Verification of the Chromosome Region 9q21 Association with Pelvic Organ Prolapse Using RegulomeDB Annotations

Maryam B. Khadzhieva,1 Dmitry S. Kolobkov,2,3 Svetlana V. Kamoeva,4 Anastasia V. Ivanova,4 Serikbay K. Abilev,1 and Lyubov E. Salnikova1,5

1Laboratory of Ecological Genetics, N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, 3 Gubkin Street, Moscow 117971, Russia
2The Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University (MSU), GSP-1, Leninskiye Gory, 1-73, Office 433, Moscow 119991, Russia
3Laboratory of Animal Genetics, N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, 3 Gubkin Street, Moscow 117971, Russia
4Department of Obstetrics and Gynecology, Pirogov Russian National Research Medical University, 1 Ostrovitianov Street, Moscow 117971, Russia
5Laboratory of Molecular Immunology, Federal Research Center of Pediatric Hematology, Oncology and Immunology Named after Dmitry Rogachev, the Russian Ministry of Health and Social Development, 1 Samora Machel Street, Moscow 117198, Russia

Correspondence should be addressed to Lyubov E. Salnikova; salnikovalyubov@gmail.com

Received 12 June 2015; Accepted 28 July 2015

1. Introduction

Pelvic organ prolapse (POP) is the dropping of the pelvic organs caused by weakness or damage to the normal support of the pelvic floor. Prevalence of a disease state (stage II prolapse or greater) varies by data source from 3% of parous women [1] to 19% of women with advanced disease requiring surgery [2]. As many as 50% of women may have some degree of POP. Clinical manifestations related to POP often become evident after menopause [3], though it is becoming a serious health problem for women of all ages with first symptoms that might be experienced before age 30 [4]. POP rarely occurs as a separate condition and often correlates with urinary and faecal incontinence, sexual dysfunction, psychological, and social disadaptation [5]. Risk factors, which predispose to POP, include pelvic floor injury (vaginal parity and pelvic floor trauma during childbirth), lifestyle and health conditions (mainly, menopause, body mass index (BMI), chronic cough, constipation, and heavy lifting), genetic-related conditions (ethnicity, connective tissue disorders, and...
familial character of prolapse) [6, 7]. Forty-three percent of the variation in susceptibility for POP may be explained by genetic effects [8].

The majority of multifactorial disorders are characterized by a large spectrum of genetic variations in disease modifying genes, whereas information about causative polymorphic genes is scarce. In contrast, the genetic studies of POP have been mainly focused on a limited number of causative genes, among them are the genes controlling the collagen and elastin synthesis and remodeling [9–12], extracellular matrix metabolism [9, 13–16], and hormone receptors [17–19]. There are also three genome-wide linkage studies of the same group of researchers that have determined chromosome 9q21 [20], six other loci [21], and chromosomes 10q24–26 and 17q25 [22] as the regions associated with a predisposition for advanced POP in European pedigrees. It is known that family-based genetic studies may be unsuccessful for complex traits in general population [23] and the results should be validated in independent studies. No candidate gene studies have been performed yet from these genetic regions.

Functional SNPs in a specified chromosome region can be chosen with a powerful tool, RegulomeDB, a database which provides functional annotations of SNPs in the human genome using data sets from ENCODE and other sources [24]. These annotations include data on chromatin structure, methylation, protein motifs, and binding. RegulomeDB presents a scoring system, with categories ranging from 1 to 7, where category 7 variants lack evidence of regulatory function, while category 1 variants are those “likely to affect binding and linked to expression of a gene target” [24]. Category 1 is further divided into 1a–1f subcategories. A variant scored as 1a has the highest confidence on functionality. To date, functional annotation data have been mostly used for a post hoc analysis of GWAS data [25–28] (Hong et al., 2014; Rajkumar et al., 2014; Rosenthal et al., 2014; Zia et al., 2015) while selection of putative functional SNPs may also be useful in future studies.

To verify GWAS findings on POP association with the 9q21 chromosome region [20] and six other loci (rs1455311 (4q21), rs1036819 (8q24), rs430794 (9q21), rs10868525 (Assembly GRCh37.p13, 9: 81163473–89618028 bp)) has shown an evidence for a predisposition to POP [20]. The top effect was found for rs4077632 and rs10868525 (Assembly GRCh37.p13, 9: 81163473–89618028 bp) which is associated with an increased risk for pelvic organ prolapse. This finding was replicated in a larger cohort of patients with POP (rs4077632, 9q21, [21]).

2. Materials and Methods

2.1. Subjects. The study was approved by the Ethics Committee of the Pirogov Russian National Research Medical University (university review board approval number 117 (April 16, 2012)) and was conducted according to the principles of the Declaration of Helsinki; all patients provided written informed consent.

Based on the POP-Q examination, POP patients were diagnosed with stages III-IV prolapse and controls without prolapse and prior history of prolapse surgery (stage 0).

Subjects were recruited from the clinical bases of the Department of Obstetrics and Gynecology of the Pirogov Russian National Research Medical University (RNRMU), Moscow, Russia (from December 2011 to September 2013). Exclusion criteria for the group under study included family history of pelvic organ prolapse, lack of informed consent, pregnancy, gynecologic malignancy, and hereditary (Marfan or Ehlers-Danlos syndromes) or acquired connective tissue disorders (rheumatoid arthritis, systemic lupus erythematosus, scleroderma, polymyositis, and dermatomyositis). We also carefully collected data on perineal trauma and episiotomy (surgical incision, which is made to prevent more extensive vaginal traumas during childbirth). These data were collected based on the copies of the hospital medical records and the results of physical examination (the scars). Women with perineal trauma and episiotomy were combined in the group “perineal trauma in childbirth,” since both conditions lead to pelvic muscle weakness, higher risk of third- and fourth-degree traumas, and pelvic floor disorders [29].

To compare the studied genotypes distribution in Russian population and HapMap populations, we included the following populations from the project [30]: Caucasian populations, Utah residents with Northern and Western European ancestry (CEU), and Toscans in Italy (TSI); Asian populations, Chinese in Metropolitan Denver, Colorado (CHD), and Japanese in Tokyo (JPT); African populations, Yoruba in Ibadan, Nigeria (YRI), and Luhya in Webuye, Kenya (LWK); and two other populations, Gujarati Indians in Houston, Texas (GIH), and Mexican ancestry in Los Angeles, California (MEX). SNP genotype data were downloaded from the HapMap project site (The International HapMap Consortium, 2003), all HapMap samples were divided by gender, and only female samples (CEU, n = 89; TSI, n = 44; CHD, n = 44; JPT, n = 45; YRI, n = 80; LWK, n = 45; GIH, n = 43; MEX, n = 43) were included in a principal component analysis.

2.2. SNPs Selection and Genotyping. The purpose of this study was to replicate the results of genome-wide linkage and genome-wide association studies for pelvic organ prolapse. In the linkage study of high-risk POP pedigrees, 9q21 region (between rs4077632 and rs10868525 (Assembly GRCh37.p13, 9: 81163473–89618028 bp)) has shown an evidence for a predisposition to POP [20]. The top effect was found for rs1139451. Some other intergenic markers (e.g., rs4077632) and possible genes candidates in the region (e.g., TLE1 and TLE4) have been mentioned. Additionally to the SNPs cited by the authors (rs4077632 and rs1139451), we chose four markers utilizing RegulomeDB functional annotations for the region 9q21.

To compare the studied genotypes distribution in Russian population and HapMap populations, we included the following populations from the project [30]: Caucasian populations, Utah residents with Northern and Western European ancestry (CEU), and Toscans in Italy (TSI); Asian populations, Chinese in Metropolitan Denver, Colorado (CHD), and Japanese in Tokyo (JPT); African populations, Yoruba in Ibadan, Nigeria (YRI), and Luhya in Webuye, Kenya (LWK); and two other populations, Gujarati Indians in Houston, Texas (GIH), and Mexican ancestry in Los Angeles, California (MEX). SNP genotype data were downloaded from the HapMap project site (The International HapMap Consortium, 2003), all HapMap samples were divided by gender, and only female samples (CEU, n = 89; TSI, n = 44; CHD, n = 44; JPT, n = 45; YRI, n = 80; LWK, n = 45; GIH, n = 43; MEX, n = 43) were included in a principal component analysis.
DNA was isolated from 200 μL of blood using gDNA purification kit Diatom DNA Prep 200 (Isogene laboratory, Moscow, Russia). The genotyping was performed with a PCR-CTPP (polymerase chain reaction with confronting two-pair primers) [31]. Amplification was carried out using an ABI thermal cycler with two external and two internal sequence-specific primers (Supplementary Table 1, in Supplementary Material available online at http://dx.doi.org/10.1155/2015/837904) and tubes PCR MasterMix (Isogene laboratory, Moscow, Russia). The PCR products were analyzed in ethidium bromide-stained 2% agarose gel. Blinded duplicates (10%) of randomly taken DNA samples for each SNP were genotyped once more with 100% concordance in genotype calling.

2.3. Statistical Analysis. Categorical demographic and clinical variables were assessed using Fisher exact test (two-tailed). Since continuous demographic and clinical variables did not assume a normal distribution, the Mann-Whitney U nonparametric test was used to compare such variables. Deviation from Hardy-Weinberg equilibrium was assessed by $\chi^2$ analysis. Logistic regression analysis implemented in SNPStats package [32] was performed to assess associations between studied SNPs and POP. In multivariate analysis, we adjusted adjusted for main covariates: age, BMI, vaginal parity, and vaginal trauma in childbirth. Genetic model was selected using the Akaike Information Criterion (AIC) value with the lowest AIC value considered corresponding to the best-fitting model for the fitted variant. For genotypes with minor allele frequencies <5% only dominant and additive genetic models were evaluated. A two-tailed $P$ value < 0.05 was considered statistically significant.

Pairwise linkage disequilibrium (LD) values were measured as Lewontin’s $D’$-values and estimated from genotype data using the expectation-maximization algorithm implemented in SNPStats software. The SNPStats software was also used to evaluate haplotype association with POP by individual haplotype $P$ values.

Gene-gene interaction were estimated using “SNPassoc” package [33] within R statistical software [34] performing log-likelihood ratio tests (LRTs).

WINPEPI test power calculator [35] was used to assess sample size and test power. Assuming that odds ratio (OR) $\geq 2.0$ is clinically meaningful [36], we established test power in the range of 70.30% to 96.06% (at two-sided significance level) for this effect size, sample size $n = 500$ (300 controls and 200 cases), and risk genotype frequencies in a range of 10%–50%.

Successful replication study presents a statistically significant association of the same genetic variant in the same direction [37]. Our study cannot be considered a classic replication study since prior results have been obtained in a genome-wide linkage analysis for a gene region. Nevertheless, our results provide verification of the association of pelvic organ prolapse with the 9q21 haplotype in Caucasians. It is accepted that replication studies do not require genome-wide multiple testing correction [38]. In our study, we also did not adjust for multiple comparisons since we tested robust statistical associations [39, 40]. The Bonferroni correction was applied when testing for SNP interaction.

To detect admixture, we investigated Russian and HapMap populations structure via the discriminant analysis of principal components (DAPC) [41]. It was evaluated using the adegenet package in R statistical software [42].

Interpopulation differences in allele frequencies were assessed with pairwise Weir and Cockerham’s Fst estimator (Fst), using package diversity [43].

3. Results

3.1. SNPs Characteristics. SNPs considered in the study are described in Table 1. Of the 12 studied SNPs, five SNPs are intergenic, while seven SNPs are located in the genes. Among these seven SNPs, five SNPs are related to the protein-coding genes, TLE1 (rs2777781), TLE4 (rs2807303), COL18A1 (rs2236479), ZFAT (rs1036819), and FRMD3 (rs12551710),

<table>
<thead>
<tr>
<th>rs</th>
<th>Chromosome region</th>
<th>Chromosome position assembly GRCh37.p13</th>
<th>Gene</th>
<th>Function</th>
<th>RegulomeDB score</th>
<th>Alleles</th>
<th>MAF (Allele)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1455311</td>
<td>4q21</td>
<td>79964587</td>
<td>LINC01088</td>
<td>Intron</td>
<td>5</td>
<td>C:T</td>
<td>0.217 (C)</td>
</tr>
<tr>
<td>rs1036819</td>
<td>8q24</td>
<td>13561945</td>
<td>ZFAT-ASI, ZFAT</td>
<td>NearGene-3</td>
<td>4</td>
<td>A:C</td>
<td>0.128 (C)</td>
</tr>
<tr>
<td>rs4077632</td>
<td>9q21</td>
<td>81163473</td>
<td>—</td>
<td>—</td>
<td>No data</td>
<td>A:G</td>
<td>0.310 (G)</td>
</tr>
<tr>
<td>rs2807303</td>
<td>9q21</td>
<td>82187095</td>
<td>TLE4</td>
<td>NearGene-5</td>
<td>2a</td>
<td>C:T</td>
<td>0.336 (T)</td>
</tr>
<tr>
<td>rs2777781</td>
<td>9q21</td>
<td>84215034</td>
<td>TLE1</td>
<td>Intron</td>
<td>3a</td>
<td>A:T</td>
<td>0.309 (T)</td>
</tr>
<tr>
<td>rs1139451</td>
<td>9q21</td>
<td>84495608</td>
<td>—</td>
<td>—</td>
<td>5</td>
<td>C:T</td>
<td>0.221 (C)</td>
</tr>
<tr>
<td>rs12237222</td>
<td>9q21</td>
<td>85834743</td>
<td>LOC102723989</td>
<td>Intron</td>
<td>1a</td>
<td>G:T</td>
<td>0.425 (G)</td>
</tr>
<tr>
<td>rs12551710</td>
<td>9q21</td>
<td>86088295</td>
<td>FRMD3</td>
<td>Intron</td>
<td>1d</td>
<td>C:T</td>
<td>0.042 (T)</td>
</tr>
<tr>
<td>rs430794</td>
<td>9q22</td>
<td>93852815</td>
<td>—</td>
<td>—</td>
<td>5</td>
<td>A:C</td>
<td>0.272 (A)</td>
</tr>
<tr>
<td>rs8027714</td>
<td>15q11</td>
<td>24964597</td>
<td>—</td>
<td>—</td>
<td>6</td>
<td>A:G</td>
<td>0.040 (A)</td>
</tr>
<tr>
<td>rs1810636</td>
<td>20p13</td>
<td>2654925</td>
<td>—</td>
<td>—</td>
<td>4</td>
<td>G:T</td>
<td>0.332 (T)</td>
</tr>
<tr>
<td>rs2236479</td>
<td>21q22</td>
<td>46919132</td>
<td>COL18A1</td>
<td>Intron</td>
<td>1f</td>
<td>A:G</td>
<td>0.339 (A)</td>
</tr>
</tbody>
</table>
## Table 2: The distribution of RegulomeDB category 1a–1f SNPs in the whole human genome and in the region 9q21.

<table>
<thead>
<tr>
<th>RegulomeDB category</th>
<th>Category description</th>
<th>Number of SNPs with RegulomeDB score 1a–1f in the whole human genome</th>
<th>Number of SNPs with RegulomeDB score 1a–1f in the region 9q21</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>eQTL + TF binding + matched TF motif + matched DNase footprint + DNase peak</td>
<td>352</td>
<td>1</td>
</tr>
<tr>
<td>1b</td>
<td>eQTL + TF binding + any motif + DNase footprint + DNase peak</td>
<td>2568</td>
<td>17</td>
</tr>
<tr>
<td>1c</td>
<td>eQTL + TF binding + any motif + DNase footprint + DNase peak</td>
<td>85</td>
<td>0</td>
</tr>
<tr>
<td>1d</td>
<td>eQTL + TF binding + any motif + DNase peak</td>
<td>1668</td>
<td>7</td>
</tr>
<tr>
<td>1e</td>
<td>eQTL + TF binding + matched TF motif</td>
<td>54</td>
<td>0</td>
</tr>
<tr>
<td>1f</td>
<td>eQTL + TF binding/DNase peak</td>
<td>34706</td>
<td>127</td>
</tr>
</tbody>
</table>

## Table 3: The distribution of rs12237222 genotypes among cases and controls.

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Control (n = 291)</th>
<th>POP (n = 210)</th>
<th>Crude P value, OR (95% CI)</th>
<th>Adjusted P value, OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12237222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/T</td>
<td>104 (35.7)</td>
<td>55 (26.2)</td>
<td>0.21</td>
<td>0.12 (dom)</td>
</tr>
<tr>
<td>T/G</td>
<td>136 (46.7)</td>
<td>114 (54.3)</td>
<td>0.023 (dom)</td>
<td>1.57 (1.06–2.31)</td>
</tr>
<tr>
<td>G/G</td>
<td>51 (17.5)</td>
<td>41 (19.5)</td>
<td>1.41 (0.91–2.18)</td>
<td></td>
</tr>
</tbody>
</table>

HWP: Hardy-Weinberg probability; OR: odds ratio; CI: confidence interval; dom: dominant model.

*Adjusted by age, body mass index (BMI), perineal trauma in childbirth, and vaginal parity. In multivariate analysis, there were 271 controls and 198 cases.

and two SNPs are mapped to the RNA gene LINC01088 (rs1455311) and uncharacterized gene LOC102723989 (rs12237222). With few exceptions (rs2807303 and rs1036819), the majority of the SNPs assigned to the genes lie within intron regions. The choice of the genetic markers rs12237222 (score 1a) and rs12551710 (score 1d) in the region 9q21 was based on the RegulomeDB annotations. We present here the comparative qualitative and quantitative characteristics of the RegulomeDB SNPs in the categories 1a–1f in the whole human genome and in the region 9q21 (Table 2). Among the relatively small number of the SNPs scored 1a in the whole human genome, only one SNP with a score 1a was found in the region 9q21. This is the SNP rs12237222.

### 3.2. Characteristics of the Study Population

In total, 210 patients with pelvic organ prolapse (stages III-IV) and 292 control subjects with no even minimal prolapse were included in the study. The description of the study population is presented in Supplementary Table 2. Briefly, our cases and controls were similar in ethnicity (Caucasians from the European region of Russian Federation) and age (57.65 ± 10.80 and 57.25 ± 12.70, resp.). Higher BMI (29.17 ± 5.85 and 27.46 ± 6.56, P = 5.6 × 10⁻⁴), vaginal parity (1.69 ± 0.65 and 1.46 ± 0.54, P = 0.014), and perineal trauma in childbirth (52.45% versus 32.75%, P = 1.7 × 10⁻⁷) were found in POP patients than in matched controls.

### 3.3. Single-Locus Analysis

All SNPs were in Hardy-Weinberg equilibrium in the cases and in the controls. In the crude analysis, the minor rs12237222-G allele frequency was found to be higher in the cases than in the controls under the dominant model (Table 3). This effect disappeared in the multivariate analysis. Other SNPs were not associated with POP neither in the crude nor in the multivariate analysis (Supplementary Table 3).

### 3.4. Haplotype Analysis

For the haplotype analysis, we included all studied SNPs located in the region 9q21 (Table 4). Haplotype number four comprising rs12237222-G allele was significantly associated with POP in the analysis adjusted for age, BMI, vaginal parity, and perineal trauma in childbirth. When we excluded this SNP from the haplotype analysis, we found that haplotype number one that included risk alleles from the remaining five SNPs became the reference haplotype (the most common). The pronounced effect was also found for the two-SNP-haplotype rs2777781-A/rs12237222-G with total haplotype frequency 0.3087 (Table 4). No significant association was observed between other haplotypes and POP development.

It is interesting to note that rs12237222 is in weak but highly significant linkage disequilibrium with rs11139451 (D’ = 0.257, P = 1.45 × 10⁻⁶) that has been found to show the maximum HLOD score in the study of Allen-Brady et al. [20] (Supplementary Table 4).

*Adjusted by age, body mass index (BMI), perineal trauma in childbirth, and vaginal parity. In multivariate analysis, there were 271 controls and 198 cases.

### 3.5. Analysis of SNP Interaction

Figure 1 presents a heat plot color map for an analysis of SNP interaction. Cells are colored according to interaction P value (less significant effects, light yellow; more significant effects, green). The most significant interactions (dark green cells) were observed for two SNP pairs, rs12551710 (RegulomeDB score 1d) and rs1455311 (RegulomeDB score 5) (P = 7.3 × 10⁻⁵), from first
Table 4: The distribution of 9q21 haplotypes among cases and controls.

<table>
<thead>
<tr>
<th>Number</th>
<th>rs4077632</th>
<th>rs2807303</th>
<th>rs2777781</th>
<th>rs1139451</th>
<th>rs12237222</th>
<th>rs12551710</th>
<th>Frequencies</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>POP</td>
<td>Control</td>
<td>POP</td>
<td></td>
<td></td>
<td>P value</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>A T A T T C</td>
<td>0.1334</td>
<td>0.0985</td>
<td>—</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>A C A T T C</td>
<td>0.1185</td>
<td>0.0943</td>
<td>0.38</td>
<td>0.59 (0.19–1.90)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>G C A T G C</td>
<td>0.0436</td>
<td>0.0937</td>
<td>0.24</td>
<td>1.92 (0.65–5.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A C A T G C</td>
<td>0.059</td>
<td>0.1019</td>
<td>0.029</td>
<td>3.36 (1.13–9.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only haplotypes with total frequencies ≥10% are considered. Analysis is adjusted by age, body mass index (BMI), peritoneal trauma in childbirth, and vaginal parity. Significant results are in bold.

3.6. Analysis of Population Structure. The discriminant analysis of principal components showed that Russian population has no internal stratification (Figure 2). Russian population was located close to other Caucasian populations, CEU and TSI. African YRI and LWK populations and Asian CHD and JPT populations were far from Caucasian populations, while GIH and MEX populations took an intermediate position (Figure 2).

Fst statistics are presented in Supplementary Table 5. The pairwise Fst values between the examined populations were in line with the data of DAPC. The lowest pairwise Fst values (Fst = 0) were revealed for CEU-RUS and CEU-TSI populations.

4. Discussion

In this gene candidate study, we tested the association of the region 9q21 and six other loci [20, 21] with advanced POP in Caucasians. We used RegulomeDB resource to select regulatory SNPs in the region 9q21. There was only one SNP rs12237222 in the region with the best RegulomeDB score 1a. This SNP appeared to be the key SNP for the 9q21 haplotype associated with POP. Two other category 1 variants, which are also characterized as eQTL, rs2236479 (21q22), and rs12551710 (9q21) exhibited epistatic (nonlinear interaction) effects with rs2777781 and rs1455311, respectively. To control for the confounding effect of known clinical covariates (age, BMI, vaginal parity, and pelvic floor trauma in childbirth), we adjusted for these factors. To assess hidden population stratification, we compared the genotypic frequencies of 12 studied SNPs to those of HapMap ethnic groups and found that our sample showed little structure.

In terms of ancestry, Russians are Caucasians. The study of Russian population was representative, from an ethnic standpoint of the other Caucasian populations (CEU and TSI). Other HapMap populations in this study, YRI, LWK, JPT, and CHD populations differed significantly from all three Caucasian populations, Russian, CEU, and TSI. It is worth noting that RUS, CEI, and TSI populations and relatively close to them GIH and MEX populations are characterized by high prevalence of POP [4, 44, 45]. In contrast, African and Asian women presented in our study by YRI, LWK, JPT, and CHD populations are less prone to POP development [4, 44, 46, 47]. This may be explained mainly by anatomical features [48, 49]. Nevertheless, it is an interesting challenge for future research directions to determine whether ethnic-specific genotype frequencies of the studied SNPs make any contribution to higher prevalence of POP in Caucasian, GIH, and MEX females compared with Asian and African females.

All our patients had an advanced prolapse (stages III-IV on the POP-Q examination), while our controls had not even minimal prolapse. Consideration of more extreme phenotypes is recommended for the genetic association studies [50]. Utilizing extreme phenotypes increases the association effects, with improved statistical power of relatively modest sample sizes [51]. In relation to POP, this direction could
be taken into account not only for patients but for controls as well. In a recent study of Wu et al. [15], their control group included controls with no or minimal prolapse, who were older than patients. We applied another approach for recruitment of an appropriate comparison group. Should POP arise, we would expect that subclinical manifestation progresses to more severe symptoms. The rate of this progress depends on many variables including the history of vaginal childbirth and other clinical conditions (e.g., hysterectomy), age, frequent heavy lifting, and intense physical activity leading to the repetitive increase of pelvic pressure [52]. For this reason, we purposely included women without any subclinical POP in our control group.

It is known that over 98% of the human genome comprises noncoding DNA. However, the Encyclopedia of DNA elements (ENCODE) project has recently demonstrated that a biochemical function could be assigned to 80% of the human genome [53]. Regulatory activity of noncoding DNA explains the fact that a very high amount (90%) of GWAS-identified SNPs associated with multifactorial diseases and complex traits are located in noncoding regions [54]. RegulomeDB includes data sets from ENCODE and other sources to predict regulatory function of a SNP in a score based system. Lower scores mean increasing evidence that a variant is located in a functional region: category 1 indicates that a variant has the highest likelihood of being linked to the expression of a target gene and category 1a is characterized by a full set of features (eQTL + TF binding + matched TF motif + matched DNase Footprint + DNase peak). RegulomeDB project provides an important tool for both tasks, planning an association study and post hoc analysis; its data are useful for generating hypotheses and explaining possible epistasis interactions.

SNP interaction might be common among eQTL SNPs that influence gene expression. Epistatic interactions involve local or distant factors acting in cis or trans. Cis-interaction usually occurs within 500 kb region and may sometimes reflect an underlying haplotype effect, while transinteraction indicates interaction between SNPs from distant regions [55]. Loci involved in gene-gene interactions may not show associations on their own. It is also known that in replication studies some loci associations could not be verified, but the corresponding regions might show clear interaction signals [56]. We could not replicate all associations from the studies of Allen-Brady et al. [20, 21], but the interaction effects revealed in our study for several SNPs are in line with the associations found in the above studies for the relevant SNPs.

It is proposed that epistasis among regulatory variants can have strong phenotypic effects, since gene regulation has
an important role in adaptation to different harmful stimuli. Regulatory SNPs often lie in noncoding sequences and affect transcription of coding and noncoding RNAs or RNA functions and processing [57]. In our study, among SNPs associated with POP in the haplotype and/or interaction analysis, three SNPs (TLE1 rs2777781, FRMD3 rs12551710, and COL18A1 rs2236479) were located within the genes with possible impact upon development and organization of muscle and connective tissue of the pelvic floor [20, 21]. Their effects may be realized based on regulation of the expression of these genes or via transregulatory activity. Two other SNPs, rs145531I and rs12237222 (the top SNP in this assay), within the nonprotein coding genes were also associated with POP in the haplotype or interaction analysis. Based on data of SCAN SNP and CNV Annotation Database [58] assayed in lymphoblastoid cells, these SNPs appeared to exhibit transregulatory activity in CEU population: rs145531I was associated with the expression of the NAPRTI (nicotinate phosphoribosyl transferase domain containing 1) gene and rs12237222 influenced the expression of the CAT (catalase) and GYPE (glycoporphin E) genes. NAPRTI and CAT genes encode antioxidant enzymes and control response to oxidative stress, which is known to play a role in the development of many chronic or late-onset diseases including POP [59]. Regulatory SNPs may exert a tissue-dependent effect on gene expression but in general, “allele specific expression differences between individuals dominate over tissue-specific effects” [60]. Basing on these data, we can speculate that the above SNPs may be associated with POP in haplotype or interaction dependent mode, in particular through influence on efficacy of oxidative stress response.

A limitation of our study is that, firstly, our sample size was modest and the study is powered to detect only relatively large effect sizes (minimum detectable OR≈1.7). Secondly, the study is limited to the population of Russian Federation. The strength of the study is in the sampling of extreme phenotypes, the inclusion of all main risk factors for POP in multivariate regression analysis, little population structure, and using of RegulomeDB resource to select functional SNPs.

5. Conclusions

The findings of the present study verified the results of the genome-wide linkage analysis in European pedigrees with POP. We found that the key SNP in the haplotype 9q21 associated with POP in our Russian population had the best RegulomeDB score 1a and two other SNPs category 1 were involved in SNP-SNP interaction associations. To the best of our knowledge, this is the first successful research which was performed utilizing RegulomeDB annotations a priori, for planning investigations. RegulomeDB annotations can be useful for designing an association study of the chromosome region.

Conflict of Interests

The authors declare no potential conflict of interests.

Acknowledgment

This work was supported by the Russian Foundation for Basic Research, Project 15-04-02378.

References


