Human Reproduction Update, Vol.18, No.2 pp. 198–210, 2012
Advanced Access publication on January 18, 2012 doi:10.1093/humupd/dmr050

human reproduction update

Genetics of age at menarche: a systematic review

Volodymyr Dvornyk* and Waqar-ul-Haq

School of Biological Sciences, University of Hong Kong, Pokfulam Road, Pokfulam, Hong Kong SAR, P.R. China

*Correspondence address. Tel: +852-2299-0611; Fax: +852-2559-9114; E-mail: dvornyk@hku.hk

Submitted on June 26, 2011; resubmitted on October 19, 2011; accepted on November 9, 2011

TABLE OF CONTENTS

BACKGROUND: Menarche is the first menstrual period of a girl at puberty. The timing of menarche is important for health in later life. Age at menarche is a complex trait and has a strong genetic component. This review summarizes the results of the genetic studies of age at menarche conducted to date, highlights existing problems in this area and outlines prospects of future studies on genetic factors for the trait.

METHODS: PubMed and Google Scholar were searched until May 2011 using the keywords: 'menarche', 'puberty' and 'age at menarche' in combination with the keywords 'polymorphism', 'candidate gene', 'genome-wide association study' and 'linkage'.

RESULTS: Our search yielded 170 papers, 35 of which were selected for further analysis. Several large-scale genome-wide association studies along with a powerful meta-analysis of their aggregated data identified about 50 candidate genes for the trait. Some genes were replicated in different studies of Caucasians (e.g. *LIN28B*, *TMEM38B*) or in different ethnicities (e.g. *SPOCK*, *RANK* and *RANKL*). However, despite the large volume of results obtained, there is a huge gap in relevant data on ethnic groups other than Caucasians.

CONCLUSIONS: The reviewed studies laid a solid basis for future research on genetics of age at menarche. However, as yet specific genes for this trait have not been identified consistently in all ethnicities and types of studies. We suggest expanding the research to different ethnicities and propose several methodologies to increase the efficiency of studies in this area, including a systems approach, which combines existing high-throughput methods in a single pipeline.

Key words: age at menarche / genetic variants / association study / linkage analysis / complex trait

Introduction

Menarche is the first menstrual bleeding that marks a beginning of the female's reproductive life. It is considered as one of the most important events in female puberty. There is a secular trend in the mean age at menarche with a steady decline over the last several decades (Hwang *et al.*, 2003; Herman-Giddens, 2007). Age at menarche has important implications in female fertility (McKibben and Poston, 2003; Pascual *et al.*, 2005) and, furthermore, it may serve as an indicator of possible health complications in later life. Earlier menarche is associated with an increased risk of some diseases, such as breast cancer (Schatzkin *et al.*, 1987; Peeters *et al.*, 1995), gynecological cancers (Marshall *et al.*, 1998; Fujita *et al.*, 2008) and various cardiovascular diseases (Cooper *et al.*, 1999; Lakshman *et al.*, 2009). Girls with early menarche exhibited elevated blood pressure and glucose intolerance compared with girls who matured later (Remsberg *et al.*, 2005). Overall, earlier menarche (before 12 years) results in higher mortality (Lakshman *et al.*, 2009). On the other hand, there is a positive correlation between early menarche and high bone mineral density (Ito *et al.*, 1995). Recent data suggest that age at menarche is significantly associated with body composition, insulin sensitivity and blood lipid levels (Feng *et al.*, 2008).

Like many other complex traits, age at menarche is determined by both genetic and environmental factors and their interactions. The observed secular trend in age at menarche (Hwang *et al.*, 2003; Herman-Giddens, 2007) may suggest that environmental rather than

genetic factors are the major contributors to the observed phenotypic variance of the trait. However, this suggestion is inconclusive. Twin and familial studies of age at menarche indicated that 57–82% of the variance in timing of puberty can be explained by heritable factors (Kaprio *et al.*, 1995; Anderson *et al.*, 2007; Morris *et al.*, 2011). Despite the fairly large range of the estimates, these data support the hypothesis that genetic factors play an important role in determining age at menarche. Among a wide variety of environmental factors that may contribute to age at menarche, weight, height, stressful life events, family relations, absence or presence of an adult male in the household, psychological adjustment (Graber *et al.*, 1995), nutrition, living standards (Graham *et al.*, 1999), physical activity (Chavarro *et al.*, 2004) and ethnicity (Herman-Giddens *et al.*, 1997; Anderson *et al.*, 2003; Chumlea *et al.*, 2003) were reported.

Identification of the environmental factors and genes that contribute to puberty and the timing of menarche may add to our understanding of the physiological mechanism of this trait and the associated fertility and health risks. In this paper, we review the current status of genetic studies conducted on age at menarche in humans. After providing an overview of the results of the genetic studies performed thus far, we aim to give recommendations for future studies to identify genetic factors involved in determining the variation in the timing of age at menarche.

Methods

The databases of MEDLINE (PubMed) and Google Scholar were searched electronically until May 2011 using the keywords: 'menarche', 'age at menarche' and 'puberty' in combination with the keywords 'polymorphism', 'candidate gene', 'genome-wide association study' and 'linkage'. All retrieved articles were screened for relevance to the topic of this manuscript following the HuGE guidelines (Bray *et al.*, 2006; Ioannidis *et al.*, 2008). The reference lists of the retrieved articles were also screened for appropriate sources. Some of the retrieved articles, although formally matching some of the search keywords, did not contain data, that would be appropriate for the goals of our study. Therefore, such articles were excluded from the further analysis. The final set of the literature sources included all articles on genome-wide linkage mapping and genome-wide or candidate gene association studies of age at menarche (Fig. 1).

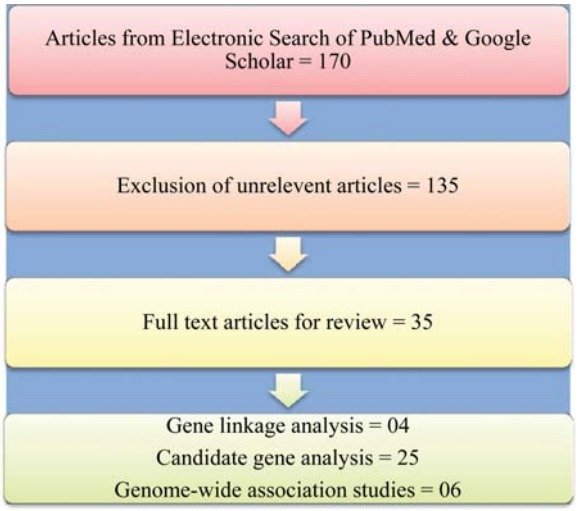

Figure 1 Flowchart of selection of genetic studies for age at menarche.

Results

Our literature search yielded in total 35 articles on the genetics of age at menarche, including 4 articles on genome-wide linkage analysis, 6 papers on genome-wide association studies and 25 papers on association studies of various candidate genes.

Candidate gene association studies

Candidate genes related to estrogen metabolism

All genes mentioned in the article are described in Supplementary data, Table S1. Since menarche is associated with significant changes in hormone levels, particularly estrogen (Sizonenko, 1989; Bandini *et al.*, 2008), genes involved in hormone metabolism were among the first candidates for analysis of their possible association with age at menarche. The first report about the association of estrogen-metabolizing genes with age at menarche was published in 1997, when the A1/A1 genotype of the common MspI polymorphism of the *CYP17* gene was suggested to confer a later age at menarche (Feigelson *et al.*, 1997). The subsequent reports about a possible effect of this polymorphism on age at menarche were contradictory (Table I). Finally, a meta-analysis of 11 case–control studies and 5 studies with continuous outcome and aggregate sample size of >11 000 individuals did not reveal any effect of this polymorphism on age at menarche ($P > 0.05$; Pei *et al.*, 2008).

Estrogen receptor α (*ESR1*) is a pleiotropic gene commonly studied for its association with various phenotypes. Two frequently analyzed polymorphisms of this gene are XbaI (rs9340799) and PvuII (rs2234693), which are in strong linkage disequilibrium. Except for a study by Stavrou *et al.* (2002), all the others conducted so far reported no association of these single nucleotide polymorphisms (SNPs) with age at menarche (Table I). However, one study (Manuck *et al.*, 2011) suggested that the XbaI genotypes may affect age at menarche through interaction with the quality of family environment ($P = 0.007$). Likewise, Stavrou *et al.* (2006) reported a combined effect of *ESR1* and *ESR2* on AAM. They also found that girls with the AA genotype of the *ESR1* 1730A→G polymorphism had menarche a 7 months later than girls with the AG genotype ($P = 0.005$).

The data on the possible association of other estrogen-metabolizing genes with age at menarche are scarce. Lai *et al.* (2001) analyzed the *CYP3A4*, *CYP17*, *CYP1B1* and *CYP1A2* genes in a multiethnic sample but found no association with age at menarche (Table I). Three studies reported association of several SNPs and variable number of tandem repeats of the *CYP19* gene with the onset of menarche (Guo *et al.*, 2006b; Mitchell *et al.*, 2008; Xita *et al.*, 2010). However, two of these studies (Mitchell *et al.*, 2008; Xita *et al.*, 2010) utilized very small samples, so that their results should be treated cautiously.

Finally, one study determined a significant association of the functional polymorphism (rs1042838, Val660Leu) of the progesterone receptor gene (*PGR*) with age at menarche (Taylor *et al.*, 2010).

Table I Association of candidate genes involved in steroid metabolism with age at human menarche.

Candidate gene	Polymorphism	*n*[a]	Ethnicity	*P*	References
ESR1	rs9340799 (XbaI)	145	Greek	0.017	Stavrou *et al.* (2002)
		152	Caucasian	0.54	Mitchell *et al.* (2008)
		273	Multiethnic	>0.05	Silva *et al.* (2010)
		317	Japanese	0.456	Gorai *et al.* (2003)
		455	Caucasian	0.17	Manuck *et al.* (2011)
	rs2234693 (PvuII)	145	Greek	0.21	Stavrou *et al.* (2002)
		152	Caucasian	0.14	Mitchell *et al.* (2008)
		273	Multiethnic	>0.05	Silva *et al.* (2010)
		317	Japanese	0.622	Gorai *et al.* (2003)
		455	Caucasian	0.11	Manuck *et al.* (2011)
	rs2228480	397	Caucasian	0.03[b]	Long *et al.* (2005)
	rs3778082			0.03[b]	Long *et al.* (2005)
ESR2	1082G→A	145	Greek	N.S.	
	1730A→G			0.005	Stavrou *et al.* (2006)
CYP1A1	MspI	317	Japanese	0.473	Gorai *et al.* (2003)
COMT	Hsp92II	317	Japanese	0.828	Gorai *et al.* (2003)
CYP17	MspAI	459	Multiethnic	0.047	Feigelson *et al.* (1997)
		351	Caucasian	0.8	Lai *et al.* (2001)
		82	Black	0.3	
		76	Asian	0.7	
		28	Indian–Pakistan	0.2	
		46	Others	0.2	
		583	All	0.8	
		317	Japanese	0.028	Gorai *et al.* (2003)
		124	Caucasian	>0.058	Kulik-Rechberger *et al.* (2007)
		152	Caucasian	0.26	Mitchell *et al.* (2008)
	rs10883783 rs619824	1048	Caucasian	>0.05	Guo *et al.* (2006b)
CYP19	rs2470144	1048	Caucasian	<0.05	Guo *et al.* (2006b)
	rs2445761				
	rs2470176				
	rs8029807				
	rs2255192				
	3′-UTR	152	Caucasian	0.95	Mitchell *et al.* (2008)
	$(TTTA)_7$	152	Caucasian	0.04	Mitchell *et al.* (2008)
		130	Greek	<0.025	Xita *et al.* (2010)
	$(TTTA)_{11}$	152	Caucasian	0.42	Mitchell *et al.* (2008)
		130	Greek	<0.001	Xita *et al.* (2010)
CYP1A2	CYP1A2*1F	351	Caucasian	0.7	Lai *et al.* (2001)
		82	Black	0.5	
		76	Asian	0.3	
		28	Indian–Pakistan	0.6	
		46	Others	0.4	
		583	All	0.8	
CYP3A4	CYP3A4*1B	351	Caucasian	N.S.	Lai *et al.* (2001)
		82	Black	0.1	
		76	Asian	N.S.	
		28	Indian–Pakistani	N.S.	
		46	Others	1.0	
		583	All	0.3	

Continued

Table I *Continued*

Candidate gene	Polymorphism	n[a]	Ethnicity	P	References
CYP1B1	CYP1B1*3 (Leu432Val)	351	Caucasian	0.8	Lai *et al.* (2001)
		82	Black	0.9	
		76	Asian	N.S.	
		28	Indian–Pakistani	N.S.	
		46	Others	0.3	
		583	All	0.9	
	CYP1B1*2 (Ala119Ser)	152	Caucasian	0.03	Mitchell *et al.* (2008)
HSDB1	rs2830	152		0.56	Mitchell *et al.* (2008)
	rs615942			0.52	Mitchell *et al.* (2008)
	rs592389			0.42	Mitchell *et al.* (2008)
SHBG	(TAAAA)n	130	Greek	<0.05	Xita *et al.* (2005)
PGR	rs1042838 (Val660Leu)	444	Caucasian	0.03	Taylor *et al.* (2010)

UTR, untranslated; N.S., Not specified.
[a]Number of subjects genotyped.
[b]Non-significant after correction for multiple testing.

Miscellaneous genes

In addition to the estrogen-metabolizing genes, several genes involved in other metabolic processes were examined for their possible association with age at menarche (Table II).

Chemokine (C-C-motif) receptor 3 (*CCR3*), which plays an important role in metabolic pathways related to endometrial function (Zhang *et al.*, 2000), manifested significant association and linkage to age at menarche in a sample of 1048 Caucasian females from 354 nuclear families (Yang *et al.*, 2007). This finding was later supported at the genome-wide level by the genome-wide association study, which also identified Sparc/Osteonectin, CWCV and Kazal-like domains proteoglycan (*SPOCK*) as a novel candidate gene for age at menarche (Liu *et al.*, 2009).

Obesity was previously reported as a factor influencing the onset of puberty (Matkovic *et al.*, 1997; Lin-Su *et al.*, 2002), and serum leptin levels inversely correlated with age at menarche (Matkovic *et al.*, 1997). However, analyses of two leptin gene polymorphisms did not detect the association of the gene with the timing of menarche (Comings *et al.*, 2001; Rothenbuhler *et al.*, 2009; Table II). On the contrary, a replacement polymorphism (Q223R) of the leptin receptor gene (*LEPR*) was significantly associated with age at menarche (Riestra *et al.*, 2011).

Likewise, possible contribution to the age at menarche was recently suggested for two important pleiotropic genes, *TNFRSF11A* and *TNFSF11* (also known as *RANK* and *RANKL*, respectively; Lu *et al.*, 2010). They have been acknowledged as key players in a wide variety of processes regulating cell death and proliferation, immunity, morphogenesis of the lymphoid tissue (Anderson *et al.*, 1997; So *et al.*, 2006) and, specifically, development of a lactating mammary gland during pregnancy (Theill *et al.*, 2002).

Some SNPs may not be associated with age at menarche directly but may influence it through SNP/SNP or/and SNP/environment interactions, as was recently reported for methylenetetrahydrofolate reductase (*MTHFR*), an important gene for homocysteine metabolism (Liu *et al.*, 2010b).

Genome-wide linkage analysis

Four genome-wide linkage analyses on age at menarche have been conducted to date (Table III). Two studies (Guo *et al.*, 2006a; Pan *et al.*, 2008) were performed by the same research group on largely the same sample of about 2500 Caucasian females from >400 pedigrees, but the design of the studies was different. In Guo *et al.* (2006a), a univariate linkage analysis was performed. In Pan *et al.* (2008), a bivariate linkage analysis was conducted to identify loci that contribute to both age at menarche and bone mineral density. Both studies reported logarithms of the odd (LOD) > 3 for the same region, 22q13 (Table III). This region was also identified by Rothenbuhler *et al.* (2006) although with lower LOD = 1.09–1.63. The region spans about 1 Mb and harbors nearly 50 genes, some of which may be candidates genes for age at menarche. The authors suggested two genes, E1A binding protein p300 (*EP300*) and sterol regulatory element-binding transcription factor-2 (*SREBF2*), as probable contributors to age at menarche (Guo *et al.*, 2006a; Pan *et al.*, 2008). *EP300* was shown to regulate interaction between steroidogenic factor-1 and early growth response-1 in their activating LH-β subunit gene, which is essential for differentiation of reproductive organs (Mouillet *et al.*, 2004). *SREBF2* plays an important role in the transcription of sterol-regulated genes and cholesterol homeostasis (Ettinger *et al.*, 2004). Interestingly, the microRNA miR33 located within intron 16 of the *SREBF2* gene also participates in the control of cholesterol homeostasis (Najafi-Shoushtari *et al.*, 2010). This fact is in further support of the observed linkage of the 22q13 region with age at menarche.

A few more candidate genes for age at menarche were suggested in the other genomic regions identified by the studies of Guo *et al.* (2006a) and Pan *et al.* (2008) (Table III). Catechol-*O*-methyltransferase (*COMT*) located on 22q11 is important for estrogen metabolism; *PGR* located on 11q23 is expressed in endometrium (Attia *et al.*, 2000) and contributes to mammary gland morphogenesis (Brisken *et al.*, 2000). A candidate gene at 3p25 is peroxisome proliferator-activated receptor γ (*PPARG*), which may influence the onset of menarche through modulation of estrogenic actions (Suzuki *et al.*, 2006). In addition, Guo *et al.*

Table II Data about association of miscellaneous genes with age at human menarche.

Candidate gene	Polymorphism	n[a]	Ethnicity	P	References
CCR3	rs6441948	1048	Caucasian	0.009[b]	Yang *et al.* (2007)
	rs3091309			0.006[b]	
BRCA1	deleterious mutation	2662		0.53	Kotsopoulos *et al.* (2005)
BRCA2	deleterious mutation	1285		0.30	
FGFR2	rs2420946	1368	Japanese	0.019	Kawase *et al.* (2009)
IGF1	rs6214	1048	Caucasian	0.0153	Zhao *et al.* (2007)
LEPR	Q223R	338	Spanish	<0.05	Riestra *et al.* (2011)
LEP	D7S1875	183	Caucasian	0.28	Comings *et al.* (2001)
	−2459	247	Caucasian	0.9	Rothenbuhler *et al.* (2009)
NPY1R	rs7687423			0.8	
GPR54	rs350132			0.6	
MTHFR	5 SNPs[c]	306		<0.05	Liu *et al.* (2010b)
SPOCK	rs13357391	477	Caucasian	<0.05	Liu *et al.* (2009)
	rs1859345			<0.05	
TNFSF11 (RANKL)					
	rs2200287	306	Caucasian	0.005	Lu *et al.* (2010)
	rs9525641			0.039	
	rs1054016			0.047	
	3 haplotypes	825	Chinese	<0.05	Pan *et al.* (2011)
TNFRSF11A (RANK)	rs3826620			0.022	
		825	Chinese	0.018	Pan *et al.* (2011)
	rs9956850			0.046	
	rs7239261			0.006	
	rs8094884			0.009	
	rs8089829			0.034	
	7 haplotypes			<0.05	
VDR	BsmI	1058	Caucasian	0.9	Grimm *et al.* (2005)

[a]Number of subjects genotyped.
[b]Significant after correction for multiple testing.
[c]No direct association but three SNP/SNP interactions (rs2066470/rs1476413, rs2066470/rs4846049 and rs17037390/rs4846049).

(2006a) reported significant epistatic interaction between 22q13 and 3q13.

Three linkage signals with LOD > 2 on chromosomes 8 and 16 (Table III) were reported by Rothenbuhler *et al.* (2006) who incorporated adjustment for menarcheal weight in their analysis. However, the sample size was much smaller, only 98 sister pairs.

In the largest genome-wide linkage study to date, three cohorts of Caucasian females from Australia, the UK and The Netherlands were combined in one large sample (Anderson *et al.*, 2008). Despite the large sample size and the large number of sibling pairs, the analysis of the combined sample did not identify genomic regions with significant (LOD > 3.0) linkage; only a peak with suggestive linkage (LOD = 2.0) was determined on chromosome 12.

Genome-wide association studies

Despite the wide use of genome-wide association mapping in modern epidemiology, this approach was introduced to studies on genetics of age at menarche only in 2009, when results of four genome-wide association studies were published (He *et al.*, 2009; Liu *et al.*, 2009; Ong *et al.*, 2009; Sulem *et al.*, 2009). Three of these studies (He *et al.*, 2009; Ong *et al.*, 2009; Sulem *et al.*, 2009) reported a strong signal in the 6q21 region, where the *LIN28B* (lin-28 homolog B) gene is located (Table IV). The LIN28B protein controls expression of mature miRNAs in the LET7 family (Viswanathan *et al.*, 2009). A major allele of this gene confers a 0.12 years earlier menarche (Ong *et al.*, 2009). In addition, *LIN28B* was associated with several other pubertal characteristics and height (Ong *et al.*, 2009; Widen *et al.*, 2010).

One of the recently identified important candidate genes for age at menarche is *SPOCK* (Liu *et al.*, 2009), which controls expression of *MMP-2* (matrix metalloproteinase-2), a key gene involved in initiation of menstrual bleeding (Irwin *et al.*, 1996; Nakada *et al.*, 2001). This study utilized a two-step design with genome-wide association studies at the first stage followed by candidate gene association and linkage mapping at the second stage.

One more candidate gene, transmembrane protein 38B (*TMEM38B*), was suggested based on the strong association signals for several SNPs located near this gene (He *et al.*, 2009; Table IV).

Table III Results of genome-wide linkage analyses of age at human menarche.

Linkage region	LOD Score	Candidate genes	*n*[a]	Reference
22q13	3.70	*EP300*, *SREBF2*	2461	Guo *et al.* (2006a)
22q11	2.68	*COMT*		
11q23	1.98	*PGR*		
16q21	3.33	Not specified	98[b]	Rothenbuhler *et al.* (2006)
16q12	3.12	Not specified		
22q13	1.09–1.63	Not specified		
8p12	2.18	*GNRH1*, *LEPROTL1*		
22q13	3.33	*EP300*, *SREBF2*	2522	Pan *et al.* (2008)
3p25	3.36	*PPARG*		
3q13	2.31	*CASR*		
7p15	2.44	*AHR*		
15q13	2.97	*ATD*		
12q	2.0	*IGF1*(rs6124)	13 697	Anderson *et al.* (2008)

LOD, logarithms of the odd.
[a]Number of subjects genotyped.
[b]Number of sister pairs.

There is not much information about the functions of this gene. However, experiments in animal models showed that this protein has two subtypes with complementary physiological functions related to intracellular Ca^{2+} transport (Yazawa *et al.*, 2007).

Two other genome-wide association studies (Perry *et al.*, 2009; Elks *et al.*, 2010) confirmed the previously determined genes for age at menarche (*LIN28B* and *TMEM38B*) and identified 30 new loci with high statistical power ($P < 5 \times 10^{-8}$). In addition, 10 loci with suggestive association ($P < 1.9 \times 10^{-6}$) were determined. Both studies utilized a meta-analysis approach that made it possible to significantly increase the power of analysis. To date, the largest genome-wide association study of age at menarche by Elks *et al.* (2010) employed a two-stage design. At the first stage, a meta-analysis of 32 genome-wide association studies in 87 802 Caucasian women of European ancestry was performed, and then a follow-up replication association mapping in up to 14 731 women was conducted.

The study by Elks *et al.* (2010) yielded other important results. It was previously suggested that age at menarche may have a shared genetic basis or correlate with some other complex traits, for example, obesity (Wang *et al.*, 2006), bone mineral density (Guo *et al.*, 2005) and height (Onland-Moret *et al.*, 2005). Indeed, the Elks *et al.* (2010) study identified miscellaneous obesity candidate genes (e.g. *FTO*, *SEC16B*, *TRA2B* and *TMEM18*), bone metabolism genes (e.g. *ESR1*, *BMP2* and *BMP6*), height-related genes (e.g. *LIN28B*, *HMGA2* and *PPARD*) as associated with age at menarche. It also provided evidence that genes involved in energy homeostasis (*BSX*, *CRTC1* and *MCHR2*) and hormonal regulation (*INHBA*, *PCSK2* and *RXRG*) are also associated with age at menarche. Furthermore, using the ingenuity and gene-set enrichment pathway analyses, they suggested that co-enzyme A and fatty acid biosynthesis may be related to the onset of menarche.

One more SNP, rs13281615, significantly associated with age at menarche, was reported by a genome-wide association study of the breast cancer susceptibility loci in a cohort of 1002 Chinese females (Jiang *et al.*, 2011). This SNP is located on 8q24.21 with the closest annotated gene being POU class 5 homeobox 1B. However, this association might be a false-positive because the genome-wide significance was not high ($P = 0.023$).

Discussion

Strengths, weaknesses, opportunities and threats

Although a strengths, weaknesses, opportunities and threats (SWOT) analysis has been commonly used mostly in business, examples of its successful application in biology and medicine are available (Garnick, 2007; Miller, 2007; Boytsov and van de Werf, 2011). With respect to genetics of menarcheal age, the SWOT approach could help to evaluate current knowledge in this area and outline future efforts (Table V). The strengths of this research area are largely outlined above, in Results. The main points of three other SWOT components are discussed below.

Problems in research on genetics of age at menarche

Owing to their importance for female health in later life, the research on the timing of onset of menarche and menopause has gained increased attention during the last decade. The number of published reports on genetics of these traits increases every year. A comprehensive review of the genetic studies on age at natural menopause was published recently (Voorhuis *et al.*, 2010). One of the main problems indicated in that review is inconsistency in results of different studies and failure to replicate previous findings. Only in a very few studies has association of the same polymorphisms or genomic regions with menopausal age been reported for different populations. Also, there is almost no overlap between results from the different types of study (i.e. candidate gene association studies, genome-wide linkage and genome-wide association analyses). The same issues may be applied to the studies on age at menarche, albeit to the lesser extent.

The inconsistency in the results of association studies may be caused by several factors. One of the most common of those is the different ethnic background of the studied populations (cohorts). It is an acknowledged fact that significant inter-ethnic differences exist in allele frequencies of many candidate genes for various complex traits (Beavan *et al.*, 1998; Dvornyk *et al.*, 2003; Lei *et al.*, 2003). These differences may underlie the well-known inter-ethnic dissimilarities in prevalence or characteristics of the traits, e.g. mean age at menarche (Herman-Giddens *et al.*, 1997; Anderson *et al.*, 2003; Chumlea *et al.*, 2003). Even individuals declaring themselves as belonging to the same ethnic group may have a different ethnic background. For example, although Finns and Italians are both considered as Caucasians, they have fairly different population genetic structures (Nelis *et al.*, 2009). Recruiting subjects of different ethnicities to a single

Table IV Results of genome-wide association studies of age at human menarche.

Chromosome location	SNP	*n*[a]	*P*	Ethnicity	Candidate gene(s)	References
5q31	rs2348186	477	4.92×10^{-7}	Caucasian	*SPOCK*	Liu *et al.* (2009)
	rs7701979		8.03×10^{-6}			
	rs13357391		5.77×10^{-6}			
		1387	5.09×10^{-3}	Chinese		
	rs1859345	477	1.58×10^{-5}	Caucasian		
		1387	4.37×10^{-3}	Chinese		
	rs10054991	477	1.20×10^{-5}	Caucasian		
	rs12653349		1.61×10^{-5}			
	rs17779700		4.81×10^{-6}			
6q21	rs314276	4714	1.5×10^{-8}	Caucasian	*LIN28B*	Ong *et al.* (2009)
6q21	rs314277	17 438	2.7×10^{-13}	Caucasian	*LIN28B*	He *et al.* (2009)
	rs314263		3.2×10^{-13}			
	rs369065		2.4×10^{-11}			
	rs314280		2.3×10^{-8}			
	rs4946651		3.1×10^{-8}			
	rs314262		9.7×10^{-8}			
	rs7759938	17 510	7.0×10^{-9}	Caucasian	*LIN28B*	Perry *et al.* (2009)
		87 802	5.4×10^{-60}	Caucasian	*LIN28B*	Elks *et al.* (2010)
	rs314280	10 040	1.8×10^{-14}	Caucasian	*LIN28B*	Sulem *et al.* (2009)
9q31.2	rs7861820	17 438	3.4×10^{-9}	Caucasian	*TMEM38B*, *FKTN*, *FSD1L*, *TAL2* & *ZNF462*	He *et al.* (2009)
	rs12684013		3.6×10^{-8}			
	rs4452860		7.9×10^{-8}			
	rs7028916		9.7×10^{-8}			
	rs2090409	17 510	1.7×10^{-9}	Caucasian		Perry *et al.* (2009)
		87 802	2.2×10^{-33}	Caucasian		Elks *et al.* (2010)
30 New discovered loci						
1q22–q23	rs466639	87 802	1.3×10^{-13}		*RXRG*	Elks *et al.* (2010)
1q25.2	rs633715		2.1×10^{-8}		*SEC16B*	
2q33	rs12617311		6.0×10^{-13}		*PLCL1*	
2q16.1	rs17268785		9.7×10^{-11}		*CCDC85A*	
2q22.3	rs17188434		1.1×10^{-9}		*NR4A2*	
2q25.3	rs2947411[b]		1.7×10^{-8}		*TMEM18*	
3q13.32	rs6438424		1.4×10^{-13}		N.S.	
3q26.2–q27	rs2002675		1.2×10^{-9}		*TRA2B*, *ETV5*	
3q22.3	rs6439371[b]		1.3×10^{-8}		*TMEM108*	
	rs6439371		1.3×10^{-8}		*NPHP3*	
3q21.31	rs7617480		2.8×10^{-8}		*KLHDC8B*	
3q21.3	rs6762477		2.8×10^{-8}		*RBM6*	
5q31.1	rs13187289		1.9×10^{-10}		*PHF15*	
6q22.32	rs1361108		1.7×10^{-8}		*C6orf173*, *TRMT11*	
6q16	rs4840086		2.4×10^{-8}		*PRDM13*, *MCHR2*	
7p13	rs1079866		5.5×10^{-14}		*INHBA*	
8q21.13	rs7821178		3.0×10^{-9}		*PXMP3*	
9q31.3	rs10980926		4.2×10^{-11}		*ZNF483*	
11q24.3	rs6589964		1.9×10^{-12}		*BSX*	
11q13.5	rs10899489		8.1×10^{-9}		*GAB2*	
11q15.4	rs4929923[b]		1.2×10^{-8}		*TRIM66*	
11q15	rs900145		1.6×10^{-8}		*ARNTL*	
14q32.2	rs6575793		1.2×10^{-8}		*BEGAIN*	
16p13.1	rs1659127		4.0×10^{-9}		*MKL2*	
16q22.1	rs1364063[b]		1.8×10^{-8}		*NFAT5*	
16q12.2	rs9939609		3.1×10^{-8}		*FTO*	
17q21.33	rs9635759		7.3×10^{-13}		*CA10*	
18q21.1	rs1398217		2.3×10^{-13}		*FUSSEL18*	
19q13.11	rs10423674		5.9×10^{-9}		*CRTC1*	
19q13.2	rs76422134		3.5×10^{-10}		*VGLL3*	
20p11.2	rs852069		3.3×10^{-8}		*PCSK2*	

Continued

Table IV *Continued*

Chromosome location	SNP	n^a	*P*	Ethnicity	Candidate gene(s)	References
10 Possible loci						
2q21.2	rs12472911	87802	1.5×10^{-7}		*LRP1B*	Elks *et al.* (2010)
3q21.3	rs2687729		1.3×10^{-7}		*EEFSEC*	
3q27.1	rs3914188		2.6×10^{-7}		*ECE2*	
5q31	rs757647		5.4×10^{-8}		*KDM3B*	
11p11.2	rs16938437		5.6×10^{-8}		*PHF21A*	
13q34	rs9555810		5.6×10^{-8}		*C13orf16, ARHGEF7*	
5q22.2	rs3743266		8.0×10^{-7}		*RORA*	
15q23	Rs7359257		1.9×10^{-6}		*IQCH*	
18q21.1	rs2243803		3.4×10^{-7}		*SLC14A2*	
19p13.2	rs1862471		1.5×10^{-7}		*OLFM2*	

N.S., Not specified.
[a]Number of subjects genotyped.
[b]Significant after correction for multiple testing.

cohort results in population stratification, which is common, for example, in the studies employing European Americans (Campbell *et al.*, 2005) and may result in biased results (Marchini *et al.*, 2004).

Another possible factor for the inconsistency is different sample size and therefore different statistical power of studies. Samples of some of the studies were quite small, i.e. <200 individuals (Kulik-Rechberger *et al.*, 2007; Mitchell *et al.*, 2008; Xita *et al.*, 2010). Since complex traits are determined by many genes/polymorphisms, any single genetic variant contributes just a small proportion to the total variation of the trait. Therefore, small-scale studies may underestimate a real effect of polymorphisms (Ioannidis *et al.*, 2003; Ambrosius *et al.*, 2004; Hattersley and McCarthy, 2005).

Differences in study design, for example, lack of control for covariates, may also cause the inconsistency of the results. Complex traits, such as age at menarche and age at natural menopause, are influenced by multiple environmental factors (Petridou *et al.*, 1996; Chie *et al.*, 1997; Dvornyk *et al.*, 2006), which may interact with candidate genes. The number of such environmental covariates is apparently large and many of them are still unknown, so that it is hardly possible to account for all of them. For that reason, an association analysis of complex traits always yields somewhat biased results. The bias may be quite noticeable if significant covariates were not incorporated into the analysis. Lack of control for at least the known significant covariates for age at menarche and age at natural menopause is often related to the fact that a great majority of published genetic studies on these traits were not specifically designed for this purpose but rather were a by-product of epidemiological studies of other complex traits, for example, osteoporosis, cardiovascular diseases and cancers. It should be admitted, however, that the number of possible environmental covariates is apparently smaller for age at menarche than for age at natural menopause simply because of the fact that a life span before the former is much shorter than before the latter. Therefore, many factors with a potential significant effect may either not take place or have a much smaller effect during adolescence. For example, smoking, alcohol intake and duration of breastfeeding were determined among such significant factors for age at natural menopause (Lu *et al.*, 2010, 2010a, b) because of their estrogen-related effect (Harlow and Signorello, 2000; Ma *et al.*, 2006). However, they likely do not affect age at menarche because they usually occur in adulthood.

One more source for the inconsistency is a problem of multiple testing. It is thought that not adjusting for multiple testing may yield more false-positive results. On the other hand, strict adjustments for multiple testing (e.g. Bonferroni correction) are notoriously too conservative and in many cases reject the existing association, especially if it is weak, which is common for the candidate genes of complex traits. This problem is of particular concern in large-scale studies, for example, genome-wide association studies or genome-wide gene expression profiling, when thousands of markers are analyzed simultaneously (Frazer *et al.*, 2009).

Some non-overlapping results of genome-wide linkage studies and genome-wide association studies may be attributed to differences in their concept and design. Linkage analysis is thought to be more appropriate for finding genes either contributing to monogenic traits or having 'a major gene effect' on a complex trait, i.e. accounting for a relatively large proportion to the overall genetic variance of the trait (Cui *et al.*, 2010). In contrast, the design of genome-wide association studies is usually based on the common gene: common disease hypothesis (Iyengar and Elston, 2007; Iles, 2008).

The genome-wide linkage studies of age at menarche conducted so far by different groups of researchers yielded inconsistent results (Table I). The possible reasons for that may lie in the complex genetic basis of age at menarche, between-study differences in design, sample size, statistical analyses used etc. Altogether, these factors affect the statistical power of the analysis, which is of key importance when considering the observed inconsistency between the studies (Liu *et al.*, 2003). In most cases, the power of linkage analysis is low unless a locus/loci accounts for a substantial proportion of trait heritability (Risch and Merikangas, 1996). For example, if a quantitative trait locus (QTL) has heritability of 10%, more that 800 000 random sib pairs should be analyzed to detect a linkage signal of LOD score ~3.3 with 80% power at the significance level of 10−4 (Risch and Merikangas, 1996). Of course, such a sample size is far beyond that any linkage study can afford.

Table V Strengths, weaknesses, opportunities and threats for studies on genetics of age at human menarche.

Strengths	Weaknesses
• Several candidate genes and genomic regions were identified • Extensive data on physiology of puberty and reproduction are available • High-throughput technologies, i.e. genome-wide sequencing, genotyping, transcriptomics and proteomics are already available • Advanced bioinformatics methods and computer technologies for analysis of large volume of data are in place • Critical mass in genetic and reproductive epidemiology • Evidence for shared genetic basis of age at menarche and other complex traits—a good start-up for future research	• Largely unknown etiology of age at menarche • Inconsistencies in results from different studies and failure to replicate results • Lack of data from ethnicities other than Caucasians • Absence of data about gene and protein expression in relation to age at menarche • Rare and potentially functionally important genetic variants may be overlooked • Insufficient data about gene–gene and gene–environment interactions and their effects on age at menarche • Lack of the systems approach
Opportunities	**Threats**
• Integration of genetic mapping and functional studies of candidate genetic variants • Metabolomic profiles of cells and tissues to identify genes and pathways contributing to age at menarche • New hypotheses about mechanisms of the onset of menarche and its possible health consequences from large-scale research • Use animal models to understand mechanisms of age at menarche • Efficiency of the epidemiological research may be increased by organizing international consortia and combining efforts and data of different research groups • Utilization of the systems approach integrating high-throughput genomic and metabolomic methods in a single pipeline • Implementation of the results in express diagnostics and preventive medicine	• Given that menarche is a complex trait and therefore the number of contributing genes is likely large, is it possible to solve this problem? • Environmental factors may be more significant than expected and their number may be very large, if not indefinite • How to bring the data from different studies and approaches to the common denominator? • Technological ability to detect candidate genetic variants does exist but what do they mean functionally? • How to use the data to control age at menarche? If it is technically possible, is it indeed necessary?

In addition to the above-mentioned inconsistency of results, there are two more gaps in genetic studies on menarcheal age: almost no results from ethnicities other than Caucasians and lack of genome-wide gene expression profiling. Since age at menarche is a trait with some ethnic background (Herman-Giddens *et al.*, 1997; Anderson *et al.*, 2003; Chumlea *et al.*, 2003), genetic variants contributing to it may not be the same in different ethnicities (Jackson, 2004).

It should be noted here that a weakness of focusing on menarche is that it is the arbitrary 'end-point' of an enormous cascade of events, called puberty. Therefore, a possibly limited contribution of genetic studies on menarche seen from the perspective of puberty must also be taken into account.

Menarcheal age and other complex traits: evidence for a shared genetic basis

The results of the genetic studies of age at menarche provide strong evidence for its shared genetic basis with many other traits, particularly those that are either associated with pubertal development (e.g. height, BMI, bone mineral density) or influenced by menarcheal age in later life (various disease phenotypes). For example, Elks *et al.* (2010) determined several obesity- and height-related genes to be associated with age at menarche (Supplementary data, Tables S2 and S3). Many genes were chosen for the association analysis based on the available information about the possible influence of age at menarche on the respective phenotypes. For example, estrogen-metabolizing genes, in addition to their well-known importance for reproductive function, have been extensively studied for their association with breast cancer (e.g. Andersen *et al.*, 1994; Ye and Parry, 2002) and osteoporosis (see Liu *et al.*, 2003 for review). One of the reviewed articles provided direct evidence about two genomic regions (22q13 and 3p25), which may harbor a QTL for both age at menarche and bone mineral density (Pan *et al.*, 2008).

Another interesting question is whether the onsets of menarche and menopause are determined by the same genes. So far, attempts to determine correlations between these two phenotypes have yielded inconclusive results (e.g. Whelan *et al.*, 1990; van Noord *et al.*, 1997; Do *et al.*, 1998; Peccei, 2000; Kaczmarek, 2007). Estimates from a number of twin- and family-based studies suggest that the contribution of genetic factors to age at natural menopause is similar to that of age at menarche, i.e. 57–82% (Snieder *et al.*, 1998; de Bruin *et al.*, 2001; Murabito *et al.*, 2005). The estrogen-dependent etiology of these traits suggests that genes involved in steroid metabolism may contribute to both. Indeed, results from genetic association studies suggest that at least some genes likely underlie both traits (e.g. Lu *et al.*, 2010; Liu *et al.*, 2010b). On the other hand, the more pronounced secular trend of age at menarche may result from differences in gene–environment interactions. Currently, the data about these effects, as well as the combined effect of various genes (epistasis), on the timing of menarche are scarce. Therefore, putting more efforts into research on these effects will help to advance our knowledge about a genetic basis of menarcheal age.

Prospects for research on genetics of age at menarche

Significant modifications of the existing methodology for studies on genetics of age at menarche are needed to solve the problems outlined above. They can be summarized as follows.

Designs of the studies should be modified in order to increase their statistical power. This may include (but not be limited to) modifications commonly applied to population- and family-based association

studies, genome-wide association studies, linkage studies and microarray expression profiling, such as increasing sample size, better controlling for various confounding factors, decreasing sample heterogeneity etc. (Risch and Merikangas, 1996; Hattersley and McCarthy, 2005; Lee and Saeed, 2007). A powerful method is a meta-analysis, which makes it possible to significantly increase the sample size and the statistical power of analysis (Elks *et al.*, 2010).

Another modification is shifting the research focus from the common gene: common disease hypothesis to the rare gene hypothesis (Frazer *et al.*, 2009; Gorlov *et al.*, 2011). Rare variants (with population frequencies <5%) were suggested to account for up to several percent of differences between phenotypes of carriers and non-carriers (Liu *et al.*, 2005a). However, genome-wide association studies and candidate gene association studies are not quite appropriate for detecting such rare variants because their methodology is grounded on the common variant hypothesis, as outlined above. To overcome these limitations, several approaches have been proposed recently (Bodmer and Bonilla, 2008; Zhang *et al.*, 2010).

Extensive studies on the ethnicities other than Caucasian are needed to elucidate a genetic basis of the inter-ethnic differences in age at menarche.

Microarrays are a very powerful method for genome-wide gene expression profiling and identification of genes, which may contribute to the given trait. Microarrays have been utilized extensively and efficiently for studying a genetic basis of various complex traits and diseases (e.g. Yanagawa *et al.*, 2001; Arimoto *et al.*, 2003; Liu *et al.*, 2005b), including menopause (Dvornyk *et al.*, 2007; Gomez-Santos *et al.*, 2011). Although the studies on menopause were of a small scale and thus had limited power, they nonetheless demonstrated the feasibility of this approach and its appropriateness for the identification of candidate genes for this trait. Microarrays seem to be more suitable for searching for genes that underlie menarche/menopause *per se* rather than the timing of onset of these traits (Dvornyk *et al.*, 2007). On the other hand, a certain proportion of genes for menarche (or menopause) most likely contribute to the timing of these traits.

A real breakthrough in genetics of menarcheal age may be accomplished by implementing the systems approach, which incorporates candidate gene association studies, genome-wide association studies, linkage analysis, microarray gene expression profiling and proteomics in a single pipeline (Dvornyk *et al.*, 2004). This approach is based on a concept that the molecular mechanism underlying a trait consists of three regulatory levels, which correspond to DNA, mRNA and protein. Candidate QTL can be determined at each of these levels by using the respective methods: association and linkage, at the DNA level; genome-wide gene expression microarrays, at the mRNA level; and proteomics, at the protein level. Identification of the QTL for menarcheal age may follow either forward (DNA → RNA → protein) or backward (protein → RNA → DNA) strategies but the latter looks more promising because it provides a possibility for systemic analysis of the gene expression, respective regulatory networks, and effect of gene–gene and gene–environment interactions on the onset of a complex trait (Lei *et al.*, 2005). Ultimately, such a combined approach should open a new avenue in the efforts towards determining a genetic basis of menarche and its timing.

Several studies reviewed here have successfully demonstrated the power and efficiency of the systems approach by incorporating some of its elements. For example, identification of *SPOCK* as a candidate gene for age at menarche utilized genome-wide association studies and follow-up candidate SNP and haplotype-based association analysis (Liu *et al.*, 2009). Likewise, the design of the study, which replicated previous results on association of several candidate genes with menarcheal age and identified about 40 novel loci, incorporated a meta-analysis of several large-scale genome-wide association studies and the follow-up replication analysis (Elks *et al.*, 2010).

Threats

As age at menarche is a complex trait, the potential number of contributing genes may be very large. For example, a Skeletal Gene Database lists several hundred genes related to bone biology (Ho *et al.*, 2000). A task of identifying all candidate genes for a complex trait is therefore quite challenging, especially given that the effect of the most of them is quite weak. This is equally applicable to environmental factors. Another challenge is bringing the continuously growing volume of data (often controversial) from different studies to the common denominator, to make meaningful conclusions. Last but not least is to find out how to implement technically the results in order to accomplish a possible goal: managing age at menarche to prevent associated health complications later in life.

Conclusion

Research into the genetics of menarche and its timing is in its early stages. Despite great efforts having been made during the last decade and significant results obtained, the problem of determining the genetic factors underlying this trait is far from being solved. However, with the advances in modern high-throughput technologies and implementation of the systems approach, the ultimate goal of identifying genes and their variants contributing to menarche and its onset becomes achievable.

Supplementary data

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

Authors' role

V.D. designed the study, wrote the major part of the manuscript and made major revisions; W.H. collected the data, and contributed to drafting the initial version of the manuscript and to revisions.

Funding

No specific funding was obtained to support this study.

References

Ambrosius WT, Lange EM, Langefeld CD. Power for genetic association studies with random allele frequencies and genotype distributions. *Am J Hum Genet* 2004; **74**:683–693.

Andersen TI, Heimdal KR, Skrede M, Tveit K, Berg K, Borresen AL. Oestrogen receptor (*ESR*) polymorphisms and breast cancer susceptibility. *Hum Genet* 1994;**94**:665–670.

Anderson DM, Maraskovsky E, Billingsley WL, Dougall WC, Tometsko ME, Roux ER, Teepe MC, DuBose RF, Cosman D, Galibert L. A homologue of the

TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* 1997;**390**:175–179.

Anderson SE, Dallal GE, Must A. Relative weight and race influence average age at menarche: results from two nationally representative surveys of US girls studied 25 years apart. *Pediatrics* 2003;**111**:844–850.

Anderson CA, Duffy DL, Martin NG, Visscher PM. Estimation of variance components for age at menarche in twin families. *Behav Genet* 2007;**37**:668–677.

Anderson CA, Zhu G, Falchi M, van den Berg SM, Treloar SA, Spector TD, Martin NG, Boomsma DI, Visscher PM, Montgomery GW. A genome-wide linkage scan for age at menarche in three populations of European descent. *J Clin Endocrinol Metab* 2008;**93**:3965–3970.

Arimoto T, Katagiri T, Oda K, Tsunoda T, Yasugi T, Osuga Y, Yoshikawa H, Nishii O, Yano T, Taketani Y *et al.* Genome-wide cDNA microarray analysis of gene-expression profiles involved in ovarian endometriosis. *Int J Oncol* 2003; **22**:551–560.

Attia GR, Zeitoun K, Edwards D, Johns A, Carr BR, Bulun SE. Progesterone receptor isoform A but not B is expressed in endometriosis. *J Clin Endocrinol Metab* 2000; **85**:2897–2902.

Bandini LG, Must A, Naumova EN, Anderson S, Caprio S, Spadano-Gasbarro JI, Dietz WH. Change in leptin, body composition and other hormones around menarche—a visual representation. *Acta Paediatr* 2008;**97**:1454–1459.

Beavan S, Prentice A, Dibba B, Yan L, Cooper C, Ralston SH. Polymorphism of the collagen type I a1 gene and ethnic differences in hip-fracture rates. *N Engl J Med* 1998;**339**:351–352.

Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;**40**:695–701.

Boytsov S, van de Werf F. Regional challenges and opportunities in cardiovascular research: The Russian Federation 'SWOT' analysis. *Am Heart J* 2011; **161**:427–430.

Bray M, Higgins JP, Ioannidis JP, Khoury MJ, Little J, Manolio TA, Smeeth T, Sterne J. (2006) The HuGENet™ HuGE Review Handbook, version 1.0., http://www.hugenet.ca (28 February 2006, date last accessed).

Brisken C, Heineman A, Chavarria T, Elenbaas B, Tan J, Dey SK, McMahon JA, McMahon AP, Weinberg RA. Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. *Genes Dev* 2000; **14**:650–654.

Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, Altshuler D, Ardlie KG, Hirschhorn JN. Demonstrating stratification in a European American population. *Nat Genet* 2005;**37**:868–872.

Chavarro J, Villamor E, Narvaez J, Hoyos A. Socio-demographic predictors of age at menarche in a group of Colombian university women. *Ann Hum Biol* 2004; **31**:245–257.

Chie WC, Liu YH, Chi J, Wu V, Chen A. Predictive factors for early menarche in Taiwan. *J Formos Med Assoc* 1997;**96**:446–450.

Chumlea WC, Schubert CM, Roche AF, Kulin HE, Lee PA, Himes JH, Sun SS. Age at menarche and racial comparisons in US girls. *Pediatrics* 2003;**111**:110–113.

Comings DE, Gade R, Muhleman D, Peters WR, MacMurray JP. The *LEP* gene and age of menarche: maternal age as a potential cause of hidden stratification in association studies. *Mol Genet Metab* 2001;**73**:204–210.

Cooper GS, Ephross SA, Weinberg CR, Baird DD, Whelan EA, Sandler DP. Menstrual and reproductive risk factors for ischemic heart disease. *Epidemiology* 1999;**10**:255–259.

Cui Y, Li G, Li S, Wu R. Designs for linkage analysis and association studies of complex diseases. *Methods Mol Biol* 2010;**620**:219–242.

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

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

Dvornyk V, Liu H, Shen H, Lei SF, Zhao L, Huang QR, Qin Y, Jiang DK, Long J, Zhang Y *et al.* Differentiation of Caucasians and Chinese at bone mass candidate genes: implication for ethnic difference of bone mass. *Ann Hum Genet* 2003;**67**:216–227.

Dvornyk V, Xiao P, Liu YJ, Shen H, Deng HW. Systemic approach to the study of complex bone disorders at the whole-genome level. *Curr Genomics* 2004;**5**:93–108.

Dvornyk V, Long JR, Liu PY, Zhao LJ, Shen H, Recker RR, Deng HW. Predictive factors for age at menopause in Caucasian females. *Maturitas* 2006;**54**:19–26.

Dvornyk V, Liu Y, Lu Y, Shen H, Lappe JM, Lei S, Recker RR, Deng H. Effect of menopause on gene expression profiles of circulating monocytes: a pilot in vivo microarray study. *J Genet Genomics* 2007;**34**:974–983.

Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, Lunetta KL, Visser JA, Byrne EM, Cousminer DL *et al.* Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet* 2010;**42**:1077–1085.

Ettinger SL, Sobel R, Whitmore TG, Akbari M, Bradley DR, Gleave ME, Nelson CC. Dysregulation of sterol response element-binding proteins and downstream effectors in prostate cancer during progression to androgen independence. *Cancer Res* 2004;**64**:2212–2221.

Feigelson HS, Coetzee GA, Kolonel LN, Ross RK, Henderson BE. A polymorphism in the *CYP17* gene increases the risk of breast cancer. *Cancer Res* 1997; **57**:1063–1065.

Feng Y, Hong X, Wilker E, Li Z, Zhang W, Jin D, Liu X, Zang T, Xu X, Xu X. Effects of age at menarche, reproductive years, and menopause on metabolic risk factors for cardiovascular diseases. *Atherosclerosis* 2008;**196**:590–597.

Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet* 2009;**10**:241–251.

Fujita M, Tase T, Kakugawa Y, Hoshi S, Nishino Y, Nagase S, Ito K, Niikura H, Yaegashi N, Minami Y. Smoking, earlier menarche and low parity as independent risk factors for gynecologic cancers in Japanese: a case–control study. *Tohoku J Exp Med* 2008;**216**:297–307.

Garnick MB. Redefining the refined retroperitoneal lymph node dissection in testis cancer: a SWOT analysis of nonrandomized data. *J Clin Oncol* 2007; **25**:4337–4338.

Gomez-Santos C, Hernandez-Morante JJ, Margareto J, Larrarte E, Formiguera X, Martinez CM, Garaulet M. Profile of adipose tissue gene expression in premenopausal and postmenopausal women: site-specific differences. *Menopause* 2011;**18**:675–684.

Gorai I, Tanaka K, Inada M, Morinaga H, Uchiyama Y, Kikuchi R, Chaki O, Hirahara F. Estrogen-metabolizing gene polymorphisms, but not estrogen receptor-a gene polymorphisms, are associated with the onset of menarche in healthy postmenopausal Japanese women. *J Clin Endocrinol Metab*, 2003; **88**:799–803.

Gorlov IP, Gorlova OY, Frazier ML, Spitz MR, Amos CI. Evolutionary evidence of the effect of rare variants on disease etiology. *Clin Genet* 2011;**79**:199–206.

Graber JA, Brooks-Gunn J, Warren MP. The antecedents of menarcheal age: heredity, family environment, and stressful life events. *Child Dev* 1995; **66**:346–359.

Graham MJ, Larsen U, Xu X. Secular trend in age at menarche in China: a case study of two rural counties in Anhui Province. *J Biosoc Sci* 1999;**31**:257–267.

Grimm C, Tempfer CB, Walch K, Reinthaller A, Tomovski C, Huber JC, Leodolter S, Hefler LA. The influence of a vitamin D receptor gene polymorphism on the timing of female reproductive functions in humans. *Maturitas*, 2005;**51**:135–139.

Guo Y, Zhao LJ, Shen H, Guo Y, Deng HW. Genetic and environmental correlations between age at menarche and bone mineral density at different skeletal sites. *Calcif Tissue Int* 2005;**77**:356–360.

Guo Y, Shen H, Xiao P, Xiong DH, Yang TL, Guo YF, Long JR, Recker RR, Deng HW. Genomewide linkage scan for quantitative trait loci underlying variation in age at menarche. *J Clin Endocrinol Metab* 2006a;**91**:1009–1014.

Guo Y, Xiong DH, Yang TL, Guo YF, Recker RR, Deng HW. Polymorphisms of estrogen-biosynthesis genes *CYP17* and *CYP19* may influence age at menarche: a genetic association study in Caucasian females. *Hum Mol Genet* 2006b; **15**:2401–2408.

Harlow BL, Signorello LB. Factors associated with early menopause. *Maturitas* 2000; **35**:3–9.

Hattersley AT, McCarthy MI. What makes a good genetic association study? *Lancet* 2005;**366**:1315–1323.

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

Herman-Giddens ME. The decline in the age of menarche in the United States: should we be concerned? *J Adolesc Health* 2007;**40**:201–203.

Herman-Giddens ME, Slora EJ, Wasserman RC, Bourdony CJ, Bhapkar MV, Koch GG, Hasemeier CM. Secondary sexual characteristics and menses in

young girls seen in office practice: a study from the Pediatric Research in Office Settings network. *Pediatrics* 1997;**99**:505–512.

Ho NC, Jia L, Driscoll CC, Gutter EM, Francomano CA. A skeletal gene database. *J Bone Miner Res* 2000;**15**:2095–2122.

Hwang JY, Shin C, Frongillo EA, Shin KR, Jo I. Secular trend in age at menarche for South Korean women born between 1920 and 1986: the Ansan Study. *Ann Hum Biol* 2003;**30**:434–442.

Iles MM. What can genome-wide association studies tell us about the genetics of common disease? *PLoS Genet* 2008;**4**:e33.

Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. *Lancet* 2003; **361**:567–571.

Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, Balding DJ, Chokkalingam A, Dolan SM, Flanders WD *et al.* Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008; **37**:120–132.

Irwin JC, Kirk D, Gwatkin RB, Navre M, Cannon P, Giudice LC. Human endometrial matrix metalloproteinase-2, a putative menstrual proteinase. Hormonal regulation in cultured stromal cells and messenger RNA expression during the menstrual cycle. *J Clin Invest* 1996;**97**:438–447.

Ito M, Yamada M, Hayashi K, Ohki M, Uetani M, Nakamura T. Relation of early menarche to high bone mineral density. *Calcif Tissue Int* 1995;**57**:11–14.

Iyengar SK, Elston RC. The genetic basis of complex traits: rare variants or 'common gene, common disease'?. *Methods Mol Biol* 2007;**376**:71–84.

Jackson FL. Human genetic variation and health: new assessment approaches based on ethnogenetic layering. *Br Med Bull* 2004;**69**:215–235.

Jiang Y, Han J, Liu J, Zhang G, Wang L, Liu F, Zhang X, Zhao Y, Pang D. Risk of genome-wide association study newly identified genetic variants for breast cancer in Chinese women of Heilongjiang Province. *Breast Cancer Res Treat* 2011;**128**:251–257.

Kaczmarek M. The timing of natural menopause in Poland and associated factors. *Maturitas* 2007;**27**:139–153.

Kaprio J, Rimpela A, Winter T, Viken RJ, Rimpela M, Rose RJ. Common genetic influences on BMI and age at menarche. *Hum Biol* 1995;**67**:739–753.

Kawase T, Matsuo K, Suzuki T, Hiraki A, Watanabe M, Iwata H, Tanaka H, Tajima K. *FGFR2* intronic polymorphisms interact with reproductive risk factors of breast cancer: results of a case control study in Japan. *Int J Cancer*, 2009;**125**: 1946–1952.

Kotsopoulos J, Lubinski J, Lynch HT, Neuhausen SL, Ghadirian P, Isaacs C, Weber B, Kim-Sing C, Foulkes WD, Gershoni-Baruch R *et al.* Age at menarche and the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers. *Cancer Causes Contr*, 2005;**16**:667–674.

Kulik-Rechberger B, Skorupski P, Bogusiewicz M, Miotla P, Rechberger T. Polimorfizm genu *CYP17* a wiek menarche [Polymorhism of *CYP17* gene and age of menarche]. *Ginekol Pol* 2007;**78**:929–932.

Lai J, Vesprini D, Chu W, Jernstrom H, Narod SA. *CYP* gene polymorphisms and early menarche. *Mol Genet Metab* 2001;**74**:449–457.

Lakshman R, Forouhi NG, Sharp SJ, Luben R, Bingham SA, Khaw KT, Wareham NJ, Ong KK. Early age at menarche associated with cardiovascular disease and mortality. *J Clin Endocrinol Metab* 2009;**94**:4953–4960.

Lee NH, Saeed AI. Microarrays: an overview. *Methods Mol Biol* 2007;**353**:265–300.

Lei SF, Deng FY, Liu XH, Huang QR, Qin Y, Zhou Q, Jiang DK, Li YM, Mo XY, Liu MY *et al.* Polymorphisms of four bone mineral density candidate genes in Chinese populations and the comparison with the other populations of different ethnicity. *J Bone Miner Metab* 2003;**21**:34–42.

Lei SF, Wu S, Dvornyk V, Deng HW. Two strategies to identify genes underlying complex diseases. *Curr Genomics* 2005;**6**:551–561.

Lin-Su K, Vogiatzi MG, New MI. Body mass index and age at menarche in an adolescent clinic population. *Clin Pediatr* 2002;**41**:501–507.

Liu YZ, Liu YJ, Recker RR, Deng HW. Molecular studies of identification of genes for osteoporosis: the 2002 update. *J Endocrinol* 2003;**177**:147–196.

Liu PY, Zhang YY, Lu Y, Long JR, Shen H, Zhao LJ, Xu FH, Xiao P, Xiong DH, Liu YJ *et al.* A survey of haplotype variants at several disease candidate genes: the importance of rare variants for complex diseases. *J Med Genet* 2005a;**42**:221–227.

Liu YZ, Dvornyk V, Lu Y, Shen H, Lappe JM, Recker RR, Deng HW. A novel pathophysiological mechanism for osteoporosis suggested by an in vivo gene expression study of circulating monocytes. *J Biol Chem* 2005b;**280**:29011–29016.

Liu YZ, Guo YF, Wang L, Tan LJ, Liu XG, Pei YF, Yan H, Xiong DH, Deng FY, Yu N *et al.* Genome-wide association analyses identify *SPOCK* as a key novel gene underlying age at menarche. *PLoS Genet* 2009;**5**:e1000420.

Liu PY, Lu Y, Recker RR, Deng HW, Dvornyk V. *ALOX12* gene is associated with the onset of natural menopause in white women. *Menopause* 2010a;**17**:152–156.

Liu PY, Lu Y, Recker RR, Deng HW, Dvornyk V. Association analyses suggest multiple interaction effects of the methylenetetrahydrofolate reductase polymorphisms on timing of menarche and natural menopause in whites. *Menopause* 2010b;**17**:185–190.

Long JR, Xu H, Zhao LJ, Liu PY, Shen H, Liu YJ, Xiong DH, Xiao P, Liu YZ, Dvornyk V *et al.* The oestrogen receptor a gene is linked and/or associated with age of menarche in different ethnic groups. *J Med Genet*, 2005;**42**:796–800.

Lu Y, Liu P, Recker RR, Deng HW, Dvornyk V. *TNFRSF11A* and *TNFSF11* are associated with age at menarche and natural menopause in white women. *Menopause* 2010;**17**:1048–1054.

Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res* 2006;**8**:R43.

Manuck SB, Craig AE, Flory JD, Halder I, Ferrell RE. Reported early family environment covaries with menarcheal age as a function of polymorphic variation in estrogen receptor-a. *Dev Psychopathol* 2011;**23**:69–83.

Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. *Nat Genet* 2004;**36**:512–517.

Marshall LM, Spiegelman D, Goldman MB, Manson JE, Colditz GA, Barbieri RL, Stampfer MJ, Hunter DJ. A prospective study of reproductive factors and oral contraceptive use in relation to the risk of uterine leiomyomata. *Fertil Steril* 1998;**70**:432–439.

Matkovic V, Ilich JZ, Skugor M, Badenhop NE, Goel P, Clairmont A, Klisovic D, Nahhas RW, Landoll JD. Leptin is inversely related to age at menarche in human females. *J Clin Endocrinol Metab* 1997;**82**:3239–3245.

McKibben SL, Poston DL Jr. The influence of age at menarche on the fertility of Chinese women. *Soc Biol* 2003;**50**:222–237.

Miller MG. Environmental metabolomics: a SWOT analysis (strengths, weaknesses, opportunities, and threats). *J Proteome Res* 2007;**6**:540–545.

Mitchell ES, Farin FM, Stapleton PL, Tsai JM, Tao EY, Smith-Dijulio K, Woods NF. Association of estrogen-related polymorphisms with age at menarche, age at final menstrual period, and stages of the menopausal transition. *Menopause* 2008;**15**:105–111.

Morris DH, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Familial concordance for age at menarche: analyses from the Breakthrough Generations Study. *Paediatr Perinat Epidemiol* 2011;**25**:306–311.

Mouillet JF, Sonnenberg-Hirche C, Yan X, Sadovsky Y. p300 regulates the synergy of steroidogenic factor-1 and early growth response-1 in activating luteinizing hormone-beta subunit gene. *J Biol Chem* 2004;**279**:7832–7839.

Murabito JM, Yang Q, Fox C, Wilson PW, Cupples LA. Heritability of age at natural menopause in the Framingham Heart Study. *Obstet Gynecol Surv* 2005; **60**:656–657.

Najafi-Shoushtari SH, Kristo F, Li Y, Shioda T, Cohen DE, Gerszten RE, Naar AM. MicroRNA-33 and the *SREBP* host genes cooperate to control cholesterol homeostasis. *Science* 2010;**328**:1566–1569.

Nakada M, Yamada A, Takino T, Miyamori H, Takahashi T, Yamashita J, Sato H. Suppression of membrane-type 1 matrix metalloproteinase (MMP)-mediated MMP-2 activation and tumor invasion by testican 3 and its splicing variant gene product, N-Tes. *Cancer Res* 2001;**61**:8896–8902.

Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D, Karachanak S, Piskackova T, Balascak I, Peltonen L *et al.* Genetic structure of Europeans: a view from the North-East. *PLoS One* 2009;**4**:e5472.

Ong KK, Elks CE, Li S, Zhao JH, Luan J, Andersen LB, Bingham SA, Brage S, Smith GD, Ekelund U *et al.* Genetic variation in *LIN28B* is associated with the timing of puberty. *Nat Genet* 2009;**41**:729–733.

Onland-Moret NC, Peeters PH, van Gils CH, Clavel-Chapelon F, Key T, Tjonneland A, Trichopoulou A, Kaaks R, Manjer J, Panico S *et al.* Age at menarche in relation to adult height: the EPIC study. *Am J Epidemiol* 2005; **162**:623–632.

Pan R, Liu YZ, Deng HW, Dvornyk V. Association analyses suggest the effects of *RANK* and *RANKL* on age at menarche in Chinese women. *Climacteric*, 2011 Oct 24. [Epub ahead of print].

Pan F, Xiao P, Guo Y, Liu YJ, Deng HY, Recker RR, Deng HW. Chromosomal regions 22q13 and 3p25 may harbor quantitative trait loci influencing both age at menarche and bone mineral density. *Hum Genet* 2008;**123**:419–427.

Pascual J, Garcia-Moro CE, Hernandez M. Biological and behavioral determinants of fertility in Tierra del Fuego. *Am J Phys Anthropol* 2005;**127**:105–113.

Peccei JS. Genetic correlation between the ages of menarche and menopause. *Hum Nat* 2000;**11**:43–63.

Peeters PH, Verbeek AL, Krol A, Matthyssen MM, de Waard F. Age at menarche and breast cancer risk in nulliparous women. *Breast Cancer Res Treat* 1995; **33**:55–61.

Pei YF, Zhang L, Deng HW, Dvornyk V. *CYP17 Msp*A1 polymorphism and age at menarche: a meta-analysis. *Dis Markers* 2008;**25**:87–95.

Perry JR, Stolk L, Franceschini N, Lunetta KL, Zhai G, McArdle PF, Smith AV, Aspelund T, Bandinelli S, Boerwinkle E *et al.* Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. *Nat Genet* 2009;**41**:648–650.

Petridou E, Syrigou E, Toupadaki N, Zavitsanos X, Willett W, Trichopoulos D. Determinants of age at menarche as early life predictors of breast cancer risk. *Int J Cancer* 1996;**68**:193–198.

Remsberg KE, Demerath EW, Schubert CM, Chumlea WC, Sun SS, Siervogel RM. Early menarche and the development of cardiovascular disease risk factors in adolescent girls: the Fels Longitudinal Study. *J Clin Endocrinol Metab* 2005; **90**:2718–2724.

Riestra P, Garcia-Anguita A, Torres-Cantero A, Bayonas MJ, de Oya M, Garces C. Association of the Q223R polymorphism with age at menarche in the leptin receptor gene in humans. *Biol Reprod* 2011;**84**:752–755.

Risch NJ, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;**273**:1516–1517.

Rothenbuhler A, Fradin D, Heath S, Lefevre H, Bouvattier C, Lathrop M, Bougneres P. Weight-adjusted genome scan analysis for mapping quantitative trait loci for menarchal age. *J Clin Endocrinol Metab* 2006;**91**:3534–3537.

Rothenbuhler A, Lotton C, Fradin D. Three common variants of *LEP*, *NPY1R* and *GPR54* show no association with age at menarche. *Horm Res* 2009;**71**:331–335.

Schatzkin A, Palmer JR, Rosenberg L, Helmrich SP, Miller DR, Kaufman DW, Lesko SM, Shapiro S. Risk factors for breast cancer in black women. *J Natl Cancer Inst* 1987;**78**:213–217.

Silva IV, Rezende LC, Lanes SP, Souza LS, Madeira KP, Cerri MF, Paes MF, Daltoe RD, Chambo-Filho A, Guimaraes MC *et al.* Evaluation of *Pvu*II and *Xba*I polymorphisms in the estrogen receptor a gene (*ESR1*) in relation to menstrual cycle timing and reproductive parameters in post-menopausal women. *Maturitas*, 2010;**67**:363–367.

Sizonenko PC. Physiology of puberty. *J Endocrinol Invest* 1989;**12**:59–63.

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

So T, Lee SW, Croft M. Tumor necrosis factor/tumor necrosis factor receptor family members that positively regulate immunity. *Int J Hematol* 2006;**83**:1–11.

Stavrou I, Zois C, Ioannidis JP, Tsatsoulis A. Association of polymorphisms of the oestrogen receptor a gene with the age of menarche. *Hum Reprod* 2002; **17**:1101–1105.

Stavrou I, Zois C, Chatzikyriakidou A, Georgiou I, Tsatsoulis A. Combined estrogen receptor a and estrogen receptor b genotypes influence the age of menarche. *Hum Reprod* 2006;**21**:554–557.

Sulem P, Gudbjartsson DF, Rafnar T, Holm H, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Alexandersen P, Feenstra B, Boyd HA *et al.* Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. *Nat Genet* 2009;**41**:734–738.

Suzuki T, Hayashi S, Miki Y, Nakamura Y, Moriya T, Sugawara A, Ishida T, Ohuchi N, Sasano H. Peroxisome proliferator-activated receptor g in human breast carcinoma: a modulator of estrogenic actions. *Endocr Relat Cancer* 2006; **13**:233–250.

Taylor KC, Small CM, Epstein MP, Sherman SL, Tang W, Wilson MM, Bouzyk M, Marcus M. Associations of progesterone receptor polymorphisms with age at menarche and menstrual cycle length. *Horm Res Paediatr* 2010;**74**:421–427.

Theill LE, Boyle WJ, Penninger JM. RANK-L and RANK: T cells, bone loss, and mammalian evolution. *Annu Rev Immunol* 2002;**20**:795–823.

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

Viswanathan SR, Powers JT, Einhorn W, Hoshida Y, Ng TL, Toffanin S, O'Sullivan M, Lu J, Phillips LA, Lockhart VL *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat Genet* 2009;**41**: 843–848.

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

Wang W, Zhao LJ, Liu YZ, Recker RR, Deng HW. Genetic and environmental correlations between obesity phenotypes and age at menarche. *Int J Obes* 2006;**30**:1595–1600.

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

Widen E, Ripatti S, Cousminer DL, Surakka I, Lappalainen T, Jarvelin MR, Eriksson JG, Raitakari O, Salomaa V, Sovio U *et al.* Distinct variants at *LIN28B* influence growth in height from birth to adulthood. *Am J Hum Genet* 2010; **86**:773–782.

Xita N, Chatzikyriakidou A, Stavrou I, Zois C, Georgiou I, Tsatsoulis A. The (TTTA)n polymorphism of aromatase (*CYP19*) gene is associated with age at menarche. *Hum Reprod* 2010;**25**:3129–3133.

Xita N, Tsatsoulis A, Stavrou I, Georgiou I. Association of *SHBG* gene polymorphism with menarche. *Mol Hum Reprod*, 2005;**11**:459–462.

Yanagawa R, Furukawa Y, Tsunoda T, Kitahara O, Kameyama M, Murata K, Ishikawa O, Nakamura Y. Genome-wide screening of genes showing altered expression in liver metastases of human colorectal cancers by cDNA microarray. *Neoplasia* 2001;**3**:395–401.

Yang F, Xiong DH, Guo Y, Shen H, Xiao P, Zhang F, Jiang H, Recker RR, Deng HW. The chemokine (C-C-motif) receptor 3 (*CCR3*) gene is linked and associated with age at menarche in Caucasian females. *Hum Genet* 2007;**121**:35–42.

Yazawa M, Ferrante C, Feng J, Mio K, Ogura T, Zhang M, Lin PH, Pan Z, Komazaki S, Kato K *et al.* TRIC channels are essential for Ca2+ handling in intracellular stores. *Nature* 2007;**448**:78–82.

Ye Z, Parry JM. The *CYP17 Msp*A1 polymorphism and breast cancer risk: a meta-analysis. *Mutagenesis* 2002;**17**:119–126.

Zhang J, Lathbury LJ, Salamonsen LA. Expression of the chemokine eotaxin and its receptor, CCR3, in human endometrium. *Biol Reprod* 2000;**62**:404–411.

Zhang L, Pei YF, Li J, Papasian CJ, Deng HW. Efficient utilization of rare variants for detection of disease-related genomic regions. *PLoS One* 2010;**5**:e14288.

Zhao J, Xiong DH, Guo Y, Yang TL, Recker RR, Deng HW. Polymorphism in the insulin-like growth factor I gene is associated with age at menarche in Caucasian females. *Hum Reprod*, 2007;**22**:1789–1794.