ISSN 1022-7954, Russian Journal of Genetics, 2017, Vol. 53, No. 7, pp. 785–794. © Pleiades Publishing, Inc., 2017.
Original Russian Text © R.I. Khusainova, A.V. Tyurin, D.A. Shapovalova, E.K. Khusnutdinova, 2017, published in Genetika, 2017, Vol. 53, No. 7, pp. 816–826.

HUMAN GENETICS

Search for Osteoarthritis Genetic Markers in Women with Undifferentiated Connective Tissue Dysplasia

R. I. Khusainova[a], *, A. V. Tyurin[b], D. A. Shapovalova[a], and E. K. Khusnutdinova[a]

[a] *Institute of Biochemistry and Genetics, Ufa Scientific Center, Russian Academy of Sciences, Ufa 450054 Russia*
[b] *Department of Hospital Therapy No. 2, Bashkir State Medical University, Ufa, 450076 Russia*
**e-mail: ritakh@mail.ru*

Received June 26, 2016

Abstract—We conducted an association study of ten polymorphisms in six candidate genes of OA (*rs1799750* (*MMP1*), *rs35068180* (*MMP3*), *rs2252070* (*MMP13*), *rs63118460* and *rs2276455* (*COL2A1*), *rs143383* (*GDF5*), *rs1544410*, *rs7975232*, *rs731236*, and *rs2228570* (*VDR*)) with the development of osteoarthritis (OA) in 333 women taking into account the localization of the pathological process, the age of disease manifestation, and ethnicity and investigated the presence of signs of undifferentiated connective tissue dysplasia (UCTD). On the basis of clinical and genetic data, we revealed statistically significant models to predict the development of osteoarthritis of various localizations (knee, hip, and generalized forms).

Keywords: osteoarthritis, arthrosis, genetic predisposition, connective tissue dysplasia, candidate genes
DOI: 10.1134/S1022795417060072

INTRODUCTION

Osteoarthritis (OA, arthrosis in the Russian literature (ICD-10:M15-M19)) is a progressive joint disease of unknown etiology characterized by damage to all parts of the joints—cartilage, subchondral bone, synovium, ligaments, capsule, and periarticular muscle [1]. The prevalence of OA in the world is 6.43% on average; it correlates with age and reaches the maximum values (13.9%) in individuals older than 45 years [2]. The disease is characterized by a significant decrease in quality of life and high initial disability of the patients.

There is evidence that patients with OA display high prevalence of various phenotypic markers of connective tissue dysplasia (CTD) [3]. The basis of CTD involves defects in the synthesis or catabolism of extracellular matrix components and regulators of the connective tissue morphogenesis leading to disruption of the connective tissue structure. Considering that the destructive changes in the connective tissue structure have been revealed in OA and CTD, it is possible to suggest the existence of common pathogenetic mechanisms in the formation of these pathologies.

In the International Classification of Diseases (ICD-10), the term "connective tissue dysplasia" is not a nosological unit; individual syndromes such as differentiated and undifferentiated CTD occur in different chapters of the classification (chapters XIII and XVII). Differentiated CTD includes hereditary syndromes (Ehlers–Danlos, Marfan, Stickler, etc.). Undifferentiated connective tissue dysplasia (UCTD) is a multifactor form of inherited connective tissue disorders that do not fit into the structure of hereditary syndromes; it is characterized by genetic heterogeneity and diverse clinical manifestations. UCTD is commonly a progressive disorder, underlies the formation of a significant number of somatic diseases, and determines the course of the underlying disease [4]. It is believed that the phenotypic signs of UCTD with various intensities are found in 10–22.5% of the world's population [5]. In recent years, the issue of CTD and its role in formation of diseases of various organs and systems has received great attention, but this task is far from complete [6, 7].

Given the multifactor nature of OA, the genetic component of the disease is currently under active search. The results of twin studies have shown that the genetic component of OA can vary from 50 to 65% [8, 9]. Molecular genetic studies allowed the identification of several candidate genes of OA; however, significant progress in identifying the genetic basis of OA was achieved with the technology of genome-wide analysis of associations (GWAS) of hundreds of thousands of single nucleotide polymorphisms (SNPs) and the study of large samples (>100000 samples) in the framework of international consortia, such as TREAT-OA and ArcOGEN [10, 11]. These studies confirmed the contribution of a number of known genes in the development of OA and identified many new candidate loci. However, it became apparent during GWAS that the number of the OA associated loci does not increase proportionally to the increase in the volume of samples, which may indicate a significant clinical

and genetic heterogeneity of the disease. The main problem of interpreting the results of genome-wide studies is the determination of the required gene in the associated region of the genome and the mechanism of its contribution to the pathogenesis of the disease. The second important matter is the population factor. Loci that influence the development of OA in some populations are not associated with disease in others. Gender differences have also been revealed during the identification of genetic markers of OA.

Thus, significant progress has been made in identifying the genetic predisposition to OA; however, there are many unresolved key issues. The issue of OA comorbid with phenotypic manifestations of UCTD is extremely relevant; it is a fundamental and practical concern, the solution of which will contribute to the development of approaches for early diagnosis of OA based on understanding of molecular pathogenesis of the disease.

The purpose of this study was to search for genetic markers of predisposition to OA in women taking into account the localization of the pathological process, the presence of signs of UCTD in general and its individual phenotypes, and also the ethnicity of the patients.

MATERIALS AND METHODS

The study material was DNA from 333 women who were examined for signs of OA and UCTD at the Therapeutic Department of the Municipal Clinical Hospital No. 18 and at Polyclinics Nos. 2, 18, and 38 of Ufa, in compliance with standards of the Helsinki Declaration developed by the World Medical Association (WMA) "Ethical Principles for Medical Research Involving Human Subjects" and with the approval of the local bioethics committees of the Bashkir State Medical University (Protocol No. 28 of October 29, 2012) and the Institute of Biochemistry and Genetics, Ufa Scientific Center, Russian Academy of Sciences (Protocol No. 4 of November 15, 2012).

The group of female patients with osteoarthritis (158 women) was formed using the criteria of the American Association of Rheumatology (1995), classification by V.A. Nasonova and M.G. Astapenko [12] with the onset of the disease under the age of 55 years and also the presence of radiologic osteoarthritis. Among them, polyarthrosis (POA) was diagnosed in 41 women (23.4%), gonarthrosis (GA) in 80 women (50.6%), and coxarthrosis (CA) in 37 women (26%). Radiologic grade II (according to the Kellgren–Lawrence grading system) was in 97 patients (61.4%), grade III in 40 (25.3%), and grade IV in 21 (13.3%). The disease duration was 3 to 18 years (4.7 ± 1.8 years on average).

The presence of phenotypic signs of UCTD was estimated in scores according to the technique by T.I. Kadurina and V.N. Gorbunova [5]. UCTD was identified as mild grade at the total score of 9 to 14 and severe at the total score higher than 15. The symptom complex of UCTD was identified in 156 people, including in 125 women (80.1%) with mild UCTD and in 31 (19.9%) with severe UCTD.

The cases of differentiated connective tissue diseases, age over 65 years, active infection, and traumatic injury to the joints in the anamnesis were excluded during the sample selection.

According to the ethnic composition, the sample was composed of 121 women (36.3%) of Russian origin, 129 (38.7%) of Tatar origin, and 30 (9.01%) of Bashkir origin; 36 (10.8%) were mestizos and 17 (5.19%) were women of other ethnic groups.

To study the comorbidity of OA and UCTD, four comparison groups were formed. The first group included 92 women with a combination of OA and UCTD (OA+ UCTD+); the second group included 66 women with isolated OA (OA+ UCTD–), the third group included 64 women with isolated UCTD (OA– UCTD+). The control group consisted of 111 women without signs of OA and UCTD (OA– UCTD–). The characteristics of the study groups are given in Table 1.

DNA was extracted from peripheral blood using a standard method of phenol-chloroform extraction [13]. Genotyping was performed using PCR/RFLP techniques and using restriction endonucleases manufactured by Fermentas (Lithuania). The restriction of *rs1799750* (*MMP1*), *rs35068180* (*MMP3*), *rs2252070* (*MMP13*), *rs63118460* and *rs2276455* (*COL2A1*), *rs143383* (*GDF5*), and *rs1544410* and *rs7975232* (*VDR*) polymorphic loci was performed using *Alu*I, *Tth*III, *Bst*NI, *Pvu*II, *Hin*dIII, *Bsi*EI endonucleases, respectively, at 37°C, and the *rs2228570* and *rs731236* (*VDR*) polymorphic loci were restricted using *Bse*GI and *Taq*I, respectively, at a temperature of 55°C. To separate DNA fragments after PCR and RFLP, 7% polyacrylamide gel (PAGE) electrophoresis was used and the results were evaluated using the Geldokulant system (France).

The results were statistically processed using generally accepted methods of descriptive and variational statistics using standard Microsoft Excel 2007, Statistica 6.0, MedCalc software packages. The heterogeneity of the investigated samples was assessed with the principal component analysis using the Smatpca software program. Linkage disequilibrium and haplotype analysis were conducted using the Haploview 4.2 software program (http://www.chgb.org.cn/lda/lda.htm). Intergroup comparison was conducted by calculating the χ^2 index with Yates correction for continuity. The degree of association was measured using the value of the odds ratio. Linkage disequilibrium (LD) between pairs of SNPs was assessed using the coefficient D' suggested by Lewontin and the Pearson coefficient of correlation r^2. Prediction models were constructed using multivariable logistic regression with ROC curve

Table 1. The main characteristics of the studied groups

Group	*N*	Age		
		$M \pm m$	min	max
Total	333	50.13 ± 12.60	15	65
OA in general	158	54.72 ± 8.50	20	65
CA	37	53.66 ± 9.57	28	65
GA	80	55.64 ± 8.15	26	65
POA	41	53.07 ± 8.78	20	65
UCTD in general	156	49.46 ± 12.94	15	65
UCTD mild	125	49.44 ± 12.76	15	65
UCTD severe	31	48.19 ± 14.30	18	65
OA+ UCTD+	92	54.40 ± 9.56	20	65
OA+ UCTD–	66	42.35 ± 13.89	40	65
OA– UCTD+	64	54.51 ± 7.29	15	64
OA– UCTD–	111	47.09 ± 10.95	19	65

Symbol "+" indicates the presence of the sign; "–" indicates the absence of the sign; *N* is the number of individuals; $M \pm m$ is the mean value and standard error.

construction. The performance of prediction models was assessed by calculating the area under the curve (AUC). Power calculation of the study was performed using the statistical calculator http://osse.bii.a-star.edu.sg/calculation2.php for alleles and sensitivity of contingency tables at a significance level of 0.05 and the intergroup number of degrees of freedom was 3 for genotypes and 5 for haplotypes. Correction for multiple comparisons was calculated using the Benjamini–Hochberg procedure (false discovery rate, FDR) and permutation test.

RESULTS

Ten polymorphisms of six OA candidate genes were studied: *rs1799750* (*c.-1673delG*) of the metalloproteinase *1* gene (collagenase, *MMP1*), *rs35068180* (*c.-1671_1670insT*) of the metalloproteinase *3* gene (stromelysine 1, *MMP3*), *rs2252070* (*c.-105G>A*) of the metalloproteinase *13* gene (collagenase 3, *MMP13*), *rs63118460* (*c.1531-108A>G*) and *rs2276455* (*c.2304+148C>T*) of the collagen type II gene (*COL2A1*), *rs143383* (*c.-275T>C*) of the growth/differentiation factor 5 gene (*GDF5*), *rs1544410* (c.*1024+283G>A*; *Bsm*I), *rs7975232* (*c.1025-49G>T*, *Apa*I), *rs731236* (*c.1056T>C*, *Taq*I) and *rs2228570* (*c.2T>A/G/C, p.Met1/Lys/Arg/Thr*; *Fok*I, *rs10735810*) of the vitamin D receptor gene (*VDR*). The studied loci were searched for the associations with OA in general, the localization of the pathological process, and the presence of UCTD and its specific phenotypic signs, as well as OA comorbid with UCTD. A gene-candidate approach was used and the genes were selected on the basis of the involvement of the protein products in the connective tissue formation and metabolism and their suggested role in the development of both UCTD and OA. The characteristics of the studied loci are presented in Table 2.

The distribution of genotype frequencies conform to the Hardy–Weinberg equilibrium ($p > 0.05$) for all the studied loci, except for *rs1544410* (*VDR*), *rs7975232* (*VDR*), and *rs35068180* (*MMP3*). The observed frequency of heterozygotes exceeded the expected frequency for the *rs1544410* and *rs35068180* loci and it was lower than the expected heterozygocity for the *rs7975232* locus (Table 2). In addition, in the control group, the genetic equilibrium conformed to the Hardy–Weinberg law of equilibrium for all the studied loci, which indicates sufficient genetic heterogeneity and the absence of errors in the formation of the study sample and during genotyping. Minor allele frequencies varied from 0.314 in the *rs731236* (*VDR*) locus to 0.479 in the *rs35068180* (*MMP3*) locus.

Taking into account the ethnic heterogeneity of the sample, we assessed the heterogeneity of the studied polymorphic loci using the principal component analysis and revealed no population and genetic differentiation. However, given the multifactor nature of the studied pathology and the presence of two numerous ethnically homogeneous samples (Russian and Tatar), we considered it possible to perform additional evaluations in the main comparison groups based on the ethnic factor.

The search for associations of the studied loci with the development of OA in general and taking into account the localization of the pathological process did not show statistically significant results after the use of correction for multiple comparisons, which probably indicates the phenotypic heterogeneity of the disease.

We also compared the distribution of allele and genotype frequencies of the studied polymorphisms of

Table 2. Characteristics of the study loci of candidate genes

No.	Locus	Location	Gene	Chromosomal localization of the gene	H_{pred}	H_{obs}	HW_{pval}		MAF
							in total sample	in control	
1	*rs143383*	5'UTR	*GDF5*	20q11.22	0.481	0.481	0.217	0.839	0.402
2	*rs731236*	Exon 9	*VDR*	12q13.11	0.408	0.431	0.462	0.859	0.314
3	*rs7975232*	Intron 8	*VDR*	12q13.11	0.587	0.495	0.004	0.169	0.448
4	*rs1544410*	Intron 8	*VDR*	12q13.11	0.385	0.467	0.001	0.189	0.371
5	*rs2228570*	Exon 2	*VDR*	12q13.11	0.456	0.474	0.636	0.488	0.386
6	*rs2276455*	Intron 33	*COL2A1*	12q13.11	0.401	0.450	0.062	0.156	0.343
7	*rs63118460*	Intron 23	*COL2A1*	12q13.11	0.472	0.492	0.513	0.252	0.437
8	*rs1799750*	5'UTR	*MMP1*	11q22.2	0.520	0.494	0.415	0.466	0.445
9	*rs2252070*	5'UTR	*MMP13*	11q22.2	0.464	0.440	0.425	0.792	0.327
10	*rs35068180*	5'UTR	*MMP3*	11q22.2	0.429	0.499	0.015	0.228	0.479

H_{obs} is the observed heterozygocity; H_{pred} is the predicted heterozygocity; HW_{pval} is the p value for deviation from the Hardy–Weinberg equilibrium (supported at the level of $p > 0.05$); MAF is the minor allele frequency; 5'UTR is the 5'-untranslated region.

OA candidate genes in relation to the presence or absence of phenotypic signs of UCTD in general, as well as on the basis of the severity of UCTD. The **G* allele of the *rs1544410* polymorphic variant of the *VDR* gene was found to be associated with UCTD in general (χ^2 = 9.360; p = 0.002; OR = 1.77; 95% CI 1.22–2.56) and mild UCTD (χ^2 = 8.069; p = 0.005; OR = 1.78; 95% CI 1.19–2.66) and the statistical significance remained after the use of correction for multiple comparisons (FDR) (Table 3). The patients with severe UCTD showed a similar tendency; however, the differences in the **G* allele frequency between the comparison groups did not reach statistical significance. In addition, the **G*G* homozygous genotype was associated with severe UCTD (χ^2 = 5.97; p = 0.016 (p^* = 0.018); OR = 2.76; 95% CI 1.19–6.39), which supports the significance of this locus in the formation of the UCTD phenotype of different grades of manifestation. However, the subsequent power evaluation of the study revealed low power values, which ranged from 52.6 to 58.4% for significant associations of allele **G* with UCTD. To achieve the 80% power value, it is necessary to form comparison groups consisting of at least 200 subjects each and to conduct further studies to prove these associations.

The symptom complex of UCTD included 25 phenotypic characteristics, according to the table by T.I. Kadurina and V.N. Gorbunova [5]. Given the clinical heterogeneity of this condition, we analyzed the associations of the studied loci with the specific manifestations of UCTD.

The significance of polymorphic variants of the metalloproteinases genes (*MMP*) in the formation of gallbladder deformities was revealed (**2G* of the *rs1799750* locus of the *MMP1* gene (χ^2 = 6.48; p = 0.01; OR = 2.10; 95% CI 1.16–3.47) and **G* of the *rs2252070* locus of the *MMP13* gene (χ^2 = 4.92; p = 0.020; OR = 1.8; 95% CI 1.01–3.08)), in the formation of gastrooesophageal reflux disease (GORD) (**G* of the *rs2252070* locus of the *MMP13* gene (χ^2 = 5.87; p = 0.01; OR = 2.18; 95% CI 1.14–4.15)), in hernias (**A* of the *rs2252070* locus of the *MMP13* gene (χ^2 = 6.18; p = 0.01; OR = 2.30; 95% CI 1.2–4.55)), in arterial hypertension (**A* of the *rs1799750* locus of the *MMP1* gene (χ^2 = 6.0; p = 0.014; OR = 1.51; 95% CI 1.08–2.11)), and in severe joint hypermobility syndrome (JHS) (**A* of the *rs2252070* locus of the *MMP13* gene (χ^2 = 46.0; p = 0.001; OR = 10.7; 95% CI 4.7–24.2)).

Polymorphic variants of the collagen type II gene (*COL2A1*) were found to be associated with development of severe joint hypermobility syndrome (**A* of the *rs2276455* locus (χ^2 = 4.77; p = 0.02; OR = 1.89; 95% CI 1.05–3.37)), arterial hypertension (**A*A* of the *rs2276455* locus (χ^2 = 4.71; p = 0.029; OR = 2.01; 95% CI 1.06–3.76) and **C* of the *rs63118460* locus (χ^2 = 4.61; p = 0.03; OR = 1.44; 95% CI 1.01–2.02)), dolichostenomelia (**C* of the *rs63118460* locus (χ^2 = 6.93; p = 0.008, OR = 4.01; 95% CI 1.40–11.90)), and hemorrhagic condition (**A*A* of the *rs2276455* locus (χ^2 = 3.99; p = 0.045, OR = 1.88; 95% CI 1.01–3.52)).

An association of the vitamin D receptor (*VDR*) gene polymorphisms was revealed with the formation of mild grade joint hypermobility (**G* of the *rs7975232* locus (χ^2 = 4.08; p = 0.043; OR = 1.52; 95% CI 1.01–2.28)), visceroptoses (**G* of the *rs1544410* locus (χ^2 = 6.62; p = 0.01; OR = 1.92; 95% CI 1.16–3.18) and **C* of the *rs731236* locus (χ^2 = 3.92; p = 0.04; OR = 1.68; 95% CI 1.01–2.82)), chest deformity (**G* of the *rs1544410* locus (χ^2 = 4.74; p = 0.02; OR = 3.09; 95% CI 1.16–8.22)), and also paradontosis (**T* of the

Table 3. Results of allele and genotype frequency distribution of the *rs1544410* locus of the *VDR* gene in women with UCTD in general and of different grade of manifestation

Sample	*N*	Power of the study, %	Alleles		Genotypes		
			**G*	**A*	**G*G*	**G*A*	**A*A*
UCTD in general	118	58.4	165 (0.699) $\chi^2 = 9.36$; $p = 0.002$; $p^* = 0.009$; OR = 1.77 (1.22–2.56)	71 (0.301)	62 (0.525)	41 (0.347)	15 (0.128) $\chi^2 = 4.005$; $p = 0.045$; $p^* = 0.049$ OR = 0.51 (0.26–0.99)
UCTD mild	90	52.6	126 (0.700) $\chi^2 = 8.069$; $p = 0.005$; $p^* = 0.014$; OR = 1.78 (1.19–2.66)	54 (0.300)	45 (0.500) $\chi^2 = 4.460$; $p = 0.035$; $p^* = 0.065$	36 (0.400)	9 (0.100)
UCTD severe	28	24.8	39 (0.696) $\chi^2 = 3.19$; $p = 0.073$; $p^* = 0.064$	17 (0.304)	17 (0.607) $\chi^2 = 5.97$; $p = 0.016$; $p^* = 0.018$; OR = 2.76 (1.19–6.39)	5 (0.179) $\chi^2 = 5.65$; $p = 0.01$; $p^* = 0.015$ OR = 0.30 (0.1–0.84)	6 (0.214)
Absence of UCTD	134		152 (0.567)	116 (0.433)	48 (0.358)	56 (0.418)	30 (0.224)

N is the number of individuals. Absolute values indicate the number of cases; allele and genotype frequency are given in parentheses; χ^2 is the Pearson chi-square test; *p* is the significance level, indicated only for statistical significance (at $p < 0.05$) or tendency; OR is the odds ratio and 95% confidence interval (in parentheses); p^* is with correction for multiple comparisons of FDR; OR was calculated only for significant differences between the comparison groups.

rs1544410 locus ($\chi^2 = 3.97$; $p = 0.046$; OR = 1.96; 95% CI 1.1–3.86)).

The **C* allele of the *rs143383* polymorphic variant of the growth/differentiation factor 5 gene (*GDF5*) was found to be significant in the development of myopia ($\chi^2 = 4.55$; $p = 0.032$; OR = 1.56; 95% CI 1.01–2.35).

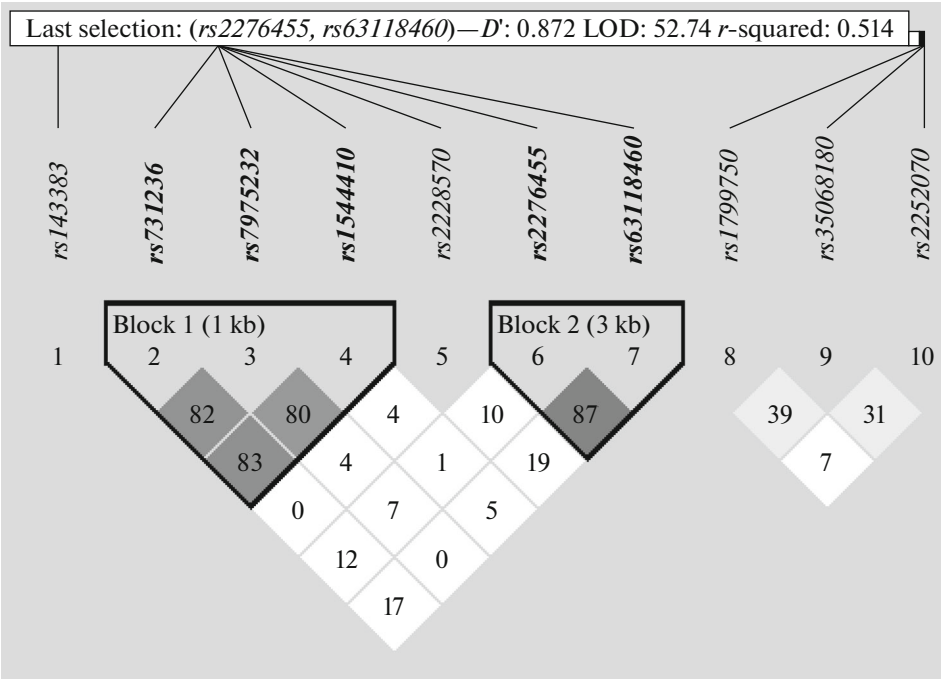
Thus, all the studied OA candidate genes were found to be involved in the formation of certain phenotypic signs of UCTD.

During the subsequent analysis, only one association met 80% power of the study—the **A* allele of the *rs2252070* locus of the *MMP13* gene with severe grade joint hypermobility (99.9%), which makes it possible to consider this polymorphism as a marker of an increased risk of developing severe JHS. Homozygous genotypes of two polymorphic loci of the *VDR* gene associated with visceroptoses had a power of 78–79%; the power of the study was lower in the remaining loci.

To identify possible molecular-genetic aspects of OA and UCTD comorbidity, we compared the distribution of allele and genotype frequencies of the studied loci between groups of women with OA in conjunction with UCTD and without it. It was revealed that the **G*G* genotype of the *rs1544410* polymorphic variant and the **G*T* genotype of the *rs7975232* locus of the *VDR* gene enhance the risk of OA development in conjunction with UCTD ($\chi^2 = 4.200$; $p = 0.040$; OR = 0.50; 95% CI 0.26–0.99 and $\chi^2 = 3.80$; $p = 0.049$; OR = 1.90; 95% CI 1.01–3.63, respectively); however, the power of the study was only 54.3% and we consider the result to be a tendency, which needs further confirmation.

Given the multifactor nature of OA and UCTD, we analyzed our samples taking into account the ethnic factor and searched for the associations of the studied loci in the most representative groups of women of Russian and Tatar ethnicity. Several statistically significant results were obtained indicating the involvement of the polymorphisms of the *GDF5* gene in the risk of developing OA in women of Tatar ethnicity and of the *MMP13*, *COL2A1* and *VDR* genes in UCTD symptom complex in women of Russian ethnicity, but the power of the study ranged from 3.2 to 58.9%, which does not allow us to define these data as significant, and it is necessary to increase the power of the study via expansion of ethnically homogeneous samples.

We evaluated the linkage disequilibrium of the studied loci of the *VDR*, *COL2A1*, *MMP1*, *MMP3*,



Analysis of linkage disequilibrium of the loci under study. Two blocks of haplotypes are selected according to the revealed linkage disequilibrium of the loci shown in the figure; the numbers correspond to the coefficient *D*'.

MMP13, GDF5 genes with further comparative analysis of haplotypes; correction for multiple comparisons was conducted using the method of random permutations (permutation test); and all the calculations were performed using the Haploview 4.2 software program. It was found that the *rs1544410*, *rs7975232*, and *rs731236* polymorphic variants of the *VDR* gene are in the linkage disequilibrium ($D' \geq 0.79$) and compose one linkage cluster; the *rs63118460* and *rs2276455* polymorphic variants of the *COL2A1* gene compose the second cluster ($D' = 0.85$) (figure). There was no significant association of haplotypes with OA and the formation of the symptom complex of UCTD after the use of corrections for multiple comparisons. The carriers of the **CTA* haplotype have a reduced risk of developing OA and UCTD in general, as well as of mild UCTD. The power of the study ranged from 40 to 65%, which does not allow drawing final conclusions about the involvement of the vitamin D receptor and collagen II type genes in the pathogenesis of OA and UCTD (Table 4).

To assess the diagnostic significance of markers of the disease, a multivariable regression analysis was conducted. The equation for logistic regression included 35 clinical predictors (25 phenotypes of UCTD, the presence of UCTD in general and mild and severe grades of UCTD, the age of OA manifestation before 50 and after 50 years, five groups of different ethnic origin (Russians, Tatars, Bashkirs, mestizos, and others), and 10 genetic loci (2 alleles in each). The consistency of the models was evaluated using the ROC (receiver operating characteristics) analysis and calculating the AUC (area under curve), which is a numerical indicator under the ROC curve, the values of which in the range 0.6–0.7 were considered to be an average result, 0.7–0.8 were a good result, and 0.8–0.9 were a very good result. On the basis of all the generated models, the three most important predicting OA development of different localizations were chosen (coxarthrosis, gonarthrosis, and polyarthrosis). The characteristics of the models are presented in Table 5.

The model for the OA diagnosis of arthrosis of the hip (coxarthrosis) included the presence of spinal pathology (kyphosis/lordosis), varicose veins, visceroptosis, gastrooesophageal reflux disease, the**A*A* genotype of the *rs2252070* locus of the *MMP13* gene, and the **T* allele of the *rs63118460* locus of the *COL2A1* gene. This model had a medium diagnostic value, and the sensitivity and specificity amounted to 66.5 and 68%, respectively.

The model for the OA diagnosis of the knee joints (gonarthrosis) included UCTD, varicose veins, GORD, spinal pathology (kyphosis/lordosis), the **T* allele of the *rs63118460* locus of the *COL2A1* gene, and the **C*C* genotype of the *rs143383* locus of the *GDF5* gene and had a higher sensitivity—71%; the specificity was 62%, and the prediction consistency reached the good level.

The model for the diagnosis of generalized OA (polyarthrosis) consisted of connective tissue dysplasia in general, GORD, severe myopia, spinal deformities in the form of kyphosis/lordosis, gallbladder

Table 4. Analysis of haplotype frequencies of the *rs1544410*, *rs7975232,* and *rs731236* polymorphic loci of the *VDR* gene

Sampling	Power of the study, %	Haplotype frequencies					
		**CGG*	**TTA*	**CTG*	**CTA*	**CGA*	**TTG*
Absence of OA $N = 134$		0.381	0.284	0.208	0.090	0.030	0.007
OA in general $N = 146$	40	0.432	0.274	0.218	0.034 $\chi^2 = 8.16$; $p = 0.004$; $p^* = 0.0325$; OR = 0.27 (0.11–0.68)	0.021	0.021
Absence of UCTD		0.351	0.299	0.194	0.104	0.037	0.015
UCTD mild $N = 89$	65	0.438	0.281	0.237	0.011 $\chi^2 = 14.92$; $p = 0.001$; $p^* = 0.001$; OR = 0.07 (0.01–0.40)	0.011 $\chi^2 = 3.74$; $p = 0.053$	0.022
UCTD severe $N = 28$	<30	0.464 $\chi^2 = 3.47$; $p = 0.062$	0.214	0.179	0.107	0.036	0.0
UCTD in general $N = 117$	50	0.453 $\chi^2 = 4.83$; $p = 0.027$; $p^* = 0.214$	0.256	0.223	0.034 $\chi^2 = 10.45$; $p = 0.001$; $p^* = 0.006$; OR = 0.27 (0.11–0.65)	0.017	0.017

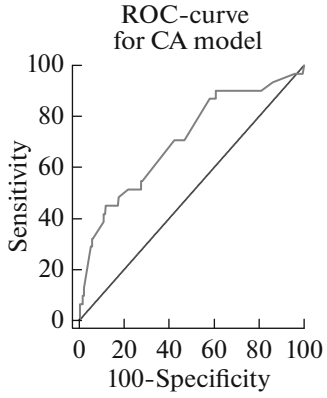
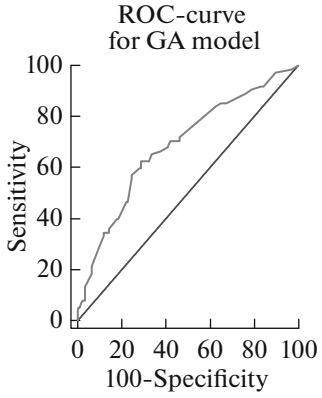
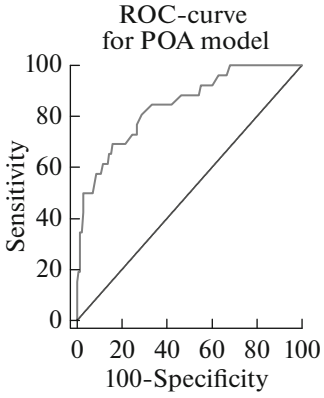
N is the number of individuals. χ^2 is the Pearson chi-square test; *p* is the significance level, indicated only for statistical significance (at $p < 0.05$) or tendency; OR is the odds ratio and 95% confidence interval (in parentheses); *p** is with correction for multiple comparisons (permutation test); OR was calculated only for significant differences between the comparison groups.

deformity, snapping from the temporomandibular joint, the **C*C* genotype of the *rs143383* locus of the *GDF5* gene, the **G*G* genotype of the *rs1544410* locus, and the **G*T* genotype of the *rs7975232* locus of the *VDR* gene. This model had the best prediction consistency and statistical significance. The sensitivity of this model was the highest and amounted to 92.3%; the specificity was somewhat lower—56.5%.

Analyzing the models, we can mention several patterns. The smallest number of markers was included in the model for OA diagnosis of the hip joints, which is probably due to the importance of external factors such as physical activity and increased body weight in the formation of osteoarthritis of this localization. Varicose veins disease included in the model of coxarthrosis and gonarthrosis risk diagnostics reflects modern concepts on the role of congestive venous congestion in the development of pathology of the joints of the lower extremities. Genetic markers in these models represent mainly the loci responsible for the morphology of the connective tissue. The model to diagnose the risk of polyarthrosis includes markers of UCTD and bone metabolism, suggesting a possible role of the subchondral bone pathology in the development of osteoarthritis that affects several joints.

Thus, a comprehensive analysis of the associations of alleles, genotypes, and haplotypes of polymorphic loci of metalloproteinases (*MMP1*, *MMP3*, *MMP13*), type II collagen (*COL2A1*), growth/differentiation factor 5 (*GDF5*), and vitamin D receptor (*VDR*) genes with osteoarthritis in women taking into account the localization of the pathological process, the presence of UCTD signs in general and its specific phenotypic signs, and ethnicity of patients has been performed. Using clinical and genetic data with multivariable logistic regression, statistically significant models to predict the development of osteoarthritis of the knee joints, hip joints, and multiple localizations have been identified. The most studied candidate genes were significant in the development of OA and UCTD; how-

Table 5. Clinical genetic models of OA risk prediction

Coxarthrosis	Gonarthrosis	Polyarthrosis
Kyphosis/lordosis Varicose veins GORD Ptoses *MMP13* *COL2A1*	UCTD Varicose veins GORD Kyphosis/lordosis COL2A1 GDF5	UCTD GORD Severe myopia Kyphosis/lordosis Gallbladder deformities Snapping from the supramaxillary joint *VDR* *GDF5*
		
$\chi^2 = 16.93$; $p = 0.039$; AUC = 0.686	$\chi^2 = 23.43$; $p = 0.009$; AUC = 0.707	$\chi^2 = 41.89$; $p = 0.000079$; AUC = 0.842

χ^2 is the Pearson chi-square test; p is the achieved significance level (statistically significant at $p < 0.05$); AUC is area under curve.

ever, it is necessary to validate the results considering the power of the study in order to obtain definitive conclusions about the contribution of each candidate gene to the development of isolated and comorbid cases of these pathologies.

DISCUSSION

The available literature contains quite a large number of studies devoted to the search for the genetic basis of OA pathogenesis, whereas UCTD has been insufficiently addressed and the genetic markers have been identified only for monogenic forms of CTD.

In the present study, the greatest significance of polymorphic variants of the vitamin D receptor gene (*VDR*) were revealed, which were associated with the symptom complex of UCTD, its specific clinical manifestations, and comorbid cases of OA and UCTD and became part of the clinical and genetic prediction models in the analysis of multiple clinical and genetic factors.

The vitamin D receptor (*VDR*) belongs to the family of trans-active regulatory transcription factors, which act as ligand-dependent transcription factors for several genes related to the secretion of parathyroid hormone, calcium-phosphorus mineral metabolism, and growth and differentiation of cells and is similar to the receptors of steroid and thyroid hormones [14]. As a transcription factor, it exerts considerable influence on the homeostasis of the human body. According to the published data, the **B* (**A*) allele of the *Bsm*I locus of the *VDR* gene was associated with osteophyte severity in spondylarthritis in the British population [15]. In female patients older than 50 years with severe gonarthrosis from Russia, the accumulation of the **B*B* (**A*A*) genotype of the *Bsm*I locus of the *VDR* gene was found compared to the control group [16]. According to our findings, the **G* allele of the *Bsm*I (*rs1544410*) locus is a risk allele, which does not agree with the indicated published data. The **b*b* (**G*G*) genotype of the *Bsm*I polymorphic variant was more frequently found in individuals with congenital aplasia of the knee joint in an Italian population [17], which is partly consistent with the results of our studies. The role of the *VDR* gene polymorphic variants with the formation of phenotypic UCTD signs was studied mainly by Russian scientists. An association of the *Bsm*I and *Tag*I polymorphic variants with the size of the eye was established, which together with external factors can contribute to development of myopia [18]. On the basis of our findings, the **G* allele and the **G*G* genotype of the *rs1544410*

(*Bsm*I) locus is associated with an increased risk of UCTD development.

Among the OA candidate genes which we studied, the role of the collagen type II gene in OA and UCTD phenotype development is addressed in the largest number of publications in the literature. In the study of the *rs3737548* and *rs2276455* polymorphisms of the *COL2A1* gene in women from Finland suffering from POA, it was shown that carriage of at least one minor allele **A* of the *s2276455* polymorphism enhances the risk of development of the disease; the **TG* haplotype of these polymorphic variants was also found to be a risk factor for POA development [19]. Moreover, no association of the *COL2A1* gene polymorphisms with OA in general was shown, or with its radiologic characteristics in Belgian postmenopausal women suffering from OA of the knee joints [20]. Our study also found no statistically significant associations of the *rs2276455* polymorphic variant with the risk of OA development, while the **T* allele of the *rs63118460* locus of the *COL2A1* gene was a significant predictor of OA development of the knee (gonarthrosis) and hip (coxarthrosis) joints in clinical and genetic prediction models.

In the study of the significance of the *COL2A1* gene polymorphic variants in the development of certain manifestations of UCTD, the significance of the *rs1635529* polymorphism of the *COL2A1* gene in the development of myopia in patients from Europe [21] and the *rs1793949* polymorphic variant of the *COL2A1* gene in the development of cleft palate in children from the Baltic States [22] was established. In this study, the *COL2A1* gene loci were also associated with certain manifestations of UCTD; however, the low power does not allow making unambiguous conclusions on the significant role of the collagen type II gene in the formation of the symptom complex of UCTD.

The *GDF5* gene is one of the genes with proven association with OA in genome-wide association studies (GWAS) [23]. When studying the functional features of its polymorphic variants, it was revealed that the **T* allele of the *rs143383* locus, which is located in the promoter region of the gene, reduces the expression of the growth/differentiation factor 5, which probably influences cartilage and connective tissue metabolism in general. According to a meta-analysis, the **C* allele and the **C*C* genotype of the *rs143383* polymorphic variant of the *GDF5* gene demonstrate a weak but statistically significant association with OA, and the association was more significant in Asian populations than in European ones [24]. On the basis of the results of our studies, the *rs143383* polymorphic variant of the *GDF5* gene within a prediction model was a predictor of osteoarthritis development, which is consistent with the published data.

The available literature also contains publications on the association of the *GDF5* polymorphic gene variants with different symptoms of UCTD [25, 26]. In this study, the *rs143383* polymorphic variant of the *GDF5* gene is associated with the development of myopia, but the association shows an insignificant power.

Matrix metalloproteinases genes (*MMP*) encode the enzymes of the extracellular matrix degeneration. The *rs1799750*, *rs35068180*, and *rs2252070* loci under study localize in the regulatory regions of the *MMP1*, *MMP3*, and *MMP13* genes, respectively, and influence the expression of these genes. Matrix metalloproteinase genes are involved in catabolic processes in the connective tissue and are actively studied in connection with various manifestations of UCTD and OA; however, unambiguous results still have not been obtained. An association of minor alleles of the *rs1799750* and *rs35068180* polymorphisms of the *MMP1* and *MMP3* genes, respectively, with development of GORD was established [27]. Our study revealed a tendency in the association of the *MMP1* and *MMP13* gene polymorphisms with certain phenotypic manifestations of UCTD.

Thus, the role of metalloproteinases, collagen type II, growth/differentiation factor 5 (*GDF5*), and vitamin D receptor genes, separately and in their combinations, in the formation of predisposition to OA and the symptom complex of UCTD remains to be concluded. Our study is one of the stages of the search for genetic markers of these pathologies and determines the relevance of further studies, taking into account these findings and current trends in the study of the molecular-genetic basis of multifactor diseases.

ACKNOWLEDGMENTS

This study was supported by the Russian Federation for Basic Research, grant no. 16-04-01487_a.

REFERENCES

1. Bruyère, O., Cooper, C., Arden, N., et al., Can we identify patients with high risk of osteoarthritis progression who will respond to treatment? A focus on epidemiology and phenotype of osteoarthritis, *Drugs Aging*, 2015, vol. 32, no. 3, pp. 179–187. doi 10.1007/s40266-015-0243-3
2. Alekseeva, L.I. and Tsvetkova, E.S., Osteoarthrosis: from the past to the future, *Nauchno-Prakt. Revmatol.*, 2009, no. 2, pp. 31–37.
3. Alekseenko, E.Yu. and Govorin, A.V., Clinical manifestations of connective tissue dysplasia in patients with osteoarthritis, *Kuban. Nauch. Med. Vestn.*, 2009, no. 6, pp. 7–9.
4. Zemtsovskii, E.V., Undifferentiated connective tissue dysplasia: a new interpretation of the concept, *Med. Vestn. Sev. Kavk.*, 2008, vol. 2, no. 10, pp. 8–14.
5. Kadurina, T.I. and Gorbunova, V.N., Differentiated and undifferentiated connective tissue dysplasia, in *Displaziya soedinitel'noi tkani: rukovodstvo dlya vrachei* (Connective Tissue Dysplasia: A Guide for Physicians), St. Petersburg: ELBI, 2009, pp. 29–33.

6. *Nasledstvennye narusheniya soedinitel'noi tkani v kardiologii: diagnostika i lechenie* (Hereditary Disorders of Connective Tissue in Cardiology: Diagnosis and Treatment), *Ross. Kardiol. Zh.*, 2013, no. 1(99), suppl. 1.

7. National recommendations of the Russian Scientific Medical Society of Physicians for the diagnosis, treatment, and rehabilitation of patients with dysplasia of connective tissue, *Med. Vestn. Sev. Kavk.*, 2016, vol. 11, no. 1, pp. 2–76.

8. Zhai, G., Hart, D.J., Kato, B.S., et al., Genetic influence on the progression of radiographic knee osteoarthritis: a longitudinal twin study, *Osteoarthr. Cartil.*, 2007, vol. 15, no. 2, pp. 222–225.

9. Ishimori, M.L., Altman, R.D., Cohen, M.J., et al., Heritability patterns in hand osteoarthritis: the role of osteophytes, *Arthritis Res. Ther.*, 2010, vol. 12, p. R180. doi doi 10.1186/ar3144

10. Zhai, G., van Meurs, J.B.J., Livshits, G., et al., A genome-wide association study suggests that a locus within the ataxin 2 binding protein 1 gene is associated with hand osteoarthritis: the Treat-OA consortium, *J. Med. Genet.*, 2009, vol. 46, no. 9, pp. 614–616. doi 10.1136/jmg.2009.067314

11. Panoutsopoulou, K., Southam, L., Elliott, K.S., et al., Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study, *Ann. Rheum. Dis.*, 2011, vol. 70, pp. 864–867. doi 10.1136/ard.2010.141473

12. Nasonova, V.A. and Astapenko, M.G., *Klinicheskaya revmatologiya: rukovodstvo dlya vrachei* (Clinical Rheumatology: A Guide for Physicians), Moscow: Meditsina, 1989.

13. Mathew, C.C., The isolation of high molecular weight eukaryotic DNA, *Methods Mol. Biol.*, 1985, vol. 2, pp. 31–34. doi doi 10.1385/0-89603-064-4:31

14. Barthel, T.K., Mathern, D.R., Whitfield, G.K., et al., 25-Dihydroxyvitamin D3/VDR-mediated induction of FGF23 as well as transcriptional control of other bone anabolic and catabolic genes that orchestrate the regulation of phosphate and calcium mineral metabolism, *J. Steroid Biochem. Mol. Biol.*, 2007, vol. 103, nos. 3–5, pp. 381–388. doi 10.1016/j.jsbmb.2006.12.054

15. Jordan, K.M., Syddall, H., Dennison, E.M., et al., Birthweight, vitamin D receptor gene polymorphism, and risk of lumbar spine osteoarthritis, *J. Rheumatol.*, 2005, vol. 32, no. 4, pp. 678–683.

16. Krylov, M.Yu., Myakotkin, V.A., Kolesnik, T.V., and Alekseeva, L.I., *VDR* gene polymorphism in osteoarthritis of the knee, *Nauchno-Prakt. Revmatol.*, 2006, no. 5, pp. 15–19.

17. Granchi, D., Stea, S., Sudanese, A., et al., Association of two gene polymorphisms with osteoarthritis secondary to hip dysplasia, *Clin. Orthopaed. Relat. Res.*, 2002, vol. 403, pp. 108–117.

18. Bannur, R., Voitovich, A.N., and Larionova, V.I., The role of the receptor for vitamin D and its genetic polymorphism in predicting the course of myopia in children, *Oftal'mol. Vedomosti*, 2010, no. 3, pp. 27–33.

19. Hämäläinen, S., Solovieva, S., Hirvonen, A., et al., COL2A1 gene polymorphisms and susceptibility to osteoarthritis of the hand in Finnish women, *Ann. Rheum. Dis.*, 2009, vol. 68, no. 10, pp. 1633–1637. doi 10.1136/ard.2008.094706

20. Baldwin, C.T., Cupples, L.A., Joost, O., et al., Absence of linkage or association for osteoarthritis with the vitamin D receptor/type II collagen locus: the Framingham Osteoarthritis Study, *J. Rheumatol.*, 2002, vol. 29, no. 1, pp. 161–165.

21. Metlapally, R., Li, Y.J., Tran-Viet, K.N., et al., COL1A1 and COL2A1 genes and myopia susceptibility: evidence of association and suggestive linkage to the COL2A1 locus, *Invest. Ophthalmol. Vis. Sci.*, 2009, vol. 50, no. 9, pp. 4080–4086. doi 10.1167/iovs.08-3346

22. Nikopensius, T., Jagomagi, T., Krjutskov, K., et al., Genetic variants in COL2A1, COL11A2, and IRF6 contribute risk to nonsyndromic cleft palate, *Birth Defects Res., Part A*, 2010, vol. 88, no. 9, pp. 748–756. doi 10.1002/bdra.20700

23. Kerkhof, H.J.M., Lories, R.J., Meulenbelt, I., et al., A Genome-Wide Association Study identifies a locus on chromosome 7q22 to influence susceptibility for osteoarthritis, *Arthritis Rheum.*, 2010, vol. 62, no. 2, pp. 499–510. doi 10.1002/art.27184

24. Pan, F., Tian, J., Winzenberg, T., et al., Association between GDF5 rs143383 polymorphism and knee osteoarthritis: an updated meta-analysis based on 23995 subjects, *BMC Musculoskelet. Disord.*, 2014, vol. 15, pp. 404–415. doi 10.1186/1471-2474-15-404

25. Xiao, J.L., Meng, J.H., Gan, Y.H., et al., Association of GDF5, SMAD3 and RUNX2 polymorphisms with temporomandibular joint osteoarthritis in female Han Chinese, *J. Oral Rehabil.*, 2015, vol. 42, no. 7, pp. 529–536. doi 10.1111/joor.12286

26. Williams, F.M., Popham, M., Hart, D.J., et al., GDF5 single-nucleotide polymorphism rs143383 is associated with lumbar disc degeneration in Northern European women, *Arthritis Rheum.*, 2011, vol. 63, no. 3, pp. 708–712. doi 10.1002/art.30169

27. Cheung, W.Y., Zhai, R., Bradbury, P., et al., Single nucleotide polymorphisms in the matrix metalloproteinase gene family and the frequency and duration of gastroesophageal reflux disease influence the risk of esophageal adenocarcinoma, *Int. J. Cancer*, 2012, vol. 131, no. 11, pp. 2478–2486. doi 10.1002/ijc.27541

Translated by M. Novikova