Gene 686 (2019) 228–236

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Association of genetic polymorphisms with age at menarche in Russian women

Irina Ponomarenko[a], Evgeny Reshetnikov[a,*], Oksana Altuchova[b], Alexey Polonikov[c], Inna Sorokina[a], Anna Yermachenko[d,e], Volodymyr Dvornyk[f], Oleg Golovchenko[a], Mikhail Churnosov[a]

[a] Department of Medical Biological Disciplines, Belgorod State University, 308015 Belgorod, Russia
[b] Department of Obstetrics and Gynecology, Belgorod State University, 308015 Belgorod, Russia
[c] Department of Biology, Medical Genetics and Ecology, Kursk State Medical University, 305041 Kursk, Russia
[d] Department of Social Epidemiology, Pierre Louis Institute of Epidemiology and Public Health, 75571 Paris, France
[e] Sorbonne Universités, 75320 Paris, France
[f] Department of Life Sciences, College of Science and General Studies, Alfaisal University, 11533 Riyadh, Saudi Arabia

ARTICLE INFO

Keywords:
Age at menarche
Association
Gene-gene interaction
Single nucleotide polymorphism

ABSTRACT

Objectives: Examine the association of genetic polymorphisms with age at menarche (AAM) in Russian women.
Study design: A total of 1613 Russian females were recruited for the study.

Fifty two polymorphisms were analyzed for their association with AAM, height, and BMI.

The associations were analyzed assuming the additive, dominant, and recessive models and using the log-linear regression as implemented in PLINK v. 2.050. The 2-, 3-, and 4-loci models of gene-gene interactions were analyzed using the MB-MDR method and validated by the permutation test.
Main outcome measures: Genetic polymorphism rs6438424 *3q13.32* was independently associated with AAM in Russian women. In addition, 14 SNPs were determined as possible contributors to this trait through gene-gene interactions.
Results: The obtained results suggest that 14 out of 52 studied polymorphisms may contribute to AAM in Russian women. The rs6438424 *3q13.32* polymorphism was associated with AAM according to both additive and dominant models ($p_{perm} = 0.005$). In total 12 two-, three-, and four-locus models of gene-gene interactions were determined as contributing to AAM ($p_{perm} \leq 0.006$). Nine of the 14 AAM-associated SNPs are also associated with height and BMI ($p_{perm} \leq 0.003$). Among 14 AAM-associated SNPs (a priori all having regulatory significance), the highest regulatory potential was determined for rs4633 *COMT*, rs2164808 *POMC*, rs2252673*INSR*, rs6438424 *3q13.32*, and rs10769908 *STK33*. Eleven loci are *cis*-eQTL and affect expression of 14 genes in various tissues and organs (FDR < 0.05).

The neuropeptide-encoding genes were overrepresented among the AAM-associated genes ($p_{bonf} = 0.039$).
Conclusions: The rs6438424 polymorphism is independently associated with AAM in Russian females in this study. The other 14 SNPs manifest this association through gene-gene interactions.

1. Introduction

Age at menarche (AAM) is one of the important characteristics of the pubertal development and marks a beginning of the reproductive period. AAM is controlled by the hypothalamic–pituitary–ovarian axis and associated with female fertility and possible health complications in the later life. Early menarche may increase a risk for obesity (Guo and Ji, 2011), uterine myoma (Wise and Laughlin-Tommaso, 2016), endometriosis (Nnoaham et al., 2012), breast cancer (Yermachenko and Dvornyk, 2014), cardiovascular diseases (Feng et al., 2008), type 2 diabetes mellitus, infertility and psychological problems (Yermachenko and Dvornyk, 2014).

Abbreviations: AAM, age at menarche; GWAS, genome-wide association studies; SNP, single nucleotide polymorphism; BMI, body mass index; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; HWE, Hardy-Weinberg equilibrium; MB-MDR, Model Based Multifactor Dimensionality Reduction; MDR, Multifactor Dimensionality Reduction; FDR, False Discovery Rate; LD, linkage disequilibrium; β, coefficient of the linear regression

* Corresponding author at: 85, Pobedy St., Belgorod State Universite, Department of Medical Biological Disciplines, Belgorod 308015, Russia.
E-mail address: reshetnikov@bsu.edu.ru (E. Reshetnikov).

https://doi.org/10.1016/j.gene.2018.11.042
Received 12 June 2018; Received in revised form 19 October 2018; Accepted 15 November 2018
Available online 16 November 2018

AAM has significant between-ethnic variation (Yermachenko and Dvornyk, 2014). The variation in AAM was reported between populations of the same ethnicity too. For example, the mean AAM in females from South Europe (12.0–12.6 years) is lower than in the Northern European countries (13.0–13.3 years) (Yermachenko and Dvornyk, 2014; Parent et al., 2003). Twin and family studies suggested that genetic factors account up to 53–74% of AAM variance (Kaprio et al., 1995). Several genome-wide association studies (GWAS) having conducted so far determined over 100 polymorphisms, which might contribute to the trait (He et al., 2009; Ong et al., 2009; Perry et al., 2009; Sulem et al., 2009; Elks et al., 2010; Demerath et al., 2013; Perry et al., 2014). However, the reproducibility of these results in various populations and ethnicities is low. For example, in a sample of 6269 Chinese females, Delahanty et al. (Delahanty et al., 2013) confirmed association with AAM of only 9 out of 37 single nucleotide polymorphism (SNPs) previously reported by Elks et al. (Elks et al., 2010) in their GWAS. Pyun et al. (Pyun et al., 2014) failed to reproduce association with AAM of 42 SNPs from the GWAS by He et al. (He et al., 2010) in the Korean sample of 3452 females. Out of 33 SNPs significantly associated with AAM in women of the European ancestry (He et al., 2009; Ong et al., 2009; Perry et al., 2009; Sulem et al., 2009; He et al., 2010), only two were later confirmed in the sample of 15,000 Japanese females (Tanikawa et al., 2013).

There is evidence that candidate genes for menarche (e.g., *LIN28B, FTO, TNNI3K, MAP2K5, FANCL, STK33, GPRC5B, POMC/RBJ* and the others) may also be associated with various anthropometric characters (e.g., height, body mass index (BMI), and the others) and thus suggest a common genetic basis for these traits (Ong et al., 2009; Elks et al., 2010; Perry et al., 2014; Fernandez-Rhodes et al., 2013). Replication studies are important for understanding a role of specific candidate genes for AAM in populations of different ethnicity, history, and genetic structure. With reference to this, genetic factors for AAM in Russian females have not been studied so far.

This article presents the results of a replication analysis of the previously reported AAM-associated polymorphisms in Russian women. In addition, several SNPs, which demonstrated non-significant associations with AAM but were involved in the menarche-related pathways (e.g., vitamin D metabolism, polycystic ovary syndrome, etc.), were also analyzed for their association with AAM.

2. Material and methods

2.1. Study subjects

The participants were recruited through the Perinatal Centre of St. Joasaph Belgorod Regional Clinical Hospital during 2009–2013. The eligible participants were unrelated women of Russian descent (self-declared) born in Central Russia (Rudyh and Sirotina, 2015). The following exclusion criteria were adopted: non-Russian descent, a birthplace outside of Central Russia, malignant tumors of a small pelvis and breast, chronic severe diseases of the vital organs (heart, respiratory or renal failure), severe autoimmune diseases. The study was approved by the Regional Ethics Committee of Belgorod State University. All participants signed an informed consent prior to the enrolment to this study.

The information about AAM was obtained through the questionnaire. AAM was defined as an age (full years) of first menses. Each participant was asked a question: "Your age of first menses - _ full years?" Women with AAM ≥ 18 years (n = 7) or refusing to answer (n = 21) were excluded from the study. The anthropometric characteristics were collected by standard methods: height was measured to the nearest 0.1 cm using the portable stadiometer; weight was then measured in an upright position, to the nearest 0.1 kg, using a calibrated balance beam scale. All participants were clinically examined for presence of benign tumors and hyperplastic disorders of the reproductive organs in women (leiomyoma, endometriosis, and endometrial hyperplasia).

The study sample consisted in total of 1700 females, including 966 with various isolated or combined benign tumors and hyperplastic disorders of the reproductive organs and 734 otherwise healthy.

2.2. Blood sample collection and DNA handling

Blood (5 ml) was drawn by a certified nurse from the ulnar vein of each participant to a plastic vial (Vacutainer®) containing 0.5 M EDTA solution (pH = 8.0). Total genomic DNA was isolated from buffy coat using the standard phenol-chloroform method and then checked for quality using Nanodrop 2000 spectrophotometer (Thermo Scientific, Inc.). Only samples with A260/A280 = 1.7–2.0 were included in the analysis. The isolated DNA was stored at −80°C.

2.3. SNP selection

The 53 SNPs for the analysis were selected based on the following criteria (Ponomarenko, 2018): 1) Previously reported associations with AAM or traits having common biological pathways with menarche (anthropometric characteristics, obesity, vitamin D metabolism, etc.), 2) Regulatory potential (regSNP), 3) Effect on gene expression (eSNP), and 4) Tag value (tagSNP).

The regulatory potential of the SNPs and effect on gene expression was estimated using the online tools HaploReg (v4.1) (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php), RegulomeDB (Version 1.1) (http://regulome.stanford.edu/), rSNPBase (http://rsnp.psych.ac.cn/index.do), SNPinfo Web Server – SNP Function Prediction (FuncPred) (https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html), Blood eQTL browser (http://genenetwork.nl/bloodeqtlbrowser/), and GTExportal (http://www.gtexportal.org/).

The tagSN*P* values were determined using the SNPinfo Web Server of the National Institute of Environmental Health Sciences (NIEHS) – LD TAG SNP Selection (TagSNP) (https://snpinfo.niehs.nih.gov/snpinfo/snptag.html) and the data from the International HapMap (phase III) and dbSNP. The LD between SNP pairs was estimated assuming the maximum between-SNP distance of 250 kb and $r^2 \geq 0.8$ (± 10–20 kb and ± 500–700 kb at 5`- and 3`-end of a gene, respectively).

The information about the selected SNPs is given in Supplementary Table 1. All SNPs appear to have a significant regulatory potential (Supplementary Table 2), 44 of them (83.02%) are eSNPs (Supplementary Table 3), 30 are tagSNPs, 17 are associated with various anthropometric characteristics (Supplementary Table 4).

Out of the 53 selected SNPs, 14 were associated with AAM according to the GWAS results (He et al., 2009; Ong et al., 2009; Perry et al., 2009; Sulem et al., 2009; Elks et al., 2010; Demerath et al., 2013; Perry et al., 2014) and 28 – according to the candidate gene association studies (Delahanty et al., 2013; Pyun et al., 2014; He et al., 2010; Tanikawa et al., 2013; Fernandez-Rhodes et al., 2013; Stavrou et al., 2006) (Supplementary Table 4). In addition, 11 SNPs, which did not demonstrated significant association with AAM, but were associated or tag with the traits related to the menarche (e.g., vitamin D metabolism, polycystic ovary syndrome development, anthropometric characters, etc., Supplementary Table 4). These SNPs included rs1884051 *ESR1*, rs3020394 *ESR1*, rs12324955 *FTO*, rs4633 *COMT*, rs11724758 *FABP2*, rs222020 *GC*, rs222003 *GC*, rs1544410 *VDR*, rs3756261 *EGF*, rs7766109 *F13A1*, rs2252673 *INSR*. All these SNPs have a significant regulatory potential, ten of them are eSNPs and nine are tagSNPs.

2.4. SNP genotyping

DNA samples were genotyped using the Sequenom MassARRAY® iPLEX platform, which is based on MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometry at the Centre of Genomic Sciences, the University of Hong Kong. The analyzed DNA samples had concentration 10–15 ng/ml. The MassARRAY® assay

design software Assay Design Suite 1.0 (http://agenabio.com/assay-design-suite-10-software) was used to create a single well iPLEX SNP genotyping assay. For this purpose, 53 SNPs of interest were retrieved from dbSNP of NCBI and imported according to their IDs to Assay Design Suite 1.0. After completing the consecutive automatic steps, the genotyping assay was successfully generated and tested for cross amplification.

2.5. Data quality control

Quality of the genotypic data was assessed according to the missing call rate defined as a fraction of missing calls per SNP for all samples. The MassARRAY® Typer 4.0 software was used for cluster analysis of the genotype calls. All samples successfully passed the quality control with the following parameters: SNPs with call rate > 80%, the success rate of duplicate check > 99.5%, and the success rate of the blank check > 90%. One SNP, rs11724758, did not pass the quality control (call rate < 80%) and was excluded from the further analysis. The individuals with the proportion of determined genotypes < 95% out of the maximum possible number were also excluded from the analysis (n = 87).

The final sample used in the analyses included 1613 female participants (917 women with benign tumors and hyperplastic disorders of the reproductive organs and 696 otherwise healthy women). The proportion of the determined genotypes for the 52 SNPs was 98.86%. The general characteristics of the study participants are given in Table 1.

2.6. Statistical analysis

Differences in means of AAM between the analyzed groups were estimated by the Kruskal–Wallis test. All polymorphisms were checked for their correspondence to the Hardy-Weinberg equilibrium (HWE) using the chi-square test. Since the distribution of AAM, height, and BMI values in the sample was not normal (according to the Shapiro-Wilk test), they were transformed using the QQ-plot function in the R programming environment.

Association between a SNP minor allele and AAM was analyzed using log-linear regression assuming the three principal models (additive, recessive, and dominant). The year of birth, presence of isolated or combined benign tumors and hyperplastic disorders of the reproductive organs in women were transformed into discrete variables and used as covariates (four and eight classes, respectively, Table 2). The correction for multiple comparisons was made using the adaptive permutation test. The significance level was set at $p_{perm} < 0.01$. For the given sample size (n = 1613), the expected detectable differences in AAM were 0.13–0.29 year under the additive model, 0.21–0.30 year under the dominant model, and 0.21–1.80 year under the recessive model (at 80% power, α = 0.05 for 2-sided test). The haplotype blocks were determined using the 'confidence intervals' algorithm at D' > 0.8 as implemented in Haploview v.4.2 (https://www.broadinstitute.org/haploview/haploview). Statistical power for each SNP was computed using Quanto 1.2.4 (http://hydra.usc.edu/gxe 2009). The association analyses were conducted using the PLINK v. 2.050 software (available at: http://zzz.bwh.harvard.edu/plink/).

Table 1
Anthropometric characteristics of the participants.

Parameters	$\bar{X}$ ± SD(min-max)/%(n)
n	1613
Age, yrs	39.53 ± 8.98 (18–77)
Age at menarche, yrs	13.32 ± 1.28 (9–17)
Proportion of the participants by relative age at menarche, % (n)	
Early (< 12 yrs)	6.32 (102)
Average (12–14 yrs)	79.85 (1288)
Late (> 14 yrs)	13.83 (223)
Height, m	1.65 ± 0.06 (1.48–1.99)
Weight, kg	72.33 ± 14.01 (42–132)
BMI, kg/m^2	26.53 ± 5.03 (15.22–48.91)
Proportion of the participants by relative BMI, % (n)	
Underweight (< 18.50)	2.23 (36)
Normal weight (18.50–24.99)	42.96 (693)
Overweight (25.00–29.99)	31.06 (501)
Obese (> 30.00)	23.75 (383)

Table 2
Characteristics of age at menarche in the participants.

Parameters	n	Age at menarche, yrs ($\bar{X}$ ± SD)	P*
Year of birth			
1930–1960	301	13.69 ± 1.38	< 0.001
1961–1970	639	13.41 ± 1.18	
1971–1980	505	13.11 ± 1.26	
1981–1990	168	12.91 ± 1.34	
Body mass index			
Underweight (< 18.50)	36	13.33 ± 1.39	0.99
Normal weight (18.50–24.99)	693	13.33 ± 1.30	
Overweight (25.00–29.99)	501	13.31 ± 1.25	
Obesity (> 30.00)	383	13.31 ± 1.30	
Presence of benign tumors and hyperplastic disorders of the reproductive organs			
Isolated uterine leiomyoma	192	13.55 ± 1.39	0.04
Isolated endometriosis	103	13.07 ± 1.33	
Isolated endometrial hyperplasia	167	13.22 ± 1.31	
Combined uterine leiomyoma and endometriosis	109	13.36 ± 1.26	
Combined uterine leiomyoma and endometrial hyperplasia	170	13.42 ± 1.30	
Combined endometriosis and endometrial hyperplasia	85	13.35 ± 1.26	
Combined uterine leiomyoma, endometriosis and endometrial hyperplasia	98	13.39 ± 1.23	
No benign tumors and hyperplastic disorders of the reproductive organs	696	13.26 ± 1.25	

P, level of significance according to the Kruskal-Wallis test.

The gene-gene interactions were analyzed for the two-, three-, and four-locus models using MB-MDR (Model Based Multifactor Dimensionality Reduction) (Calle et al., 2008; Calle et al., 2010) as implemented in the namesake software (v. 2.6) for the R programming environment. MB-MDR is a modification of MDR and makes it possible to analyze gene-gene interactions with adjustment for covariates and validation by the permutation test with 1000 permutations (Mahachie et al., 2012). The significance level was set at $p_{perm} < 0.01$. The gene-gene interactions and their relative contribution to the total variance of the trait within the 2-, 3-, and 4-locus models were visualized using the MDR method (http://www.multifactordimensionalityreduction.org/), as implemented in MDR v. 3.0.2 (http://sourceforge.net/projects/mdr).

2.6.1. Analysis of association of the AAM candidate genes with body size

Since the candidate genes for AAM were previously suggested to contribute to height and weight characteristics of females (Ong et al., 2009, Elks et al., 2010, Perry et al., 2014, Fernandez-Rhodes et al., 2013 and others), we analyzed all studied SNPs for such association. The BMI association analyses were conducted using age and hyperplastic disorders of the reproductive organs in women (see above) as covariates.

2.6.2. Functional SNPs

The AAM-associated and strongly linked to them SNPs were analyzed for their functional significance (non-synonymous SNPs, regulatory potential, and eQTLs). The SNPs in strong linkage disequilibrium (LD) ($r^2 \geq 0.8$) with the AAM associated ones were determined using the online version of HaploReg (v4.1). The linkage

disequilibrium was estimated using the data of the European population from the 1000 Genomes Project Phase.

Non-synonymous SNPs and their predictive potential were analyzed using SIFT (http://sift.jcvi.org/).

2.6.3. Regulatory effects

The in silico analysis of the regulatory potential of the candidate SNPs for AAM was conducted using HaploReg (v4.1) (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php), RegulomeDB (Version 1.1) (http://regulome.stanford.edu/), rSNPBase (http://rsnp.psych.ac.cn/index.do), and SNP Function Prediction (FuncPred) (https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html). The possible regulatory effects of polymorphisms in a strong linkage disequilibrium ($r^2 \geq 0.8$) with the AAM associated SNPs were analyzed using HaploReg (v4.1).

2.6.4. Expression QTLs

The effect of the candidate SNPs for AAM on gene expression level (*cis*- and *trans*-eQTL) was estimated in peripheral blood using the data from the Blood eQTL browser (http://genenetwork.nl/bloodeqtlbrowser/), and in other organs and tissues using the GTExportal data (http://www.gtexportal.org/) as of 25.08.2017. To determine significant eQTLs, the False Discovery Rate (FDR) ≤0.05 was applied. Likewise, eQTL values of SNPs in the strong LD ($r^2 \geq 0.8$) with the AAM associated polymorphisms were estimated.

2.6.5. Pathway analyses

The functional significance of the candidate genes for AAM in the various biological pathways was studied using the Gene Ontology Portal tools available at http://geneontology.org. The results of multiple comparisons were adjusted with the Bonferroni test. The gene interaction networks were constructed using GeneMANIA (version 3.5.0) available at http://genemania.org.

3. Results

3.1. Study participants characteristics

The mean AAM was 13.32 ± 1.28 years (Table 1) and depended on the year of birth and presence of benign tumors and hyperplastic disorders of the reproductive organs in women (Table 2). There was a clear secular trend in AAM ($p < 0.001$). The largest mean AAM difference (0.78 yr) was observed between the participants born in 1930–1960 and 1981–1990. The mean AAM was different also between women with and those without the various reproductive pathologies ($p = 0.04$).

3.2. SNP association analysis

The data about the studied SNPs are given in (Ponomarenko et al.,2018a). All polymorphisms had MAF > 5% and corresponded to the HWE ($p_{bonf} < 0.001$).

Polymorphism rs6438424 *3q13.32* was associated with AAM according to the additive ($\beta = 0.067 \pm 0.024$, $p = 0.005$, $p_{perm} = 0.005$) and dominant ($\beta = 0.102 \pm 0.037$, $p = 0.006$, $p_{perm} = 0.005$) models (Table 3). The individuals with genotype CC had menarche at age 13.49 ± 1.28 years that is 0.17 year later then in individuals with genotype AC (13.32 ± 1.28 years) and 0.33 year later than in individuals of the AA genotype (13.16 ± 1.29 years) ($p = 0.002$, Fig. 1).

The power of the additive model was 94.44% (differences in AAM per allele ≥0.16 year), and of the dominant one 83.85% (the AAM difference between genotypes AC/CC and AA was 0.21 year, $p = 0.004$). None of the haplotypes showed significant association with AAM (Ponomarenko et al., 2018a).

3.3. SNP × SNP interactions

In total 12 significant gene-gene interactions affecting AAM were determined for the 2-, 3-, and 4-locus models (Table 4). Out of the 14 SNPs involved in the models, rs7579411 *LHCGR* and rs1073768 *GHRH* contributed to the largest number of the models (seven and six, respectively). Polymorphisms rs6438424 *3q13.3* and rs7759938 *LIN28B* contribute to the models at all levels studied (2-, 3-, and 4-locus).

Three polymorphisms of the *LIN28B* gene (rs7759938, rs314276, rs4946651) contributed to the most significant gene-gene interactions affecting AAM (in total 6 of the 12 models). Genotypic combinations associated with AAM according to the 12 significant models are given in (Ponomarenko et al., 2018a). The most significant association were determined for the following combinations: rs2252673 GG *INSR* × rs4946651 GG *LIN28B* (*beta* = −1.029, $p = 0.00007$), rs1073768 GA *GHRH* × rs12617311 GA *PLCL1* × rs4374421 TT *LHCGR* (*beta* = −0.324, $p = 0.00003$), rs6438424 AA *3q13.32* × rs1073768 GA *GHRH* × rs4633 TT *COMT* × rs7579411 CC *LHCGR* (*beta* = −1.116, $p = 0.00003$).

The gene-gene interactions among 14 SNPs, which are included in the best gene-gene interaction models associated with AAM, are visualized in Fig. 2. Despite the large number of the existing interactions, the contribution of any of them to the trait was quite modest. The largest value did not exceed 0.39%.

3.4. Association of the candidate genes for AAM with body size

Anthropometric characteristics of the participants are given in (Ponomarenko et al., 2018a). The older individuals had a higher BMI ($p < 0.001$). The participants with isolated endometriosis, isolated endometrial hyperplasia, and combined endometriosis and endometrial hyperplasia had the lowest BMI (25.15–25.88), while those with isolated uterine leiomyoma, combined uterine leiomyoma, endometriosis and endometrial hyperplasia had the highest BMI (27.95–28.10) ($p < 0.001$). The mean height in all age groups was not significantly different ($p > 0.09$).

Fifteen polymorphisms were associated with BMI within 11 significant multilocus models of gene-gene interactions (Ponomarenko et al., 2018a). The specific genotype combinations are given in (Ponomarenko et al., 2018a). Neither SNPs nor haplotypes were individually associated with BMI ($p_{perm} > 0.01$).

Sixteen polymorphisms were associated with height within 13 significant multilocus models of gene-gene interactions (Ponomarenko et al., 2018a). None of the SNPs or haplotypes was individually associated with height ($p_{perm} > 0.01$).

In summary, among the 14 AAM-associated polymorphisms, six SNPs were associated with height and six – with BMI (Supplementary Table 5). Three SNPs (rs12617311 *PLCL1,* rs1073768 *GHRH* и rs4633 *COMT*) were significantly associated with all three traits (AAM, BMI, and height).

3.5. Functional SNP

3.5.1. Non-synonymous SNPs

None of the AAM-associated SNPs was replacement. However, one of these SNPs, rs4633, was linked ($r^2 = 0.99$) to rs4680, which results in a Val/Met replacement in the COMT protein. This amino acid change has SIFT Score = 0.02 that corresponds to the predictive value “deleterious” (i.e., ≤0.05).

3.5.2. Regulatory effects

The results of the in silico analysis of the regulatory potential according to the five databases are given in Supplementary Table 2. The most pronounced effects were determined for rs4633 *COMT*; the significant regulatory potential was suggested for rs2164808 *POMC*, rs2252673*INSR*, rs6438424 *3q13.32* and rs10769908 *STK33*.

Table 3
Associations of the 52 SNPs with age at menarche.

Chr	SNP	n	Additive model			Dominant model			Recessive model		
			β	SE	P	β	SE	P	β	SE	P
1	rs1514175	1607	−0.011	0.024	0.638	0.001	0.034	0.980	−0.044	0.046	0.346
1	rs466639	1608	0.030	0.036	0.402	0.053	0.040	0.191	−0.175	0.136	0.198
1	rs7538038	1606	−0.030	0.029	0.314	−0.044	0.034	0.198	0.024	0.085	0.781
2	rs713586	1605	0.012	0.024	0.629	0.005	0.036	0.885	0.030	0.043	0.483
2	rs2164808	1609	0.012	0.024	0.616	0.011	0.038	0.774	0.022	0.041	0.590
2	rs7589318	1604	0.041	0.026	0.116	0.053	0.034	0.112	0.046	0.059	0.438
2	rs4374421	1548	0.019	0.026	0.473	0.034	0.035	0.321	−0.003	0.056	0.956
2	rs7579411	1593	0.032	0.024	0.176	0.015	0.036	0.685	0.081	0.042	0.055
2	rs6729809	1541	0.022	0.026	0.398	0.024	0.035	0.489	0.040	0.057	0.486
2	rs4953616	1598	0.021	0.027	0.431	0.038	0.034	0.262	−0.017	0.064	0.796
2	rs6732220	1604	−0.026	0.028	0.341	−0.028	0.034	0.413	−0.052	0.071	0.465
2	rs4953655	1606	−0.032	0.028	0.255	−0.037	0.034	0.276	−0.047	0.072	0.517
2	rs887912	1529	−0.003	0.029	0.903	0.016	0.035	0.641	−0.098	0.075	0.189
2	rs12617311	1605	0.008	0.025	0.758	−0.009	0.034	0.781	0.060	0.054	0.266
3	rs6438424	1594	0.067	0.024	0.005	0.102	0.037	0.006	0.076	0.040	0.058
4	rs2013573	1607	0.015	0.032	0.634	0.006	0.036	0.856	0.126	0.114	0.269
4	rs13111134	1608	0.020	0.029	0.488	0.022	0.034	0.521	0.035	0.083	0.675
4	rs222003	1609	0.022	0.047	0.640	0.026	0.048	0.585	−0.120	0.302	0.692
4	rs222020	1609	0.011	0.037	0.755	0.017	0.040	0.665	−0.058	0.156	0.708
4	rs3756261	1597	0.027	0.046	0.553	0.030	0.047	0.533	−0.011	0.276	0.968
5	rs757647	1591	0.041	0.028	0.149	0.046	0.034	0.180	0.066	0.076	0.380
6	rs7766109	1608	0.031	0.024	0.203	0.026	0.038	0.493	0.058	0.041	0.156
6	rs4946651	1610	0.017	0.024	0.476	0.013	0.035	0.703	0.038	0.045	0.401
6	rs7759938	1607	0.032	0.027	0.239	0.042	0.034	0.212	0.029	0.065	0.658
6	rs314280	1588	0.008	0.025	0.758	0.007	0.035	0.839	0.015	0.047	0.754
6	rs314276	1560	0.039	0.026	0.132	0.051	0.034	0.134	0.047	0.059	0.431
6	rs3020394	1608	0.047	0.026	0.067	0.036	0.034	0.287	0.133	0.058	0.022
6	rs1884051	1610	0.039	0.026	0.138	0.035	0.033	0.301	0.093	0.059	0.119
6	rs7753051	1607	−0.007	0.026	0.805	−0.002	0.034	0.950	−0.029	0.062	0.643
7	rs1079866	1608	−0.002	0.031	0.949	0.000	0.036	0.995	−0.021	0.097	0.829
8	rs2288696	1606	−0.031	0.031	0.309	−0.044	0.035	0.211	0.028	0.100	0.776
9	rs2090409	1516	−0.018	0.026	0.493	−0.013	0.036	0.722	−0.044	0.051	0.393
9	rs10980926	1601	0.008	0.026	0.743	0.021	0.034	0.533	−0.020	0.058	0.733
9	rs10441737	1558	0.001	0.026	0.967	0.008	0.034	0.807	−0.019	0.058	0.746
11	rs10769908	1585	0.007	0.024	0.768	0.009	0.038	0.808	0.010	0.041	0.810
11	rs555621	1607	0.054	0.025	0.030	0.074	0.035	0.036	0.062	0.047	0.188
11	rs11031010	1588	0.027	0.037	0.463	0.035	0.042	0.395	−0.016	0.139	0.905
11	rs1782507	1602	−0.028	0.025	0.272	−0.034	0.034	0.320	−0.038	0.051	0.454
11	rs6589964	1609	−0.020	0.024	0.389	−0.039	0.038	0.299	−0.014	0.040	0.721
12	rs1544410	1601	−0.009	0.024	0.721	−0.014	0.034	0.686	−0.007	0.047	0.886
14	rs999460	1607	0.020	0.025	0.431	0.039	0.034	0.247	−0.008	0.053	0.877
14	rs4986938	1608	0.021	0.025	0.402	0.058	0.034	0.087	−0.045	0.052	0.385
15	rs2241423	1600	0.006	0.031	0.839	0.033	0.036	0.354	−0.181	0.096	0.061
16	rs12444979	1601	−0.031	0.034	0.354	−0.029	0.038	0.451	−0.107	0.120	0.372
16	rs9939609	1606	0.018	0.024	0.456	0.006	0.036	0.872	0.049	0.043	0.251
16	rs12324955	1607	−0.005	0.026	0.862	−0.007	0.034	0.833	−0.001	0.062	0.984
18	rs1398217	1596	−0.038	0.025	0.131	−0.054	0.036	0.136	−0.040	0.045	0.374
19	rs2252673	1602	−0.048	0.030	0.107	−0.048	0.035	0.167	−0.109	0.086	0.207
20	rs1073768	1607	0.012	0.024	0.620	0.013	0.038	0.728	0.019	0.041	0.644
22	rs4633	1609	0.025	0.024	0.281	0.031	0.038	0.422	0.037	0.039	0.337
23	rs5930973	1589	0.043	0.050	0.390	NA	NA	NA	NA	NA	NA
23	rs3092921	1608	0.011	0.044	0.803	NA	NA	NA	NA	NA	NA

The results were obtained by the linear regression analysis with adjustment for covariates (year of birth, benign tumors and hyperplastic disorders of the reproductive organs in women).
β, coefficient of the linear regression (change of the transformed age at menarche per minor allele).
SE, standard error.

In addition, we analyzed 329 polymorphisms, which were linked ($r^2 \geq 0.8$) to the 14 AAM-associated SNPs (Ponomarenko et al., 2018b). Among those, three SNPs (one nonsynonymous and two synonymous) were located in exons, one in 5′-UTR, 156 in introns, and 183 in intergenic regions. Eighteen SNPs were located in evolutionarily conserved regions.

Several SNPs of these appeared to have significant regulatory potential (Ponomarenko et al., 2018b). For example, rs7766336 (linked to rs7759938 and rs314276 of *LIN28B*) has a promoter histone mark in 19 tissues, is located in a DNase-1 hypersensitive region in 13 tissues, a binding region for four regulatory, and a region of 19 regulatory motifs. Polymorphism rs5878829 linked to the above mentioned SNPs ($r^2 = 0.98$ и $r^2 = 0.95$, respectively) is located in the region of 30 regulatory motifs.

The AAM-associated rs4946651 *LIN28B* is strongly linked to rs11754600 ($r^2 = 0.95$), which has a histone mark in 24 tissues and is located in a DNase-1 hypersensitive region in 13 tissues, a binding region for two regulatory proteins, and a region of 13 regulatory motifs (Ponomarenko et al., 2018b).

The significant regulatory potential was also detected for several SNPs strongly linked ($r^2 \geq 0.8$) to the AAM-associated SNPs (Ponomarenko et al., 2018b). These SNPs manifested their regulatory effects in organs and tissues, which are pathogenetically important for menarche: brain, ovaries, adipose tissue, liver, fetal adrenals, and others.

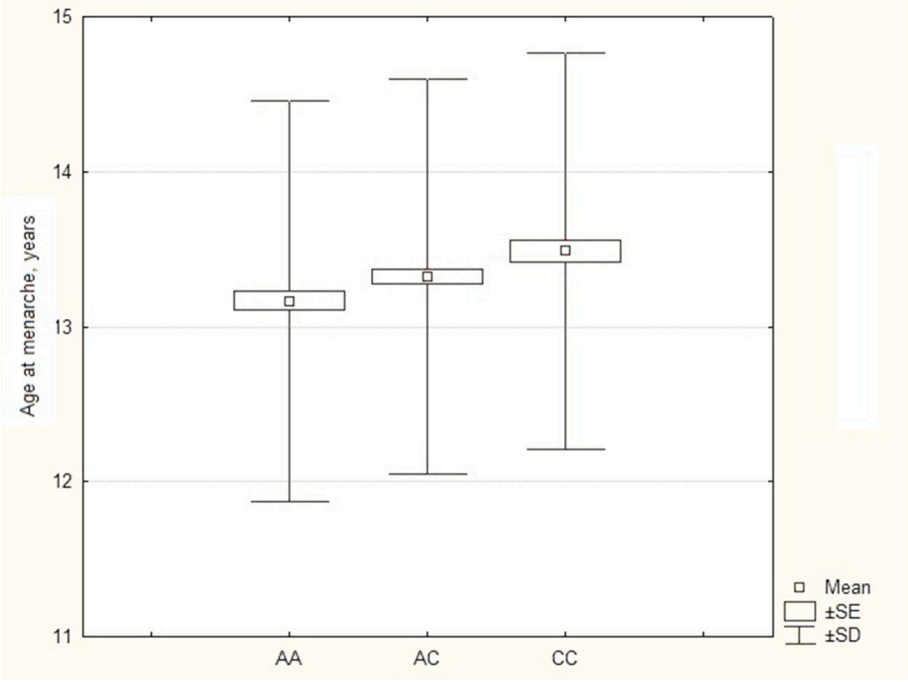

Fig. 1. Age at menarche depending on the genotypes of the rs6438424 *3q13.32* locus, years.

Table 4
The most significant models of SNP-SNP interactions associated with AAM.

N	Models of SNP × SNP interactions	NH	*beta*H	WH	NL	*beta*L	WL	Pperm
Two-locus models ($p < 1*10^{-4}$)								
1	rs314276 *LIN28B* × rs7579411 *LHCGR*	2	0.345	15.94	1	−0.140	4.24	< 0.001
2	rs6438424 *3q13.32* × rs7579411 *LHCGR*	2	0.245	13.83	2	−0.245	16.80	0.004
3	rs7759938 *LIN28B* × rs7579411 *LHCGR*	2	0.340	15.31	1	−0.115	3.02	0.004
4	rs2252673 *INSR* × rs4946651 *LIN28B*	1	0.146	3.14	1	−1.029	15.92	0.006
Three-locus models ($p < 1*10^{-7}$)								
1	rs6438424 *3q13.32* × rs1073768 *GHRH* × rs4374421 *LHCGR*	5	0.318	30.24	3	−0.269	22.31	< 0.001
2	rs6438424 *3q13.32* × rs1073768 *GHRH* × rs7579411 *LHCGR*	5	0.375	28.74	2	−0.403	17.68	< 0.001
3	rs1073768 *GHRH* × rs12617311 *PLCL1* × rs4374421 *LHCGR*	2	0.203	5.76	4	−0.365	30.52	< 0.001
4	rs12617311 *PLCL1* × rs7759938 *LIN28B* × rs7579411 *LHCGR*	5	0.367	30.27	2	−0.189	7.03	0.001
Four-locus models ($p < 1*10^{-12}$)								
1	rs2090409 *TMEM38B* × rs1073768 *GHRH* × rs2164808 *POMC* × rs4633 *COMT*	5	0.458	22.20	9	−0.588	54.70	< 0.001
2	rs6589964 *BSX* × rs1073768 *GHRH* × rs10769908 *STK33* × rs7759938 *LIN28B*	11	0.550	56.90	5	−0.540	23.33	< 0.001
3	rs314276 *LIN28B* x rs12617311 *PLCL1* × rs2164808 *POMC* × rs7579411 *LHCGR*	10	0.579	52.67	5	−0.335	18.28	< 0.001
4	rs6438424 *3q13.32* × rs1073768 *GHRH* × rs4633 *COMT* × rs7579411 *LHCGR*	9	0.525	42.91	10	−0.494	57.15	< 0.001

The results were obtained using the MB-MDR method with adjustment for covariates (year of birth, benign tumors and hyperplastic disorders of the reproductive organs in women); NH, number of significant genotypic combinations associated with the late menarche; *beta* H, coefficient of the linear regression for significant genotypic combinations associated with the late menarche (change of the transformed age at menarche); WH, the Wald test value for significant genotypic combinations associated with the late menarche; NL, number of significant genotypic combinations associated with the early menarche; *beta* L, coefficient of the linear regression for significant genotypic combinations associated with the early menarche (change of the transformed age at menarche); WL, the Wald test value for significant genotypic combinations associated with the early menarche; P_{perm}, P value for the permutation test (1000 permutations).

3.5.3. Expression QTLs

Three AAM-associated loci (rs2164808, rs1073768, and rs4633) were found to affect significantly ($p < 5 \times 10^{-5}$, FDR < 0.05) the mRNA expression level of some genes (*DNAJC27*, *MANBAL,* and *COMT*) in peripheral blood *(cis*-eQTL) (Ponomarenko et al., 2018b). In addition, two of these SNPs are in strong LD with the other *cis*-eQTL SNPs affecting mRNA expression in peripheral blood: rs2164808 is linked to rs4665765 ($r^2 = 0.99$), which affects expression of *DNAJC27* ($p = 4.5 \times 10^{-5}$, FDR = 0.02), and rs4633 is linked to rs4680 ($r^2 = 0.99$) which affects expression of *COMT* ($p = 1.9 \times 10^{-5}$, FDR = 0.01). No *trans*-eQTL SNPs were determined (FDR > 0.05).

According to the data of the Genotype-Tissue Expression (GTEx) Project, 9 out of the 14 AAM-associated SNPs are *cis*-eQTLs, i.e., significantly associated ($p < 8 \times 10^{-5}$, FDR ≤ 0.05) with the expression level of 12 genes in various tissues and organs (Ponomarenko et al., 2018b).

The most notable associations were determined for allele C of rs6438424 (alt), which appears to increase expression of *RP11-384F7.2* in the visceral fat ($\beta = 0.41$, $p = 3.5 \times 10^{-6}$, FDR ≤ 0.05) and the adrenal glands ($\beta = 0.70$, $p = 1.5 \times 10^{-11}$, FDR ≤ 0.05), and *LSAMP* ($\beta = 0.74$, $p = 6.3 \times 10^{-13}$, FDR ≤ 0.05) in the adrenal glands. Polymorphism rs10769908 had the highest *cis*-eQTL value and was associated with the expression level of *STK33* in five tissues and *TRIM66* in three tissues (Ponomarenko et al., 2018b).

Eight of the 14 AAM-associated SNPs were in strong LD ($r^2 \geq 0.8$) with > 200 SNPs significantly ($p < 8.5 \times 10^{-5}$, FDR ≤ 0.05) associated with the mRNA expression in various tissues (Ponomarenko et al., 2018b). Specifically, rs10769908 is linked ($r^2 \geq 0.8$) to over 120

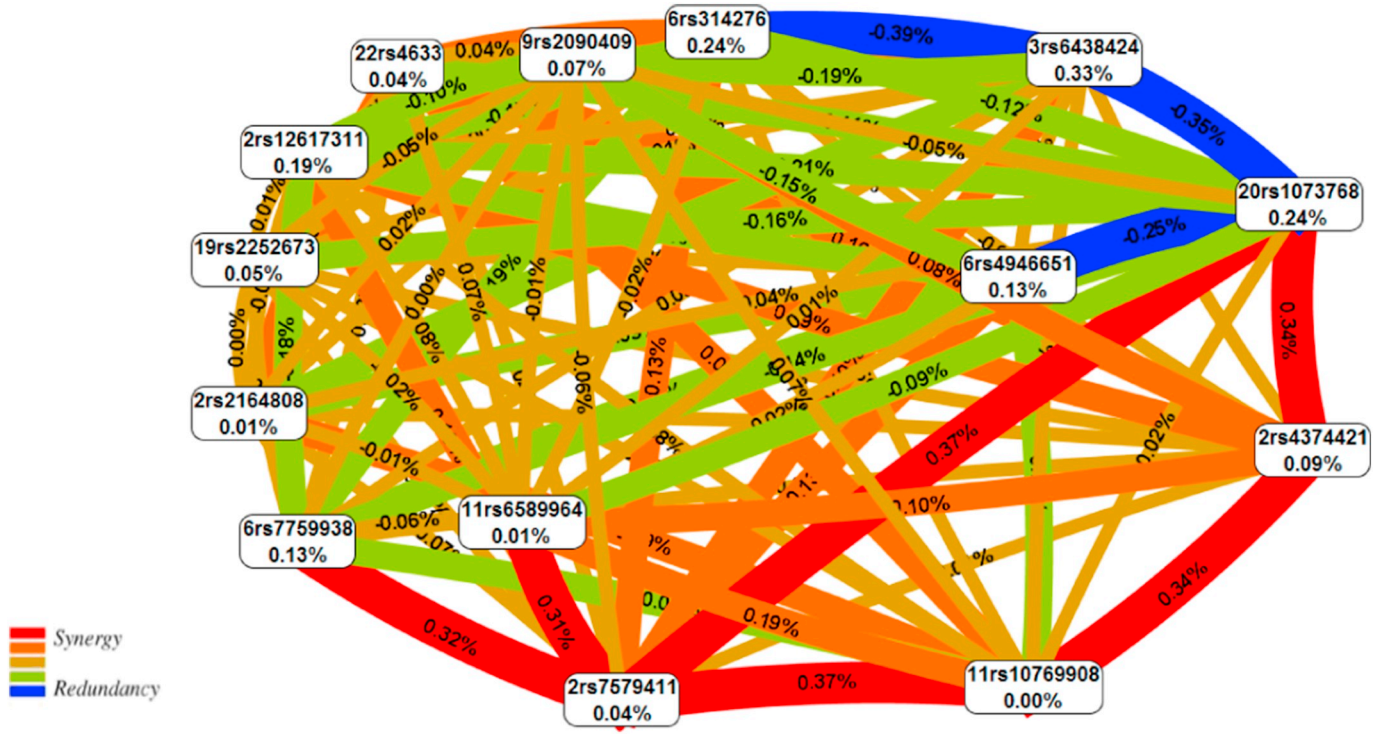

Fig. 2. The interaction entropy graph of SNP-SNP interactions with AAM based on the multifactor dimensionality reduction (MDR) analysis. Positive values of entropy indicate synergistic interactions while the negative values indicate redundancy. The polymorphisms are denoted by the chromosome number and rs SNP ID. The red and orange colors denote strong and moderate synergism, respectively, brown color denotes the independent effect, green and blue colors denote moderate and strong antagonism. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

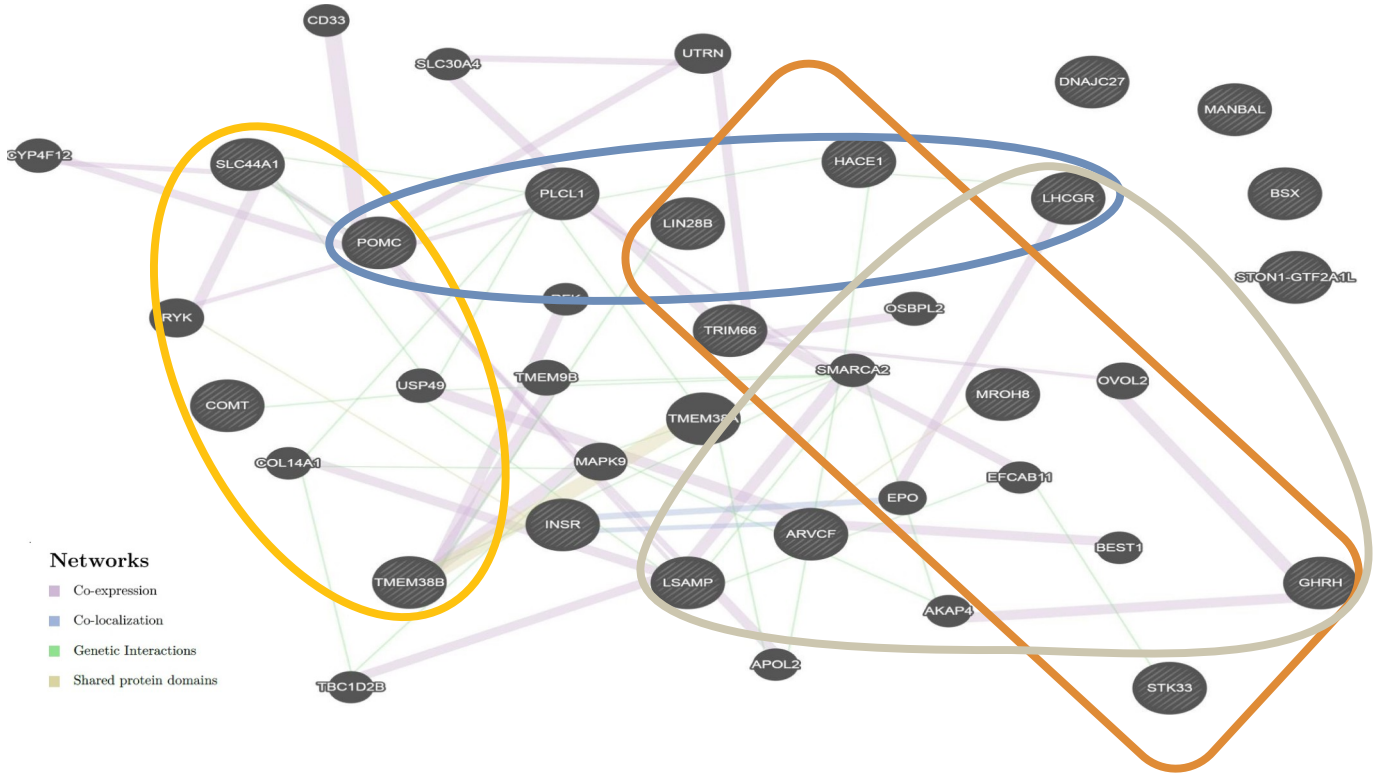

Fig. 3. The interaction networks of the candidate genes for age at menarche inferred using GeneMANIA (http://genemania.org). The candidate genes determined in the present study are cross-shaded; the colored figures outline the SNP × SNP interactions within the 4-locus models obtained by the MB-MDR method.

SNPs associated with the transcription level of the *STK33* and *TRIM66* genes in various tissues.

In addition, rs6438424 is in strong LD ($r^2 \geq 0.8$) with 8 SNPs, which are associated with the expression levels of *RP11-384F7.2* (visceral fat and adrenals) and *LSAMP* (adrenals), rs4946651 – with 14 SNPs affecting expression of *LINC00577* in the brain, rs7759938 and rs314276 – with 12 SNPs affecting expression of *HACE1* in EBV-transformed lymphocytes, rs2090409 – with 25 SNPs that influence the *SLC44A1* expression in the colon (Ponomarenko et al., 2018b).

Overall, 11 of the 14 AAM-associated loci had the *cis*-eQTL value (affected expression of 14 genes): three SNPs were independently associated with the mRNA expression levels, seven SNPs were both independently associated with the mRNA expression levels and linked to other *cis*-eQTLs, one SNP is linked to polymorphisms with *cis*-eQTL value.

3.6. Pathway analyses

The in silico analysis of the functional significance was conducted for the 10 AAM-associated genes (*POMC, LHCGR, PLCL1, LIN28B, TMEM38B, STK33, BSX, INSR, GHRH, COMT*) and for 14 genes whose expression is affected by the AAM-associated SNPs according to the eQTL analysis (*DNAJC27, STON1-GTF2A1L, LSAMP, RP11-384F7.2, LINC00577, HACE1, TRIM66, CTB-133G6.2, MROH8, MANBAL, ARVCF, SLC44A1, STK33, COMT*). *STK33* and *COMT* are present in both lists. The information about six of these (*STON1-GTF2A1L, RP11-384F7.2, LINC00577, CTB-133G6.2*) was not available in the Gene Ontology databases.

According to the PANTHER protein class and PANTHER molecular function databases, neuropeptide coding genes (Fold Enrichment – 97.23, $p_{bonf} = 0.039$) and neuropeptide hormone activity (Fold Enrichment - > 100.00, $p_{bonf} = 0.032$), respectively, are over-represented in the above-mentioned list of genes.

Using GeneMANIA (http://genemania.org), a network of the gene-gene interactions between 15 AAM-associated genes and 20 other genes was inferred (Fig. 3). These gene-gene interactions are realized through co-expression (62.57%), co-localization (30.04%), genetic interactions (5.61%), and common protein domains (1.79%). The gene-gene interactions of the AAM-associated genes may be either direct or via intermediate genes (e.g., *USP49*, *SMARCA2*) (Ponomarenko et al., 2018b).

4. Discussion

This study reports 14 loci associated with AAM in women of Russia. Polymorphism rs6438424 located in region 3q13.32 demonstrated the most significant association; its polymorphic allele C was associated with later menarche according to the additive and dominant models, whereas reference allele A was associated with early menarche. The latter is in support of the previously reported results (Elks et al., 2010; Delahanty et al., 2013). Elks et al. (Elks et al., 2010) also suggested contribution of this SNP to height based on the GIANT consortium data. Oh et al. (Oh et al., 2016) reported association of rs6438424 with the breast tissue development.

According to the GTExportal database, polymorphism rs6438424 and eight polymorphisms linked to it have the *cis*-eQTL value and may affect expression of *RP11-384F7.2* and *LSAMP* in the adrenal glands and visceral fat tissue. The *RP11-384F7.2* gene belongs to the genes controlling synthesis of long non-coding RNAs (lncRNAs) (http://www.ensembl.org/), which contribute to numerous processes: histone modification, DNA methylation, chromatin remodeling, and the others, which are important for embryonic development, reproduction, and oncogenesis (Schmitz et al., 2016). *LSAMP* (*limbic system-associated membrane protein*) encodes a protein of the immunoglobulin LAMP, OBCAM and neurotrimin (IgLON) family. The LSAMP protein mediates selective growth of neurons and axon targeting, and contributes to the axon developing and remodeling of the limbic system. (http://www.genecards.org/).

According to the results, locus rs7579411 of *LHCGR* was present in the largest number of the most significant models of gene-gene interactions associated with AAM (seven out of twelve). Another *LHCGR* polymorphism, rs4374421, was involved in the two models. The association of these polymorphisms with AAM was previously reported by He et al. (He et al., 2010). In the present study, rs7579411 was also associated with height, and rs4374421 with BMI (involved in 10 models of gene-gene interactions). The GTExportal database suggests that rs7579411 may be associated with the expression of the *STON1-GTF2A1L* gene in the thyroid gland.

The *LHCGR* gene encodes a receptor for luteinizing hormone (LH) and gonadotropin (http://www.genecards.org/). LH produces multiple effects on the female organism during prepuberty and puberty: induces follicle development and androgen synthesis in the ovaries, initiates ovulation and menarche, and the others (Abreu and Kaiser, 2016; Plant, 2015).

Polymorphism rs1073768 of the *GHRH* gene was involved in the largest number of gene-gene interaction models (6 out of 12) associated with AAM and was associated with height and BMI. The HaploReg (v4.1) database suggests that rs1073768 has a significant regulatory potential: it contains an enhancer histone mark in various cell types. The significant association of rs1073768 with AAM was previously reported by He et al. (He et al., 2010). The *GHRH* (*growth hormone-releasing hormone*) gene controls release of the growth hormone, which has a profound effect in the pubertal age. The GHRH and GH levels are inversely related to BMI, accumulation of visceral fat, obesity, and metabolic disorders (Stanley and Grinspoon, 2015).

Three polymorphisms of the *LIN28B* gene (rs7759938, rs314276, rs4946651) were presented in 6 out of 12 most significant models of gene-gene interactions. All these polymorphisms were previously reported for their association with AAM: rs7759938 (Perry et al., 2009; Elks et al., 2010; Perry et al., 2014; Delahanty et al., 2013), rs314276 (Ong et al., 2009), and rs4946651 (He et al., 2009). There are data suggesting that rs7759938 and rs314276 may contribute to the pubertal development and height (Ong et al., 2009; Perry et al., 2014; Lango Allen et al., 2010; Cousminer et al., 2014). Ong et al. (Ong et al., 2011) reported association of rs314276 with BMI and weight in females, but not in males.

The GTExportal data suggest that the AAM-associated polymorphisms of *LIN28B* have important *cis*-eQTL values: rs4946651 and 14 linked to it are associated with transcription level of *LINC00577* in the basal ganglia, rs7759938 and rs314276 and 12 linked to them SNPs may affect expression of the *HACE1* gene in the EBV-transformed lymphocytes.

LIN28B (*lin-28 homolog B*) encodes a protein of the LIN28 family, which represses the *let-7* family of miRNA (Tsialikas and Romer-Seibert, 2015). There are data that the let-7 miRNAs are involved in cell cycle control and oncogenesis (Tsialikas and Romer-Seibert, 2015). The targets of the *let-7* miRNA, *Myc, Kras, Igf2bp1* and *Hmga2*, control body size and metabolism in mammals (Zhu et al., 2011).

This study firstly reports two SNPs, rs4633 *COMT* and rs2252673 *INSR*, as possible contributors to AAM (through gene-gene interactions). In addition, rs4633 manifested significant association with height and BMI (through gene-gene interactions). Three SNPs, rs4680, rs165656, and rs165722, which were in strong LD with rs4633, manifested significant regulatory effects: all influenced expression of *ARVCF* in the thyroid gland. The rs4633 and rs4680 polymorphisms are associated with the *COMT* expression level in peripheral blood. The rs4680 polymorphism encodes a replacement substitution Val/Met in COMT and according to the SIFT database has predictive value “deleterious” (SIFT Score ≤ 0.05): the Met variant has 40% lower activity as compared to the Val one (Chen et al., 2004). Moreover, rs4680 is associated with the level of some metabolites in blood (e.g., X-11593-*O*-methylascorbate, X-01911, (Shin et al., 2014)). COMT (catechol-*O*-methyltransferase) is an enzyme, which participates in the metabolism of catecholamines and estrogens (Sannino et al., 2017). Polymorphism

rs2252673 of *INSR*, which was associated with AAM in the present study, was previously associated with polycystic ovary syndrome (Goodarzi et al., 2011).

5. Conclusions

The rs6438424 polymorphism is independently associated with AAM in Russian females in this study. The other 14 SNPs manifest this association through gene-gene interactions. The extensive in silico analysis suggested that these associations might have more pronounced effect through about 300 SNPs linked to the above polymorphisms and implicated in multiple regulatory mechanisms in menarche-related organs and tissues.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2018.11.042.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Abreu, A.P., Kaiser, U.B., 2016. Pubertal development and regulation. Lancet Diabetes Endocrinol. 3, 254–264.
Calle, M.L., Urrea, V., Vellalta, G., Malats, N., Steen, K.V., 2008. Improving strategies for detecting genetic patterns of disease susceptibility in association studies. Stat. Med. 27, 6532.
Calle, M.L., Urrea, V., Malats, N., Van Steen, K., 2010. Mbmdr: an R package for exploring gene-gene interactions associated with binary or quantitative traits. Bioinformatics (17), 2198–2199.
Chen, J., Lipska, B.K., Halim, N., Ma, Q.D., Matsumoto, M., Melhem S Kolachana, B.S., Hyde, T.M., Herman, M.M., Apud, J., et al., 2004. Functional analysis of genetic variation in catechol-o-methyltransferase (COMT): effects on mRNA, protein enzyme activity in postmortem human brain. Am. J. Hum. Genet. (5), 807–821.
Cousminer, D.L., Stergiakouli, E., Berry, D.J., Ang, W., Groen-Blokhuis, M.M., Körner, A., Siitonen, N., Ntalla, I., Marinelli, M., Perry, J.R., et al., 2014. Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. Hum. Mol. Genet. (16), 4452–4464.
Delahanty, R.J., Beeghly-Fadiel, A., Long, J.R., Gao, Y.T., Lu, W., Xiang, Y.B., Zheng, Y., Ji, B.T., Wen, W.Q., Cai, Q.Y., et al., 2013. Evaluation of GWAS-identified genetic variants for age at menarche among Chinese women. Hum. Reprod. (4), 1135–1143.
Demerath, E.W., Liu, C.T., Franceschini, N., Chen, G., Palmer, J.R., Smith, E.N., Chen, C.T., Ambrosone, C.B., Arnold, A.M., Bandera, E.V., et al., 2013. Genome-wide association study of age at menarche in African-American women. Hum. Mol. Genet. 16, 3329–3346.
Elks, C.E., Perry, J.R.B., Sulem, P., Chasman, D.I., Franceschini, N., He, C., Lunetta, K.L., Visser, J.A., Byrne, E.M., Cousminer, D.L., et al., 2010. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nat. Genet. 12, 1077–1085.
Feng, Y., Hong, X., Wilker, E., Li, Z., Zhang, W., Jin, D., Liu, X., Zang, T., Xu, X., Xu, X., 2008. Effects of age at menarche, reproductive years, and menopause on metabolic risk factors for cardiovascular diseases. Atherosclerosis (2), 590–597.
Fernandez-Rhodes, L., Demerath, E.W., Cousminer, D.L., Tao, R., Dreyfus, J.G., Esko, T., Smith, A.V., Gudnason, V., Harris, T.B., Launer, L., et al., 2013. Association of adiposity genetic variants with menarche timing in 92,105 women of European descent. Am. J. Epidemiol. 3, 451–460.
Goodarzi, M.O., Louwers, Y.V., Taylor, K.D., Jones, M.R., Cui, J., Kwon, S., Chen, Y.D., Guo, X., Stolk, L., Uitterlinden, A.G., et al., 2011. Replication of association of a novel insulin receptor gene polymorphism with polycystic ovary syndrome. Fertil. Steril. (5), 1736–1741.e11.
Guo, X., Ji, C., 2011. Earlier menarche can be an indicator of more body fat: study of sexual development and waist circumference in Chinese girls. Biomed. Environ. Sci. 5, 451–458.
He, C., Kraft, P., Chen, C., Buring, J.E., Paré, G., Hankinson, S.E., Chanock, S.J., Ridker, P.M., David, J., Chasman, D.I., 2009. Genome-wide association studies identify novel loci associated with age at menarche and age at natural menopause. Nat. Genet. (6), 724–728.
He, C., Kraft, P., Buring, J.E., Chen, C., Hankison, S.E., Pare, G., Chanock, S., Ridker, P.M., Hunter, D.J., 2010. A large-scale candidate-gene association study of age at menarche and age at natural menopause. Hum. Genet. (5), 515–527.
Kaprio, J., Rimpela, A., Winter, T., Viken, R.J., Rimpela, M., Rose, R.J., 1995. Common genetic influences on BMI and age at menarche. Hum. Biol. 5, 739–753.
Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al., 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 7317, 832–838.
Mahachie, J.M., Cattaert, T., Van Lishout, F., Gusareva, E.S., Van Steen, K., 2012. Lower-order effects adjustment in quantitative traits model-based multifactor dimensionality reduction. PLoS One 7, e29594.
Nnoaham, K.E., Webster, P., Kumbang, J., Kennedy, S.H., Zondervan, K.T., 2012. Is early age at menarche a risk factor for endometriosis? A systematic review and meta-analysis of case-control studies. Fertil. Steril. (3), 702–712.e6.
Oh, H., Bodelon, C., Palakal, M., Chatterjee, N., Sherman, M.E., Linville, L., Geller, B.M., Vacek, P.M., Weaver, D.L., Chicoine, R.E., et al., 2016. Ages at menarche- and menopause-related genetic variants in relation to terminal duct lobular unit involution in normal breast tissue. Breast Cancer Res. Treat. (2), 341–350.
Ong, K.K., Elks, C.E., Li, S., Zhao, J.H., Luan, J., Andersen, B., Bingham, S.A., Brage, S., Smith, G.D., Ekelund, U., et al., 2009. Genetic variation in LIN28B is associated with the timing of puberty. Nat. Genet. (6), 729–733.
Ong, K.K., Elks, C.E., Wills, A.K., Wong, A., Wareham, N.J., Loos, R.J., Kuh, D., Hardy, R., 2011. Associations between the pubertal timing-related variant in LIN28B and BMI vary across the life course. J. Clin. Endocrinol. Metab. (1), E125–E129.
Parent, A.S., Teilmann, G.J., Juul, A., Skakkebaekn, N.E., Toppari, J., Bourguignon, J.P., 2003. The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. Endocr. Rev. (5), 668–693.
Perry, J.R., Stolk, L., Franceschini, N., Lunetta, K.L., Zhai, G., McArdle, P.F., Smith, A.V., Aspelund, T., Bandinelli, S., Boerwinkle, E., et al., 2009. Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. Nat. Genet. (6), 648–650.
Perry, J.R., Day, F., Elks, C.E., Sulem, P., Thompson, D.J., Ferreira, T., He, C., Chasman, D., Esko, T., Thorleifsson, G., et al., 2014. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 7520, 92–97.
Plant, T.M., 2015. Neuroendocrine control of the onset of puberty. Front. Neuroendocrinol. 38, 73–88.
Ponomarenko, I.V., 2018. Selection of polymorphic loci for association analysis in genetic-epidemiological studies. Res. Result. Med. Pharm. 2, 40–54 (In Russian).
Ponomarenko, I.V., Reshetnikov, E.A., Altuchova, O.B., Golovchenko, O.V., Polonikov, A.V., Sorokina, I.N., Yermachenko, A., VYa, Dvornyk, Churnosov, MI., 2018a. Dataset on statistical analysis of genetic polymorphisms association with age at menarche, height and body mass index in Russian women. Data in Brief.
Ponomarenko, I.V., Reshetnikov, E.A., Altuchova, O.B., Golovchenko, O.V., Polonikov, A.V., Sorokina, I.N., Yermachenko, A., VYa, Dvornyk, Churnosov, M.I., 2018b. The data about functional significance of the SNPs associated with age at menarche in Russian women. Data in Brief.
Pyun, J.A., Kim, S., Cho, N.H., Koh, I., Lee, J.Y., Shin, C., Kwack, K., 2014. Genome-wide association studies and epistasis analyses of candidate genes related to age at menarche and age at natural menopause in a Korean population. Menopause 5, 522–529.
Rudyh, N.A., Sirotina, S.S., 2015. Genetic interrelations of Russian and Ukrainian populations of Belgorod region. Res. Result Med. Pharm. 3, 72–79 (In Russian).
Sannino, S., Padula, M.C., Managò, F., Schaer, M., Schneider, M., Armando, M., Scariati, E., Sloan-Bena, F., Mereu, M., Pontillo, M., et al., 2017. Adolescence is the starting point of sex-dichotomous COMT genetic effects. Transl. Psychiatry(5), e1141.
Schmitz, S.U., Grote, P., Herrmann, B.G., 2016. Mechanisms of long noncoding RNA function in development and disease. Cell. Mol. Life Sci. (13), 2491–2509.
Shin, S.Y., Fauman, E.B., Petersen, A.K., Krumsiek, J., Santos, R., Huang, J., Arnold, M., Erte, I., Forgetta, V., Yang, T.P., et al., 2014. An atlas of genetic influences on human blood metabolites. Nat. Genet. (6), 543–550.
Stanley, T.L., Grinspoon, S.K., 2015. Effects of growth hormone releasing hormone on visceral fat,metabolic and cardiovascular indices in human studies. Growth Hormon. IGF Res. (2), 59–65.
Stavrou, I., Zois, C., Chatzikyriakidou, A., Georgiou, I., Tsatsoulis, A., 2006. Combined estrogen receptor α and estrogen receptor β genotypes influence the age of menarche. Hum. Reprod. (2), 554–557.
Sulem, P., Gudbjartsson, D.F., Rafnar, T., Holm, H., Olafsdottir, E.J., Olafsdottir, G.H., Jonsson, T., Alexandersen, P., Feenstra, B., Boyd, H.A., et al., 2009. Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. Nat. Genet. (6), 734–738.
Tanikawa, C., Okada, Y., Takahashi, A., Oda, K., Kamatani, N., Kubo, M., Nakamura, Y., Matsuda, K., 2013. Genome wide association study of age at menarche in the Japanese population. PLoS One(5), e63821.
Tsialikas, J., Romer-Seibert, J., 2015. LIN28: roles and regulation in development and beyond. Development (14), 2397–2404.
Wise, L.A., Laughlin-Tommaso, S.K., 2016. Epidemiology of uterine fibroids: from menarche to menopause. Clin. Obstet. Gynecol. (1), 2–24.
Yermachenko, A., Dvornyk, V., 2014. Nongenetic determinants of age at menarche: a systematic review. Biomed. Res. Int. 2014, 371583.
Zhu, H., Shyh-Chang, N., Segre, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G., et al., 2011. The Lin28/let-7 axis regulates glucose metabolism. Cell (1), 81–94.