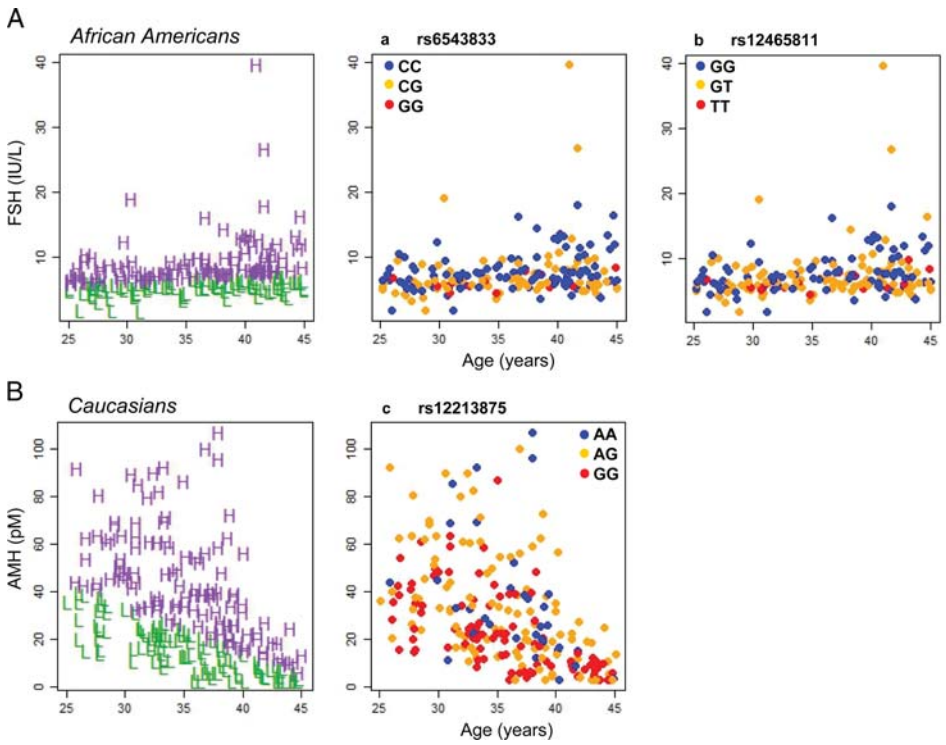

Figure 4 Phenotypes and genotypes of SNPs associated with FSH and AMH. Shown are scatter plots depicting high (H) and low (L) FSH and AMH by age in African American (**A**) and Caucasian (**B**) women. Shown is the distribution of genotypes among the women with respect to age and FSH or AMH for the most significant variants: (a) rs6543833 and (b) rs12465811 (for FSH) and (c) rs12213875 (for AMH). The three genotypes are designated by blue, orange and red points ($n = 200–232$ women).

other ethnic cohort (Caucasians). Additionally, there were nine adjacent SNPs in the nearby upstream region with significantly low *P*-values, and all were within or near the *MYADML* gene. We also found that several top hits associated with FSH or AMH were also significantly associated with follicle number, with similar magnitudes and directions of effects. These results provide validation of the association and evidence for underlying genetic associations between these three phenotypes, as well as their ability to serve as independent but also overlapping ovarian reserve markers. Moreover, several variants associated with AMH or FSH (this study), and follicle count (unpublished data), localized to the same genomic regions of 12p13.1–13.2 and 13q12.13, which may be a hot spot for causative alleles linked with ovarian reserve.

None of the variants were within or close to known ovarian genes and therefore identify excellent candidate genes for further investigations. While these findings are interesting and the largest study to date on genetic markers of FSH and AMH, future meta-analyses and additional validation in independent studies are needed. To our knowledge, this is the only study to genotype and examine ovarian reserve hormones in a fertile population of women. All previous cohort studies that have analyzed some aspect of ovarian reserve have utilized either post-menopausal women (over age 45 years) or women recruited through infertility/reproductive clinics and are therefore not reproductively 'normal' populations. This work is the first to analyze a cohort of women (non-menopausal) of reproductive age recruited from the general population and not seeking infertility treatments. This study is also the first to examine these reproductive traits in women of African as well as European ancestry, and may help to provide a baseline and foundation for future study of these ethnic groups.

As indicated by our previous work in a larger population of 245 Caucasian women, and validated here in both Caucasians and African Americans, of all hormonal, anthropometric and environmental variables, AMH and FSH were most closely associated with female age and follicle number. AMH paralleled the pattern observed between follicle number and age both *in vivo* and in histologic ovarian specimens (Hansen *et al.*, 2008), while FSH was inversely related. Consistent with the role of AMH, we found that in Caucasian and African American women AMH and follicle number were closely related and both declined with female age. Interestingly, AMH levels were more variable among African American women compared with Caucasian women, and showed a weaker correlation with age. As predicted, FSH levels increased with both female age and decreasing follicle number in both ethnic groups, and surprisingly, were also more variable and less strongly correlated with age among African American than Caucasian women. It is interesting to note that none of the subjects had polycystic ovary syndrome or other reproductive disorders, which could be associated with elevated AMH or FSH levels. All women also had regular cycles and no hyperandrogenemia. Therefore, the greater variability cannot be explained by ovarian or other hormonal disorders. To our knowledge there is no other study with which to compare these results and no other study to

Table III Variants associated with AMH levels.

SNP	Chr	Gene	Position	Cytoband	Alleles[a]	MAF	AA effect[b]	AB effect[b]	BB effect[b]	P-value[c]	Corrected P-value[d]
African Americans											
rs9875589	3	*TPRXL*	139611157	P25.1	A/C	0.46	−4.17 ± 1.82	4.96 ± 2.35	14.02 ± 3.72	3.70×10^{-6}	0.15 (1)
rs12295403	11	*TMEM86A*	18689653	P15.1	T/G	0.10	45.99 ± 0.00	16.67 ± 5.11	2.98 ± 1.68	5.07×10^{-6}	0.17 (1)
Caucasians											
rs12213875	6	*JARID2*	14694982	P23	A/G	0.37	13.82 ± 4.18	9.91 ± 2.07	−1.25 ± 1.41	2.96×10^{-7}	0.011 (0.20)
rs9396503	6	*JARID2*	14644068	P23	T/G	0.37	14.92 ± 4.60	8.26 ± 1.97	−0.07 ± 1.50 0	4.28×10^{-6}	0.16 (1)

Genes: *TPRXL*, tetra-peptide repeat homeobox-like; *TMEM86A*, transmembrane protein 86A; *JARID2*, jumonji, *AT* rich interactive domain 2.
[a]Alleles are shown as minor/major allele.
[b]The difference in mean (±SEM) serum AMH levels (pM) for a given genotype. 'A' denotes the minor allele and 'B' denotes the major allele.
[c]P-values are based on the Fisher's exact test for SNP association with AMH.
[d]P-values corrected at the chromosome-level (and genome-wide level) by permutation testing and Bonferroni correction.

assess these hormones in normal fertile women of African and European ancestry.

Several hormones, including FSH, AMH, E_2 and inhibin B, among others, have been correlated with ovarian response to controlled ovarian stimulation or pregnancy outcomes in IVF cycles (Eldar-Geva *et al.*, 2005; Wunder *et al.*, 2008; Wu *et al.*, 2009). Studies have reported that AMH may be the best marker of ovarian reserve as it can be measured throughout the menstrual cycle and is more closely associated with the number of oocytes retrieved and reproductive outcomes in IVF (Durlinger *et al.*, 2002; Wunder *et al.*, 2008; Wu *et al.*, 2009). AMH also has a reported lower variability within and between cycles in the same women (van Disseldorp *et al.*, 2010). Previous work has also identified a few variants in the genes encoding the receptors for AMH and FSH, AMHR2 (Kevenaar *et al.*, 2007) and FSHR (Gromoll and Simoni, 2005; Loutradis *et al.*, 2006), and found correlations with various reproductive parameters including age at menopause, or FSH sensitivity and menstrual cycle length, respectively. Interestingly, a recent large-scale candidate gene association study found that the group of genes involved with steroid hormone metabolism and biosynthesis pathways including FSHB, was significantly associated with age of menopause (He *et al.*, 2010). This indicates that several reproductive (i.e. FSH) and sex steroid hormones may share underlying genetic associations that are linked with the reproductive lifespan. However, many of these studies relied on comparisons of hormone profiles and their ability to predict reproductive outcomes in the context of small populations of infertile or subfertile (or post-menopausal) women. Thus, we now identify genetic variants associated with markers of ovarian reserve in a large community-based population.

None of the genes associated with the identified variants have demonstrated roles in ovarian biology and are therefore interesting candidate genes. As FSH and AMH are related to ovarian function and reflective of ovarian reserve, the genes and variants associated with them might have roles in the brain-pituitary-ovary axis and be expressed within the ovary or brain—and some of them are. The two top variants, as well as other nearby variants associated with FSH are located upstream or within the *MYADML* gene at 2p22.3. *MYADML* has almost exclusive expression in the human testis and limited expression in the brain (www.ncbi.nlm.nih.gov/unigene). However, the role of this gene in ovarian biology is unknown. The *GRIN2B* gene containing the top SNP at 12p13.1 associated with FSH in Caucasian women, encodes one subunit of the ionotropic NMDA glutamate receptor, which is involved in learning and memory and expressed within the brain and testis. Genetic association studies have implicated this gene in several neurological diseases (Allen *et al.*, 2008), however studies of this receptor in ovarian function are lacking. The two variants marginally associated with AMH in Caucasian women are located upstream of the *JARID2* gene at 6p23. *JARID2* is an ortholog of the mouse jumonji gene, which encodes a DNA-binding nuclear protein and regulator of histone methyltransferases that is essential for embryogenesis and differentiation. JARID2 negatively regulates cell growth and proliferation and is ubiquitously expressed, with expression in both human and mouse oocytes and ovary, and human embryonic stem cells (Ovarian Kaleidoscope: http://ovary.stanford.edu/) (Ben-Shlomo *et al.*, 2002). Therefore, JARID2 could play a role in germ cell and/or somatic cell growth within the developing ovary.

Table IV Associations between FSH- and AMH-related SNPs and AFC in African American and Caucasian women.

SNP[a]	Race	Gene	Alleles[b]	MAF	FSH/AMH genotype effect (concentration)[c]	AFC genotype effect (follicles)[d]	*P*-value[e]
rs6543833	Afr Amer	*MYADML*	**G**/C	0.27	−0.62 ± 0.34 (FSH)	+1.09 ± 0.32	1.85×10^{-3} (9.25×10^{-3})*
rs12465811	Afr Amer	*MYADML*	**T**/G	0.32	−0.38 ± 0.24 (FSH)	+0.94 ± 0.25	1.31×10^{-3} (6.55×10^{-3})*
rs9875589	Afr Amer	*TPRXL*	A/**C**	0.46	+14.02 ± 3.72 (AMH)	+4.88 ± 0.63	3.68×10^{-5} (1.84×10^{-4})*
rs12295403	Afr Amer	*TMEM86A*	**T**/G	0.10	+16.67 ± 5.11 (AMH)	+5.17 ± 0.86	3.34×10^{-3} (1.67×10^{-2})*
rs11255291	Afr Amer	*ITIH2*	**T**/C	0.25	−1.34 ± 0.22 (FSH)	+1.31 ± 0.36	0.211 (1)
rs6488619	Caucasian	*GRIN2B*	C/**T**	0.25	+1.16 ± 0.33 (FSH)	+0.55 ± 0.049	0.104 (0.416)
rs10061804	Caucasian	*NPR3*	**T**/C	0.28	+2.05 ± 0.80 (FSH)	+1.44 ± 0.30	0.310 (1)
rs12213875	Caucasian	*JARID2*	**A**/G	0.37	+13.82 ± 4.18 (AMH)	+3.08 ± 0.52	0.0514 (0.206)
rs9396503	Caucasian	*JARID2*	**T**/G	0.37	+14.92 ± 4.60 (AMH)	+3.74 ± 0.66	0.0415 (0.166)

$P < 0.05$ for four out of five FSH/AMH-associated SNPs also associated with AFC, compared with 0.34 (<1) expected by chance.
[a]SNPs at each locus are those associated with serum FSH or AMH levels, rather than those with the strongest signal for follicle number.
[b]Alleles are shown as minor/major allele; bold indicates the effect allele.
[c]The difference in mean (±SEM) serum FSH levels (IU/l) or AMH levels (pM) for a given genotype.
[d]The difference in mean (±SEM) number of follicles for the corresponding genotype calculated from the regression analysis.
[e]*P*-values are based on the Fisher's exact test for SNP association with the hormone; parentheses denote corrected *P*-values; *$P < 0.05$.

Interestingly, several variants associated with FSH or AMH were also associated with AFC and variants associated with all three ovarian reserve markers localized to some of the same genomic regions. A variant associated with AMH (rs17084228) and a variant associated with AFC, were located within the same 30-kb genomic region of 13q12.13, where the nearest genes are *GPR12* and *WASF3*. Both genes are predominantly expressed in the brain, with *WASF3* also highly expressed in the ovary. One SNP associated with FSH (rs6488619), and two other SNPs associated with AFC, were also located in a similar genomic region at 12p13.1 and 12p13.2, respectively. Therefore, these genomic regions on chromosomes 12 and 13, which also comprised long stretches of associated SNPs in LD, may be hot spots for causative alleles or genes associated with ovarian reserve. Further, as all of these ovarian reserve markers had both distinct and shared genetic loci, this emphasizes the multifactorial and polygenic nature of ovarian reserve and female fertility. Future functional studies may examine the expression of these genes in human ovarian tissue and female germ cells, altered gene expression in ovarian tissue from women with different genotypes, and genetic over-expression and silencing studies in ovarian and stem cell lines. This work will help to elucidate the expression patterns and functional roles of these candidate genes in human ovarian biology.

Environmental factors associated with ovarian reserve

Environmental factors can affect human reproductive traits. We discovered modest effects of smoking and OCP use on clinical markers of ovarian reserve, with alterations in several reproductive and sex steroid hormones, as well as follicle numbers. Smoking has adverse effects on fertility and pregnancy. There is continuing debate about the relationship between cigarette smoke and ovarian failure. Studies have reported an association between smoking and an earlier age at natural menopause, with women who smoke entering menopause 1–2 years earlier than comparable non-smokers (Torgerson *et al.*, 1994; Bromberger *et al.*, 1997; van Noord *et al.*, 1997; Cooper *et al.*, 1999; Gold *et al.*, 2001). Former smokers also reportedly have a slightly earlier age of menopause than non-smokers (Willett *et al.*, 1983; Cooper *et al.*, 1999) and women prenatally exposed to maternal cigarette smoke may enter menopause sooner (Strohsnitter *et al.*, 2008). Limited studies on special patient/infertility populations have suggested that smoking is not associated with AFC or AMH levels (Kinney *et al.*, 2007; Nardo *et al.*, 2007) but is associated with elevated FSH levels (Kinney *et al.*, 2007).

It is plausible that tobacco toxins have detrimental effects on the follicular pool and increase rates of follicular atrophy and atresia, leading to lower AFCs, alterations in ovarian reserve hormones, and ultimately, reduced fertility and an earlier menopause. We found that smoking was associated with increased serum levels of testosterone, E_1, E_2, AMH and a trend for increased FSH levels. This underscores the potential effects of smoking on the ovarian reserve and suggests that smoking may alter steroidogenesis, OVA and/or the hypothalamic-pituitary-ovarian axis. Indeed, previous work has found that levels of sex steroid hormones, such as testosterone, are significantly elevated in smokers (Deslypere and Vermeulen, 1984; Dai *et al.*, 1988), while other studies have reported alterations in sex steroids including lower concentrations of E_1, E_2 and estriol (Soldin *et al.*, 2011).

Other effects of tobacco toxins include increased rates of apoptosis and necrosis in various tissues of the body, as well as increased apoptosis in primordial germ cells differentiated from human stem cells *in vitro* (Kee *et al.*, 2010). These effects might also manifest as increased

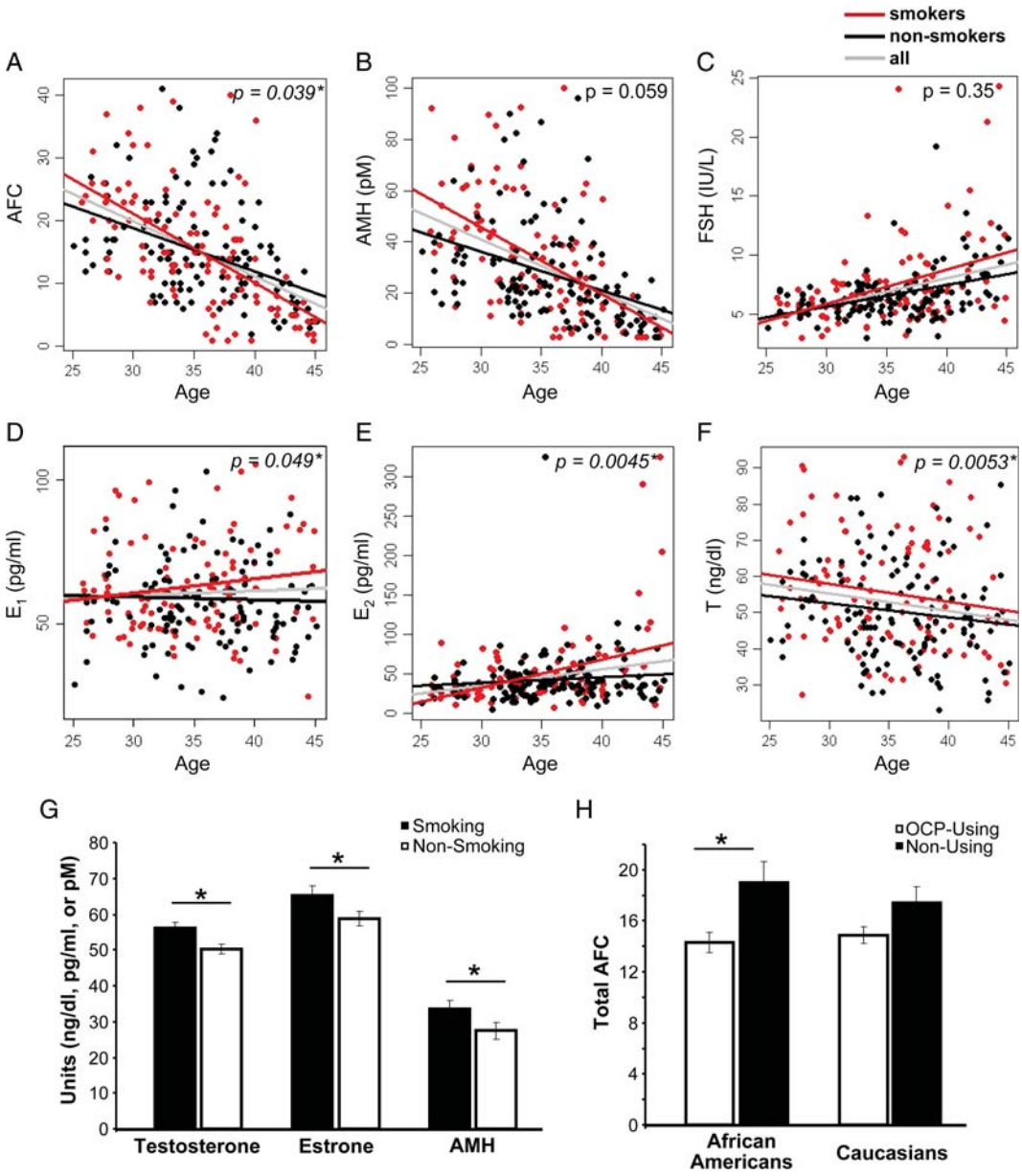

Figure 5 The association between smoking or oral contraceptive use and several reproductive hormones and follicle counts. Shown are serum hormone levels or AFCs versus age for women that were current or previous smokers (red points) or non-smokers (black points) (**A–F**). The data for all women were fit with a linear regression (gray line) and the regressions for smoking (red line) and non-smoking (black line) women were compared. *P*-values indicate a significant change in slope, intercept or both slope and intercept, associated with smoking. There were significant differences in serum levels of estrone (E_1), estradiol (E_2) and testosterone (T), and AFC between smoking and non-smoking women ($P < 0.05$). Also shown are mean serum levels of testosterone, E_1 and AMH in smoking (closed bars) and non-smoking (open bars) women (**G**) and mean AFCs in oral contraceptive-using (open bars) and non-using (closed bars) African American and Caucasian women (**H**) (*$P < 0.05$).

rates of follicular apoptosis or dysfunction within the ovaries. Indeed, polycyclic aromatic hydrocarbons in cigarette smoke are toxic to ovarian follicles in both animal models (Essenberg *et al.*, 1951) and humans (Mattison and Thorgeirsson, 1978). We found that neither previous nor current smoking was associated with overall lower follicle counts. However, smoking was associated with an increased extrapolated rate of follicular decline in Caucasian women ($P = 0.039$). There was no dose-response between the pack-years of smoking and follicle number ($P > 0.05$). Our findings are consistent with the work of Gold and co-workers, one of the largest, most comprehensive studies on menopause (Gold *et al.*, 2001). In 14 620 middle-aged women of the Study of Women's Health Across the Nation, smoking was associated with an earlier age of menopause (0.3–1.2 years earlier) but there was no dose–response effect of the amount of cigarettes smoked per day on menopausal age. Similarly, we found no dose–response of smoking but smoking was positively correlated with an

increased rate of follicular decline. There was also a positive correlation between female age and number of years of smoking, with older women on average having smoked for a greater number of years. Any effects of smoking on the follicle pool therefore, might be most evident in mid-life.

A slightly later age of menopause has been reported in women that previously used OCPs (van Noord *et al.*, 1997; Gold *et al.*, 2001), while other studies have reported no association (Bromberger *et al.*, 1997; Hardy and Kuh, 1999). It has been hypothesized that OCP use delays menopause by the hormonal actions of suppressing follicular recruitment and ovulation, thereby leading to increased numbers of follicles. In contrast, we found that OCP use did not increase follicle number. Instead, women who had taken OCPs had similar or slightly lower follicle numbers compared with women who had never taken OCPs ($P = 0.058$ in Caucasians and $P = 0.0046$ in African Americans). Similar to previous work on menopause (Gold *et al.*, 2001), we found no relation between duration of prior OCP use and follicle number or serum hormone levels. These results suggest that OCP use is not associated with increased follicle number or alterations in ovarian reserve hormones. It remains speculative that oral contraceptives may actually have some subtle inhibitory effects on follicle numbers or ovarian reserve in women of reproductive age. However, as we are limited by sample size, longitudinal studies analyzing these lifestyle factors in larger cohorts of women could provide further insight on the biological effects of these exposures on the ovarian reserve. It is likely that complex genetic and environmental influences affect follicular atresia, as well as the brain-pituitary-ovary axis and its regulation of gonadotrophins and sex steroid hormones.

Conclusion

This is the first genome-wide study to evaluate genetic associations with hormone markers of ovarian reserve. We found several differences between the ovarian reserve phenotypes of Caucasian and African American women—differences in AMH levels and subtle differences in the associations between hormonal markers with age and follicle number. Smoking and OCP use were associated with alterations in reproductive hormone profiles and/or follicle number in both ethnic groups. We identified one variant of marginal genome-wide significance and several nominally-significant variants associated with FSH levels in African American women, as well as nominal variants associated with FSH and AMH levels in both ethnic groups. These variants were validated in the Caucasian cohort. Several of the variants were also significantly associated with follicle number. Interestingly, none of the associated genes have previously demonstrated roles in ovarian function and now represent a group of candidate genes for further investigation. Future studies including multi-site meta-analyses on additional larger cohorts of women and replication in independent studies are needed. Fine mapping or sequencing and functional *in vitro* gene expression and gain- and loss-of-function studies on the identified candidate genes might reveal any effects of the genetic variants and the putative functions of these genes. Longitudinal studies are also needed to explore the effect of these variants and genes on the decline in ovarian reserve with age.

A better understanding of the genes and environmental factors that impact ovarian reserve may provide greater insight on the process of reproductive aging, the genetic requirements of human fertility and risk for development of disease. It is hoped that this work may help form a foundation for genetically identifying women at risk for early ovarian failure, fertility defects and hence associated reproductive and somatic diseases.

Supplementary data

Supplementary data are available at http://humrep.oxfordjournals.org/.

Acknowledgements

The authors express their sincere thanks to the OVA Study participants who made this work possible. The authors also thank Elizabeth Zuo and Natalia Kosovilka of the PAN Facility for their assistance with microarray processing and Marty Flores and Alison Lerner for assistance with sample maintenance, at Stanford University. The authors also sincerely acknowledge Dr Daniel McConnell at the CLASS Laboratory of the University of Michigan for assistance with serum hormone assays. The authors kindly thank all the clinical research staff not in our co-author list at both UCSF and Kaiser DOR for their efforts in facilitating the coordination of study participants and maintaining subject databases, especially Deborah Trevithick, Margaret Takeda, Tatiana Carranza, Chim Lau and Carolyne Addauan-Andersen (at UCSF) and Rosemary Murphy, Mariana Pereyra, Katy Leung, Liliana Valderramos-Metzger, Claudia Cruz and Maila Martin (at Kaiser).

Authors' roles

S.M.S.H. was involved with all aspects of this manuscript including study design, execution and all genetic and statistical analyses. She also wrote and edited the manuscript. All authors reviewed the manuscript and provided critical feedback and discussion. N.A.J. wrote custom-made script and algorithms to mediate streamlined data analyses and visualization of results, and assisted with statistical analyses. M.I.C. and M.P.R. collected all clinical data including antral follicle counts, as well as subject questionnaires and body measurements with the assistance of UCSF clinical research staff. M.I.C., R.A.R.P. and B.S. conceived and initiated the collaborative OVA Study with the goals of identifying clinical and genetic markers of ovarian aging.

Funding

This work was supported by the National Institute of Child Health and Human Development/National Institute on Aging (R01HD044876 to M.I.C., B.S. and R.A.R.P.); the National Institutes of Health/National Center for Research Resources and UCSF's Clinical and Translational Science Institute (UL1RR024131 to M.I.C.); the Ric Weiland Fellowship and the Stanford Genome Training Project Grant (to N.A.J.); the Stanford University Dean's Postdoctoral Fellowship (1048085-155-KAUMZ to S.M.S.H.) and the National Institute of Child Health and Human Development (F32HD061204 to S.M.S.H.). Funding to pay the Open Access publication charges for this article was provided by NIH grant, R01HD044876, Genetic Epidemiology of Ovarian Aging.

Conflict of interest

None declared.

References

Affymetrix. Data Sheet: Affymetrix Genome-Wide Human SNP Array 6.0. http://wwwaffymetrixcom, 2007.

Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, Khoury MJ, Tanzi RE, Bertram L. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 2008;**40**:827–834.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;**21**:263–265.

Ben-Shlomo I, Vitt UA, Hsueh AJ. Perspective: the ovarian kaleidoscope database-II. Functional genomic analysis of an organ-specific database. *Endocrinology* 2002;**143**:2041–2044.

Bromberger JT, Matthews KA, Kuller LH, Wing RR, Meilahn EN, Plantinga P. Prospective study of the determinants of age at menopause. *Am J Epidemiol* 1997;**145**:124–133.

Browning SR. Multilocus association mapping using variable-length Markov chains. *Am J Hum Genet* 2006;**78**:903–913.

Browning BL, Browning SR. Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. *Genet Epidemiol* 2007;**31**:365–375.

Cooper GS, Sandler DP. Age at natural menopause and mortality. *Ann Epidemiol* 1998;**8**:229–235.

Cooper GS, Sandler DP, Bohlig M. Active and passive smoking and the occurrence of natural menopause. *Epidemiology* 1999; **10**:771–773.

Dai WS, Gutai JP, Kuller LH, Cauley JA. Cigarette smoking and serum sex hormones in men. *Am J Epidemiol* 1988;**128**:796–805.

de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet* 2005; **37**:1217–1223.

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.

Deslypere JP, Vermeulen A. Leydig cell function in normal men: effect of age, life-style, residence, diet, and activity. *J Clin Endocrinol Metab* 1984;**59**:955–962.

Durlinger AL, Visser JA, Themmen AP. Regulation of ovarian function: the role of anti-Mullerian hormone. *Reproduction* 2002;**124**:601–609.

Eldar-Geva T, Ben-Chetrit A, Spitz IM, Rabinowitz R, Markowitz E, Mimoni T, Gal M, Zylber-Haran E, Margalioth EJ. Dynamic assays of inhibin B, anti-Mullerian hormone and estradiol following FSH stimulation and ovarian ultrasonography as predictors of IVF outcome. *Hum Reprod* 2005;**20**:3178–3183.

Essenberg JM, Fagan L, Malerstein AJ. Chronic poisoning of the ovaries and tests of albino rats and mice by nicotine and cigarette smoke. *West J Surg Obstet Gynecol* 1951;**59**:27–32.

Gentleman R, Ihaka R. and the R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008.

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.

Gordon NP. *How Does the Adult Kaiser Permanente Membership in Northern California Compare with the Larger Community?* Oakland, CA: Kaiser Permanente Division of Research, 2006.

Gromoll J, Simoni M. Genetic complexity of FSH receptor function. *Trends Endocrinol Metab* 2005;**16**:368–373.

Hansen KR, Knowlton NS, Thyer AC, Charleston JS, Soules MR, Klein NA. A new model of reproductive aging: the decline in ovarian non-growing follicle number from birth to menopause. *Hum Reprod* 2008; **23**:699–708.

Hardy R, Kuh D. Reproductive characteristics and the age at inception of the perimenopause in a British National Cohort. *Am J Epidemiol* 1999; **149**:612–620.

Hartge P. Genetics of reproductive lifespan. *Nat Genet* 2009;**41**:637–638.

He C, Kraft P, Chasman DI, Buring JE, Chen C, Hankinson SE, Pare G, Chanock S, Ridker PM, Hunter DJ. A large-scale candidate gene association study of age at menarche and age at natural menopause. *Hum Genet* 2010;**128**:515–527.

Hull MG, Glazener CM, Kelly NJ, Conway DI, Foster PA, Hinton RA, Coulson C, Lambert PA, Watt EM, Desai KM. Population study of causes, treatment, and outcome of infertility. *Br Med J* 1985; **291**:1693–1697.

International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005;**437**:1299–1320.

Kee K, Flores M, Cedars MI, Reijo Pera RA. Human primordial germ cell formation is diminished by exposure to environmental toxicants acting through the AHR signaling pathway. *Toxicol Sci* 2010; **117**:218–224.

Kevenaar ME, Themmen AP, Rivadeneira F, Uitterlinden AG, Laven JS, van Schoor NM, Lips P, Pols HA, Visser JA. A polymorphism in the AMH type II receptor gene is associated with age at menopause in interaction with parity. *Hum Reprod* 2007;**22**:2382–2388.

Kinney A, Kline J, Kelly A, Reuss ML, Levin B. Smoking, alcohol and caffeine in relation to ovarian age during the reproductive years. *Hum Reprod* 2007;**22**:1175–1185.

Koenker R. *quantreg: Quantile Regression. R Package Version 4.27*. IL, USA: Department of Economics, University of Illinois, 2009.

Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, *et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* 2008;**40**:1253–1260.

Loutradis D, Patsoula E, Minas V, Koussidis GA, Antsaklis A, Michalas S, Makrigiannakis A. FSH receptor gene polymorphisms have a role for different ovarian response to stimulation in patients entering IVF/ ICSI-ET programs. *J Assist Reprod Genet* 2006;**23**:177–184.

Mattison DR, Thorgeirsson SS. Smoking and industrial pollution, and their effects on menopause and ovarian cancer. *Lancet* 1978;**1**:187–188.

Menken J, Larsen U. Estimating the incidence and prevalence and analyzing the correlates of infertility and sterility. *Ann N Y Acad Sci* 1994; **709**:249–265.

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.

Nardo LG, Christodoulou D, Gould D, Roberts SA, Fitzgerald CT, Laing I. Anti-Mullerian hormone levels and antral follicle count in women enrolled in *in vitro* fertilization cycles: relationship to lifestyle factors, chronological age and reproductive history. *Gynecol Endocrinol* 2007; **23**:486–493.

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;**26**:2336–2337.

Reijo R, Lee TY, Salo P, Alagappan R, Brown LG, Rosenberg M, Rozen S, Jaffe T, Straus D, Hovatta O, *et al.* Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet* 1995;**10**:383–393.

Rosen MP, Johnstone EB, Gillham SJ, Modan AE, Lipshutz AK, Reijo-Pera R, Cedars MI. Is antral follicle count a genetic trait?. *Menopause* 2010a;**17**:109–113.

Rosen MP, Sternfeld B, Schuh-Huerta SM, Reijo Pera RA, McCulloch CE, Cedars MI. Antral follicle count: absence of significant midlife decline. *Fertil Steril* 2010b;**94**:2182–2185.

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.

Soldin OP, Makambi KH, Soldin SJ, O'Mara DM. Steroid hormone levels associated with passive and active smoking. *Steroids* 2011; **76**:653–659.

Strohsnitter WC, Hatch EE, Hyer M, Troisi R, Kaufman RH, Robboy SJ, Palmer JR, Titus-Ernstoff L, Anderson D, Hoover RN, *et al*. The association between in utero cigarette smoke exposure and age at menopause. *Am J Epidemiol* 2008;**167**:727–733.

Torgerson DJ, Avenell A, Russell IT, Reid DM. Factors associated with onset of menopause in women aged 45–49. *Maturitas* 1994;**19**:83–92.

van Disseldorp J, Lambalk CB, Kwee J, Looman CW, Eijkemans MJ, Fauser BC, Broekmans FJ. Comparison of inter- and intra-cycle variability of anti-Mullerian hormone and antral follicle counts. *Hum Reprod* 2010;**25**:221–227.

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.

van Rooij IA, Tonkelaar I, Broekmans FJ, Looman CW, Scheffer GJ, de Jong FH, Themmen AP, te Velde ER. Anti-Mullerian hormone is a promising predictor for the occurrence of the menopausal transition. *Menopause* 2004;**11**:601–606.

Willett W, Stampfer MJ, Bain C, Lipnick R, Speizer FE, Rosner B, Cramer D, Hennekens CH. Cigarette smoking, relative weight, and menopause. *Am J Epidemiol* 1983;**117**:651–658.

Wu CH, Chen YC, Wu HH, Yang JG, Chang YJ, Tsai HD. Serum anti-Mullerian hormone predicts ovarian response and cycle outcome in IVF patients. *J Assist Reprod Genet* 2009;**26**:383–389.

Wunder DM, Guibourdenche J, Birkhauser MH, Bersinger NA. Anti-Mullerian hormone and inhibin B as predictors of pregnancy after treatment by in vitro fertilization/intracytoplasmic sperm injection. *Fertil Steril* 2008;**90**:2203–2210.