

# Linkage Analysis of Extremely Discordant and Concordant Sibling Pairs Identifies Quantitative Trait Loci Influencing Variation in Human Menopausal Age

Kristel M. van Asselt,<sup>1,2,3,\*</sup> Helen S. Kok,<sup>1,2,3,\*</sup> Hein Putter,<sup>4</sup> Cisca Wijmenga,<sup>2</sup> Petra H. M. Peeters,<sup>1</sup> Yvonne T. van der Schouw,<sup>1</sup> Diederick E. Grobbee,<sup>1</sup> Egbert R. te Velde,<sup>3</sup> Sietse Mosselman,<sup>5</sup> and Peter L. Pearson<sup>2</sup>

<sup>1</sup>Julius Center for Health Sciences and Primary Care, <sup>2</sup>Department of Biomedical Genetics, and <sup>3</sup>Department of Reproductive Medicine, University Medical Center Utrecht, and <sup>4</sup>Department of Medical Statistics, Leiden University Medical Center, Utrecht; and <sup>5</sup>N. V. Organon, Oss, The Netherlands

Age at natural menopause may be used as parameter for evaluating the rate of ovarian aging. Environmental factors determine only a small part of the large variation in menopausal age. Studies have shown that genetic factors are likely to be involved in variation in menopausal age. To identify quantitative-trait loci for this trait, we performed a genomewide linkage study with age at natural menopause as a continuous quantitative phenotype in Dutch sister pairs, through use of a selective sampling scheme. A total of 165 families were ascertained using extreme selected sampling and were genotyped for 417 markers. Data were analyzed by Haseman-Elston regression and by an adjusted variance-components analysis. Subgroup analyses for early and late menopausal age were conducted by Haseman-Elston regression. In the adjusted variance-components analysis, 12 chromosomes had a LOD score of  $\geq 1.0$ . Two chromosomal regions showed suggestive linkage: 9q21.3 (LOD score 2.6) and Xp21.3 (LOD score 3.1). Haseman-Elston regression showed rather similar locations of the peaks but yielded lower LOD scores. A permutation test to obtain empirical *P* values resulted in a significant peak on the X chromosome. To our knowledge, this is the first study to attempt to identify loci responsible for variability in menopausal age and in which several chromosomal regions were identified with suggestive and significant linkage. Although the finding of the region on the X chromosome comes as no surprise, because of its widespread involvement in premature ovarian failure, the definition of which particular gene is involved is of great interest. The region on chromosome 9 deserves further consideration. Both findings require independent confirmation.

## Introduction

On average, women maintain optimal fertility until 30 years of age, after which fertility steadily decreases, reaching 50% by age 35 years and approaching essentially zero by age 41 years (Noord-Zaadstra et al. 1991). This latter average is confirmed by several studies, both in industrialized populations and in populations that do not use contraception, showing that the average maternal age at birth of the last child is 41–42 years (Bongaarts 1982; Wood 1989). In many Western societies, during the last 3 decades there has been an increasing delay in childbearing. A major consequence of this trend is a progressive inability to reproduce. In particular, women

who exhibit early ovarian failure run the risk of involuntary infertility by delaying commencement of childbearing. It has been suggested that the age at which female infertility arises as a result of ovarian aging correlates with the age at natural menopause (te Velde et al. 1998) and, therefore, that variation in age at natural menopause is a suitable parameter for assessing the rate of ovarian aging.

The absence of any clear single causative environmental factor to account for variation in the age of menopause suggests a major involvement of genetic factors (Cramer et al. 1995*b*; van Noord et al. 1997). It has been reported that common environmental factors associated with age at menopause explain only a small percentage of the total variation in menopausal age (van Noord et al. 1997). Arguments in favor of the involvement of genetic factors have been provided by several studies. Family history has been shown to be a predictor for early menopause (Cramer et al. 1995*b*), and a strong association was found between mothers' and daughters' menopausal ages (Torgerson et al. 1997). Twin and family studies have confirmed a substantial genetic contri-

Received September 5, 2003; accepted for publication December 12, 2003; electronically published February 4, 2004.

Address for correspondence and reprints: Dr. C. Wijmenga, University Medical Center Utrecht, Department of Biomedical Genetics/Complex Genetics Group, Stratenum 2.112, PO Box 80030, 3508 TA Utrecht, The Netherlands. E-mail: t.n.wijmenga@med.uu.nl

\* These authors contributed equally to this work.

© 2004 by The American Society of Human Genetics. All rights reserved. 0002-9297/2004/7403-0013\$15.00

bution to the variation in age at menopause. Snieder et al. (1998) estimated a relative contribution of genetic factors to the timing of natural menopause of 63% in a large twin sample collected in the United Kingdom. In an Australian twin sample, heritability was estimated at 31%–53% (Treloar et al. 1998). In a study of a Dutch population, heritability estimates of 70% and 85% were reported in twin and sister pairs, respectively (de Bruin et al. 2001).

In ~1% of women, menopause occurs before age 40, which is defined arbitrarily as premature ovarian failure (POF) (Coulam 1982; Coulam et al. 1986). Monogenic causes are reported for a portion of the women exhibiting POF (Aittomaki et al. 1996; Gromoll et al. 1996; Beau et al. 1998; Davison et al. 1998; Murray et al. 1998; Allingham-Hawkins et al. 1999; Touraine et al. 1999; Davis et al. 2000; Marozzi et al. 2000b; Crisponi et al. 2001; Doherty et al. 2002; Schlessinger et al. 2002).

To date, however, the genes involved in the normal range in menopause remain largely unclear. The only gene variants reported so far to be associated with variation in age of menopause are a noncoding polymorphism in the estrogen receptor 1 ( $\alpha$ ) (ER) gene and a mutation in the factor V Leiden gene. The demonstrated effect of the ER polymorphism was small; homozygosity for one of the two alleles reduced the age of menopause by 1 year (Weel et al. 1999). Moreover, two independent studies were unable to replicate this finding (Gorai et al. 2003; H. S. Kok, N. C. Onland, K. M. van Asselt, C. H. van Gils, Y. T. van der Schouw, D. E. Grobbee, and P. H. M. Peeters, unpublished data). Recently, we have shown that the factor V Leiden mutation is associated with an earlier age at menopause (van Asselt et al. 2003); women who were heterozygous for the mutation had an onset of menopause that occurred 3 years earlier than in women without the mutation. Given the large variation in menopausal age, the absence of clear common causative environmental factors, and a high heritability, age at menopause meets all criteria of a (quantitative) complex trait.

The amount of genotyping required to detect a locus can be reduced considerably, with retention or improvement of power, by selecting those sib pairs most likely to show deviation from the expected proportion of allele sharing. Pairs that are genetically informative for linkage are those with concordant late or early age at menopause and discordant pairs in which one member has an early menopause and the other a late menopause. Each group has a different power to detect loci, depending on the underlying genetic model, which is as yet unknown and can only be assumed in power studies.

To identify susceptibility loci determining variation in menopausal age, we performed the first genomewide linkage study of age at natural menopause as a con-

tinuous quantitative phenotype in Dutch sister pairs, through use of a selective sampling scheme.

## Methods

### *Phenotype and Inclusion Criteria*

The phenotype, or trait, of interest is age at natural menopause. We defined “natural menopause” as proposed by the World Health Organization (World Health Organisation Scientific Group 1996)—namely, as  $\geq 12$  consecutive months of amenorrhea not due to surgery or other obvious causes. Accordingly, women who were premenopausal at the time of hysterectomy or either bilateral or unilateral ovariectomy were not considered eligible as probands; unilateral ovariectomy was regarded as an exclusion criterion, because it is known to be associated with an earlier age at menopause (Cramer et al. 1995a; Melica et al. 1995; Hardy and Kuh 1999). Women who had started using oral contraceptives (OCs) or hormonal replacement therapy (HRT) within 1 year after their last period were also considered ineligible, since age at menopause cannot be determined with certainty in such cases.

In QTL analysis, power can be increased considerably by sampling sib pairs with phenotypic values in the extreme parts of the trait distribution (Risch and Zhang 1995). Accordingly, we aimed to restrict our sampling to female sib pairs with menopausal ages of  $\leq 45$  years and/or  $\geq 54$  years (corresponding to the 12.5th percentiles of the distribution of age at menopause). When possible, the parents or additional sisters or brothers were included, to reconstruct the missing parental genotypes for phase determination.

### *Family Ascertainment*

We used several strategies to ascertain families. Two large-scale population-based cohorts, the Diagnostisch Onderzoek Mammacarcinoom (DOM) and Prospect-EPIC cohorts, were used to identify possible probands. The design and sampling strategy of these cohorts have been described in detail elsewhere (de Waard et al. 1984; Boker et al. 2001). In brief, both cohorts consist of women who were invited to join the respective studies through an existing regional population-based program of breast cancer screening. For these two cohorts, information on menopausal age and whether or not menopause had occurred naturally was available for ~40,000 women in the region of Utrecht. Probands fulfilling the inclusion criteria were approached, and inquiries were made about the presence of sisters and, if alive, about their menopausal status. Probands were sent pedigree forms to provide information on family structure and menopausal status of sister(s). Physicians clarified apparent inconsistencies in the data through a tel-

ephone call. If at least two sisters complied with the inclusion criteria, the family was asked to participate. Blood samples (20 ml of EDTA-treated blood) were taken, and female participants were requested to fill out a questionnaire on reproductive history. A total of 126 families was included through use of these cohorts. Further, through press releases in a newspaper, women's magazines, the Dutch Twin Registry, and outpatient clinics of the University Medical Center in Utrecht and Nijmegen, 39 additional families were included. In total, information and DNA from 583 individuals from 165 families was collected. The institutional review board of the University Medical Center Utrecht approved the study protocol. Written informed consent was obtained from all participants.

From the questionnaire data, age at start and stop of HRT, age at start and stop of OC use, age at hysterectomy and/or ovariectomy, and age at final cessation of menstruation were obtained. It was decided that, in specific situations, age at menopause might be determined with some degree of acceptable uncertainty in the case of additional sisters only: (a) women who stopped using OCs at age  $\leq 45$  years and had no periods after that were assigned a menopausal age of 1 year earlier ( $n = 16$ ); (b) women who had a hysterectomy or bilateral ovariectomy at age  $\geq 54$  years and were still premenopausal at that time were assigned a menopausal age of 1 year later ( $n = 7$ ); (c) women who were aged  $\geq 54$  years at ascertainment and were still premenopausal were assigned a menopausal age of 1 year later ( $n = 7$ ); (d) for women who reported starting HRT 1 year later than their menopausal age, it remained uncertain whether 12 mo of amenorrhea had passed ( $n = 11$ ), and therefore reported age at last menstruation was accepted as menopausal age; and (e) for women aged  $> 54$  years for whom the age of starting HRT was equal to the age of last menstruation, menopausal age was assigned to be the age at which HRT was started ( $n = 10$ ). In 88% of women, age at natural menopause was unambiguously determined.

### Genotyping

DNA was extracted from peripheral blood lymphocytes by high salt extraction (Miller et al. 1988). DNA concentration was measured with a spectrophotometer (Tecan), and samples were diluted with distilled water to a concentration of 15 ng/ $\mu$ l. The DNA samples were divided into 96-well microtiter plates. Every DNA plate contained as many as 86 unique DNA samples, 6 blind duplicate samples, 3 CEPH controls, and 1 negative control.

The Mammalian Genotyping Service of the Marshfield Medical Research Foundation performed the genotyping for the initial screen. The marker set was based on Marshfield Screening Set 11 and consisted of 408

polymorphic microsatellite markers with an average spacing of 10 cM and an average heterozygosity of 0.75. Genotypes with a reported Marshfield quality score  $< 0.99$  were recoded as unknown. Six markers located on the Y chromosome were excluded from all analyses.

Fifteen additional markers were used to increase the information content for some chromosomal regions. Marker positions in these regions were verified using the Ensembl and Decode databases (Kong et al. 2002). Either the forward or the reverse oligonucleotide primer was labeled with 6-FAM, HEX, or NED fluorescent dyes (Biolegio and Applied Biosystems). PCRs were performed in a 10- $\mu$ l volume containing 30 ng of template DNA, 25 ng of each oligonucleotide primer, 200 mM of each dNTP, and 0.4 U AmpliTaq Gold (Applied Biosystems), in 1  $\times$  PCR buffer II with 2.5 mM MgCl<sub>2</sub> (Applied Biosystems). DNA was initially denatured at 94°C for 7 min and was then subjected to 33 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s, followed by a final extension step of 30 min at 72°C. PCR products were diluted four times with distilled water, and 1 ml of diluted product was mixed with 4 ml of HiDi (Applied Biosystems) and 0.1 ml of GS-500 size standard (Applied Biosystems) and was separated in POP6 polymer on an ABI 3700 capillary DNA sequencer (Applied Biosystems). Data were analyzed using Genescan 3.6 and Genotyper 3.5 for Windows NT (Applied Biosystems). Two independent raters genotyped the additional markers and checked results for inconsistency. If they disagreed, the genotypes were set to unknown.

### Study Subjects

Relationships between siblings were verified using the program GRR (Abecasis et al. 2001). This resulted in the identification of three half sibs and one MZ twin pair. The three half sibs were excluded, as was one sister from the MZ twin pair.

In total, 579 individuals from 165 families were included in the final study population. Four hundred thirty-one women met the inclusion criteria and were assigned a phenotypic value; the remaining individuals were parents, brothers, and sisters included for additional phase information. Inheritance within families was verified using the Pedcheck program (O'Connell and Weeks 1998). If there were apparent inheritance errors due to presumptive mistyping or misidentification, the complete family was excluded from the analysis for that specific marker. Allele frequencies were calculated from all available genotypes.

### Statistical Analysis

Two methods were applied to analyze the data; the original Haseman-Elston regression method and a variation of the variance-components (VC) method. Non-normality of the data as a result of the extreme sampling

scheme requires special attention in the analyses. Although the Haseman-Elston regression analysis is robust to deviations from normality (Haseman and Elston 1972), it exhibits loss of power in this study design, since it depends strongly on the magnitude of the phenotypic difference of a sib pair, which is, by definition, small in extremely concordant sib pairs. Single- and multipoint LOD scores were computed by the Mapmaker/Sibs program (Kruglyak and Lander 1995), using all possible distinct pairs.

Since standard VC QTL analysis is susceptible to non-random ascertainment, we used a modification of the likelihood function by correcting for the ascertainment bias introduced by the extreme sampling (see, e.g., Morton [1959]); the likelihood function is shown in the appendix. For this analysis, identity-by-descent (IBD) probabilities, as derived from the Mapmaker/Sibs program (Kruglyak and Lander 1995), were used.

For chromosomes with suggestive evidence for linkage, we performed simulations to obtain significance thresholds. Empirical *P* values were obtained by permuting the IBD probabilities over the sib pairs. A genomewide *P* value <.05 was considered to be of genomewide significance.

The mean and SD of menopausal age were estimated from a population-based study, Prospect-EPIC, and were 50.2 years and 4.2 years, respectively. Since smoking status at menopause was available for all sib pairs, an analysis with this information as a covariate was performed.

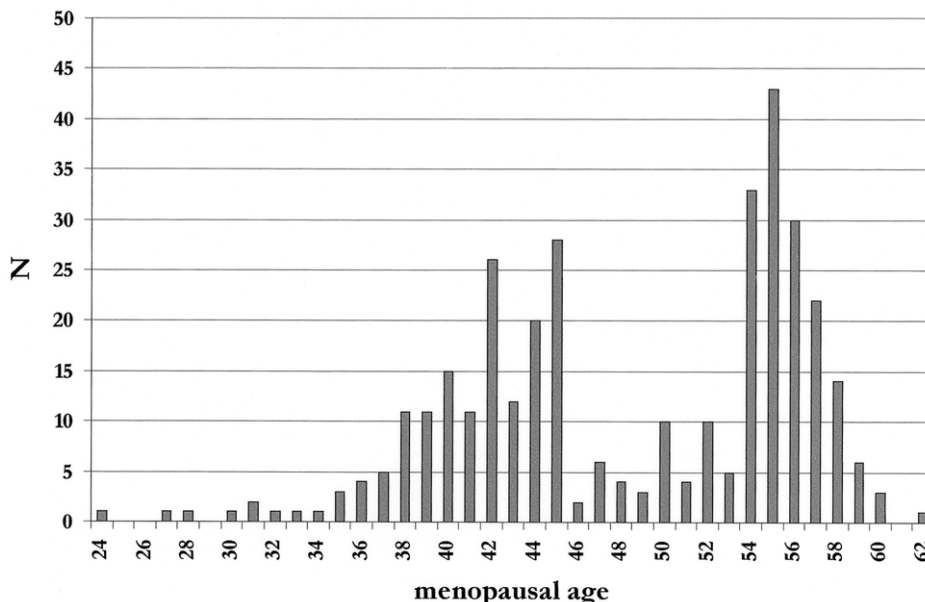
Figure 1 shows the distribution of age at menopause

in the study population. Although participants were essentially sampled from extremes of the distribution, it turned out that some women had experienced natural menopause at a less extreme age. The ages at menopause of these women were also included in the Haseman-Elston analyses because, although relatively few in number, they expanded the range of phenotypic variation being analyzed. However, the trait values of these women could not be included in the adjusted VC analyses without caution, because, in this model, the likelihood is divided by the probability that a particular family is selected. Therefore, the ascertainment scheme requires clear cutoff points, and a family was included only when it contained at least two sisters with menopausal ages in the 12.5th percentiles. Therefore, in this analysis, women with menopausal ages outside the extremes participated only when they had at least two sisters who met the strict inclusion criteria. Of the 165 available families, 130 families were included in the adjusted VC analysis.

The overall information content of the markers was obtained by calculating the average multipoint information content for all markers.

### Subgroup Analyses

Separate analyses for families with a late menopausal age and those with an early menopausal age were conducted. The following criteria were applied to construct the respective subgroups of early and late menopausal sib pairs used for the Haseman-Elston regression. Sib



**Figure 1** Histogram of age at menopause of women included in the VC analyses

**Table 1**  
Number of Phenotyped and Genotyped Families

SIBSHIP SIZE	NO. OF FAMILIES		
	All	Late Menopause	Early Menopause
No. of phenotyped women per family:			
2	99	46	57
3	42	20	27
4	14	8	7
5	9	5	5
6	<u>1</u>	<u>1</u>	<u>0</u>
Total	165	80	96
No. of genotyped individuals per family:			
2	39	19	23
3	50	23	31
4	46	23	25
5	19	10	9
6	8	4	6
7	2	1	1
10	<u>1</u>	<u>0</u>	<u>1</u>
Total	165	80	96

pairs were included in the “early group” when at least one sister had a menopausal age of  $\leq 45$  years. For a sib pair to be included in the “late group,” one sister needed to have a menopausal age of  $\geq 54$  years. These criteria resulted in a sampling overlap between the groups—that is, 11 sib pairs were included in both groups.

## Results

### Sample Characteristics

In table 1, an overview is given of the numbers of families per size of the phenotyped and genotyped sibships for the total study population, as well as for the respective subgroups. In table 2, characteristics of the phenotyped women are given. The study subjects had a mean age of 62 years (range 26–86 years) at inclusion in the present study.

### Genotype Characteristics

For the initial screen, the overall genotype error rate was 0.2%, and the average genotyping completeness was 98.5%. When the 15 additionally typed markers were included, the average marker information content across the genome was 0.68.

### Haseman-Elston and VC Analyses

Table 3 shows the locations where linkage signals were found by the VC with LOD scores  $\geq 1$ . In figure 2, for each chromosome, the multipoint LOD scores obtained by Haseman-Elston regression and the VC model are plotted in one graph. Overall, the VC analyses yielded

higher LOD scores than did the Haseman-Elston regression. However, the genome positions of the peaks are rather similar, implying that both methods generally agree on the location but differ in the significance of the observation.

Two regions showed suggestive linkage. On chromosome 9, region 9q21.3, a LOD score of 2.6 was obtained by the VC analysis. The second region with suggestive linkage was located on the X chromosome, with the highest LOD score of 3.1 on Xp21.3. The marker information content at the peaks on chromosome 9 and X were 82% and 84%, respectively. The proportion of the total variance present in the study group explained by the QTL on chromosome region 9q21.3 was 24%, and the proportion explained by that on chromosome region Xp21.3 was 30%. Incorporating smoking at menopause as a covariate in the model did not further increase the peaks on chromosome 9 and X.

### Empirical Significance Thresholds

Permutation tests for chromosomes 9 and X were performed to assess significance of the adjusted VC analysis. We performed 1,000 simulations for both chromosomes X and 9. A Bonferroni correction was used to approximate genomewide significance levels. This may not be entirely appropriate, since the chromosomes differ in length; however, it was computationally unfeasible to perform genomewide simulations. Since marker spacings were consistent throughout the genome, and since chromosomes 9 and X are slightly above average in length, it is expected that the Bonferroni correction will be conservative rather than liberal.

The chromosomewide  $P$  value of the highest peak on chromosome 9 is .005; when the Bonferroni correction is used, this implies a genomewide  $P$  value of .115. For the X chromosome, the chromosomewide  $P$  value is .001, and the genomewide  $P$  value is .023.

### Subgroup Analyses

Two subgroups—early menopause (96 families) and late menopause (80 families)—were analyzed by the

**Table 2**  
Characteristics of the 431 Phenotyped Women

Characteristic	Value	Range or %
Mean age at inclusion (in years)	62	26–86
Median time since menopause (in years)	12	0–46
Median no. of offspring	2	0–15
No. nulliparous	54	12.5%
No. with $\geq 1$ miscarriage	103	23.9%
No. with $\geq 2$ miscarriages	32	7.4%
No. who have ever smoked	117	27.1%
No. who have ever used OCs	244	56.6%
No. who have ever used HRT	73	16.9%

**Table 3**  
Regions with Multipoint LOD Scores  $\geq 1$  by Adjusted VC Analyses

Chromosome and Position (in cM)	LOD Score	P
2:		
125	1.3	.0072
3:		
121	1.8	.0020
210	1.6	.0033
6:		
142	1.4	.0056
9:		
80	2.6	.0003
99	2.0	.0012
11:		
135	1.6	.0033
12:		
50	1.0	.0159
13:		
79	1.0	.0159
15:		
42	1.0	.0159
17:		
9	1.7	.0026
18	1.6	.0033
19:		
1	1.4	.0056
90	1.0	.0159
22:		
27	1.5	.0043
X:		
35	3.1	.0001
85	2.1	.0009
129	1.4	.0051

Haseman-Elston regression. In general, the Haseman-Elston regression analyses conducted with subgroups resulted in higher LOD scores than those obtained using all sib pairs. Figure 3 displays the results of the subgroup analyses for chromosomes 9 and X. For chromosome 9, the peak found in the overall group was increased when only the women with late menopause were analyzed, although both groups supported the linkage. On the other hand, the peak at Xp21.3 was almost entirely due to the women with early menopause.

## Discussion

We present results from the first genomewide scan for QTLs influencing menopausal age. Our analysis identified two regions, on chromosomes 9 and X, with LOD scores of 2.6 and 3.1, respectively. Empirical significance thresholds indicated that the peak on X chromosome is significant and that the peak on chromosome 9 is suggestive.

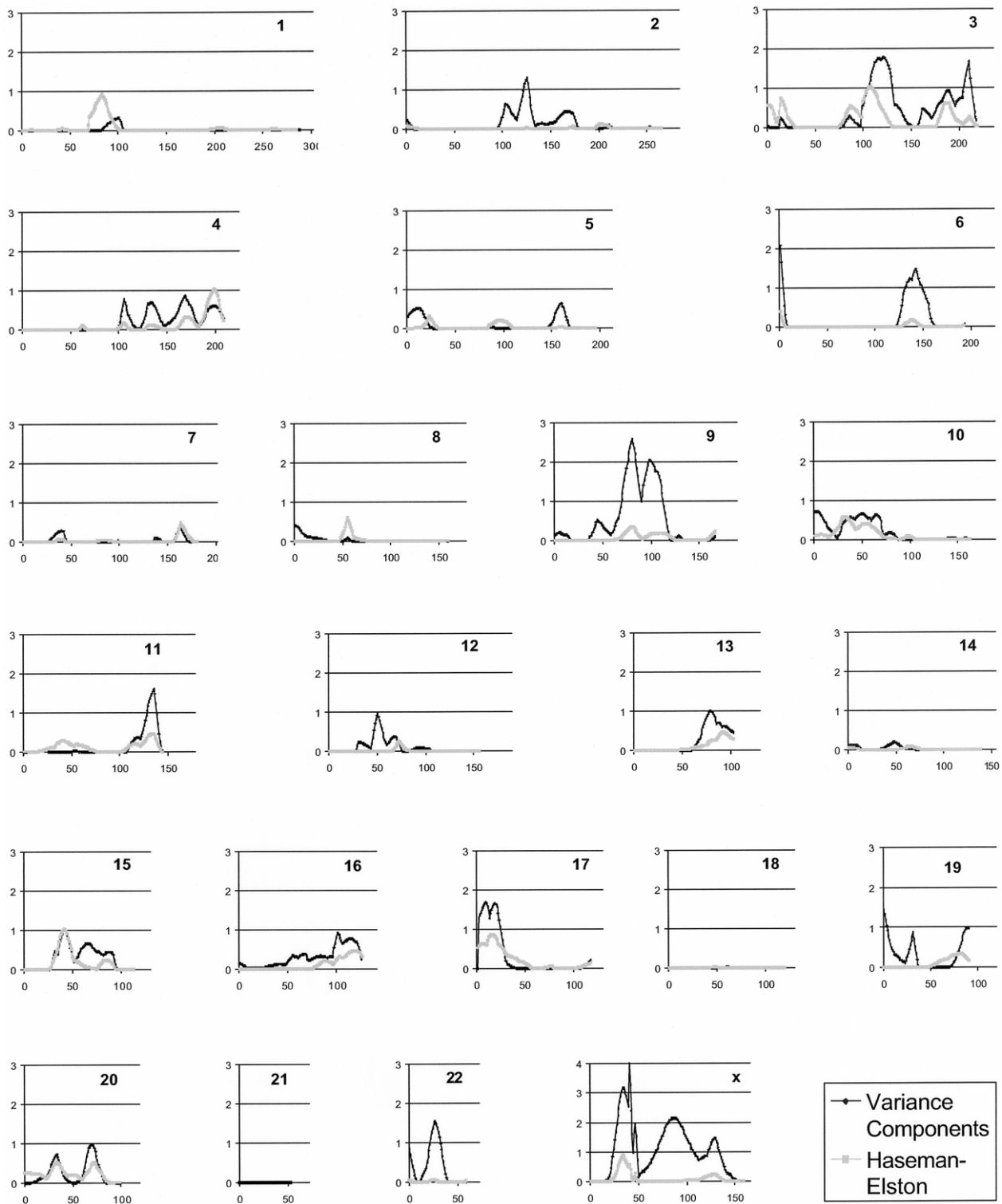
Previously, so-called “POF regions” on the X chromosome have been described (Sala et al. 1997; Marozzi

et al. 2000a; Schlessinger et al. 2002), and the peak found in our sample is located in such a region. The subgroup analyses showed that the linkage found on the X chromosome was entirely due to the women with early menopausal ages. To establish whether it is likely that the peak is based only on women with extremely early menopause, we excluded women who would be classified as having POF according to the most widely used criterion, age at menopause of  $\leq 40$  years (Coulam et al. 1986). Exclusion of those women resulted in a loss of 11 families. Although the LOD score of the peak on chromosome X dropped (LOD 2.2 on 35 cM), suggestive linkage remained. This indicates that linkage in this region is not based only on women with POF, and, stated otherwise, that the “POF region” is important not only for POF but possibly also for “early menopause.” The lower LOD score obtained after excluding women with POF may simply reflect loss of power by reducing the number of families. This hypothesis is supported by the observation that early menopause and POF sometimes segregate within the same families (Tibiletti et al. 1999; Vegetti et al. 2000), and it may be unrealistic to assume that POF and early menopause should be regarded as completely distinct entities.

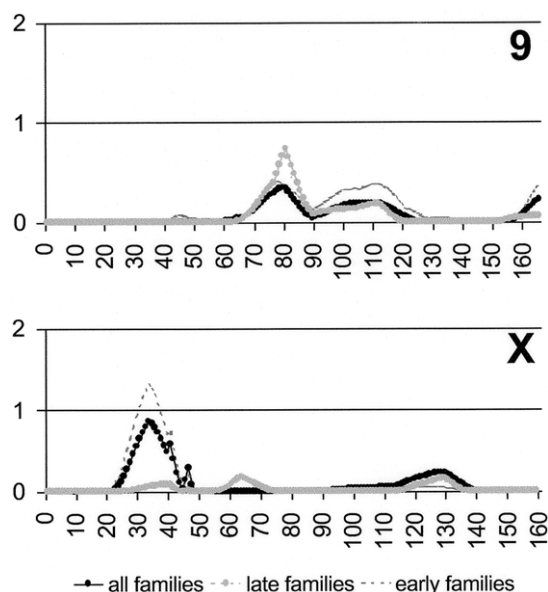
The region suggestive for linkage on chromosome 9 contains many genes. One of the genes in this region encodes for the BCL2 family, a protein involved in apoptosis. This is interesting, because apoptosis is the most common fate of the follicles in the ovaries, and the size of the follicle pool is associated with menopausal age (Hsueh et al. 1994; Billig et al. 1996; Kaipia and Hsueh 1997). Polymorphisms of BCL2 that are known to modulate the rate of human ovarian aging have not yet been described.

The two genes for which association with human menopausal age has been reported, the ER gene (chromosome 6) and the factor V Leiden polymorphism (chromosome 1), are not located in any of the regions identified with linkage in this study. The specific variances of these loci may be too low to be detected by linkage in the present study, with the low allele frequency of factor V Leiden ( $\sim 0.04$ ) and the low phenotypic impact of the ER polymorphism being likely explanations. The power analyses (van Asselt and Kok 2003) have shown that loci with an explained variance of  $\geq 20\%$  linkage could be detected with a probability of 80%.

One must keep in mind that a joint estimation of the location of genes and the estimation of their effect on the phenotype (the variance explained by the QTL) inflates the estimate of the variance attributable to each QTL (Göring et al. 2001). This is caused by the fact that the statistic—which provides evidence for the presence of a locus at a given chromosomal location and is typically given in the form of a LOD score—is itself a function of the parameters characterizing the genotype-



**Figure 2** LOD score graphs of all chromosomes. LOD scores are shown on the Y-axis, and positions in centimorgans are shown on the X-axis.



**Figure 3** LOD score graphs of subgroup analyses of chromosome 9 and X by Haseman-Elston regression. LOD scores are shown on the Y-axis, and positions in centimorgans are shown on the X-axis.

phenotype relationship. Statistical significance and the derived parameters, therefore, are not independent and are highly correlated. The extent of the bias depends on the true parameter value, sample size, and study design. The only realistic option for bias elimination appears to be the use of independent data sets for locus mapping and for estimation of the locus-effect size. Moreover, the explained variance in a selected sample cannot be extrapolated to a valid estimate in the population. Therefore, in the present study, we are unable to determine what proportion of the total variation in age at menopause is explained by the two detected QTLs.

Phenotype definition is a critical issue for genetic investigations of complex human traits. Age at natural menopause can be assessed only in retrospect, after a period of 12 mo of amenorrhea. Misclassification due to year preference is a phenomenon known to occur after some time (Hahn et al. 1997). However, the validity of self-reported menopausal age increases as the extremes are approached (den Tonkelaar 1997). Therefore, it is likely that, in a sample selected for extreme values, the degree of misclassification is low.

The use of sib pairs with extreme concordant or discordant trait values has some implications for the analyses. First, the resulting trait distribution is obviously very different from that present in the source population, resulting in an overestimation of the prevalence of extreme values of menopausal age, a possibly unrepresentative mean value, and a highly distorted variance

estimate. Accordingly, the VC analysis was performed with the mean and variance estimates derived from the Prospect-EPIC study, a population-based cohort.

In a sensitivity analysis, we set out to assess the impact of assigning different values for the mean and SD in an adjusted VC analysis, tested on the marker genotypes generated for chromosome 9. An increase of the mean menopausal age by 2% (1 year) resulted in a 4% higher LOD score. An equal decrease resulted in a 6% lower LOD score. Varying the SD of the menopausal age yielded the following results: a 25% increase in SD (~1 year) resulted in a 21% lower LOD score, and a 25% decrease (1 year) led to a 50% higher LOD score. This is reassuring because, when errors of this magnitude in the parameters are allowed for, suggestive linkage is still reached.

Subgroup analyses of women with either a very early menopause or a very late menopause have been performed by Haseman-Elston regression. Unfortunately, these subgroup analyses were not possible with the adjusted VC analyses. The ascertainment correction is sensitive to the assumption of normality of the menopausal ages in the population from which they were sampled. The smaller the selection made, the more sensitive the results of the adjusted VC analysis will be. Both subgroups supported the peak on chromosome 9. It is difficult to disentangle whether a QTL would be an “increaser” or “decreaser” when judged by which subgroup produces the highest LOD score. The reason for this is that a LOD score depends on the number of families and on the spread of menopausal age, and these are not the same for both groups.

To our knowledge, this is the first attempt to identify loci responsible for the variability of menopausal age and in which several chromosomal regions were identified with suggestive linkage. Although the involvement of the X chromosome comes as no surprise, when we consider its widespread involvement in premature ovarian failure, defining which particular gene is involved is of great interest. The region on chromosome 9 deserves further consideration. Both findings require independent confirmation.

## Acknowledgments

We wish to acknowledge the tremendous contribution of Lodewijk Sandkuijl in setting up this study. Tragically, Lodewijk died unexpectedly before completion of this work. We gratefully acknowledge the support of the Mammalian Genotyping Service of the Marshfield Medical Research Foundation and Dr. P. Nilsson, for the typing of extra markers. We thank A. Bardoel, for technical assistance with genotyping and data checking; E. Strengman and R. van 't Slot, for assistance with DNA extraction; H. van Someren, for data entry and checking; Dr. P. van Zonneveld, for ascertaining families from



outpatient clinics; Prof. D. Boomsma, for ascertaining twins from the Dutch Twin registry; and K. Witte, Y. Gents, and A. Smits, for administrative and laboratory assistance. The project was financially supported by N. V. Organon The Netherlands and by the University Medical Center Utrecht.

## Appendix

The log likelihood, corrected for ascertainment, for a sibship of size  $m$  is given by  $1 = \log[\varphi_m(y)] - \log(\text{Pr}_m)$ , where

$$\log[\varphi_m(y)] = -\frac{m}{2} \log(2\pi) - 0.5 \log[\det(\Sigma)] \\ - 0.5(y - \mu)^T \Sigma^{-1}(y - \mu) .$$

Here,  $y$  denotes the  $m$  vector of menopause ages of the family members, and  $\varphi_m(y)$  is the  $m$ -variate normal density with mean vector  $\mu$  and covariance matrix  $\Sigma$ . The  $m \times m$  covariance matrix is given by  $\Sigma_{ij} = \sigma_s^2 + \pi_{ij}\sigma_g^2$ , for  $i \neq j$ , where  $\sigma_s^2$  is the variance due to shared environment (common environment plus half of additive polygenic variance),  $\sigma_g^2$  is the variance due to the QTL, and  $\pi_{ij}$  is the IBD value. The ascertainment correction,  $\text{Pr}_m$ , is the probability that at least two sibs are in the sampling region  $R$ , with age at menopause  $<46$  years or  $>53$  years; for  $m = 2$ , this equals

$$\text{Pr}_2 = \int_R \int_R \varphi_2(y_1, y_2) dy_1 dy_2 .$$

## References

- Abecasis GR, Cherny SS, Cookson WO, Lardon LR (2001) GRR: graphical representation of relationship errors. *Bioinformatics* 17:742–743
- Aittomaki K, Herva R, Stenman UE, Juntunen K, Ylostalo P, Hovatta O, de la Chapelle A (1996) Clinical features of primary ovarian failure caused by a point mutation in the follicle-stimulating hormone receptor gene. *J Clin Endocrinol Metab* 81:3722–3726
- Allingham-Hawkins DJ, Babul-Hirji R, Chitayar D, Holden JJA, Yang KT, Lee C, Hudson R, et al (1999) Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study—preliminary data. *Am J Med Genet* 83:322–325
- Beau I, Touraine P, Meduri G, Gougeon A, Desroches A, Matuchansky C, Milgrom E, Kuttenn R, Misrahi M (1998) A novel phenotype related to partial loss of function mutations of the follicle stimulating hormone receptor. *J Clin Invest* 102:1352–1359
- Billig H, Chun SY, Eisenhauer K, Hsueh HJ (1996) Gonadal cell apoptosis: hormone-regulated cell demise. *Hum Reprod Update* 2:103–117
- Boker LK, van Noord PA, van der Schouw YT, Koot NV, de Mesquita HB, Riboli E, Grobbee DE, Peeters PHM (2001) Prospect-EPIC Utrecht: study design and characteristics of the cohort population. *European Prospective Investigation into Cancer and Nutrition. Eur J Epidemiol* 17:1047–1053
- Bongaarts CE (1982) The proximate determinants of natural marital fertility. Working paper no 89, Population Council, New York:1–43
- Coulam CB (1982) Premature gonadal failure. *Fertil Steril* 38:645–655
- Coulam CB, Adamson SC, Annegers JF (1986) Incidence of premature ovarian failure. *Obstet Gynecol* 67:604–606
- Cramer DW, Xu H, Harlow BL (1995a) Does “incessant” ovulation increase risk for early menopause? *Am J Obstet Gynecol* 172:568–573
- (1995b) Family history as a predictor of early menopause. *Fertil Steril* 64:740–745
- Crisponi L, Deiana M, Loi A, Chiappe F, Uda M, Amati P, Bisceglia L, Zelante L, Nagaraja R, Porcu S, Ristaldi MS, Marzella R, Rocchi M, Nicolino M, Lienhardt-Roussie A, Nivelon A, Verloes A, Schlessinger D, Gasparani P, Bonneau D, Cao A, Pilia G (2001) The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat Genet* 27:159–166
- Davis CJ, Davison RM, Payne NN, Rodeck CH, Conway GS (2000) Female sex preponderance for idiopathic familial premature ovarian failure suggests an X chromosome defect: opinion. *Hum Reprod* 15:2418–2422
- Davison RM, Quilter CR, Webb J, Murray A, Fisher AM, Valentine A, Serhal P, Conway GS (1998) A familial case of X chromosome deletion ascertained by cytogenetic screening of women with premature ovarian failure. *Hum Reprod* 13:3039–3041
- de Bruin JP, Bovenhuis H, van Noord PA, Pearson PL, van Arendonk JA, te Velde ER, Kuurman WW, Dorland M (2001) The role of genetic factors in age at natural menopause. *Hum Reprod* 16:2014–2018
- de Waard F, Colette HJ, Rombach JJ, Baanders-van Halewijn EA, Honing C (1984) The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. *J Chronic Dis* 37:1–44
- den Tonkelaar I (1997) Validity and reproducibility of self-reported age at menopause in women participating in the DOM-project. *Maturitas* 27:117–123
- Doherty E, Pakarinen P, Tiitinen A, Kiilavuori A, Huhtaniemi I, Forrest S, Aittomaki K (2002) A novel mutation in the FSH receptor inhibiting signal transduction and causing primary ovarian failure. *J Clin Endocrinol Metab* 87:1151–1155
- Gorai I, Tanaka K, Inada M, Morinaga H, Uchiyama Y, Kikuchi R, Chaki O, Hirahara F (2003) Estrogen-metabolizing gene polymorphisms, but not estrogen receptor-alpha gene polymorphisms, are associated with the onset of menarche in healthy postmenopausal Japanese women. *J Clin Endocrinol Metab* 88:799–803
- Göring HH, Terwilliger JD, and Blangero J (2001) Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 69:1357–1369
- Gromoll J, Simoni M, Nordhoff V, Behre HM, De Geyter C, Nieschlag E (1996) Functional and clinical consequences of

- mutations in the FSH receptor. *Mol Cell Endocrinol* 125: 177–182
- Hahn RA, Eaker E, Rolka H (1997) Reliability of reported age at menopause. *Am J Epidemiol* 146:771–775
- Hardy R, Kuh D (1999) Reproductive characteristics and the age at inception of the perimenopause in a British National Cohort. *Am J Epidemiol* 149:612–620
- Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 2:3–19
- Hsueh AJ, Billig H, Tsafiri A (1994) Ovarian follicle atresia: a hormonally controlled apoptotic process. *Endocr Rev* 15: 707–724
- Kaipia A, Hsueh AJ (1997) Regulation of ovarian follicle atresia. *Annu Rev Physiol* 59:349–363
- Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. *Nat Genet* 31:241–247
- Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 57:439–454
- Marozzi A, Manfredini E, Tibiletti MG, Furlan D, Villa N, Vegetti W, Crosignani PG, Ginelli E, Meneveri R, Dalpra L (2000a) Molecular definition of Xq common-deleted region in patients affected by premature ovarian failure. *Hum Genet* 107:304–311
- Marozzi A, Vegetti W, Manfredini E, Tibiletti MG, Testa G, Crosignani PG, Ginelli E, Meneveri R, Dalpra L (2000b) Association between idiopathic premature ovarian failure and fragile X premutation. *Hum Reprod* 15:197–202
- Melica F, Chiodi S, Cristoforoni PM, Ravera GB (1995) Reductive surgery and ovarian function in the human—can reductive ovarian surgery in reproductive age negatively influence fertility and age at onset of menopause? *Int J Fertil Menopausal Stud* 40:79–85
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215
- Morton NE (1959) Genetics tests under incomplete ascertainment. *Am J Hum Genet* 11:1–16
- Murray A, Webb J, Grimley S, Conway G, Jacobs P (1998) Studies of FraXa and FraXe in women with premature ovarian failure. *J Med Genet* 35:637–640
- Noord-Zaadstra BM, Looman CW, Alsbach H, Habbema JD, te Velde ER, Karbaat J (1991) Delaying childbearing: effect of age on fecundity and outcome of pregnancy. *BMJ* 302: 1361–1365
- O’Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266
- Risch N, Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 268: 1584–1589
- Sala C, Arrigo G, Torri G, Martinazzi F, Riva P, Larizza L, Philippe C, Jonveaux P, Sloan F, Labella T, Toniolo D (1997) Eleven X chromosome breakpoints associated with premature ovarian failure (POF) map to a 15-Mb YAC contig spanning Xq21. *Genomics* 40:123–131
- Schlessinger D, Herrera L, Crisponi L, Mumm S, Percesepe A, Pellegrini M, Pilia G, Forabosco A (2002) Genes and translocations involved in POF. *Am J Med Genet* 111:328–333
- Sniieder H, MacGregor AJ, Spector TD (1998) Genes control the cessation of a woman’s reproductive life: a twin study of hysterectomy and age at menopause. *J Clin Endocrinol Metab* 83:1875–1880
- te Velde ER, Dorland M, Broekmans FJ (1998) Age at menopause as a marker of reproductive ageing. *Maturitas* 30: 119–125
- Tibiletti MG, Testa G, Vegetti W, Alagna F, Taborelli M, Dalpra L, Bolis PF, Crosignani PG (1999) The idiopathic forms of premature menopause and early menopause show the same genetic pattern. *Hum Reprod* 14:2731–2734
- Torgerson DJ, Thomas RE, Reid DM (1997) Mothers and daughters menopausal ages: is there a link? *Eur J Obstet Gynecol Reprod Biol* 74:63–66
- Touraine P, Beau I, Gougeon A, Meduri G, Desroches A, Pichard C, Detoef M, Paniel B, Prieur M, Zorn JR, Milgrom E, Kuttann F, Misrahi M (1999) New natural inactivating mutations of the follicle-stimulating hormone receptor: correlations between receptor function and phenotype. *Mol Endocrinol* 13:1844–1854
- Treloar SA, Do KA, Martin NG (1998) Genetic influences on the age at menopause. *Lancet* 352:1084–1085
- van Asselt KM, Kok HS (2003) Age at menopause; a genetic-epidemiologic study. PhD thesis, University Medical Center Utrecht, Utrecht
- van Asselt KM, Kok HS, Peeters PHM, Roest M, Pearson PL, te Velde ER, Grobbee DE, van der Schouw YT (2003) Factor V Leiden mutation accelerates the onset of natural menopause. *Menopause* 10:477–481
- van Noord PA, Dubas JS, Dorland M, Boersma H, te Velde ER (1997) Age at natural menopause in a population-based screening cohort: the role of menarche, fecundity, and lifestyle factors. *Fertil Steril* 68:95–102
- Vegetti W, Marozzi A, Manfredini E, Testa G, Alagna F, Nicolosi A, Caliarì I, Taborelli M, Tibiletti MG, Dalpra L, Crosignani PG (2000) Premature ovarian failure. *Mol Cell Endocrinol* 161:53–57
- Weel AE, Uitterlinden AG, Westendorp IC, Burger HG, Schuit SC, Hofman A, Helmerhorst TJ, van Leeuwen JP, Pols HA (1999) Estrogen receptor polymorphism predicts the onset of natural and surgical menopause. *J Clin Endocrinol Metab* 84:3146–3150
- Wood JW (1989) Fecundity and natural fertility in humans. *Oxf Rev Reprod Biol* 11:61–109
- World Health Organisation Scientific Group (1996) Research on the menopause in the 1990s. World Health Organisation Technical Services Department series no 866. World Health Organisation, Geneva