DOI: https://doi.org/10.17816/onco642735 EDN: IKCGJW

Genetic Predisposition to Cervical Cancer and Prevalence of Oncogenic HPV Types in Female Population of the Republic of Bashkortostan

Ksenia V. Lenkova[1], Gulnara Z. Lyalina[2], Raushania K. Minyazeva[2], Vita L. Akhmetova[3], Bulat I. Yalaev[4], Irina R. Gilyazova[1, 2], Rita I. Khusainova[1, 4, 5], Ildar R. Minniakhmetov[4]

[1] Ufa Branch of the Russian Academy of Sciences, Ufa, Russia;
[2] Bashkir State Medical University, Ufa, Russia;
[3] Ufa University of Science and Technology, Ufa, Russia;
[4] Endocrinology Research Centre, Moscow, Russia;
[5] Saint-Petersburg State University, Saint Petersburg, Russia

ABSTRACT

BACKGROUND: Cervical cancer is a significant medical and social issue that greatly impacts women's quality of life and life expectancy. The primary pathogenic factor is a persistent infection with a high-risk human papillomavirus (HPV) types.

AIM: The study aimed to evaluate the prevalence and spectrum of HPV infections in the female population of the Republic of Bashkortostan, as well as to determine their molecular and genetic predisposition to cervical cancer.

METHODS: The study included 219 randomly selected HPV-positive samples to evaluate the prevalence and spectrum of HPV types. Four years later, a follow-up screening was conducted in 70 HPV-positive women. The screening included an evaluation of their HPV and cervical statuses, an examination by an obstetrician/gynecologist, cytology, colposcopy, and biopsy for histology when indicated. In the third stage, we compared polymorphisms of the *CLPTM1L* (rs27069), *PAX8* (rs10175462), and *CDC42* (rs2268177) genes in patients with histologically confirmed cervical cancer and in apparently healthy women.

RESULTS: No positive correlation was found between the number of HPV types per sample and viral load, nor between viral load and HPV clearance. Genome-wide association studies (GWAS) identified statistically significant associations between cervical cancer risk and the G allele of *CLPTM1L* rs27069 (χ^2=4.098; p=0.043), as well as with the T allele (χ^2=16.99; p=3.751e-5) and the TT genotype (χ^2=17.35; p=0.0002) of *CDC42* rs2268177. No association was found for *PAX8* (rs10175462).

CONCLUSION: This was the first Russian study to replicate GWAS results for cervical cancer. Associations were identified for *CLPTM1L* (rs27069) and *CDC42* (rs2268177), but not for *PAX8* (rs10175462). The results highlight the need for further research to confirm these associations and improve our understanding of the molecular mechanisms associated with cervical cancer risk and HPV persistence.

Keywords: human papillomavirus; HPV; cervical cancer; high-risk HPV types; genetic predisposition.

To cite this article:
Lenkova KV, Lyalina GZ, Minyazeva RK, Akhmetova VL, Yalaev BI, Gilyazova IR, Khusainova RI, Minniakhmetov IR. Genetic Predisposition to Cervical Cancer and Prevalence of Oncogenic HPV Types in Female Population of the Republic of Bashkortostan. *Russian Journal of Oncology.* 2025;30(1):5–16. DOI: 10.17816/onco642735 EDN: IKCGJW

Submitted: 09.12.2024 **Accepted:** 24.03.2025 **Published online:** 27.03.2025

DOI: https://doi.org/10.17816/onco642735 EDN: IKCGJW

Поиск генов предрасположенности к раку шейки матки с оценкой распространённости онкогенных типов вируса папилломы человека у женщин из Республики Башкортостан

К.В. Ленкова[1], Г.З. Лялина[2], Р.К. Минязева[2], В.Л. Ахметова[3], Б.И. Ялаев[4], И.Р. Гилязова[1, 2], Р.И. Хусаинова[1, 4, 5], И.Р. Минниахметов[4]

[1] Уфимский федеральный исследовательский центр Российской академии наук, Уфа, Россия;
[2] Башкирский государственный медицинский университет, Уфа, Россия;
[3] Уфимский университет науки и технологий, Уфа, Россия;
[4] Национальный медицинский исследовательский центр эндокринологии, Москва, Россия;
[5] Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

АННОТАЦИЯ

Обоснование. Рак шейки матки представляет собой серьёзную медико-социальную проблему, значительно снижающую качество и продолжительность жизни женщин. Ключевой фактор патогенеза данного заболевания — персистирующая инфекция высокоонкогенных типов вируса папилломы человека.

Цель — изучение спектра и распространённости ВПЧ-инфекции среди женщин в Республике Башкортостан, определение у них молекулярно-генетической предрасположенности к раку шейки матки.

Материалы и методы. В исследовании принимали участие 219 ВПЧ-положительных образцов, случайно отобранных в рамках скрининга, для анализа спектра и частоты типов ВПЧ. Повторный скрининг был выполнен спустя 4 года и охватил 70 ВПЧ-позитивных женщин с целью определения их ВПЧ-статуса, оценки состояния шейки матки акушером-гинекологом, проведения цитологического анализа и, при необходимости, кольпоскопии и биопсии для гистологического исследования. На третьем этапе был проведён сравнительный анализ полиморфизмов генов *CLPTM1L* (rs27069), *PAX8* (rs10175462) и *CDC42* (rs2268177) между группами пациенток с гистологически подтверждённым диагнозом и условно здоровыми женщинами.

Результаты. Прямой корреляции между количеством типов ВПЧ в образце и вирусной нагрузкой, а также между вирусной нагрузкой и элиминацией ВПЧ выявлено не было. В исследовании полиморфных вариантов, ассоциированных с риском развития рака шейки матки в рамках GWAS, были реплицированы статистически значимые ассоциации с рисковым аллелем G локуса rs27069 гена *CLPTM1L* (χ^2=4,098; *p*=0,043) и с рисковым аллелем T (χ^2=16,99; *p*=3,751e-005) и генотипом TT (χ^2=17,35; *p*=0,0002) локуса rs2268177 гена *CDC42*. В то же время, для полиморфизма rs10175462 гена *PAX8* ассоциации не выявлено.

Заключение. В рамках данного исследования нами впервые в России был проведён репликативный анализ результатов GWAS по раку шейки матки. Выявлены ассоциации для полиморфных вариантов rs27069 гена *CLPTM1L* и rs2268177 гена *CDC42*, в то время как для rs10175462 гена *PAX8* ассоциаций не обнаружено. Результаты подчёркивают необходимость дальнейших исследований для подтверждения ассоциаций и расширения понимания молекулярных механизмов, связанных с риском развития рака шейки матки и персистенцией ВПЧ-инфекции.

Ключевые слова: вирус папилломы человека; ВПЧ; рак шейки матки; онкогенные типы ВПЧ; генетическая предрасположенность.

Как цитировать:
Ленкова К.В., Лялина Г.З., Минязева Р.К., Ахметова В.Л., Ялаев Б.И., Гилязова И.Р., Хусаинова Р.И., Минниахметов И.Р. Поиск генов предрасположенности к раку шейки матки с оценкой распространённости онкогенных типов вируса папилломы человека у женщин из Республики Башкортостан // Российский онкологический журнал. 2025. Т. 30, № 1. С. 5–16. DOI: 10.17816/onco642735 EDN: IKCGJW

Рукопись получена: 09.12.2024 **Рукопись одобрена:** 24.03.2025 **Опубликована online:** 27.03.2025

BACKGROUND

Cervical cancer (CC) is one of the most common oncological diseases in women. The main predictor of CC and other anogenital carcinomas remains the human papillomavirus (HPV), which acts in combination with genetic and epigenetic factors [1, 2]. A key role in the pathogenesis of CC belongs to persistent high-risk HPV (HR HPV) infection [3, 4], which accounts for more than 99% of CC cases [5]. At present, it has been established that the correlation between HPV infection and CC is significantly stronger than that between smoking status and lung cancer [6]. Urogenital HPV infection represents a major medical and social problem, not only in terms of clinical and epidemiological impact but also with regard to demographic consequences [7].

HPV is the most common sexually transmitted infection worldwide [5]. HPV is a large family of double-stranded DNA viruses, comprising five genera (α, β, γ, μ, and ν), 48 species, and 206 types. The classification most relevant for clinical practice is based on oncogenic potential. It includes 13 HR HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68; IARC Groups 1 and 2A); 14 additional high-risk types (HPV 5, 26, 53, 66, 67, 68, 70, 73, 82, 30, 34, 69, 85, and 97; IARC Group 2B); and low-risk types (e.g., HPV 6 and 11; IARC Group 3) [8]. Globally, the five most prevalent HPV types are 16, 18, 31, 58, and 52, which together account for 50% of all HPV infections, with types 16 and 18 responsible for up to 70% of all cases [9]. The contribution of individual HPV types to cervical carcinogenesis varies. It is assumed that HPV genotype is an independent prognostic factor for cervical precancerous lesions (dysplasia), whereas for cervical cancer the key prognostic factor is the extent of tumor spread [10].

Despite the availability of local studies in the scientific data assessing the prevalence of HPV infection and its individual types in Russia, including larger-scale works such as that by Donnikov et al. [11], an important task remains not only to evaluate the prevalence of infection in Russia and its individual regions but also to monitor the course of infection in at-risk patients with consideration of their clinical data. In the absence of reliable prognostic markers for cervical cancer (CC), except for the carriage of HR HPV types, the search for genetic markers of susceptibility to CC is of particular relevance. Such markers, along with HPV infection, would allow for improved prevention, diagnosis, and disease monitoring. A recent large genome-wide association study (GWAS), which analyzed more than 9 million single nucleotide polymorphisms (SNPs), identified genetic susceptibility to CC associated with genes involved in the regulation of apoptosis and cellular immune response, as well as associations with independent loci [12]. Our study included polymorphic variants of the *CLPTM1L* (rs27069), *PAX8* (rs10175462), and *CDC42* (rs2268177) genes, selected based on scientific data on GWAS of CC. The selection criteria for the polymorphisms were their involvement in tumorigenesis, both for CC and for other malignancies, as well as their location outside *HLA* genes. The loci were chosen based on their mention in recent large publications on the subject. The polymorphisms *CLPTM1L* (rs27069) and *PAX8* (rs10175462) were selected based on one of the first major studies in this field [13]. Subsequently, to expand coverage and ensure reliability of the data, we added another locus, *CDC42* (rs2268177), which was described in a recent GWAS meta-analysis [14]. The approach to locus selection corresponds to current criteria for association analysis in genetic epidemiological studies [15], ensuring the reliability and scientific significance of our findings.

In Russia, according to our review of the scientific data, no GWAS replication studies on genetic predisposition to HPV-associated CC have previously been conducted. Therefore, the identification of prognostically significant risk markers for CC remains a pressing task requiring further research.

The study aimed to determine the prevalence of persistent HPV infection, to identify the most common oncogenic HPV types in women from the Republic of Bashkortostan, and to assess predisposition to cervical cancer based on the analysis of polymorphic variants of the *CLPTM1L* (rs27069), *PAX8* (rs10175462), and *CDC42* (rs2268177) genes.

METHODS

Study Design

The study was based on cervical epithelial cell scrape samples from 28,928 women aged 30–39 years, examined as part of a pilot cervical cancer screening project (2019) using HPV testing [16].

In the first stage, the spectrum and prevalence of HPV types were determined in women from the Republic of Bashkortostan.

In the second stage, a case group was formed, including 111 women with CC, along with a reference group.

In the third stage, associations of the polymorphic loci *CLPTM1L* (rs27069), *PAX8* (rs10175462), and *CDC42* (rs2268177) with the risk of HPV-associated CC were analyzed.

Eligibility Criteria and Subgroup Analysis

The case group consisted of 111 women with a clinical diagnosis of CC, unrelated by blood and residing in the Republic of Bashkortostan. HPV status was assumed to be positive by default. The median age of the patients was 50 years. All patients received treatment according to current clinical guidelines. The most common

histological tumor type in the cohort was squamous cell carcinoma (89.11%).

The sample with documented HPV clearance included women (51 individuals) aged 30–39 years, unrelated by blood and residing in the Republic of Bashkortostan. The group of apparently healthy women consisted of a cohort of 333 individuals, unrelated by blood and with no history of CC.

Women across all groups represented the three most common ethnicities in this region: Russians, Tatars, and Bashkirs. Sample size was not pre-calculated.

Study Setting and Duration

The study was conducted in stages from March 2019 to March 2024, beginning with the HPV screening pilot project and concluding with the analysis of associations between polymorphic variants and susceptibility to HPV-related CC.

Enrollment, questionnaires, clinical examinations, and collection of biological samples from women with documented HPV clearance were carried out at women's health clinics in Ufa in 2019, during the HPV screening pilot project, and again in 2023 to determine HPV persistence status.

Examinations, diagnosis, treatment, and collection of biological material from patients with CC were conducted between November 2020 and January 2022 at the State Autonomous Healthcare Institution Republican Clinical Oncological Dispensary of the Ministry of Health of the Republic of Bashkortostan (Ufa, Russia). At the same site, cytological analyses of cervical canal epithelial scrapings, colposcopies, biopsies, and histological examinations of biopsy material were performed.

HPV genotyping was carried out at the Center for Molecular Medicine, Ufa University of Science and Technology. DNA extraction from peripheral blood and polymorphism analyses were performed at the Laboratory of Human Genetics, Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences (Ufa, Russia).

Intervention and Outcomes Registration

At the screening stage, HPV testing was performed using the Hybrid Capture method for *in vitro* detection of 13 HR HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68), with determination of the overall viral load but without HPV type identification (Digene HPV Test, Hybrid Capture Technology, Qiagen, USA). The measurement unit for this method is RLU/COV, expressed as the luminescence intensity measured by a luminometer and reported in relative light units. Samples with an RLU/COV ratio ⩾ 1.0 were considered positive, whereas those with a ratio < 1.0 were considered negative.

To determine the spectrum and prevalence of HPV types, 219 HPV-positive samples with viral loads ranging from 1.45 to 2891.05 RLU/COV were randomly selected. Genotyping was performed using real-time PCR on a DTprime amplifier (DNA Technology, Russia) with the HPVquant-21 assay kit (DNA Technology, Russia). This method allows detection of human papillomavirus DNA of both low-risk types (HPV 6, 11, 44) and high-risk types (HPV 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, 82).

In 2023, to determine HPV persistence status, 70 women with previously positive HPV test results were invited to attend an outpatient gynecological visit at their primary healthcare facility. HPV status was assessed using cervical epithelial cell scrapings with the AmpliPrime HR HPV 14 Genotypes kit (Next BIO, Russia) on a CFX96 detector (Bio-Rad, USA). This reagent kit was chosen due to the similarity of the detectable HPV types to those assessed by the Digene test previously employed in the pilot HPV screening project. As a result of this examination, 51 women with documented HPV clearance were identified. This cohort was included in the reference group for analyzing associations between selected candidate gene loci and CC predisposition. In 19 women, HR HPV persisted for four years without CC development; these patients were not included in the molecular genetic analysis and remain under gynecological follow-up.

Association analysis of the selected polymorphic loci was carried out in a cohort of 495 women. The sample included 111 patients with cervical cancer, 51 women with HPV clearance and no cervical conditions, and 333 apparently healthy women. Genomic DNA was extracted from peripheral blood lymphocytes using the phenol–chloroform extraction method (Mathew, 1984). Genotyping of the polymorphic variants of the studied genes was performed using reagent kits for the detection of polymorphic variants by real-time PCR (DNA-Synthesis, Russia), according to the manufacturer's protocol.

Ethics Approval

The study protocol was approved by the Biomedical Ethics Committee of the Institute of Biochemistry and Genetics, Federal State Budgetary Scientific Institution Ufa Federal Research Center of the Russian Academy of Sciences (protocol No. 19, November 25, 2021). All patients provided written informed consent to participate in the study.

Statistical Analysis

For statistical data processing, the web resource PLINK, which provides a set of tools for genome-wide association analysis, was used.* The analysis was carried out using Microsoft Windows–based tools, including Excel and

* zzz.bwh.harvard.edu [Internet]. PLINK. Whole genome association analysis toolset. Available from: https://zzz.bwh.harvard.edu/plink/data.shtml

Notepad (Microsoft Office 2010, USA). A standard (basic) case–control association test based on Pearson's chi-square criterion was applied to identify associations. The degree of association was evaluated using odds ratio (OR) calculated using the following formula:

$$OR = (a \times d) / (b \times c), \quad (1)$$

where a is the frequency of the trait in the patient group; b is the frequency of the trait in the control group; c is the sum of the frequencies of all other traits in the patient group; and d is the sum of the frequencies of all other traits in the control group. Tests were performed with a two-sided significance level, and differences were considered statistically significant at $p < 0.05$. Since the objective of the study was to identify associations and their trends, we considered the use of a combination of two approaches (the basic association test and logistic regression analysis) sufficient to achieve the study goals. Logistic regression analysis was performed using MedCalc software (v. 22.016) (MedCalc Software Ltd., Belgium).

RESULTS

The primary age range of women examined within the 2019 pilot project was 30–39 years (96.1%). The mean viral load in this age group was 287.9 RLU/COV, corresponding to a high viral load [16]. Seventeen HR HPV types were detected in infected patients, with the number of simultaneously identified viral types ranging from 1 (17.6%) to 11 (0.5%) (Fig. 1; Tables 1 and 2). It was established that more than 50% of HPV-infected women carried at least three viral types. No direct correlation was observed between viral load and the number of HPV types in a sample. For instance, the viral load of a woman carrying 11 HPV types was 1810.73 RLU/COV, whereas there were cases of women infected with only one HPV type whose viral load exceeded 2197.56. Women carrying 10 HPV types had viral loads ranging from 181.08 to 1625.05 RLU/COV.

Among the 219 women included in the study, HPV type 16 was detected in 59.8% (131 patients), consistent with the scientific data indicating that this type is the most prevalent worldwide and in the studied region [17]. In contrast, HPV type 18, which is widely distributed globally, was found in only 23.7% of infected women (52 patients) in the Republic of Bashkortostan. HPV type 51 and type 56 were identified in 40.1% of women (88 patients) and in 38.8% of women (85 patients), respectively. Rare HR HPV types such as 53, 73, and 82 were detected in fewer than 10% of patients, whereas HPV type 26 was not identified (see Table 1). Low-risk HPV types 6 and 44 were detected in 2.7% and 5.5% of patients, respectively, whereas HPV type 11 was not identified in the study sample.

Among the 70 women whose viral load at the time of the 2019 screening ranged from 1.51 to 2024.23 RLU/COV, HPV clearance occurred in 51 women, whereas viral persistence was observed in 19 women. The mean viral load in the clearance group was 429.56 RLU/COV, whereas in the persistence group it averaged 307.57 RLU/COV. Thus, at the time of the pilot screening project, the mean viral load in the persistence group was lower than in the clearance group, indicating no direct effect of viral load on the likelihood of infection clearance.

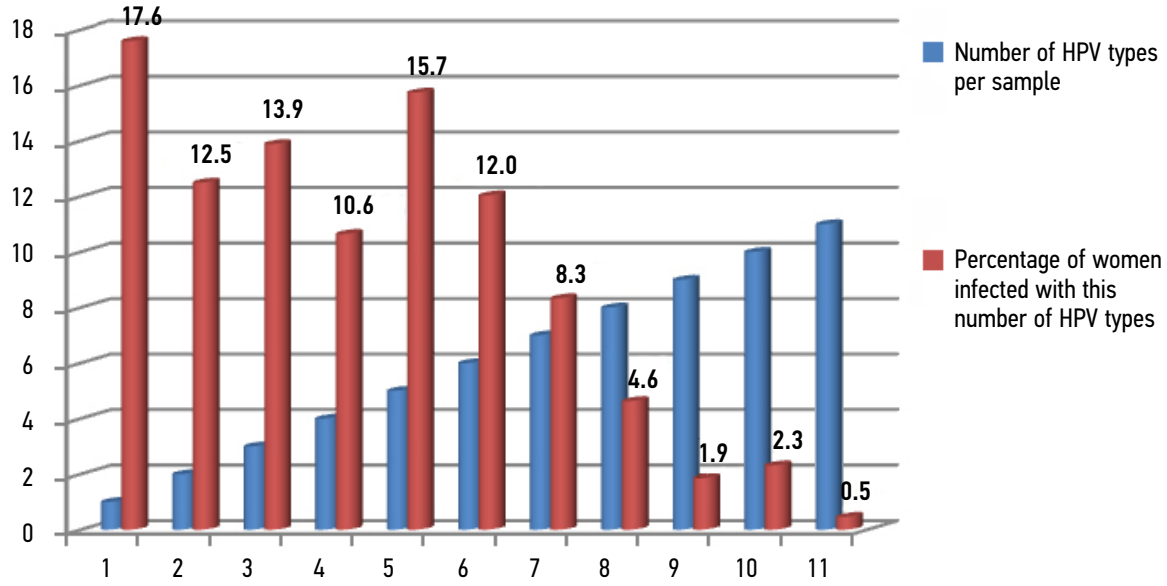

Fig 1. Number of human papillomavirus types simultaneously detected in women from the Republic of Bashkortostan.

Four years after the screening, HPV clearance was observed in 72% of women (51 of 70), which is somewhat lower compared to published data. According to WHO reports, in 80%–90% of HPV-infected women the virus clears spontaneously within an average of 1–2 years, whereas in only 10%–20% of patients the virus persists for a longer period [18].

In addition, at this stage of the study, all women invited for follow-up underwent repeat cytological examination of cervical smears to detect atypia of the stratified squamous non-keratinizing epithelium. No cytological abnormalities were found in 63 of 70 patients, whereas atypical cells were identified in 7. Cytological findings were assessed according to the Bethesda classification. All 7 women were tested positive for HPV at the time of examination. Three patients had atypical squamous cells of undetermined significance (ASC-US), with viral loads in 2019 ranging from 28.15 to 528.82 RLU/COV. In one patient, with a viral load of 1756.17 RLU/COV (2019 data), mild dysplasia (LSIL) was detected. In another patient, with a viral load of 5.15 RLU/COV (2019 data), atypical squamous cells could not exclude high-grade squamous intraepithelial lesion (ASC-H). All five women require further follow-up. High-grade squamous intraepithelial lesion (HSIL) was diagnosed in two of the 70 patients: in one case, the diagnosis was confirmed histologically (CIN II, moderate dysplasia), and in the other case, histology revealed LSIL. Thus, cytological examination revealed squamous epithelial cell atypia in 7 of 19 women (36.8%) from the high-risk group with persistent HPV infection; however, no cases of CC were identified.

The group of 70 women re-examined in 2024 included 13 patients who had previously undergone HPV genotyping in 2019 using the HPV-quant-21 kit. Four of them had a single HPV type: in one patient, HPV type 58 persisted and led to severe dysplasia, whereas in three patients the virus cleared spontaneously. Among two women whose samples initially contained two HPV types, in one case the virus cleared, and in the other it led to cervical cancer (HPV types 16 and 51 identified in 2019). The

Table 1. Spectrum and prevalence frequencies of human papillomavirus types in HPV-positive women from the Republic of Bashkortostan

No.	HPV type	Frequency, %	No.	HPV type	Frequency, %
High-risk					
1	16	131 (59.8%)	10	35	49 (22.3%)
2	51	88 (40.1%)	11	66	38 (17.3%)
3	56	85 (38.8%)	12	68	32 (14.6%)
4	31	72 (32.8%)	13	45	29 (13.2%)
5	39	67 (30.6%)	14	52	29 (13.2%)
6	58	61 (27.8%)	15	73	16 (7.3%)
7	33	60 (27.4%)	16	53	14 (6.4%)
8	59	57 (26%)	17	82	5 (2.3%)
9	18	52 (23.7%)	18	26	0
Low-risk					
1	44	12 (5.5%)	–	–	–
2	6	6 (2.7%)	–	–	–
3	11	0	–	–	–

Table 2. Results of Hardy–Weinberg equilibrium analysis for genotypes rs27069, rs10175462 and rs2268177

SNP	Minor allele	H_{obs}	H_{pred}	HW_{pval}
rs27069	A	0.450	0.458	0.768
rs10175462	A	0.493	0.491	1
rs2268177	T	0.370	0.391	0.248

Note. These results reflect data from a sample of 495 individuals, including 111 patients with cervical cancer and 384 conditionally healthy women. H_{obs} — observed heterozygosity; H_{pred} — expected heterozygosity; HW_{pval} — ***p*** value for assessing compliance with Hardy–Weinberg equilibrium.

patient underwent treatment, but at the 2024 follow-up, despite normal cytology, HPV type 16 was again detected. Three patients co-infected with four HPV types and one patient with five HPV types demonstrated complete viral clearance. Among three patients with six HPV types, viral persistence with mild dysplasia was documented in one, whereas spontaneous clearance occurred in the other two.

The results of the analysis of the selected polymorphic variants of genes are presented in Tables 3–5. All polymorphic variants were in Hardy–Weinberg equilibrium (see Table 2). The combined group of women with HPV clearance and apparently healthy women was used as the reference group for patients with CC. Thus, the study included three groups: patients with CC (N = 111), women with HPV clearance (N = 51), and apparently healthy women (N = 333). A sample of 162 individuals (the CC group and the HPV clearance group) was defined as the small cohort, whereas a sample of 495 individuals (the CC group and the combined group of women with HPV clearance and apparently healthy women) was defined as the extended cohort. At the first stage of the association analysis of the *CLPTM1L* (rs27069), *PAX8* (rs10175462), and *CDC42* (rs2268177) variants, different trends toward association between the reference groups were identified in the studied population. These findings are consistent with previously reported international data. To increase the reliability of the results, we decided to expand the control group by adding 333 women without a clinical diagnosis of cervical cancer.

In the primary analysis of the small cohort, no association was found for the rs27069 locus of the *CLPTM1L* gene with the G allele (p = 0.7649) or with the GG genotype (p = 0.067). Logistic regression analysis revealed a trend toward association for the AG genotype (p = 0.043). The comparative analysis of allele and genotype frequencies between the groups in the extended sample showed that the G allele of the rs27069 polymorphic locus of the *CLPTM1L* gene was associated with CC risk (χ^2 = 4.098; OR = 1.395; CI = 1.01–1.926; p = 0.043), although no association was found with the GG genotype (p = 0.124) (see Table 3).

For the rs10175462 locus of the *PAX8* gene, an association trend was observed for the G allele (p = 0.056; OR = 1.58; χ^2 = 3.663) in the small sample,

Table 3. Results of comparative analysis of allele and genotype frequency distribution of rs27069 alleles and genotypes of *CLTP1L* gene between the comparison groups

Alleles							
SNP	Putative risk allele	Allele frequency in CC	Healthy women	χ^2	OR	95% CI	p
rs27069 (N=162)	G	0.703	0.686	0.089	1.081	0.650–1.795	0.765
rs27069 (N=495)	G	0.703	0.629	4.098	1.395	1.01–1.926	0.043

Genotypes (N = 162)								
rs27069	N	Number of genotypes			Genotype frequencies			χ^2. p
Genotypes	162	AA	AG	GG	AA	AG	GG	
CC	111	9	48	54	0.081	0.432	0.486	χ^2 = 5.409 p = 0.067
HPV clearance	51	9	14	28	0.176	0.275	0.549	

Logistic regression for small sample (N = 162)			
Locus	Regression coefficient	Deviance residuals	p
rs27069 — *CLPTM1L* — *A/G*	1.262	–1.678–1.235	0.043
rs27069 — *CLPTM1L* — *G/G*	0.848		0.141

Genotypes (N = 495)								
rs27069	N	Number of genotypes			Genotype frequencies			χ^2. p
Genotypes	495	AA	AG	GG	AA	AG	GG	
CC	111	9	48	54	0.081	0.432	0.486	χ^2 = 4.172 p = 0.124
Apparently healthy women	384	55	175	154	0.143	0.455	0.401	

Note: The group of 162 participants included 111 women with CC and 51 women with HPV clearance, whereas the group of 495 participants included 111 women with CC and 384 apparently healthy women. SNP, single nucleotide polymorphism; CC, cervical cancer; OR, odds ratio; CI, confidence interval.

which was confirmed by logistic regression analysis for the homozygous GG genotype (p = 0.028). In the extended sample, however, no association was found for the G allele (p = 0.217) or the GG homozygous genotype (p = 0.463) (Table 4).

In the comparative analysis of allele and genotype frequencies of the rs2268177 locus of the *CDC42* gene, the small sample demonstrated a trend toward association with the risk allele T (p = 0.054; OR = 1.659; χ^2 = 3.714). In the comparative analysis of allele and genotype frequencies of the rs2268177 locus of the *CDC42* gene in the extended sample, a statistically significant association was revealed for the risk allele T (χ^2 = 16.99; OR = 1.945; CI = 1.413–2.676; p = 3.751e-005) and for the TT genotype (χ^2 = 17.35; p = 0.0002) (see Table 5).

Thus, the analysis of polymorphic variants previously associated with CC risk in GWAS replicated statistically significant associations with the risk allele G of the rs27069 locus of the *CLPTM1L* gene and with the risk allele T and TT genotype of the rs2268177 locus of the *CDC42* gene; however, no association was found for the rs10175462 polymorphism of the *PAX8* gene.

DISCUSSION

According to the scientific data, the proportion of HPV-positive women varies by age group. It amounts to approximately 30% in women under 25 years, around 12% in those aged 25–34, about 6% in those aged 35–44, and less than 5% in women aged 45–65 [19].

Thus, our findings demonstrate an average prevalence level of HPV among women aged 30–39 years, which amounted to 10.6%. HPV type 16 was detected in 59.8% of women, which, according to WHO, is the most common and high-risk type. HPV type 18 was identified in only 23.7% of women, while HPV types 51 and 56 were found in 40.1% and 38.8% of cases, respectively. Moreover, it was established that more than 50% of HPV-infected women carried at least three types of the virus, whereas approximately 17% had only one type. The sample also included a woman infected with 11 HPV types, whose viral load was 1810.73 RLU/COV, which was lower than that of a woman infected with a single HPV type, whose viral load was 2197.56 RLU/COV. No direct correlation between the number of HPV types and viral load was identified.

Table 4. Results of comparative analysis of allele and genotype frequency distribution of the rs10175462 locus of the PAX8 gene between the comparison groups

Alleles							
SNP	Putative risk allele	Allele frequency in CC	Healthy women	χ^2	OR	95% CI	p
rs10175462 (N=162)	G	0.603	0.490	3.663	1.584	0.987–2.540	0.056
rs10175462 (N=489)	G	0.603	0.557	1.526	1.212	0.893–1.643	0.217

Genotypes (N = 162)							
rs10175462	N	Number of genotypes			Genotype frequencies			χ^2. p
Генотипы	162	AA	AG	GG	AA	AG	GG	χ^2=4.514 p=0.105
РШМ	111	17	54	40	0.153	0.486	0.360	
Элиминация ВПЧ	51	15	22	14	0.294	0.431	0.274	

Logistic regression for small sample (N = 162)			
Locus	Regression coefficient	Deviance residuals	p
rs10175462 — *PAX8* — *A/G*	0.7242	–1.7272–1.1483	0.118
rs10175462 — *PAX8* — *G/G*	1.16777		0.028

Genotypes (N = 489)							
rs10175462	N	Number of genotypes			Genotype frequencies			χ^2. p
Генотипы	489	AA	AG	GG	AA	AG	GG	χ^2=1.54p p=0.463
РШМ	111	17	54	40	0.153	0.486	0.360	
Условно здоровые женщины	378	74	187	117	0.196	0.495	0.309	

Note: The group of 162 participants included 111 women with CC and 51 women with HPV clearance, whereas the group of 489 participants included 111 women with CC and 378 apparently healthy women. SNP, single nucleotide polymorphism; CC, cervical cancer; OR, odds ratio; CI, confidence interval.

Table 5. Results of the comparative analysis of the distribution of allele and genotype frequencies of the rs2268177 locus of the CDC42 gene between study groups

Alleles							
SNP	Putative risk allele	Allele frequency in CC	Healthy women	χ^2	OR	95% CI	*p*
rs2268177 (N=162)	T	0.374	0.265	3.714	1.659	0.989–2.782	0.054
rs2268117 (N=495)	T	0.374	0.235	16.99	1.945	1.413–2.676	3.751e-005

Genotypes (N = 162)								
rs2268177	N	Number of genotypes			Genotype frequencies			χ^2. *p*
Genotypes	162	TT	AT	AA	TT	AT	AA	χ^2 = N/A *p* = N/A
CC	111	18	47	46	0.162	0.423	0.414	
HPV clearance	51	4	19	28	0.078	0.372	0.549	

Logistic regression for small sample (N = 162)			
Locus	Regression coefficient	Deviance residuals	*p*
rs2268177 — *CDC42* — *A/T*	0.5857	–1.6651–1.0051	0.147
rs2268177 — *CDC42* — *T/T*	0.6788		0.286

Genotypes (N = 492)								
rs2268177	N	Number of genotypes			Genotype frequencies			χ^2. *p*
Genotypes	492	TT	AT	AA	TT	AT	AA	χ^2 = 17.35 *p* = 0.0002
CC	111	18	47	46	0.162	0.423	0.414	
Apparently healthy women	381	22	135	224	0.058	0.354	0.588	

Note: The group of 162 participants included 111 women with CC and 51 women with HPV clearance, whereas the group of 492 participants included 111 women with CC and 381 apparently healthy women. SNP, single nucleotide polymorphism; CC, cervical cancer; OR, odds ratio; CI, confidence interval.

In the cohort of women who were re-examined 4–5 years after a positive HPV test within the pilot screening project, no association was found between viral load and HPV clearance. The findings indicate that, for this region, it is not reasonable to use a screening panel focused solely on detecting HPV types 16 and 18. We also concluded that none of the currently existing and widely used diagnostic methods can serve as a reliable prognostic marker of disease progression prior to the onset of clinical manifestations.

Over the past years, several studies have been conducted to identify genetic markers of CC. Of particular interest is the comparison of international findings with the results obtained in the present study. According to Rashkin et al., rs10175462 of the *PAX8* gene on 2q13 was the first single-nucleotide substitution to show a genome-wide significant association with cervical cancer risk outside the *HLA* region in the European population (OR = 1.15; $p = 7.71 \times 10^{-14}$) [20]. The rs10175462 locus is located in the intronic region of the *PAX8* gene at position 2:113230915 (GRCh38). This single-nucleotide substitution demonstrated an association with squamous cell CC (OR = 0.80; 95% CI, 0.68–0.94; $p = 0.006$) and with other diseases at this site. RNA expression analysis in cervical samples revealed activation of *PAX8* gene transcripts in HPV-positive specimens ($p = 0.008$), whereas no activation was observed in the presence of the protective minor allele rs10175462 [21]. Although no significant associations were identified in the extended sample including both apparently healthy women and patients with CC, we observed a tendency toward an association in the group of women with documented HPV clearance for the G allele and GG genotype of the rs10175462 locus of the *PAX8* gene. In our view, this cohort is not predisposed to HPV persistence and, consequently, to HPV-associated CC. Thus, the findings are consistent with the studies by Ramachandran et al. and Rashkin et al., where the G allele is considered a risk variant [13, 20].

The rs27069 locus, located at genomic coordinate 5:1347013 (GRCh38), lies 10 kilobases (kb) upstream of

the *CLPTM1L* gene, which encodes a membrane protein involved in apoptosis [13]. A meta-analysis of studies of this locus demonstrated its significant association with cervical cancer in the European population ($p = 6.1 \times 10^{-15}$), as well as in a mixed population ($p = 1.3 \times 10^{-14}$) [14]. Approximately 50 kb from this locus is the *TERT* gene, which encodes human telomerase reverse transcriptase. The *CLPTM1L–TERT* locus has been associated with certain forms of gynecological neoplasms; however, the functional role of this substitution and its contribution to CC pathogenesis require further investigation [13]. *CLPTM1L*/CRR9 (cisplatin resistance-related protein 9) is a cytoprotective oncofetal protein expressed on the surface of tumor cells. *CLPTM1L* expression is observed on the plasma membrane of ovarian tumor cells, whereas in normal tissues this protein is virtually not expressed [22]. It should be noted that our findings are consistent with those of Bowden et al., who conducted a GWAS involving 273,377 women of European ancestry aged 40–69 years, including 4769 patients with grade 3 cervical intraepithelial neoplasia (CIN) or invasive CC [12]. A genome-wide association analysis covering 9 million polymorphic variants demonstrated that the T allele of the *CLPTM1L* rs27069 locus ($p = 2.51 \times 10^{-9}$; OR = 0.88; 95% CI, 0.84–0.92) reduces the risk of CC, whereas the common allele was identified as risk-associated, which is consistent with our findings [12]. Koel et al. performed a GWAS meta-analysis characterizing the genetic architecture of cervical phenotypes based on data from 9229 patients with CC and 490,304 controls in European and mixed populations (Meta-analysis of UK, FinnGen, Japanese RIKEN, and Estonian biobanks) [14]. The results showed that the rs27069 locus of the *CLPTM1L* gene is associated with CC ($p = 1.3 \times 10^{-14}$). In our study, the G allele of this locus reached statistical significance for association with CC ($p = 0.043$; OR = 1.395; 95% CI, 1.01–1.926), which supports the findings of Koel et al.

The rs2268177 locus is located in the intronic region of the *CDC42* gene (chr1:22088917 [GRCh38.p14]). It has been established that *CDC42* expression is significantly upregulated in CC tissues, and this protein promotes cell migration and metastasis [23]. A meta-analysis of studies on this locus demonstrated its significant association with CC both in European ($p = 3.8 \times 10^{-8}$) and mixed populations ($p = 3.1 \times 10^{-8}$) [14]. Koel's group reported that the T allele of rs2268177, located in the intron of the *CDC42* gene, is associated with CC ($p = 3.8 \times 10^{-8}$; OR = 1.12, 95% CI, 1.07–1.16). In our study, this locus was also associated with the risk allele T ($p = 3.751 \times 10^{-5}$; OR = 1.945, 95% CI, 1.413–2.676) and with the TT genotype ($p = 0.0002$; $\chi^2 = 17.35$) [14].

As a result of the analysis of the *CLPTM1L* (rs27069), *PAX8* (rs10175462), and *CDC42* (rs2268177) polymorphic variants, statistically significant associations were identified with the risk allele G of the rs27069 polymorphic locus of the *CLPTM1L* gene ($\chi^2 = 4.098$; $p = 0.043$), the risk allele T ($\chi^2 = 16.99$; $p = 3.751e-005$), and the TT genotype ($\chi^2 = 17.35$; $p = 0.00017$) of the rs2268177 locus of the *CDC42* gene. At the same time, no association was found with the rs10175462 polymorphism of the *PAX8* gene. Despite the absence of statistically significant associations for rs10175462 in the extended sample, which included both apparently healthy women and patients with CC, we observed a tendency toward association for the G allele in the group of women with documented HPV clearance ($p = 0.056$; OR = 1.58; $\chi^2 = 3.663$). This association was further confirmed by logistic regression analysis for the homozygous GG genotype ($p = 0.028$). These findings may indicate that certain risk polymorphic variants identified in GWAS (e.g., rs10175462 of the *PAX8* gene) may manifest differently across various ethnic groups and populations. Our results emphasize the importance of replicating GWAS findings in populations not included in the original studies as a necessary step to validate associations and investigate risk single-nucleotide substitutions.

Study Limitations

Several limitations should be taken into account when interpreting the obtained results. First, due to the relatively small sample size (111 women with cervical cancer and 51 women with HPV clearance), the statistical power of our study is limited, which may affect the generalizability of the findings. Second, the focus solely on spontaneous HPV clearance leaves unanswered the question of other contributing factors to CC development, such as immune status or the condition of the vaginal microbiome. These limitations highlight the need for further research to confirm and refine our conclusions.

CONCLUSION

Cervical cancer is a multifactorial malignancy. The search for novel molecular and genetic approaches to its diagnosis and prognosis is a key strategy for reducing morbidity. GWAS provide powerful tools for identifying single nucleotide polymorphisms associated with CC risk; however, replication in ethnic groups not included in the original studies remains an important task for further research.

ADDITIONAL INFORMATION

Author contributions: All the authors confirm that their authorship meets the ICMJE criteria (all authors made substantial contributions to the conceptualization, investigation, and manuscript preparation, and reviewed and approved the final version prior to publication). K.V. Lenkova: investigation, writing—original draft; G.Z. Lyalina, R.K. Minyazeva: data curation, investigation; V.L. Akhmetova, I.R. Gilyazova: investigation; B.I. Yalaev: formal analysis; R.I. Khusainova, I.R. Minniahmetov: conceptualization, writing—original draft.

Ethics approval: The study protocol was approved by the Biomedical Ethics Committee of the Institute of Biochemistry and Genetics, Federal State Budgetary Scientific Institution Ufa Federal Research Center of the Russian Academy of Sciences (protocol No. 19, November 25, 2021).
Funding sources: This work was supported by Saint Petersburg State University, project IDPURE: 103964756.

Disclosure of interests: The authors have no relationships, activities, or interests for the last three years related to for-profit or not-for-profit third parties whose interests may be affected by the content of the article, or other relationships, activities, or interests for the last three years that must be declared.

REFERENCES | СПИСОК ЛИТЕРАТУРЫ

1. Revathidevi S, Murugan AK, Nakaoka H, et al. APOBEC: A molecular driver in cervical cancer pathogenesis. *Cancer Letters*. 2021;496:104–116. doi: 10.1016/j.canlet.2020.10.004
2. Sohail F, Samar D, Anser IR, et al. Topic: Genomic characterization of Human Papillomavirus and pathogenesis leading to cervical cancer. *Global Scientific Journals*. 2021;9(10):55–70.
3. Cruz-Gregorio A, Aranda-Rivera AK, Pedraza-Chaverri J. Human Papillomavirus-related Cancers and Mitochondria. *Virus Research*. 2020;286:198016. doi: 10.1016/j.virusres.2020.198016
4. Pal A, Kundu R. Human Papillomavirus E6 and E7: The Cervical Cancer Hallmarks and Targets for Therapy. *Frontiers in Microbiology*. 2020;10. doi: 10.3389/fmicb.2019.03116
5. Yuan Y, Cai X, Shen F, Ma F. HPV post-infection microenvironment and cervical cancer. *Cancer Letters*. 2021;497:243–254. doi: 10.1016/j.canlet.2020.10.034
6. Okunade KS. Human papillomavirus and cervical cancer. *Journal of Obstetrics and Gynaecology*. 2020;40(5):602–608. doi: 10.1080/01443615.2019.1634030
7. Nesterov AS, Savchkov GV. Species spectrum of STI pathogens in women with genital papillomavirus infection. In: *Nexus Medicus: Actual problems of modern medicine: Proceedings of the All-Russian scientific and practical conference with international participation*. Ulyanovsk: Ulyanovsk State University, 2021. P. 314–316. (In Russ.) EDN PIXCLC
8. Hu Z, Ma D. The precision prevention and therapy of HPV-related cervical cancer: new concepts and clinical implications. *Cancer Medicine*. 2018;7(10):5217–5236. doi: 10.1002/cam4.1501
9. Pimple S, Mishra G. Cancer cervix: Epidemiology and disease burden. *CytoJournal*. 2022;19:21. doi: 10.25259/CMAS_03_02_2021
10. Churuksaeva ON, Kolomiets LA, Ibragimova MK, Litvyakov NV. Specificities of high-risk papillomavirus infection in patients with locally advanced cervical cancer. *Gynecology, Obstetrics and Perinatology*. 2019;18(3):21–28. doi: 10.20953/1726-1678-2019-3-21-28
11. Donnikov AE, Markelov MI, Pestrikova TY, et al. Analysis of the prevalence and viral load of different human papillomavirus types in the regions of the Russian Federation. *Obstetrics and Gynecology*. 2019;4:39–47. doi: 10.18565/aig.2019.4.39-47
12. Bowden SJ, Bodinier B, Kalliala I, et al. Genetic variation in cervical preinvasive and invasive disease: a genome-wide association study. *The Lancet Oncology*. 2021;22(4):548–557. doi: 10.1016/S1470-2045(21)00028-0
13. Ramachandran D, Dörk T. Genomic Risk Factors for Cervical Cancer. *Cancers*. 2021;13(20):5137. doi: 10.3390/cancers13205137
14. Koel M, Võsa U, Jõeloo M, et al. GWAS meta-analyses clarify the genetics of cervical phenotypes and inform risk stratification for cervical cancer. *Human Molecular Genetics*. 2023;32(12):2103–2116. doi: 10.1093/hmg/ddad043
15. Ponomarenko IV. Selection of polymorphic loci for association analysis in genetic-epidemiological studies. Research Result. *Medicine and Pharmacy*. 2018;4(2):40–54. doi: 10.18413/2313-8955-2018-4-2-0-5
16. Minniakhmetov I, Zabelin M, Olkov I, Khusainova R. Pilot project for screening cervical cancer using HPV testing. *Voprosy Onkologii*. 2020;66(6):618–624. doi: 10.37469/0507-3758-2020-66-6-618-624
17. Yashchuk AG, Rakhmatullina IR, Zainullina RM, et al. The diversity of human papillomavirus types in the female population of Ufa. *Bashkortostan Medical Journal*. 2019;14(2):25–29. (In Russ.) EDN: DMJGJA
18. World Health Organization. *Comprehensive Cervical Cancer Control. A guide to essential practice Second edition*. Geneva: World Health Organization; 2014.
19. Belokrinitskaya TE, Frolova NI, Turanova OV, et al. Results of human papillomavirus testing on self-collected versus clinician-collected samples. *Gynecology*. 2017;19(1):56–62. (In Russ.) EDN: YZGCOX
20. Rashkin SR, Graff RE, Kachuri L, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun*. 2020;11:4423. doi: 10.1038/s41467-020-18246-6
21. Ramachandran D, Wang Y, Schürmann P, et al. Association of genomic variants at PAX8 and PBX2 with cervical cancer risk. *International Journal of Cancer*. 2021;149(4);893–900. doi: 10.1002/ijc.33614
22. Parashar D, Geethadevi A, McAllister D, et al. Targeted biologic inhibition of both tumor cell-intrinsic and intercellular CLPTM1L/CRR9-mediated chemotherapeutic drug resistance. *npj Precis. Onc*. 2021;5:16. doi: 10.1038/s41698-021-00152-9
23. Ye H, Zhang Y, Geng L, Li Z. Cdc42 expression in cervical cancer and its effects on cervical tumor invasion and migration. *Int J Oncol*. 2015;46(2):757–763. doi: 10.3892/ijo.2014.2748

AUTHORS' INFO

*** Ksenia V. Lenkova;**
address: 71 October ave, Ufa, Russia, 450054;
ORCID: 0000-0001-5903-0085;
eLibrary SPIN: 8509-7901;
e-mail: ms.kv.kl@mail.ru

Gulnara Z. Lyalina;
ORCID: 0000-0001-7540-733X;
eLibrary SPIN: 7602-3416;
e-mail: davlet_g@mail.ru

Raushania K. Minyazeva
ORCID: 0000-0001-5542-9531;
eLibrary SPIN: 9660-0317;
e-mail: dr.gubaydullina@mail.ru

Vita L. Akhmetova, Cand. Sci. (Biology);
ORCID: 0000-0002-1874-0774;
eLibrary SPIN: 9716-8937;
e-mail: vita-akh@mail.ru

Bulat I. Yalaev, Cand. Sci. (Biology);
ORCID: 0000-0003-4337-1736;
eLibrary SPIN: 2546-1425;
e-mail: yalaev.bulat@endocrincentr.ru

Irina R. Gilyazova, Cand. Sci. (Biology);
ORCID: 0000-0001-9499-5632;
eLibrary SPIN: 5799-4821;
e-mail: gilyasova_irina@mail.ru

Rita I. Khusainova, Dr. Sci. (Biology);
ORCID: 0000-0002-8643-850X;
eLibrary SPIN: 4091-9326;
e-mail: khusainova.rita@endocrincentr.ru

Ildar. R. Minniakhmetov, Cand. Sci. (Biology);
ORCID: 0000-0002-7045-8215;
eLibrary SPIN: 8643-7056;
e-mail: minniakhmetov.ildar@endocrincentr.ru

ОБ АВТОРАХ

*** Ленкова Ксения Вячеславовна;**
адрес: Россия, 450054, Уфа, пр-кт Октября, д. 71;
ORCID: 0000-0001-5903-0085;
eLibrary SPIN: 8509-7901;
e-mail: ms.kv.kl@mail.ru

Лялина Гульнара Зилфировна;
ORCID: 0000-0001-7540-733X;
eLibrary SPIN: 7602-3416;
e-mail: davlet_g@mail.ru

Минязева Раушания Каримовна;
ORCID: 0000-0001-5542-9531;
eLibrary SPIN: 9660-0317;
e-mail: dr.gubaydullina@mail.ru

Ахметова Вита Леоновна, канд. биол. наук;
ORCID: 0000-0002-1874-0774;
eLibrary SPIN: 9716-8937;
e-mail: vita-akh@mail.ru

Ялаев Булат Илдусович, канд. биол. наук;
ORCID: 0000-0003-4337-1736;
eLibrary SPIN: 2546-1425;
e-mail: yalaev.bulat@endocrincentr.ru

Гилязова Ирина Ришатовна, канд. биол. наук;
ORCID: 0000-0001-9499-5632;
eLibrary SPIN: 5799-4821;
e-mail: gilyasova_irina@mail.ru

Хусаинова Рита Игоревна, д-р биол. наук;
ORCID: 0000-0002-8643-850X;
eLibrary SPIN: 4091-9326;
e-mail: khusainova.rita@endocrincentr.ru

Минниахметов Илдар Рамилевич, канд. биол. наук;
ORCID: 0000-0002-7045-8215;
eLibrary SPIN: 8643-7056;
e-mail: minniakhmetov.ildar@endocrincentr.ru

* Corresponding author / Автор, ответственный за переписку