The oestrogen receptor α gene is linked and/or associated with age of menarche in different ethnic groups


The first menstrual period, menarche, is one of the most significant milestones in a woman’s life. The age at menarche is an important anthropological variable which may influence the overall duration of tissue oestrogen exposure and then affect health in later life. Early menarche is a well established risk factor for the development of breast cancer1 and endometrial cancer.2 Women with a menarcheal age of 10 or 11 years have a 2.2 times greater risk of breast cancer than women who had their first menstrual period at the age of 12 years or more.3 Those with a menarcheal age of 17 years or more have a 45% lower risk of endometrial cancer.4 Early age at menarche is also associated with a risk for obesity5 and depression in later life.6 On the other hand, late menarche is associated with lower bone mass and an increased risk of osteoporotic fractures5 and Alzheimer’s disease.6 Therefore an understanding of potential factors responsible for the age of menarche is of considerable interest and may shed light on our understanding of these diseases.

Age at menarche is a complex trait that is determined by multiple environmental factors, including nutrition, exercise, socioeconomic conditions, psychosocial stimuli, childhood experience, general health,7–10 and genetic factors. Twin studies have shown that 53–74% of the variation in menarcheal age is attributed to genetic effects.10 Family history is a strong predictor of early menarche,10 and there are highly significant correlations between menarcheal ages of mothers and daughters.7 However, the specific genetic determinants of menarcheal age remain unclear. Oestrogen receptor α (ESR1) plays a distinct physiological role in mediating the specific effects of oestrogen;12 thus the ESR1 gene may serve as a candidate gene for age at menarche. A preliminary study in Greek adolescent girls13 suggested an association between the ESR1 gene and age at menarche. However, such an association was not observed in Japanese women14 or Dutch women.15,16

To determine the importance of the ESR1 gene for the onset of menstruation, we conducted a linkage study on a genomic region of ~80 cM centring on the ESR1 gene on chromosome 6 in 1140 white sister pairs.

METHODS

Subjects

In the linkage study, the sample was composed of 939 female offspring from 304 white families. There were two to eight female offspring in each family, giving a total of 1140 sister pairs. The available parents of the families were included for genotyping to improve the accuracy of the IBD (identity by descent) inference and thus maximise the power of linkage analysis. In the association study, 397 unrelated white females were included. From our previous epidemiological study,17 menarcheal age was available in 390 unrelated Chinese females. These were also used to test the relation between the ESR1 gene and age at menarche.

The studies were approved by the Creighton University institutional review board or the Research Administration Department of Hunan Normal University. All the study subjects signed informed consent documents before entering the project. The age of menarche for each female subject was recorded by nurse administered questionnaires.

Genotyping

DNA was extracted from whole blood employing a Puregene DNA isolation kit (Gentra Systems, Minneapolis, Minnesota, USA).

Abbreviations: ESR, oestrogen receptor; HWE, Hardy–Weinberg equilibrium; IBD, identity by descent; SNP, single nucleotide polymorphism
USA). In the white linkage study, 10 microsatellite markers—D6S287, D6S262, D6S292, D6S308, D6S441, D6S1581, D6S264, D6S1697, D6S446, and D6S281—were genotyped by using 3700 DNA analyser (Applied Biosystems, Foster City, California, USA). These markers were scattered on a region of ~80 cM centred on the ESR1 gene. The genotyping procedure for microsatellites has been described previously. PedCheck was employed to check the confirmation of Mendelian inheritance at all marker loci and to verify the siblings relations. In the white association study, seven SNPs were chosen based on their physical location within the ESR1 gene (fig 1A). Among these, three were located in the coding region (rs2077647 (10 Ser), rs1801132 (325 Pro), and rs2288480 (594 Thr)). Figure 1B shows their position in the functional domains of the ESR1 gene. Exons are depicted as filled boxes and introns as double lines. (B) Location of the three coding SNPs in the six functional domains [A to F] of the ESR1 gene.

**Statistical analyses**

**Linkage study**

A variance component linkage analysis for quantitative trait was carried out by employing the program SOLAR (http://www.sfbr.org/solar/). Two point and multipoint linkage analyses were conducted by using the maximum likelihood method, based on specifying the expected genetic covariance between sister pairs as a function of the IBD at a given marker locus. Pointwise empirical probability (p values) were estimated using the procedure “lodadj” implemented in SOLAR. This procedure samples the null distribution (the distribution of LOD (log of odds) scores obtained under the no linkage hypothesis), so that a sorted array of LOD scores is obtained; the proportion of LOD scores greater than the observed LOD score is the latter’s empirical p value. In all, 10 000 replicates were generated.

**Association study in whites**

The χ² test was employed to examine Hardy–Weinberg equilibrium (HWE) at the seven SNPs. The relation between these SNPs and the age at menarche was explored using analysis of variance (ANOVA). The frequencies of the minor allele homozygote for SNPs rs1514347, rs1801132, rs932477, rs3778082, and rs2228480 were low (4.0%, 4.1%, 0.3%, 1.3%, and 3.3%, respectively). Thus during the data analysis for these five SNPs, the subjects were divided into two groups, with and without the minor allele.

**Association study in Chinese**

The PvuII and XbaI polymorphisms are only 45 base pairs (bp) apart and they are in strong linkage disequilibrium; thus the haplotypes based on the two SNPs were tested as genetic markers. For the VDR gene, subjects were divided into two groups, with and without the minor allele, owing to the low frequency of minor allele homozygosity (6.2%). The potential interaction between the ESR1 and the VDR genes was also tested. All the above statistical analyses were performed using the program PedCheck, version 1.9.19

**Table 1** Information on the seven single nucleotide polymorphisms studied in the ESR1 gene in white women

<table>
<thead>
<tr>
<th>SNP</th>
<th>Polymorphism†</th>
<th>Location in gene</th>
<th>Allele frequency‡</th>
<th>Genotype frequency§</th>
<th>p Value¶</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2077647</td>
<td>T-C (10 Ser)</td>
<td>Exon 1</td>
<td>49.8</td>
<td>18.1</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>rs2234693</td>
<td>T-C</td>
<td>Intron 1</td>
<td>44.8</td>
<td>21.9</td>
<td>0.12</td>
</tr>
<tr>
<td>rs1514347</td>
<td>G-A</td>
<td>Intron 3</td>
<td>23.9</td>
<td>4.0</td>
<td>0.06</td>
</tr>
<tr>
<td>rs1801132</td>
<td>G-C (325 Pro)</td>
<td>Exon 4</td>
<td>20.3</td>
<td>4.1</td>
<td>0.95</td>
</tr>
<tr>
<td>rs932477</td>
<td>G-A</td>
<td>Intron 4</td>
<td>9.8</td>
<td>0.3</td>
<td>0.11</td>
</tr>
<tr>
<td>rs3778082</td>
<td>G-A</td>
<td>Intron 6</td>
<td>13.1</td>
<td>1.3</td>
<td>0.43</td>
</tr>
<tr>
<td>rs2228480</td>
<td>G-A (594 Thr)</td>
<td>Exon 8</td>
<td>19.5</td>
<td>3.3</td>
<td>0.50</td>
</tr>
</tbody>
</table>

*SNP ID in the database dbSNP, www.ncbi.nlm.nih.gov/SNP.†Bold faced letters are the minor alleles.‡The allele frequencies are for minor alleles.§Frequency of the homozygote at minor allele.¶Probability of the χ² test for Hardy–Weinberg equilibrium. SNP, single nucleotide polymorphism.
conducted employing the program SPSS V. 10 (SPSS Inc, Chicago, Illinois, USA).

RESULTS
Descriptive characteristics
In the linkage study, the sisters were on average 45.3 years old and their mean age at menarche was 13.1 years. The skewness and kurtosis values of age at menarche were fairly minor—0.45 and 0.30, respectively. Generally, the variance component analyses implemented in SOLAR are robust to deviation of normality. However, kurtosis or skewness values greater than 2 may inflate the type I error rate. The most conservative method (Bonferroni correction) was used to adjust for multiple testing using the support interval around the linkage peak was mapped to a 32 cM region. After adjusting for multiple testing using the empirical p value of 0.003, the linkage peak, about 161 cM from the p terminal of chromosome 6, was still the marker D6S1581. The closest to the location of the ESR1 gene, 6q25.1. The marker (empirical p value of 0.003) was obtained at 6q25.3, which is 0.001 for this linkage peak. A multipoint LOD score of 2.01 simulations revealed an empirical pointwise p value of 0.003. Ten thousand 155 cM from the tip of the short arm of chromosome 6, with Linkage signal was achieved at the marker D6S1581, about 13.6 years, which was significantly later than in the two white populations.

Association study in white women
The minor allele frequencies for the seven SNPs (rs2077647, PvuII, rs1514347, rs1801132, rs932477, rs3778082, and rs2228480) were 49.8%, 44.8%, 23.9%, 20.3%, 9.8%, 13.1%, and 19.5%, respectively (table 1). The χ² tests showed that except for rs2077647 all were in HWE (p > 0.05). The average distance among the SNP pairs ranged from 32 to 291 kb with an average of 133 kb. Pairwise linkage disequilibrium between the seven SNPs, calculated by D', ranged from 0.020 to 0.676, with a mean value of 0.280 (table 2). If D' = 0.7 was used as an arbitrary limit for useful linkage disequilibrium in association studies, none of the pairs would fall below this threshold. Thus the haplotype analysis was not pursued.

Association with the age at menarche was observed at two SNPs, rs3778082 and rs2228480 (table 3). For the marker rs3778082, menarche occurred about four months later in women with the minor allele A than in women without it, at (mean (SE)) 12.7 (0.1) years (p = 0.03). For the other five SNPs, we did not find any association (p > 0.05).

Association study in Chinese women
The minor allele frequencies for XbaI, PvuII, and ApaI in the Chinese sample were 22.4%, 44.8%, 23.9%, 20.3%, 9.8%, 13.1%, and 19.5%, respectively (table 1). The χ² tests showed that all were in HWE (p > 0.05). We did not find any direct association with age of menarche for the three loci in the Chinese sample; however, a potential interaction between the ESR1 and the VDR genes was observed (table 4). With genotype aa at the ApaI locus of the VDR gene, subjects carrying haplotype PX (reconstructed by XbaI and PvuII) at

---

### Table 2 Pairwise linkage disequilibrium (D') between the seven single nucleotide polymorphisms studied in the ESR1 gene

<table>
<thead>
<tr>
<th>SNPs</th>
<th>rs2077647</th>
<th>rs1514347</th>
<th>rs1801132</th>
<th>rs932477</th>
<th>rs3778082</th>
<th>rs2228480</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2077647</td>
<td>0.511</td>
<td>0.166</td>
<td>0.020</td>
<td>0.143</td>
<td>0.168</td>
<td>0.114</td>
</tr>
<tr>
<td>PvuII</td>
<td>0.345</td>
<td>0.159</td>
<td>0.270</td>
<td>0.088</td>
<td>0.131</td>
<td>0.103</td>
</tr>
<tr>
<td>rs1514347</td>
<td>0.100</td>
<td>0.661</td>
<td>0.634</td>
<td>0.576</td>
<td>0.313</td>
<td>0.116</td>
</tr>
<tr>
<td>rs1801132</td>
<td>0.654</td>
<td>0.367</td>
<td>0.370</td>
<td>0.252</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs932477</td>
<td>0.175</td>
<td>0.216</td>
<td>0.75</td>
<td>0.397</td>
<td>0.675</td>
<td>0.360</td>
</tr>
<tr>
<td>rs3778082</td>
<td>0.258</td>
<td>0.287</td>
<td>0.142</td>
<td>0.835</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>rs2228480</td>
<td>0.291</td>
<td>0.018</td>
<td>0.125</td>
<td>0.115</td>
<td>0.324</td>
<td></td>
</tr>
</tbody>
</table>

Upper right triangle, D'; lower left triangle, physical distance in the unit (base pairs).

---

### Table 3 Association analyses for the seven single nucleotide polymorphisms with the age of menarche in white women

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype 11</th>
<th>n</th>
<th>Genotype 12</th>
<th>n</th>
<th>Genotype 22</th>
<th>n</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2077647</td>
<td>12.9 (0.2)</td>
<td>72</td>
<td>12.9 (0.1)</td>
<td>251</td>
<td>12.9 (0.2)</td>
<td>74</td>
<td>0.45</td>
</tr>
<tr>
<td>PvuII</td>
<td>12.9 (0.2)</td>
<td>87</td>
<td>13.0 (0.1)</td>
<td>180</td>
<td>12.9 (0.1)</td>
<td>128</td>
<td>0.58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotypes 11 and 12</th>
<th>n</th>
<th>Genotype 22</th>
<th>n</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1514347 12.9 (0.1)</td>
<td>174</td>
<td>12.9 (0.1)</td>
<td>223</td>
<td>0.94</td>
</tr>
<tr>
<td>rs1801132 12.9 (0.1)</td>
<td>144</td>
<td>13.0 (0.1)</td>
<td>251</td>
<td>0.87</td>
</tr>
<tr>
<td>rs932477   13.1 (0.2)</td>
<td>77</td>
<td>12.9 (0.1)</td>
<td>320</td>
<td>0.36</td>
</tr>
<tr>
<td>rs3778082  13.2 (0.2)</td>
<td>99</td>
<td>12.9 (0.1)</td>
<td>298</td>
<td>0.03</td>
</tr>
<tr>
<td>rs2228480  12.7 (0.1)</td>
<td>142</td>
<td>13.1 (0.1)</td>
<td>255</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Values are mean (SE).

**11**, homozygote of the minor allele of each SNP; **12**, heterozygote; **22**, homozygote of the common allele of each SNP.

SNP, single nucleotide polymorphism.
the ESR1 gene had menarche about six months later (n = 52; 14.0 (0.2) years, mean (SE)) than in the non-carriers (n = 148; 13.5 (0.1) years, p = 0.03). This suggests a potential effect of the ESR1 gene on the onset of menstruation. However, after correction for multiple testing, this interaction effect between the two genes did not reach statistical significance.

**DISCUSSION**

To test the importance of the ESR1 gene on the age of menarche, we first undertook a linkage study on a genomic region of ~80 cM centred on the ESR1 gene. Linkage signal was achieved with a two point LOD score of 2.90 and a multipoint LOD score of 2.01. Subsequently, association was observed at two SNPs in the ESR1 gene, rs3778082 and rs2228480, in an independent unrelated white population. It is not unexpected to find the effects of the ESR1 gene on variation in menarcheal age, considering its functional importance. Oestrogen signalling is largely mediated in the oestrogen receptors, including oestrogen receptor α. The onset of menstruation is initiated by an increased amplitude of oestrogen exposure of tissues. One can therefore assume that the ESR1 gene has some effect on the onset of menstruation.

ESR1 has been categorised into six functional domains, A to F (Fig 1B), and the rs2228480 lies in the F domain. The function of the F domain is not clearly understood. It is thought to play a role in distinguishing between receptor agonist and antagonist binding to the receptor molecule, which affects the metabolism of oestrogen. SNP rs2228480 is a silent polymorphism; however, it may be correlated with oestrogenic biological activity, because gene expression could be affected through a structural change in mRNA and this may lead to an alteration in processing or efficiency of translation. SNP rs2228480 has been associated with several major diseases where lifetime oestrogen exposure is a potentially important risk modifier. For example, Roodi et al. found that the minor allele A was associated with an increased risk of breast cancer because of early exposure to oestrogen. Our observed association of the allele A with early menarche is in agreement with their results. SNP rs3778082, located in intron 6, was first reported in a Japanese population and it has never been tested for any disease trait. It may act as a regulatory element in mediating transcription or stability of mRNA, and also as a protein binding site to affect the function of this gene. The possible association of the two SNPs may come from the linkage disequilibrium with a truly causative sequence variation elsewhere in the ESR1 gene or even in another nearby gene. For the other five SNPs studied, no association was observed, partly because they may not be in linkage disequilibrium with the functional polymorphisms.

In our Chinese population, a potential interaction between the ESR1 and the VDR genes was observed, which may indirectly suggest the potential effect of the ESR1 gene. The mechanism for the interaction is unclear; however, from a physiological point of view, there are some plausible explanations. An oestrogen responsive promoter region has been characterised in the VDR gene, and the transcription of the VDR promoter is dependent on oestrogen receptor. On the other hand, vitamin D is an important factor in oestrogen biosynthesis, and may influence the balance between oestrogens and androgens. Vitamin D may act along the oestrogen response pathway, affecting the levels of oestrogen receptors as well as their function.

Several issues arise in this study. First, the age at menarche was collected retrospectively from adults, which may increase the likelihood of error. It is a limitation that there is no objective criterion to define the phenotype for these women, such as Tanner’s developmental stages. However, menarche is one of the most important milestones in a female’s life. Retrospective recall is reasonably accurate. A recent study showed a correlation of 0.79 between the original age of menarche and the information recalled 30 years later. Second, menarcheal age is affected by living environmental factors. Such information was not recorded and not used as covariates to adjust the raw data. However, in the linkage study, we limited our subjects to sister pairs, who were likely to live in more similar environments than the other relative pairs, such as grandparent–grandchild and aunt–niece. In the association study, the subjects’ age limits were similar and they may have had a similar living environment. Such subject ascertainment could improve the accuracy of the linkage and association studies. Third, most people reported menarcheal age in years and months; however, some reported it as whole years, which will affect the accuracy of the result. Fourth, the data in the Chinese study came from another project. The marker set was not the same as in the white subjects. The SNPs BsmI and TaqI were not genotyped in the white sample, and the two SNPs with significant results in white subjects—rs3778082 and rs2228480—were not genotyped in the Chinese. It would be more interesting to genotype the same markers in the two ethnic groups. In addition, in the VDR gene, only one marker (ApaI) was studied. It would be interesting to genotype more markers in the VDR gene, such as BsmI and TaqI, in a larger sample to confirm the potential effects of the ESR1 and VDR genes on age at menarche.

---

**Table 4** Potential interaction between the ESR1 and VDR genes on age at menarche in Chinese women

<table>
<thead>
<tr>
<th>VDR genotype</th>
<th>With allele A</th>
<th>n</th>
<th>p value</th>
<th>Without allele A</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESR1 px</td>
<td>0</td>
<td>13.6 (0.1)</td>
<td>127</td>
<td>0.19</td>
<td>13.5 (0.1)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13.3 (0.2)</td>
<td>63</td>
<td>1</td>
<td>14.0 (0.2)</td>
</tr>
<tr>
<td>ESR1 px</td>
<td>0</td>
<td>13.1 (0.2)</td>
<td>33</td>
<td>0.10</td>
<td>13.8 (0.2)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13.6 (0.1)</td>
<td>157</td>
<td></td>
<td>13.6 (0.1)</td>
</tr>
<tr>
<td>ESR1 Px</td>
<td>0</td>
<td>13.5 (0.1)</td>
<td>118</td>
<td>1.00</td>
<td>13.6 (0.1)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13.5 (0.2)</td>
<td>72</td>
<td></td>
<td>13.8 (0.2)</td>
</tr>
<tr>
<td>ESR1 pX</td>
<td>0</td>
<td>13.5 (0.1)</td>
<td>167</td>
<td>0.56</td>
<td>13.6 (0.1)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13.7 (0.3)</td>
<td>23</td>
<td></td>
<td>13.6 (0.3)</td>
</tr>
</tbody>
</table>

Values are mean (SE).

---

1. carriers; 0, non-carriers of the corresponding ESR1 haplotype.

---

The letter to JMG 799.
interaction effect between the ESR1 gene and the VDR gene on the age at menarche.

To our knowledge, this is the first study to show the effects of the ESR1 gene on the age at menarche based on both linkage and association studies in independent populations. For further exploration of the relation between the ESR1 gene and the age of menarche, studies with a larger sample size and denser markers are required.

ACKNOWLEDGEMENTS

The investigators were partially supported by grants from Health Future Foundation, NIH, the State of Nebraska (LBS95 and LB692), US Department of Energy, Chinese National Science Foundation, the Ministry of Education of PR China, Huo Ying Dong Education Foundation, and Hunan Normal University. The study also benefited from 211 State Key Research Fund to Xi’an Jiaotong University.

Authors’ affiliations

J-R Long, L-J Zhao, P-Y Liu, H Shen, Y-J Liu, D-H Xiong, P Xiao, Y-Z Liu, V Dvornyk, R R Recker, H-W Deng, Osteoporosis Research Centre and Department of Biomedical Sciences, Creighton University, Omaha, Nebraska, USA.

H-W Deng, The Key Laboratory of Biomedical Information Engineering of Ministry of Education and Institute of Molecular Genetics, Xi’an Jiaotong University, Xi’an, China.

J-L Li, Center for Medical Informatics, School of Medicine, Yale University, New Haven, Connecticut, USA.

H Xu, H-W Deng, Laboratory of Molecular and Statistical Genetics, College of Life Sciences, Hunan Normal University, Changsha, Hunan, China.

Competing interests: none declared

Correspondence to: Dr Hong-Wen Deng, Osteoporosis Research Center, Creighton University, 601 N 30th St, Suite 6787, Omaha, Nebraska 68131, USA; deng@creighton.edu

Received 18 October 2004

Revised version received 27 January 2005

Accepted for publication 2 February 2005

REFERENCES


