Human Reproduction Update, Vol.11, No.5 pp. 483–493, 2005 doi:10.1093/humupd/dmi024
Advance Access publication July 15, 2005

Genetic studies to identify genes underlying menopausal age

Helen S.Kok[1,2,3,4], Kristel M.van Asselt[1,2,3], Yvonne T.van der Schouw[1], Petra H.M.Peeters[1] and Cisca Wijmenga[2]

[1]Julius Center for Health Sciences and Primary Care, [2]Department of Biomedical Genetics and [3]Department of Reproductive Medicine, University Medical Center Utrecht, The Netherlands

[4]To whom correspondence should be addressed at: Department of Obstetrics & Gynaecology, Onze Lieve Vrouwe Gasthuis, P.O. Box 95500, 1090 HM Amsterdam, The Netherlands. E-mail: hskok@chello.nl

Menopausal age is important as a retrospective marker for ovarian senescence, an early menopausal age is associated with an increased risk of cardiovascular diseases and osteoporosis, whereas a later menopausal age has been associated with an increased risk of breast cancer. The worldwide average for age at natural menopause is approximately 51 years and is more or less normally distributed with a range roughly between 40 and 60 years. Environmental factors explain only a small part of the variance and it has been proposed that genetic factors are the main source of variation. Menopausal age may be considered a continuous complex trait. Complex traits are defined as traits that are influenced by both multiple genetic and environmental factors. A category of complex traits comprises those that are measured on a continuous scale. The genomic loci that make up the genetic component are called 'quantitative trait loci' or QTLs. The first linkage study on menopausal age suggests that the involvement of the X-chromosome may not be limited to premature ovarian failure (POF), but may influence the broader spectrum of menopausal age. A potentially new locus for variation in menopausal age was allocated to chromosome 9. Further studies need to identify new candidate genes to help unravel the pathophysiology of menopausal age. It is becoming increasingly clear that, in any speciality, it should be acknowledged that genetic factors are involved in many traits and that uncovering these factors may provide insight into pathogenesis and ultimately advance prevention and treatment of disease. In this review we discuss methods and basic principles of gene finding for such traits, exemplified by menopausal age as phenotype. Furthermore, we give an overview of the state of the art of candidate gene studies and linkage studies.

Key words: complex trait/genes/linkage design/menopausal age/quantitative trait loci

Introduction

Traits that are determined by the action of a single gene are usually discrete in nature. Examples are ABO blood groups and rare dominant and recessive diseases. For continuous characters, like height, the situation is less clear. Although many such characters demonstrate familial resemblance (tall parents are more likely to have tall children than short parents), they cannot be solely determined by the action of single genes. The reason for this is simply that, by definition, the genotype at a locus is not continuous but discrete in nature. Multiple genes must be involved, possibly along with environmental factors (Sham 1998) Traits that are influenced both by multiple genetic and environmental factors are called 'complex traits'. The individual loci that make up the genetic component of a quantitative trait are called 'quantitative trait loci', abbreviated as QTL.

A complete model of causation, the ultimate aim of aetiological research, would include all these factors (all genetic and non-genetic factors), their joint frequency distribution in the population and their joint effect on the trait. Although in attempting to localise a gene we are focusing on only one component of the complex causal system, it is a first step in gaining further understanding in the physiological mechanisms.

Menopausal age can be regarded as such a quantitative trait. In this article we aim to connect menopausal age and genetics in a broad sense. First, a brief overview of current knowledge on menopausal age and genetic factors will be given; second, we describe basic principles of genetic research into identification of genes involved in menopausal age (including the first genome-wide scan on menopausal age) (Van Asselt and Kok *et al*., 2004b) and third, we discuss possible future studies that may broaden and deepen genetic insight on this topic.

Menopausal age

The most widely used definition for natural menopause is as defined by the World Health Organisation as at least twelve consecutive months of amenorrhea not because of surgery or other obvious causes (World Health Organisation Scientific Group, 1996). When referring to menopausal age or onset of menopause

in this article, we mean natural menopause as defined above. The worldwide average for menopausal age is approximately 51 years and is more or less normally distributed with a range roughly between 40 and 60 years (Treloar, 1981; Bongaarts, 1982; Spira, 1988; Wood, 1989; Morabia and Costanza, 1998).

In approximately 1% of all women, menopause is reached before the age of 40 years and is defined as premature ovarian failure (POF) (Coulam, 1982; Coulam *et al*., 1986a). This definition is rather arbitrary, because it is based on age only. POF is a collective term for which proposed causes include autoimmune disease, syndromes such as fragile X, radiation, chemotherapy and certain genes.

Both isolated and familial cases of POF have been described, familial transmission is observed in 30–40% (Tibiletti *et al*., 1999; Vegetti *et al*., 2000). Only a small part can be traced to a known cause, among which are monogenetic factors. Genetic factors involved in the 'normal' range (±40–60 years) of age at natural menopause remain largely unknown.

Menopausal age is, amongst other things, important as a retrospective marker for ovarian senescence (te Velde *et al*., 1998; Marozzi *et al*., 2000a). In many industrialized countries with access to efficient contraceptive methods, a trend in delay of childbirth can be observed. In the Netherlands, the average maternal age at first childbirth has increased by 4.8 years in the last three decades, from 24.3 to 29.1 years, whereas the average age for all childbirths increased from 27.4 to 30.7 years (www.cbs.nl). As a consequence, an increasing number of women are confronted with unexpected and unwanted infertility. Furthermore, an early menopausal age is associated with an increased risk of cardiovascular diseases (van der Schouw *et al*., 1996; Wise *et al*., 1996) and osteoporosis (Kritz-Silverstein and Barrett-Connor, 1993; Barrett-Connor and Goodman-Gruen, 1995; Osei-Hyiaman *et al*., 1998), whereas a later menopausal age has been associated with an increased risk of breast cancer (Mullis *et al*., 1997).

Although this article focuses on detection of genes influencing the 'normal' variation in menopausal age, we will consider POF as well because these entities may be closely connected (*Power calculations and heterogenity*).

Environmental factors and menopausal age

Many environmental and lifestyle factors have claimed to affect menopausal age (Table I), but inconsistencies exist for the majority of factors with the exception of cigarette smoking (Midgette and Baron, 1990; Bromberger *et al*., 1997; Gold *et al*., 2001; Van Asselt *et al*., 2004a). In fact, it was the absence of any significant large environmental or lifestyle factors to explain the wide variation in age at menopause that fostered the speculation of genetic factors being the main source of variation in menopause age (van Noord *et al*., 1997). It has been estimated that environmental factors explain only about 3% of the variance (van Noord *et al*., 1997).

Genetic factors and menopausal age

Although quantitative traits are generally believed to have an important genetic component, it is desirable to establish the plausibility of involvement of genetic factors. An indication that genes are involved in a trait is the presence of familial clustering. However, it should be emphasized that familial clustering can also be determined by shared environmental factors. Heritability estimates have been derived from family and twin studies and have shown that the variation in menopausal age may be explained substantially by genetic variance. Estimates have been reported from 30 up to 85% (Snieder *et al*., 1998; Treloar *et al*., 1998; de Bruin *et al*., 2001; Kok *et al*., 2004). Because heritability is a relative measure (it is the proportion of the total phenotypic variation within a population that can be attributed to the genetic variance) the absolute genetic variance cannot be inferred. Comparisons of heritability of menopausal age across populations is not meaningful because of different nongenetic influences (measurement errors, environmental factors, etc.).

The genetic causes of POF are several X-chromosome deletions (Devi and Benn, 1999; Davis *et al*., 2000), fragile-X carriership (Allingham-Hawkins *et al*., 1999; Marozzi *et al*., 2000b; Hundscheid *et al*., 2001; Welt *et al*., 2004), FSH-receptor mutations (Aittomaki *et al*., 1995; Gromoll *et al*., 1996; Beau *et al*., 1998; Touraine *et al*., 1999; Doherty *et al*., 2002) and mutations in the FOXL2 gene (Crisponi *et al*., 2001; Schlessinger *et al*., 2002) and inhibin alpha gene (Shelling *et al*., 2000; Marozzi *et al*., 2002). Most of these factors were found by using extreme phenotypes only. It is unlikely that genetic factors involved in variation in menopausal age between 40 and 60 years will be uncovered in this manner as the relation between both phenotypes is uncertain. Compare, for example, mapping genes for body length. It is unlikely to map common genes for length by including only dwarfs or individuals with Marfan's syndrome. A useful design for exploring potential causal pathways between genetic determinants and a complex trait like menopausal age is the so–called association study using candidate genes. The principle is to study allelic associations between a trait and a marker in a set of cases (individuals with the trait) and controls (individuals without the trait) from the same population. An allele that is positively associated with the trait is considered analogous to a risk factor. These candidate genes are chosen usually because knowledge about their function may indicate a role in the trait of interest and also contain known genetic variants (polymorphisms). In this approach, it is tested whether the polymorphisms statistically associate with a defined part of the phenotypic variation of the trait under study. The estrogen receptor (ER) is a steroid transcription factor that acts as regulator of the expression of many genes and proteins (Evans, 1988; Giguere *et al*., 1988) and is also an important regulator of growth and differentiation in many tissues (Wilson *et al*., 1980; Haslam and Shyamala, 1981; Eriksen *et al*., 1988; Thomas *et al*., 1993). The ERα gene is located on the long arm of chromosome 6 and several polymorphic sites in the gene are known. Weel *et al*. (Weel *et al*., 1999) found that a non-coding polymorphism (*Pvu*II) of the ERα gene was associated with onset of menopause in a Dutch population. Women carrying the PP genotype of the *Pvu*II restriction fragment length polymorphism were found to have a 1.1-year earlier onset of menopause compared to women with the pp genotype. However, two other studies, one carried out in a Japanese population and the other in also a Dutch population, could not replicate this finding (Gorai *et al*., 2003; Kok *et al*., 2005).

The effect of the factor V Leiden mutation, a candidate gene representing a different pathway, on menopausal age was examined both individually and combined with smoking (Van Asselt *et al*., 2003). It was found that the factor V Leiden mutation and menopausal age were related, an association possibly enhanced by smoking.

Table I. Determinants of age at natural menopause reported in previous studies

Determinant	References
Age at first childbirth	Jeune (1986); Stanford *et al.* (1987); Luoto *et al.* (1994); Cramer *et al.* (1995)b; Cassou *et al.* (1997); van Noord *et al.* (1997); Whelan *et al.* (1990)
Age at last childbirth	van Keep *et al.* (1979); Jeune (1986); Luoto *et al.* (1994); Cramer *et al.* (1995b); Do *et al.* (1998)
Age at menarche	van Keep *et al.* (1979); Sherman *et al.* (1981); Neri *et al.* (1982); Stanford *et al.* (1987); Whelan *et al.* (1990); Parazzini *et al.* (1992); Cramer *et al.* (1995b); Cramer *et al.* (1995c); van Noord *et al.* (1997); Do *et al.* (1998); Kato *et al.* (1998); Rizk *et al.* (1998); Nagata *et al.* (1998); Hardy *et al.* (2000); Meschia *et al.* (2000); Cooper *et al.* (2001)
Alcohol use	Cramer *et al.* (1995a); Nilsson *et al.* (1997); Torgerson *et al.* (1997a); Do *et al.* (1998); Neslihan *et al.* (1998); Richards *et al.* (1999); Torgerson *et al.* (1994); Cooper *et al.* (2001)
Birth weight	Treloar *et al.* (2000)
Body mass index	Neri *et al.* (1982); Willett *et al.* (1983); Stanford *et al.* (1987); Brambilla and McKinlay (1989); Luoto *et al.* (1994); Cramer *et al.* (1995b); Cramer *et al.* (1995c); Bromberger *et al.* (1997); van Noord *et al.* (1997); Do *et al.* (1998); Kato *et al.* (1998); Nagata *et al.* (1998); Neslihan *et al.* (1998); Rizk *et al.* (1998); Richards *et al.* (1999); Hardy *et al.* (2000); Meschia *et al.* (2000); Cooper *et al.* (2001); de Vries *et al.* (2001); Gold *et al.* (2001)
Breastfeeding	Jeune (1986); Stanford *et al.* (1987); Whelan *et al.* (1990); Cassou *et al.* (1997)
Coffee consumption	Cramer *et al.* (1995a); Nilsson *et al.* (1997); Nagata *et al.* (1998)
Cognition	Richards *et al.* (1999)
Depression	Harlow *et al.* (1995); Richards *et al.* (1999)
Diet	Cramer *et al.* (1995a); Nagata *et al.* (1998); Bromberger *et al.* (1997)
Educational level	Neri *et al.* (1982); McKinlay *et al.* (1985); Stanford *et al.* (1987); Cramer *et al.* (1995a); Cramer *et al.* (1995c); Do *et al.* (1998); Brambilla *et al.* (1989); Parazzini *et al.* (1992);Luoto *et al.* (1994); Bromberger *et al.* (1997); Nagata *et al.* (1998); Neslihan *et al.* (1998); Hardy *et al.* (2000); Meschia *et al.* (2000); Cooper *et al.* (2001); Gold *et al.* (2001)
Ethnicity	Bromberger *et al.* (1997); Stanford *et al.* (1987); Kato *et al.* (1998); Cooper *et al.* (2001); Gold *et al.* (2001); Neri *et al.* (1982)
Employment	van Keep *et al.* (1979); Jeune (1986); Luoto *et al.* (1994); Neslihan *et al.* (1998); Cassou *et al.* (1997); Do *et al.* (1998); Gold *et al.* (2001)
Family history of Menopausal age	Torgerson *et al.* (1994); Cramer *et al.* (1995c); Torgerson *et al.* (1997a); Torgerson *et al.* (1997b); Rizk *et al.* (1998)
Height	van Keep *et al.* (1979); Stanford *et al.* (1987); Brambilla *et al.* (1989); Cramer *et al.* (1995c); van Noord *et al.* (1997)
Income	Brambilla *et al.* (1989); Stanford *et al.* (1987); Do *et al.* (1998)
Left-handedness	Leidy (1990); Pavia *et al.* (1994)
Marital status	van Keep *et al.* (1979); Neri *et al.* (1982); Stanford *et al.* (1987); McKinlay *et al.* (1985); Brambilla *et al.* (1989); Parazzini *et al.* (1992); Cramer *et al.* (1995b); Cassou *et al.* (1997); Nilsson *et al.* (1997); Neslihan *et al.* (1998); Hardy *et al.* (2000); Meschia *et al.* (2000); Gold *et al.* (2001)
Meat consumption	Torgerson *et al.* (1994); Torgerson *et al.* (1997a)
Menstrual cycle irregularity	Stanford *et al.* (1987); Parazzini *et al.* (1992); Cramer *et al.* (1995b); Bromberger *et al.* (1997); Meschia *et al.* (2000)
Menstural cycle length	Whelan *et al.* (1990); Cramer *et al.* (1995b); Cramer *et al.* (1995c); den Tonkelaar *et al.* (1998); Hardy *et al.* (1999)
Miscarriages	van Keep *et al.* (1979); Neri *et al.* (1982); Whelan *et al.* (1990); Parazzini *et al.* (1992); van Noord *et al.* (1997)
Oral contraceptive use	van Keep *et al.* (1979); Neri *et al.* (1982); Stanford *et al.* (1987); Brambilla *et al.* (1989); Cramer *et al.* (1995a); Cramer *et al.* (1995b); Bromberger *et al.* (1997); Cassou *et al.* (1997); van Noord *et al.* (1997); Neslihan *et al.* (1998); Rizk *et al.* (1998); Hardy and Kuh (1999); Richards *et al.* (1999); Meschia *et al.* (2000); Cooper *et al.* (2001); de Vries *et al.* (2001); Gold *et al.* (2001)
Parity or number of pregnancies	van Keep *et al.* (1979); Neri *et al.* (1982); Willett *et al.* (1983); Walker *et al.* (1984); McKinlay *et al.* (1985); Jeune (1986);Stanford *et al.* (1987); Brambilla *et al.* (1989); Whelan *et al.* (1990); Parazzini *et al.* (1992); Luoto *et al.* (1994); Torgerson *et al.* (1994); Cramer *et al.* (1995a); Cramer *et al.* (1995b); Cramer *et al.* (1995c); Bromberger *et al.* (1997); Torgerson *et al.* (1997a); Do *et al.* (1998); Kato *et al.* (1998); Nagata *et al.* (1998);Neslihan *et al.* (1998); Rizk *et al.* (1998); Westendorp and Kirkwood (1998); Hardy *et al.* (1999); Richards *et al.* (1999); Meschia *et al.* (2000); Cooper *et al.* (2001); Gold *et al.* (2001); de Vries *et al.* (2001);
Physical acitivity	Bromberger *et al.* (1997); Nagata *et al.* (1998); Cooper *et al.* (2001); Gold *et al.* (2001)
Psychosocial stress	Bromberger *et al.* (1997); Neri *et al.* (1982)
Rank in birth order	Van Noord *et al.* (1997)
Religion	Cramer *et al.* (1995a); Kato *et al.* (1998)
Siblings	Cramer *et al.* (1995c); van Noord *et al.* (1997)
Smoking	van Keep *et al.* (1979); Adena and Gallagher (1982); Neri *et al.* (1982); Willett *et al.* (1983); McKinlay *et al.* (1985); Jeune (1986); Stanford *et al.* (1987); Brambilla *et al.* (1989); Midgette and Baron (1990); Parazzini *et al.* (1992); Luoto *et al.* (1994); Torgerson *et al.* (1994); Cramer *et al.* (1995a); Cramer *et al.* (1995b); Cramer *et al.* (1995c); Bromberger *et al.* (1997); Cassou *et al.* (1997); Nilsson *et al.* (1997); Torgerson *et al.* (1997a); van Noord *et al.* (1997); Do *et al.* (1998); Kato *et al.* (1998); Nagata *et al.* (1998); Cooper *et al.* (1999); Richards *et al.* (1999); Hardy *et al.* (2000); Meschia *et al.* (2000); Cooper *et al.* (2001); de Vries *et al.* (2001); Gold *et al.* (2001)
Socio-economic status	McKinlay *et al.* (1985); Torgerson *et al.* (1994); van Noord *et al.* (1997); Torgerson *et al.* (1997a); Nilsson *et al.* (1997); Do *et al.* (1998); Hardy *et al.* (2000); Gold *et al.* (2001); de Vries *et al.* (2001);
Type 2 diabetes	Lopez-Lopez *et al.* (1999)
Unilateral Oophrectomy	Cramer *et al.* (1995b); Melica *et al.* (1995); Hardy *et al.* (1999)
Year of birth	Parazzini *et al.* (1992); Do *et al.* (1998)
Weight	van Keep *et al.* (1979); Jeune (1986); Stanford *et al.* (1987); van Noord *et al.* (1997)
Weight gain	Cramer *et al.* (1995b); Bromberger *et al.* (1997)
Weight reduction diet	Bromberger *et al.* (1997)

Recently, one study found a significant association between the apolipoprotein E (APOE) genotype and age at menopause, where carriers of the APOE4 allele reached menopause at an earlier age (Koochmeshgi *et al*., 2004).

An appealing characteristic of an association study involves the simple study design of collecting cases and controls and comparing both groups using basic statistical techniques. There is no need to define a mode of inheritance, and it allows the detection of genes with a weak effect. Furthermore, not only direct effects (a 'true variant'), but also indirect effects, through linkage disequilibrium (LD) mapping, can be tested using this design (LD mapping will be discussed below).

Yet, certain restrictions should be considered. A limiting factor may be the large number of case–control pairs required. In statistical terms, the problem of multiple testing arises, because of the need to test many variables. Findings of association can be influenced by the same problems present in 'classical' epidemiological studies, such as selection bias, confounding and misclassification. Most important, representing a basic condition in this design, is to be able to define the correct candidate genes. This requires a priori knowledge on underlying biology and (patho-) physiologic pathways. Therefore, a hypothesis-free approach would be preferred; ideal would be to consider all genes and all variants in all of the genome; this is, however, still not feasible. So, in absence of concretely defined candidate genes, one still has to rely on linkage studies.

If no prior hypothesis of possible genes involved in a trait or disease is available, a linkage-based genome-wide scan may be applied. The aim of such a study is to identify genomic loci that are involved in the trait under study without prior assumptions about the underlying mechanism. Because this method rather detects the loci and not necessarily the genes itself, subsequently these regions can be subjected to a search for possible genes, the positional candidate genes.

Classical linkage methods, however, do require information on parameters such as mode of inheritance and the gene frequency in the population. This information can usually be obtained for monogenetic diseases; examples of diseases in which this approach has been particularly successful are breast cancer (BRCA-1/2 gene) (Miki *et al*., 1994; Wooster *et al*., 1995; Tavtigian *et al*., 1996), Huntington's disease (Gusella *et al*., 1983) and cystic fibrosis (Knowlton *et al*. 1985). For complex traits, those parameters are usually not available. Therefore, nonparametric linkage methods (also known as model-free linkage methods) have been developed for complex traits and diseases. QTL analysis is such a nonparametric method used for quantitative traits.

Genetic approach

QTL analysis

QTL analysis is based on the number of alleles that are shared at a chromosomal locus among family members and the phenotypic similarity between these family members. Phenotypic similarity is defined as no difference or a small difference in trait-value. Usually sibling pairs (now abbreviated as sibpairs) are used. In general, sibpairs are made up of brother and sister, brother and brother or sister and sister; in case of menopausal age as trait, it obviously concerns sister pairs only.

The underlying idea of QTL analysis is that sibs showing a comparable value of a specific trait have inherited from their parents the same genetic variant and surrounding chromosomal material responsible for this trait variation.

Obviously, all sibpairs are required to have a known trait-value. In the case of natural menopause as phenotype, sibpairs with a known age at natural menopause are needed. Women who have had a surgical or chemical menopause and women, whose menopausal age is obscured by use of hormonal replacement therapy or oral contraceptives are therefore not eligible.

QTL analysis is based on the principle of 'identity by descent' (IBD).

Identity by descent

On each locus, two individuals can share 0, 1 or 2 alleles that have been inherited identically from the same parent. Sibpairs can share none of the two alleles, either the maternal or paternal allele only, or both the maternal and paternal alleles (Figure 1). An allele is considered IBD if, in two individuals, a copy of the *same* allele has been inherited from the *same* parent, as opposed to the situation in which two identical alleles have *not* been inherited from the same parent. The latter is called 'identity by state' (IBS) (Figure 2). On average, two sibs are genetically identical for 50%; in 50% of all cases they share 1 allele IBD; in 25% they do not share any identical alleles IBD and in 25% they share both alleles IBD.

If at least one of the parents is homozygous or if both parents share the same alleles, one cannot determine with certainty whether the shared allele in their offspring is an identical copy of the same allele or whether both alleles just 'look alike'. If the parents are not available for genetic testing, the IBD status of a sibpair cannot be determined with certainty either. In both cases, the IBD status is estimated on the basis of the occurrence of the alleles in the population.

For example, two sisters have the following genotypes A1A2 and A1A3 at a certain marker. Both have an allele A1, but no information on parents is available. One has to determine what the chances are of these A1 alleles being identical by descent. If the A1 allele rarely occurs in the population, it is very likely the sisters inherited the same allele from one parent. If, on the other hand, the A1 allele is very common in the given population, it is less likely the alleles of the sisters are identical.

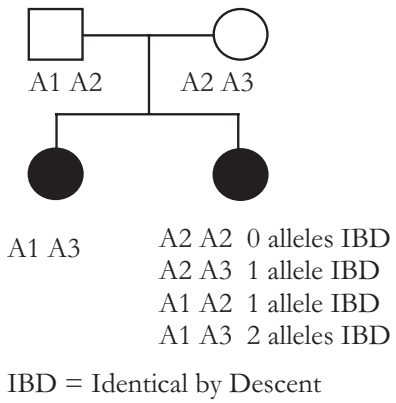

Figure 1. Examples of 'identity by descent' (IBD).

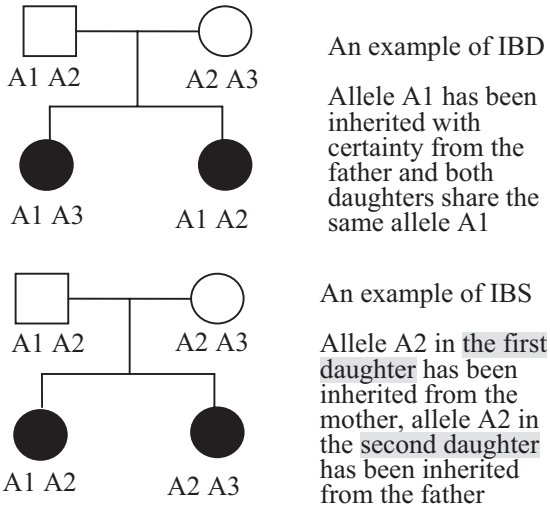

Figure 2. Comparison of 'identity by state' and 'identity by descent'.

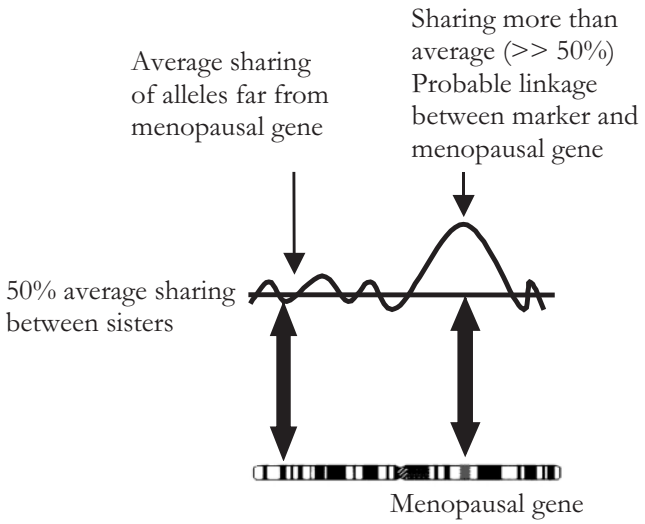

Figure 3. QTL analysis for gene sharing between siblings.

The absence of parental genotype data can sometimes be compensated – at least partially – by genotyping other brothers or sisters of the included sibpair. Their genotypes can then be used to reconstruct the genotypes of both parents, making it possible to establish the IBD status of the sibpair with a higher degree of certainty. As these extra brothers or sisters are relevant only with respect to their genotypes – as is the case with parents – menopausal status or gender are now irrelevant.

Thus for menopausal age, the assumption is that sisters who show a relatively small difference in menopausal age share more alleles IBD on a locus located near a gene that influences variation in menopausal age, than sisters whose menopausal age is very different.

Genome scan

The QTL analysis can be applied to the whole human genome. This basically means performing QTL analyses on many locations throughout the genome to locate regions where the IBD sharing between siblings is significantly higher than expected by chance only. Genetic markers are used to test the percentage of IBD sharing at a particular locus. On average siblings will share half of their genes, however, in areas near susceptibility genes, it is expected that the sharing is significantly more than 50% in siblings who have similar trait-values (Figure 3). This is indicative for a candidate gene located near these particular markers. Many markers with known chromosomal locations are now available. Microsatellite markers with a high degree of heterozygosity, ensuring the presence of many different alleles, are especially suitable for this purpose. There is an overall consensus that an average distance of 10 cM between the markers for a first scan is sufficient for detecting linkage between a marker and the phenotype. This corresponds to approximately 400 markers that are evenly distributed over all chromosomes. The plausibility of linkage between the trait and a specific marker can be expressed by means of the so-called LOD score (log of the odds that loci are linked). A LOD score of 3 on a specific marker means that supposed linkage with the marker is 1000 times more likely than the assumption that the marker is not linked. Regions with a LOD score of 1 and more may be considered potentially interesting. However, only Log Odds (ie the log of the odds that low are linked-LOD) scores of 3.3 and up are considered statistically significant (Lander and Kruglyak, 1995). The appeal of QTL analysis lies in the fact that it is model-free and requires no information on gene function; however, there is a 'hidden' assumption: only genes that are relatively rare and have a relatively large impact will be detected by QTL analysis. Common genes with a small contribution are difficult, if not impossible, to detect. It should be noted that regions are detected by QTL analysis, not genes, and those regions can be uncomfortably large, containing up to hundred genes.

Power calculations and heterogeneity

Unfortunately, no simple guidelines are available to determine the sample size needed to perform a genome-wide scan. Though, it can be inferred that quite large numbers of sibpairs are needed since the goal of QTL analysis is to determine a statistically significant deviation from the average 50% IBD sharing with only three options of IBD sharing possible, i.e. none, one or two.

Power calculations to determine the sample size needed can only be used as approximation. Paradoxically, the power calculations for a QTL analysis, which is model-free, require imputation of exactly those parameters that are unknown. Examples of such parameters are the part of the total variance in the phenotype that is attributed to genetic causes; the number of loci that possibly play a role and the frequency with which these loci occur in the population. These assumptions rather limit the practical meaning of power calculations performed prior to QTL analysis. Alternatively, power calculations can be used to calculate the maximum LOD scores that may be expected in the collected sample of sibpairs. Again, assumptions have to be made on the number of loci and the explained variance of loci. This may still prove useful, because a minimum contribution of each locus has to be assumed anyway, otherwise performing a genome-wide scan is useless as these loci will not be detected.

With respect to quantitative sibpair analysis, it has been suggested that power is enhanced when sibpairs are used that show extreme trait-values: selected sampling (Risch and Zhang, 1995).

Extreme implies that only those sibpairs are considered which have values belonging to the tails of the distribution.

The number of families needed could then be reduced accordingly. The underlying idea is that genetic changes causing a large displacement are more frequently concentrated in the extremes of the distribution and can therefore be distinguished more easily. For an analysis of variation in menopausal age, this means ascertaining families with very early or very late menopausal ages. A problem comes up because in the lower tail of the distribution of menopausal age, women could have an early menopause as a result of just being in the left tail of the normal distribution or they could have an early cessation of menstrual cycles because of the early mentioned POF.

Natural menopause follows an approximate Gaussian distribution; it exhibits a modest skewness, expressed in a long left tail. Skewness may be an indicator of heterogeneity in which two or more overlapping distributions add up to one non-normal distribution. This may be indicative of genetic heterogeneity produced by an overlap of two different genetic phenotypes and/or the presence of major genes predominately contributing to one side of the phenotypic variation.

Strong 'POF-genes' could be responsible for the described deviation of normal distribution. And that would be the trick, analyse only women within the normal variation of menopausal age. Next problem is how to identify women with POF, is the age at menopause enough to distinguish these women? Both normal early menopause (40–45 years) and POF (<40 years) have been described in the same families indicating a common causality and that an arbitrary age cut-off of 40 years does not provide a reliable method for distinguishing between the two (Tibiletti *et al*., 1999).

We have tested heterogeneity of menopausal age in two large population-based cohorts to find statistical evidence for possible admixture. Data from the Prospect and Diagnostic Investigation Mamma Cancer (DOM) cohorts were used to test our hypothesis that the distribution of age at menopause may be explained by more than one underlying distribution (de Waard *et al.*, 1984; Boker *et al*., 2001). Only women who had reached an age at study recruitment of 58 years or higher were included. At this age 98% of women had already reached menopause. This age-dependent selection should prevent an oversampling of women with an early age at menopause, which would distort the actual distribution; an excess of women with an early age at menopause because of inclusion of younger women could also be a cause of skewness. We selected women following a natural menopause, which means at least 12 months without periods and without medical or surgical interventions that affect menopausal age. Menopausal status was determined by means of questionnaires. 4037 and 9332 women were included from the Prospect and DOM cohorts, respectively.

A maximum likelihood function was used to fit two alternative models to the observed age of menopause, the first based on a single normal distribution and the second on two overlapping normal distributions to each of the cohorts and to estimate their means and standard deviations. The variance of the two components may be unequal. Hypothesis testing was carried out by comparing the models using likelihood ratio tests to select the best-fitting model. Twice the difference in log-likelihood between the two models was assumed to be distributed as chi-square, with the degrees of freedom equal to the difference in the number of parameters between the models. The distribution of age at menopause in both cohorts is skewed to the left side. The results of the test for heterogeneity are shown in Table II. Although the maximum likelihood test for admixture does not exhaustively prove that two such populations exist, two overlapping normal distributions explain the data better than one. A genome scan for menopausal age should therefore include subgroup analyses to avoid possible admixture.

Table II. Results of the mixture test of menopausal age, mean menopausal age and standard deviations (SD) are given

Prospect cohort				DOM cohort			
Mean	50.1	Mean 1	46.1	Mean	49.9	Mean1	46.7
SD	4.2	Mean 2	51.3	SD	4.0	Mean 2	51.3
		SD1	5.8			SD1	4.7
		SD2	2.9			SD2	2.7
Proportion in early group			0.21	Proportion in early group			0.31
Difference –2ln = 571				Difference –2ln = 571			

First genome-wide scan on menopausal age

A total of 165 Dutch families were ascertained using extreme selected sampling and genotyped for 417 markers. Subgroup analyses with women with either a very early menopause or a very late menopause have been performed. Twelve chromosomes had a LOD score of 1.0 or higher. Two regions showed suggestive linkage; chromosome region 9q21.3 (LOD score 2.6) and chromosome region Xp21.3 (LOD score 3.1). The finding of the region on the X-chromosome comes as no surprise, considering its widespread involvement in POF, defining which particular gene involved is of great interest. The region on chromosome 9 deserves further consideration. 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 which are known to modulate the rate of human ovarian ageing have not yet been described. None of the known candidate genes (*Genetic factors and menopausal age*) are in one of these regions. Both findings require independent confirmation.

The subgroup analyses showed that the linkage found on the X-chromosome was entirely because of the women with early menopausal ages. To establish whether it is likely that the peak is only based on women with extremely early menopause, women who may be classified as having POF according to the most widely used criterion, age at menopause at or under 40 years (Coulam *et al*., 1986b), were excluded. 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 only based on POF women, and stated otherwise, the 'POF-region' is

not only important for POF but possibly also for 'early menopause'. 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.

Factors to consider in QTL analysis

Phenotype issues

The success of QTL analysis is highly dependent on the strength of the relationship between the genotype and phenotype. The phenotype under study must therefore be well defined by stringent criteria in order to obtain a homogeneous study population. Phenotypic heterogeneity may lead to serious loss of power when the differences in phenotype have different underlying genetic causes. For menopausal age, phenotypic heterogeneity may be expected in the lower ages at menopause. Women exhibiting extremely early age at menopause because of monogenetic causes (POF) may preferably be excluded from the study aiming to identify genes involved in the normal variation of menopausal age.

Besides well defining the phenotype, an additional difficulty is the degree of certainty with which the phenotype can be determined in individuals. Specifically, menopausal age can be established only after the menstrual cycle has stopped for no less than a year. This definition, however, is subject to a certain degree of inaccuracy (Colditz *et al*., 1987; den Tonkelaar, 1997). The more the time has elapsed between menopause and the moment of inclusion, the higher this degree of inaccuracy gets. In an ideal situation, age at menopause would be determined by extensive follow-up of premenopausal women and one could argue that biochemical measures, for instance FSH-levels, are desirable. However, in practice this may prove unrealistic because of high costs and the time involved.

Family participation

As explained earlier, effort should be made to include parents or additional siblings, as the IBD status of a sibpair can be determined with more certainty. For menopause, and other phenotypes that are expressed or measured at an advanced age, the number of parents (and even siblings) available for participation is limited. Additionally, the fact that the average family size is relatively small nowadays forms a potential limitation for the availability of extra siblings.

Future prospects

The process of mapping genetic determinants for complex traits can be divided into several chronological stages which include the initial detection of linkage to a QTL, refining the chromosome localization or fine mapping and finally positionally cloning the gene concerned and detecting the genetic variant(s) contributing to the trait phenotype. Different options are possible and we will briefly discuss each of them.

Fine mapping

This is a procedure in which the linked region with suggestive or significant linkage is saturated with additional markers in the hope of increasing the LOD score and thereby reducing the error in determining the location of the putative gene.

The information content can be improved by typing additional markers, but this improvement also depends on the pedigrees and the presence of sufficient additional siblings or parents. If the IBD status cannot be determined unambiguously, extra markers are unlikely to significantly increase the information.

A simulation study was performed to investigate the added value of the fine mapping in QTLs; when the information content increased considerably, the LOD scores only slightly increased. Although the error in the estimated location diminishes, it remains largely dependent on the height of the first LOD score (Atwood and Heard-Costa, 2003). The marker bias (the phenomenon that maximum LOD scores tend to occur at marker locations) can be quite large when maximum LOD scores are below 3.0, but is reduced for higher LOD scores. Why are these results so disappointing? The relationship that is assessed in a linkage study, namely that between genetic information and phenotype, depends on two relationships: (i) between the QTL and the genetic information and (ii) between the QTL and the phenotype. For a given sample size, fine mapping yields nearly perfect genetic information and implies that the power of the linkage test is much more dependent on the strength of the relationship between the QTL and phenotype. Here, the limiting factor is the proportion of variation due to the QTL, represented by the QTL–phenotype relationship. When this proportion is small, improving the genetic information (fine mapping) is largely futile since the largest source of error lies in the weak QTL–phenotype relationship. The quality of the results will be dependent on the proportion of the total genetic variation explained by the linked QTL. On the other hand, there may be some basis for improving the quality of the segregation data by increasing the sample size (number of sibpairs) and this is a strategy that is now being implemented in some large collaborative studies in complex diseases such as diabetes types I and II, in which thousands of sibpairs are now being collected. This would also increase the likelihood of detecting QTLs with a small contribution to the phenotypic variation.

Association studies

LD mapping

Linkage disequilibrium (LD) or association mapping is based on sharing haplotypes (Cardon and Bell, 2001). The underlying assumption is that the marker alleles are closely associated with alleles of the QTLs being sought. When a set of markers is linked to the trait under study, one could test whether a certain combination of alleles (a haplotype) is associated to the trait. This could narrow down the linked region. Heterogeneous populations, populations in which multiple founder mutations segregate, are not suitable for haplotype analysis. Consider a mutation that has occurred near a marker, the marker is then associated with a certain disease allele. However, if mutations at a locus have arisen more than once in the history of a population, i.e. more than one founder mutation, the marker will be associated with different alleles for each of the different mutations. The situation where multiple ancestral haplotypes are associated with the trait is not problematic in linkage analysis, but is certainly disturbing in haplotype analysis. A homogeneous population reduces the probability of multiple haplotypes segregating. Accordingly, most

successes in haplotype analyses have been obtained in genetically isolated populations.

In complex traits it will be difficult to narrow down regions by this method for several reasons.

A mutation at a locus may have arisen more than once in the history of a population and may therefore be associated with multiple ancestral haplotypes. Secondly, the extent of linkage disequilibrium around a mutation is dependent on the history of the population. On the other hand, if the QTLs reside in chromosome regions exhibiting extensive linkage disequilibrium, which can extend up to several 100 kb they will be much easier to pinpoint by association analysis. Furthermore, the detection of single nucleotide polymorphisms (SNPs) at high densities on a genome-wide basis raises the distinct possibility of also being able to detect linkage disequilibrium over small physical distances in the coming years.

Candidate genes

With the availability of an annotated human genome sequence and a rapidly increasing amount of information on the putative function of genes, the size of a candidate region may become less of an issue than it has in the past. Database inspection combined with an insight into what genes may be critical in ovarian ageing can lead to selecting particular genes for further mutation analysis within QTLs. Several pathways might affect menopausal age. Although the involvement of general ageing genes remains under debate, some modifiers of general ageing could also affect the process of ovarian ageing. Other variants influencing the vascular pathways leading to atherosclerosis and thrombosis may be of importance. Since most of the oocytes disappear through apoptosis, genes that determine this process could also be candidates. Mutations in both pro-apoptotic and anti-apoptotic genes have been reported to be important for modulating female fertility in mice (Rucker *et al*., 2000).

Genes important in the development of the genital tract and establishing the initial ovarian follicle pool have been used as candidates in mouse models and could be of importance in humans too (Britt and Findlay, 2002).

Genome-wide association studies/genomics

The international HapMap project aims to determine the common patterns of DNA sequence variation in the human genome and to make this information available for the public domain (The International HapMap Consortium, 2003). This information will allow the indirect association approach to be applied to whole genome scans.

A SNP is a difference between chromosomes in the base present at a particular site in the DNA sequence. These SNPs, results of single historical mutational events, occur at about one in every 1200 bases on average, adding up to approximately 10 million SNPs in the human genome. Nearby SNPs show highly significant levels of linkage disequilibrium and strong associations, so in many parts of our chromosomes, only a few haplotypes are found. 'Tag' SNPs are SNPs that uniquely identify these haplotypes. The number of tag SNPs that contain most of the information about the patterns of genetic variation is estimated to be about 200,000–1,000,000, which is far more convenient than the 10 million common SNPs.

Another way to identify previously unrecognized associations between genes and a disease is by performing gene expression studies. In these studies, the expression of genes in patients' tissue (e.g. ovarian tissue) or blood can be studied by analysing mRNA levels of all the genes and comparing their levels to those of normal controls. Alleles underlying complex traits are expected to have subtle effects on disease risk. Hence, these alleles are more likely to include noncoding regulatory variants with a modest impact on mRNA expression. Recent studies have indeed shown that there is a genetic contribution to polymorphic variation in the level of gene expression (Cheung *et al*., 2003). Interestingly, gene expression will not only reveal potential genetic factors underlying the disease, but also secondary molecular factors that are a consequence of the disease. Such secondary factors are also interesting as they may point towards unrecognized pathophysiological pathways.

Replication in an independent sample

The difficulties introduced by heterogeneity, complexity of mode of inheritance and misclassification regularly lead to dubious linkages. A cautious and appropriate approach to linkage when the initial results are not straightforward is to use a first data set to try to make an appropriate hypothesis of stratifying the sample (for example subgroups) and then to replicate the study in a new cohort using the predetermined additional phenotypic ideas derived from the first analyses. Complete confidence in a finding is given by independent replication.

Conclusions

Menopausal age is genetically complex in origin. It remains unclear whether POF and menopausal age are expressions of the same process or if they are genetically separate entities. In the first case, it may be expected that genes involved overlap, in the latter it may be more feasible that completely different genes are involved. So far, candidate gene studies have not proven very successful. Other studies were not able to confirm the results found for the ER alpha gene, the Factor V Leiden mutation and APOE gene, and furthermore, the effects were small relative to the total variance observed for menopausal age. Genome-wide association studies have the potential to identify new regions harbouring candidate genes. Up to the present, linkage has been more successful as there are now two regions exhibiting suggestive linkage on chromosome 9 and the X-chromosome. Options to proceed include saturating regions appointed by linkage studies with SNPs, performing association studies and possibly also performing gene-expression studies. However, most importantly, an effort should be made to collect independent study populations.

In conclusion, the genome-wide scan on menopausal age is a first attempt to identify QTLs underlying menopausal age. It is clear that ample work still needs to be done to recognise definitive susceptibility genes for variation of menopausal age.

Acknowledgements

Financial support was obtained by N.V.Organon, The Netherlands and the University Medical Center Utrecht.

References

Adena MA and Gallagher HG (1982) Cigarette smoking and the age at menopause. Ann Hum Biol 9,121–130.

Aittomaki K, Lucena JL, Pakarinen P, Sistonen P, Tapanainen J, Gromoll J, Kaskikari R, Sankila EM, Lehvaslaiho H and Engel AR (1995) Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. Cell 82,959–968.

Allingham-Hawkins DJ, Babul-Hirji R, Chitayat D, Holden JJ, Yang KT, Lee C, Hudson R, Gorwill H, Nolin SL, Glicksman A *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.

Atwood LD and Heard-Costa NL (2003) Limits of fine-mapping a quantitative trait. Genet Epidemiol 24,99–106.

Barrett-Connor E and Goodman-Gruen D (1995) Prospective study of endogenous sex hormones and fatal cardiovascular disease in postmenopausal women. BMJ 311,1193–1196.

Beau I, Touraine P, Meduri G, Gougeon A, Desroches A, Matuchansky C, Milgrom E, Kuttenn F and 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 and Hsueh AJ (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 and Peeters PH (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, pp.1–43.

Brambilla DJ and McKinlay SM (1989) A prospective study of factors affecting age at menopause [published erratum appears in J Clin Epidemiol 1990; 43 (5): 537]. J Clin Epidemiol 42,1031–1039.

Britt KL and Findlay JK (2002) Estrogen actions in the ovary revisited. J Endocrinol 175,269–276.

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

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

Cardon LR and Bell JI (2001) Association study designs for complex diseases. Nat Rev Genet 2,91–99.

Cassou B, Derriennic F, Monfort C, Dell'Accio P and Touranchet A (1997) Risk factors of early menopause in two generations of gainfully employed French women. Maturitas 26,165–174.

Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, Morley M and Spielman RS (2003) Natural variation in human gene expression assessed in lymphoblastoid cells. Nat Genet, 33,422–425.

Colditz GA, Stampfer MJ, Willett WC, Stason WB, Rosner B, Hennekens CH and Speizer FE (1987) Reproducibility and validity of self-reported menopausal status in a prospective cohort study. Am J Epidemiol 126,319–325.

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

Cooper GS, Baird DD and Darden FR (2001) Measures of menopausal status in relation to demographic, reproductive, and behavioral characteristics in a population-based study of women aged 35–49 years. Am J Epidemiol 153,1159–1165.

Coulam CB (1982) Premature gonadal failure. Fertil Steril 38,645–655.

Coulam CB, Adamson SC and Annegers JF (1986a) Incidence of premature ovarian failure. Obstet Gynecol 67,604–606.

Coulam CB, Adamson SC and Annegers JF (1986b) Incidence of premature ovarian failure. Obstet Gynecol 67,604–606.

Cramer DW, Harlow BL, Xu H, Fraer C and Barbieri R (1995a) Cross-sectional and case-controlled analyses of the association between smoking and early menopause. Maturitas 22,79–87.

Cramer DW, Xu H and Harlow BL (1995b) Does "incessant" ovulation increase risk for early menopause? Am J Obstet Gynecol 172,568–573.

Cramer DW, Xu H and Harlow BL (1995c) 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 *et al*. (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 and chromosome defect: Opinion. Hum Reprod 15,2418–2422.

Devi A and Benn PA (1999) X-chromosome abnormalities in women with premature ovarian failure. J Reprod Med 44,321–324.

Do KA, Treloar SA, Pandeya N, Purdie D, Green AC, Heath AC and Martin NG (1998) Predictive factors of age at menopause in a large Australian twin study. Hum Biol 70,1073–1091.

Doherty E, Pakarinen P, Tiitinen A, Kiilavuori A, Huhtaniemi I, Forrest S and 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.

Eriksen EF, Colvard DS, Berg NJ, Graham ML, Mann KG, Spelsberg TC and Riggs BL (1988) Evidence of estrogen receptors in normal human osteoblast-like cells. Science 241,84–86.

Evans RM (1988) The steroid and thyroid hormone receptor superfamily. Science 240,889–895.

Giguere V, Yang N, Segui P and Evans RM (1988) Identification of a new class of steroid hormone receptors. Nature 331,91–94.

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

Gorai I, Tanaka K, Inada M, Morinaga H, Uchiyama Y, Kikuchi R, Chaki O and 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.

Gromoll J, Simoni M, Nordhoff V, Behre HM, De Geyter C and Nieschlag E (1996) Functional and clinical consequences of mutations in the FSH receptor. Mol Cell Endocrinol 125,177–182.

Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR, Sakaguchi AY *et al*. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306,234–238.

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

Hardy R, Kuh D and Wadsworth M (2000) Smoking, body mass index, socioeconomic status and the menopausal transition in a British national cohort. Int J Epidemiol 29,845–851.

Harlow BL, Cramer DW and Annis KM (1995) Association of medically treated depression and age at natural menopause. Am J Epidemiol 141,1170–1176.

Haslam SZ and Shyamala G (1981) Relative distribution of estrogen and progesterone receptors among the epithelial, adipose, and connective tissue components of the normal mammary gland. Endocrinology 108,825–830.

Hsueh AJ, Billig H and Tsafriri A (1994) Ovarian follicle atresia: a hormonally controlled apoptotic process. Endocr Rev 15,707–724.

Hundscheid RD, Braat DD, Kiemeney LA, Smits AP and Thomas CM (2001) Increased serum FSH in female fragile X premutation carriers with either regular menstrual cycles or on oral contraceptives. Hum Reprod 16,457–462.

Jeune B (1986) Parity and age at menopause in a Danish sample. Maturitas 8,359–365.

Kaipia A and Hsueh AJ (1997) Regulation of ovarian follicle atresia. Annu Rev Physiol 59,349–363.

Kato I, Toniolo P, Akhmedkhanov A, Koenig KL, Shore R and Zeleniuch-Jacquotte A (1998) Prospective study of factors influencing the onset of natural menopause. J Clin Epidemiol 51,1271–1276.

van Keep PA, Brand PC and Lehert P (1979) Factors affecting the age at menopause. J Biosoc Sci Suppl 6,37–55.

Knowlton RG, Cohen-Haguenauer O, Van Cong N, Frezal J, Brown VA, Barker D, Braman JC, Schumm JW, Tsui LC, Buchwald M *et al*. (1985) A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. Nature 318,380–382.

Kok HS, Van Asselt KM, Peeters PH, van der Schouw YT, Grobbee DE, Pearson PL and Wijmenga C (2004) Age at natural menopause is not linked with the follicle-stimulating hormone receptor region: a sib-pair study. Fertil Steril 81,611–616.

Kok HS, Onland-Moret NC, Van Asselt KM, van Gils CH, van der Schouw YT, Grobbee DE and Peeters PH (2005) No association of estrogen receptor alpha and cytochrome P450c17alpha polymorphisms with age at menopause in a Dutch cohort. Hum Reprod 20,536–542.

Koochmeshgi J, Hosseini-Mazinani SM, Morteza SS, Hosein-Pur-Nobari N and Teimoori-Toolabi L (2004) Apolipoprotein E genotype and age at menopause. Ann N Y Acad Sci 1019,564–567.

Kritz-Silverstein D and Barrett-Connor E (1993) Early menopause, number of reproductive years, and bone mineral density in postmenopausal women. Am J Public Health 83,983–988.

Lander E and Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11,241–247.

Leidy LE (1990) Early age at menopause among left-handed women. Obstet Gynecol 76,1111–1114.

Lopez-Lopez R, Huerta R and Malacara JM (1999) Age at menopause in women with type 2 diabetes mellitus. Menopause 6,174–178.

Luoto R, Kaprio J and Uutela A (1994) Age at natural menopause and sociodemographic status in Finland. Am J Epidemiol 139,64–76.

Marozzi A, Manfredini E, Tibiletti MG, Furlan D, Villa N, Vegetti W, Crosignani PG, Ginelli E, Meneveri R and Dalpra L (2000a) Molecular definition of Xq common-deleted region in patients affected by premature ovarian failure. Hum Genet 107,304–311.

Marozzi A, Manfredini E, Tibiletti MG, Furlan D, Villa N, Vegetti W, Crosignani PG, Ginelli E, Meneveri R and Dalpra L (2000b) Molecular definition of Xq common-deleted region in patients affected by premature ovarian failure. [in Process Citation] Hum Genet 107,304–311.

Marozzi A, Porta C, Vegetti W, Crosignani PG, Tibiletti MG, Dalpra L and Ginelli E (2002) Mutation analysis of the inhibin alpha gene in a cohort of Italian women affected by ovarian failure. Hum Reprod 17,1741–1745.

McKinlay SM, Bifano NL and McKinlay JB (1985) Smoking and age at menopause in women. Ann Intern Med 103,350–356.

Melica F, Chiodi S, Cristoforoni PM and 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.

Meschia M, Pansini F, Modena AB, de Aloysio D, Gambacciani M, Parazzini F, Campagnoli C, Maiocchi G and Peruzzi E (2000) Determinants of age at menopause in Italy: results from a large cross-sectional study. ICARUS Study Group Italian Climacteric Res Group Study. Maturitas 34,119–125.

Midgette AS and Baron JA (1990) Cigarette smoking and the risk of natural menopause. Epidemiology 1,474–480.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W *et al*. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266,66–71.

Morabia A and Costanza MC (1998) International variability in ages at menarche, first livebirth, and menopause. World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives. Am J Epidemiol 148,1195–1205.

Mullis PE, Yoshimura N, Kuhlmann B, Lippuner K, Jaeger P and Harada H (1997) Aromatase deficiency in a female who is compound heterozygote for two new point mutations in the P450arom gene: impact of estrogens on hypergonadotropic hypogonadism, multicystic ovaries, and bone densitometry in childhood. J Clin Endocrinol Metab 82,1739–1745.

Nagata C, Takatsuka N, Inaba S, Kawakami N and Shimizu H (1998) Association of diet and other lifestyle with onset of menopause in Japanese women. Maturitas 29,105–113.

Neri A, Bider D, Lidor Y and Ovadia J (1982) Menopausal age in various ethnic groups in Israel. Maturitas 4,341–348.

Neslihan CS, Bilge SA, Ozturk TN, Oya G, Ece O and Hamiyet B (1998) The menopausal age, related factors and climacteric symptoms in Turkish women. Maturitas 30,37–40.

Nilsson P, Moller L, Koster A and Hollnagel H (1997) Social and biological predictors of early menopause: a model for premature aging. J Intern Med 242,299–305.

van Noord PA, Dubas JS, Dorland M, Boersma H and 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.

Osei-Hyiaman D, Satoshi T, Ueji M, Hideto T and Kano K (1998) Timing of menopause, reproductive years, and bone mineral density: a cross-sectional study of postmenopausal Japanese women. Am J Epidemiol 148,1055–1061.

Parazzini F, Negri E and La Vecchia C (1992) Reproductive and general lifestyle determinants of age at menopause. Maturitas 15,141–149.

Pavia M, Hsieh CC, Ekbom A, Adami HO and Trichopoulos D (1994) Handedness, age at menarche, and age at menopause. Obstet Gynecol 83,579–582.

Richards M, Kuh D, Hardy R and Wadsworth M (1999) Lifetime cognitive function and timing of the natural menopause. Neurology 53,308–314.

Risch N and Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 268,1584–1589.

Rizk DE, Bener A, Ezimokhai M, Hassan MY and Micallef R (1998) The age and symptomatology of natural menopause among United Arab Emirates women. Maturitas 29,197–202.

Rucker EB III, Dierisseau P, Wagner KU, Garrett L, Wynshaw-Boris A, Flaws JA and Hennighausen L (2000) Bcl-x and Bax regulate mouse primordial germ cell survival and apoptosis during embryogenesis. Mol Endocrinol 14,1038–1052.

Schlessinger D, Herrera L, Crisponi L, Mumm S, Percesepe A, Pellegrini M, Pilia G and Forabosco A (2002) Genes and translocations involved in POF. Am J Med Genet 111,328–333.

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

Sham PC (1998) Statistics in human genetics. Stat Methods Med Res 7,279–300.

Shelling AN, Burton KA, Chand AL, van Ee CC, France JT, Farquhar CM, Milsom SR, Love DR, Gersak K, Aittomaki K *et al*. (2000) Inhibin: a candidate gene for premature ovarian failure. Hum Reprod 15,2644–2649.

Sherman B, Wallace R, Bean J and Schlabaugh L (1981) Relationship of body weight to menarcheal and menopausal age: implications for breast cancer risk. J Clin Endocrinol Metab 52,488–493.

Snieder H, MacGregor AJ and 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.

Spira A (1988) The decline of fecundity with age. Maturitas **(**Suppl 1**)**, 15–22.

Stanford JL, Hartge P, Brinton LA, Hoover RN and Brookmeyer R (1987) Factors influencing the age at natural menopause. J Chronic Dis 40,995–1002.

Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S, Merajver S, Thorlacius S, Offit K, Stoppa-Lyonnet D *et al*. (1996) The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet 12,333–337.

The International HapMap Consortium (2003) The International HapMap Project. Nature 426,789–796.

Thomas ML, Xu X, Norfleet AM and Watson CS (1993) The presence of functional estrogen receptors in intestinal epithelial cells. Endocrinology 132,426–430.

Tibiletti MG, Testa G, Vegetti W, Alagna F, Taborelli M, Dalpra L, Bolis PF and Crosignani PG (1999) The idiopathic forms of premature menopause and early menopause show the same genetic pattern. Hum Reprod 14,2731–2734.

den Tonkelaar I (1997) Validity and reproducibility of self-reported age at menopause in women participating in the DOM-project. Maturitas 27,117–123.

den Tonkelaar I, te Velde ER, Looman CW (1998) Menstrual cycle length preceding menopause in relation to age at menopause. Maturitas 29,115–123.

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

Torgerson DJ, Thomas RE, Campbell MK and Reid DM (1997a) Alcohol consumption and age of maternal menopause are associated with menopause onset. Maturitas 26,21–25.

Torgerson DJ, Thomas RE and Reid DM (1997b) 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, Detoeuf M, Paniel B, Prieur M, Zorn JR *et al*. (1999) New natural inactivating mutations of the follicle-stimulating hormone receptor: correlations between receptor function and phenotype. Mol Endocrinol 13,1844–1854.

Treloar AE (1981) Menstrual cyclicity and the pre-menopause. Maturitas 3,249–264.

Treloar SA, Do KA and Martin NG (1998) Genetic influences on the age at menopause. Lancet 352,1084–1085.

Treloar A, Sadrzadeh S, Do KA, Martin G and Lambalk B (2000) Birth weight and age at menopause in Australian female twin pairs: exploration of the fetal origin hypothesis. Hum Reprod 15,55–59.

Van Asselt KM, Kok HS, Peeters PH, Roest M, Pearson PL, te Velde ER, Grobbee DE and van der Schouw YT (2003) Factor V Leiden mutation accelerates the onset of natural menopause. Menopause 10,477–481.

Van Asselt KM, Kok HS, Der Schouw YT, Grobbee DE, te Velde ER, Pearson PL and Peeters PH (2004a) Current smoking at menopause rather than duration determines the onset of natural menopause. Epidemiology 15,634–639.

Van Asselt KM, Kok HS, Putter H, Wijmenga C, Peeters PH, van der Schouw YT, Grobbee DE, te Velde ER, Mosselman S and Pearson PL (2004b) Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative trait loci influencing variation in human menopausal age. Am J Hum Genet 74,444–453.

Vegetti W, Marozzi A, Manfredini E, Testa G, Alagna F, Nicolosi A, Caliari I, Taborelli M, Tibiletti MG, Dalpra L *et al*. (2000) Premature ovarian failure. Mol Cell Endocrinol 161,53–57.

te Velde ER, Dorland M and Broekmans FJ (1998) Age at menopause as a marker of reproductive ageing. Maturitas 30,119–125.

de Vries ETI, van Noord PA, van der Schouw YT, te Velde ER and Peeters PH (2001) Oral contraceptive use in relation to age at menopause in the DOM cohort. Hum Reprod 16,1657–1662.

de Waard F, Collette HJ, Rombach JJ, Baanders-van Halewijn EA and Honing C (1984) The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chronic Dis 37,1–44.

Walker AR, Walker BF, Ncongwane J and Tshabalala EN (1984) Age of menopause in black women in South Africa. Br J Obstet Gynaecol 91,797–801.

Weel AE, Uitterlinden AG, Westendorp IC, Burger H, Schuit SC, Hofman A, Helmerhorst TJ, van Leeuwen JP and Pols HA (1999) Estrogen receptor polymorphism predicts the onset of natural and surgical menopause. J Clin Endocrinol Metab 84,3146–3150.

Welt CK, Smith PC and Taylor AE (2004) Evidence of early ovarian aging in fragile X premutation carriers. J Clin Endocrinol Metab, 89,4569–4574.

Westendorp RG and Kirkwood TB (1998) Human longevity at the cost of reproductive success. Nature 396,743–746.

Whelan EA, Sandler DP, McConnaughey DR and Weinberg CR (1990) Menstrual and reproductive characteristics and age at natural menopause. Am J Epidemiol 131,625–632.

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

Wilson EA, Yang F and Rees ED (1980) Estradiol and progesterone binding in uterine leiomyomata and in normal uterine tissues. Obstet Gynecol 55,20–24.

Wise PM, Krajnak KM and Kashon ML (1996) Menopause: the aging of multiple pacemakers. Science 273,67–70.

Wood JW (1989) Fecundity and natural fertility in humans. Oxf Rev Reprod Biol 11,61–109.

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C and Micklem G (1995) Identification of the breast cancer susceptibility gene BRCA2. Nature 378,789–792.

World Health Organisation Scientific Group (1996) Research on the menopause in the 1990s. 866.

Submitted on February 28, 2005; revised on April 27, 2005; accepted on May 20, 2005