SMJ Saratov Medical Journal
ISSN 2712-8253, Saratov Medical Journal
2026. Volume 7. Issue 3. Article CID e0305
DOI: 10.15275/sarmj.2026.0305

Original article

Reprint

SNP-PCa test system: Allele-specific real-time PCR for genotyping prostate cancer risk markers

Roman N. Fomkin, Timofey E. Pylaev, Vladimir M. Popkov ✉, Georgy G. Mezirov, Olga A. Fomkina, Alla B. Bucharskaya, Natalya V. Polukonova
✉ popkov.vm@staff.sgmu.ru

Razumovsky State Medical University of Saratov, Saratov, Russia

Received 20 July, 2026, Accepted 29 August 2026

© This article is an open access publication. Russian Text. Published in *Saratov Journal of Medical Scientific Research*, 2026; 22(3): 327-336. https://doi.org/10.15275/ssmj2203327. ISSN 1995-0039.

Abstract:
Objective: To develop a prototype of an allele-specific real-time polymerase chain reaction (PCR) for genotyping four single-nucleotide polymorphisms (SNPs) associated with prostate cancer risk and to evaluate its laboratory characteristics.
Materials and methods: Twelve oligonucleotides were synthesized: four universal forward and eight allele-specific reverse primers. A control series of 12 genomic deoxyribonucleic acid preparations was verified by Sanger sequencing. Pilot reactions were assessed by the threshold cycle, curve shape, melting temperature, and electrophoresis. Extended testing of three loci was performed on 154 samples.
Results: Sanger sequencing confirmed 10 of the 12 archival PCR genotypes, followed by restriction fragment length polymorphism analysis. The pilot characteristics were locus-dependent; research work on the *HNF1B* component is ongoing. In the extended series, a genotype assignment was established for *CASC17, CASC8* and *CCAT2* in 154/154, 154/154, and 153/154 tests, respectively; a complete three-locus profile was formed in 153/154 samples (99.35%; 95% confidence interval 96.44-99.98; p=0.0034 relative to the 95% benchmark value). The *CASC17* and *CCAT2* genotypes are consistent with the Hardy–Weinberg equilibrium; a significant deviation was revealed for *CASC8* (p=0.0008).
Conclusion: In the extended series, 3 of 4 prototype components ensured an interpretable genotyping result in 99.78% of tests. Sequence verification, repeatability, and reproducibility assessment are required to complete the analytical evaluation of the test system.

Keywords: prostate cancer, single-nucleotide polymorphism, genotyping, real-time polymerase chain reaction, Hardy-Weinberg equilibrium.

Cite as: Fomkin RN, Pylaev TE, Popkov VM, Mezirov GG, Fomkina OA, Bucharskaya AB, Polukonova NV. SNP-PCa test system: Allele-specific real-time PCR for genotyping prostate cancer risk markers. *Saratov Medical Journal* 2026; 7(3): e0305. https://doi.org/10.15275/sarmj.2026.0306.

Introduction

Among the male population in Russia, prostate cancer (PCa) constitutes a significant proportion of cancer incidence [1]. Hereditary predisposition to the disease is polygenic in nature: the largest trans-ethnic genome-wide association study (GWAS) to date has identified 269 variants independently associated with the risk [2]. Hence, a practical implication for the interpretation of any individual marker is as follows: each marker reflects only a portion of the heritable risk and cannot be considered an independent diagnostic indicator of PCa.

This study included four single-nucleotide polymorphisms (SNPs). The *HNF1B* (HNF1 homeobox B) variant rs4430796 is located in the 17q12 region, and its association with PCa risk has been confirmed by pooled genetic studies [3]. The polymorphism rs1859962, located in the 17q24 region, is designated as *CASC17* (cancer susceptibility 17) in the nomenclature used here; the association of this variant with the disease has been evaluated in a systematic review and meta-analysis [4]. *CASC8* (cancer susceptibility 8) variant rs1447295 and *CCAT2* (colon cancer-associated transcript 2) variant rs6983267 are located in the 8q24 region, which contains several independent loci for hereditary susceptibility to PCa [5]. For the 8q24 variants, population-based heterogeneity in both allele frequencies and effect sizes has been described, both of which must be taken into account when developing genetic risk panels [2, 6].

These four loci were previously investigated by our research group using polymerase chain reaction (PCR) with a subsequent restriction fragment length polymorphism (RFLP) analysis [7]. The corresponding combination of markers is incorporated into a patented method of risk prediction [8]. In this article, the PCR-RFLP results are used only for between-method comparison, while the new data pertain to the sequence-based verification of control samples, the development of allele-specific real-time PCR (qPCR), and its laboratory validation.

Allele-specific PCR moves genotyping to an instrumentally recorded format; however, the quality of allelic discrimination depends on primer sequences, their 3′

© Fomkin R.N., Pylaev T.E., Popkov V.M., Mezirov G.G., Fomkina O.A., Bucharskaya A.B., Polukonova N.V., 2026

end stability, annealing temperatures, reaction mixture composition, and locus-specific characteristics [9]. The current MIQE 2.0 (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines require a transparent description of sample preparation, control reactions, the criteria for determining the threshold cycle, technical replicates, and reproducibility [10]. That is why, during the development, it is desirable to evaluate operational performance separately from analytical accuracy and clinical validity.

We assumed that transitioning the four selected markers to an allele-specific qPCR format would facilitate obtaining instrumentally interpretable genotyping results, but that the analytical characteristics of the reaction would vary across loci, thereby necessitating locus-specific evaluation criteria.

The *objective* was to develop an allele-specific qPCR prototype for genotyping SNPs associated with the risk of PCa, and to evaluate laboratory performance characteristics of this prototype.

Materials and Methods

Study design and material

We employed a sequential design for our laboratory technology-based study and used materials from 2025-2026, obtained as part of a government procurement project on developing an in vitro diagnostic medical device (SNP-PCa Test System). The units of observation were as follows:

- In the control series, a genomic DNA sample;
- In the pilot series, an individual allele-specific reaction or a paired wild type/mutant type (WT/MT) setup for a single sample (reactions using primer mixes for the reference and alternative alleles, respectively: see the subsection *Allele-Specific Real-Time Polymerase Chain Reaction* for further details;
- In the extended validation phase of the study, the unit was the genotyping of one locus within a single DNA sample.

For the integrated analysis of the three scaled components, the sample was treated as an independent unit.

The control series comprised 12 previously genotyped DNA samples (three for each of locus). Based on archival PCR-RFLP analysis, they were considered heterozygous. The extended series included all 154 available DNA samples, which were tested across three scaled loci: 86 from patients with PCa and 68 from participants in a comparison group. Technically invalid results were not excluded from the denominator. No selection based on clinical characteristics was performed for this technical analysis: disease stage, prostate-specific antigen concentration, and tumor morphological features were not compared in this study. The *HNF1B* component was not included in the extended validation phase of the study; its optimization and analysis are ongoing outside the scope of this publication and will be the subject of a separate report.

Our study was conducted in accordance with the principles of the Declaration of Helsinki, the standards of Good Clinical Practice, and the requirements of the Russian Federation legislation regarding public health, personal data protection, medical confidentiality, and the handling of human genetic data. The study was approved by the Ethics Committee of Razumovsky State Medical University of Saratov, Russia (Protocol No. 7 of January 14, 2025).

Prior to enrollment, all participants signed an informed voluntary consent to their participation, the collection and use of biological material, molecular genetics testing, the processing of personal data and health-related information, and the publication of anonymized, aggregated results. The requirements of Articles 5 and 10 of the Russian Federal Law No. 152-FZ of July 27, 2006, *On Personal Data*, were observed. Information constituting medical confidentiality was not disclosed in an identifiable form and was handled in accordance with Part 3 of Article 13 of the Russian Federal Law No. 323-FZ of November 21, 2011, *On the Basic Principles of Citizens' Health Protection in the Russian Federation*. In compliance with Articles 13.1 and 13.2 of the Russian Federal Law No. 86-FZ of July 5, 1996, *On State Regulation in the Field of Genetic Engineering* (as amended by the Russian Federal Law No. 43-FZ of February 20, 2026), individual genetic data of the participants (such as analysis results, genotypes, and other identifying information) were neither published nor made publicly available, nor were they transferred to foreign entities or states. This publication presents only aggregated, anonymized results of statistical analysis.

Target selection and primer design

We investigated the variants *HNF1B* rs4430796, *CASC17* rs1859962, *CASC8* rs1447295, and *CCAT2* rs6983267. The allelic pairs are presented according to the orientation used in the primer systems, which does not imply a designation of REF (reference allele) or ALT (alternative allele) status as defined in the public Database of Single-Nucleotide Polymorphisms (dbSNP). Reference sequences were obtained from the publicly available GenBank database, maintained by the U.S. National Center for Biotechnology Information (NCBI). Primers were designed using the software tools Primer3Plus and OligoCalc; specificity was verified via the Basic Local Alignment Search Tool (BLAST), while melting temperatures and secondary structures were calculated using OligoAnalyzer 3.1 (Integrated DNA Technologies, USA).

For each locus, we developed one universal forward primer and two allele-specific reverse primers, totaling 12 synthesized oligonucleotides (*Table* 1). The specific sequences and working concentrations of the primers are included in the confidential, protected technical documentation of the product and are not disclosed: the combination of markers is protected by Russian Federation patent RU 2827104 C1 [8].

Table 1. Key characteristics of primer systems

Locus	SNP	Allele pair*	Number of primers	Primer length, nt	Purity, %	Calculated product length, bp
HNF1B	rs4430796	C/T	3	20-22	>98	123
CASC17	rs1859962	A/C	3	20-23	>98	115
CASC8	rs1447295	G/T	3	24	>98	92
CCAT2	rs6983267	G/T	3	20-25	>98	55

* Allele pairs are indicated according to the orientation used in the primer systems; nt, nucleotides; bp, base pairs.

Oligonucleotide synthesis and DNA preparation

Unmodified oligonucleotides were synthesized via the phosphoramidite method on a controlled-pore glass (CPG-500, 500 Å) solid support using a Polygen 12 automated DNA synthesizer (PolyGen GmbH, Germany). The product was separated from the support in a 37% aqueous ammonia solution at 55 °C for 24 hours. After desalting and reprecipitation, the oligonucleotides were purified via denaturing electrophoresis in a 20% polyacrylamide gel. The certified purity exceeded 98%.

Venous blood was collected prior to surgery in 4 mL tubes containing the tripotassium salt of ethylenediaminetetraacetic acid (K_3-EDTA). Within 24 hours, the samples were was distributed into 500 µL aliquots and stored at -20 °C. Genomic DNA was extracted using the D-blood kit (Biolabmix LLC, Russia); DNA concentration and optical purity were measured using a NanoDrop OneC spectrophotometer (Thermo Fisher Scientific, USA) at wavelengths of 230, 260, and 280 nm. The DNA was stored at -20 °C. A concentration of 10 ng/µL was used for pilot experiments; the limit of detection was not determined. Pilot runs used 10 ng/µL. The limit of detection was not determined.

Sanger sequencing-based validation of the control set

For Sanger sequencing, we amplified 453 bp fragments of HNF1B, 317 bp fragments of *CASC17*, and 350 bp fragments of *CCAT2*. Sanger sequencing was performed by Genterra LLC (Russia). The resulting sequences were compared with reference sequences from GenBank. Sequencing was employed as an independent method to verify the genotypes of the control samples.

Allele-specific real-time polymerase chain reaction

For each locus, two parallel allele-specific PCRs were performed using the intercalating fluorescent dye SYBR Green I (Thermo Fisher Scientific, USA). We used the terms “WT” and “MT” with respect to primer mixes designed for the preferential amplification of the reference allele or the alternative allele, respectively. We should point out that these terms were not used to denote normal or mutant genotypes. Each 20 µL reaction mixture contained 5 µL of DNA (10 ng/µL), 10 µL of 2× HS-qSYBR-blue (Biolabmix LLC, Russia), and the appropriate primers. Each reaction was performed in three technical replicates.

Amplification was carried out using a LightCycler 96 qPCR thermal cycler (Roche, Switzerland) according to the following protocol: 95 °C for 2 min, followed by 45 cycles of 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 15 s, with fluorescence detection via the FAM (6-carboxyfluorescein) fluorescence channel. Melting of the products was carried out from 65 to 95 °C in 1 °C increments every 10 sec. The quantification cycle (Cq) was determined at the point where the amplification curve intersected the threshold level, and the melting temperature (*Tm*) was identified by the maximum of the corresponding peak. The term “ΔCq” in our study represents the difference between the quantification cycles; in the pilot series, the absolute difference (|ΔCq|) between the WT and MT reactions for a given sample was analyzed.

During the pilot phase, the presence of a sigmoidal amplification curve, characterized by an obvious exponential phase followed by a plateau, was assessed. This feature served as an exploratory characteristic rather than a formalized validation criterion. In the extended dataset, while the final genotype assignments were retained, the full set of Cq values from the three technical replicates was not; consequently, intra-run and inter-run variabilities were not calculated.

Electrophoretic evaluation of products

PCR products were separated via horizontal electrophoresis in a 2.5% agarose gel in single-shot Tris-acetate-EDTA (TAE) buffer at 130 V for 90 min. The Step50 bp Plus DNA molecular mass marker (50-1,500 bp) was employed, and visualization was performed on a BlueCube-300 gel documentation system (SERVA, Germany). We assessed the presence of a product of the expected length and additional bands as a sign of possible nonspecific amplification.

Endpoints and statistical analysis

The primary technical endpoint for the extended validation was the proportion of locus-level analyses that yielded a valid genotype assignment, defined as an unambiguous biallelic result recorded in the primary laboratory database upon completion of the genotyping workflow. This metric reflects operational success – i.e., the possibility of obtaining an interpretable final genotype assignment rather than analytical accuracy. The operational research criterion was a proportion exceeding 95%; however, this threshold was not a regulatory characteristic of the product.

Secondary endpoints were as follows: concordance between archival genotyping and Sanger sequencing; plateau of pilot reactions and curve compliance with the adopted quality criterion; Cq, *Tm*, and |ΔCq|; presence of the target amplicon and additional electrophoretic bands; concordance between PCR-RFLP and qPCR. For *CASC17*, *CASC8*, and *CCAT2*, the proportion of samples yielding a complete three-locus profile was determined, and the genotypes from the extended dataset were tested for adherence to Hardy-Weinberg equilibrium. PCR-RFLP was not considered a reference method.

Categorical data are presented here as counts and proportions. For the latter, exact two-sided 95% Clopper–Pearson confidence intervals (CI) were calculated. Between-method concordance was assessed using the proportion of exact matches and Cohen’s κ coefficient, interpreted according to the following scale proposed by J.R. Landis and G.G. Koch [11]: <0, poor (concordance worse than chance); 0-0.20, slight; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, substantial; and >0.80, almost perfect. In the pilot series, Cq, *Tm*, and |ΔCq| values were characterized by the median and range, while parametric comparisons and normality testing were not performed, because there were only 2-4 independent paired observations per locus across different genotypic classes.

Compliance of the genotype distribution with Hardy-Weinberg equilibrium was assessed using an asymptotic χ^2 test (with one degree of freedom) and an exact test based on the conditional distribution of the number of heterozygotes, given fixed allele frequencies [12].

We tested the assumption regarding operational performance using an independent sample. For a complete three-locus profile, where π represents the true probability of

obtaining such a profile, the hypotheses H0: $\pi \leq 0.95$ vs. H1: $\pi > 0.95$ were tested using a one-sided exact binomial test ($\alpha = 0.05$). The actual *p*-value and the one-sided 95% lower confidence bound are reported. As only one integrated hypothesis was tested, no adjustments for multiple comparisons were made. Locus-specific metrics were evaluated descriptively, with two-sided 95% CI provided.

The 461/462 locus-level conclusions were used for descriptive purposes only. Three loci were examined within the same 154 DNA samples; hence, these observations were statistically dependent. The sample size was determined by the available set of completed studies. We performed no a priori sample size calculation or retrospective power analysis, and the uncertainty of the estimates was characterized by CIs.

Statistical analyses were performed using *R* version 4.6.0 (R Core Team) within the RStudio Desktop environment (version 2025.09.2+418). Specifically, the exact binomial test and Clopper–Pearson CIs were calculated using functions from the *stats* package; the Hardy–Weinberg equilibrium test was performed using the *HardyWeinberg* package (version 1.7.9); Cohen's κ coefficient was computed based on the formula for a contingency table; and forest plots were generated using the *meta* (version 8.3-0) and *metafor* (version 5.0-1) packages.

Results

Prototype components and control set

For the 4 loci, we synthesized 12 oligonucleotides: 4 universal forward primers and 8 allele-specific reverse primers. The expected product lengths were 123 bp for *HNF1B*, 115 bp for *CASC17*, 92 bp for *CASC8*, and 55 bp for**CCAT2* (*Table* 1). Among 12 samples classified beforehand via PCR-RFLP as heterozygous, Sanger sequencing confirmed the original genotype in 10 cases (83.3%; 95% CI: 51.6-97.9). For *HNF1B*, sample 0037 was reclassified from CT to CC; for *CASC17*, sample 0126 was reclassified from AC to AA; for *CASC8* and *CCAT2*, all three control samples matched the sequencing results (*Table* 2, *Figure* 1).

Table 2. Comparison of archival polymerase chain reaction (PCR) with subsequent restriction fragment length polymorphism (RFLP) analysis and Sanger sequencing in a control series

No.	Sample Code	Locus	PCR-RFLP	Sanger sequencing	Result
1	0024	*HNF1B* rs4430796	CT	CT	Match
2	0037	*HNF1B* rs4430796	CT	CC	Mismatch
3	0044	*HNF1B* rs4430796	CT	CT	Match
4	0101	*CASC17* rs1859962	AC	AC	
5	0108	*CASC17* rs1859962	AC	AC	
6	0126	*CASC17* rs1859962	AC	AA	Mismatch
7	0045	*CASC8* rs1447295	GT	GT	Match
8	0049	*CASC8* rs1447295	GT	GT	
9	0060	*CASC8* rs1447295	GT	GT	
10	0024	*CCAT2* rs6983267	GT	GT	
11	0037	*CCAT2* rs6983267	GT	GT	
12	0044	*CCAT2* rs6983267	GT	GT	

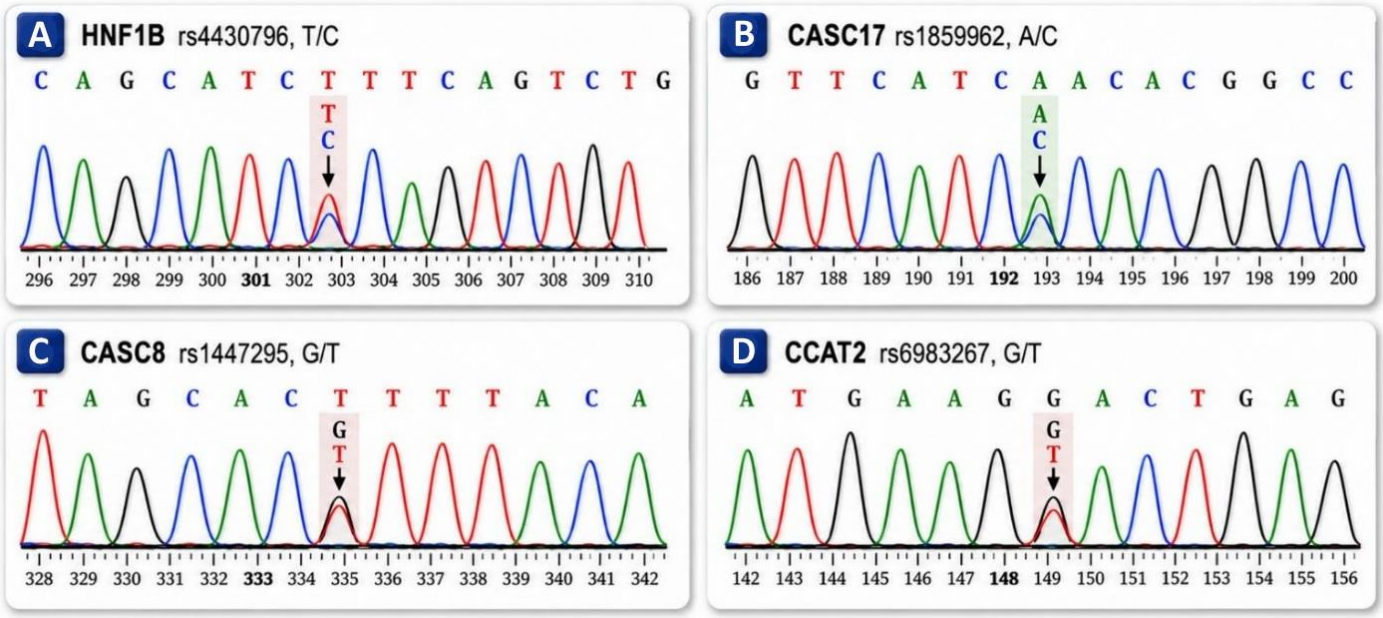

Figure 1. Representative results of Sanger sequencing for control samples of the single-nucleotide polymorphisms under study: A – *HNF1B* rs4430796, genotype T/C; B – *CASC17* rs1859962, genotype A/C; C – *CASC8* rs1447295, genotype G/T; D – *CCAT2* rs6983267, genotype G/T. The nucleotide positions studied are shown by arrows.

Pilot evaluation of four loci

In the *HNF1B* reaction, none of the four runs met the accepted pilot criterion of a full curve reaching a plateau, and additional electrophoretic products were detected in all four reactions. The *Tm* ranged from 82-84 °C, and the median |ΔCq| was 4.5 cycles. For this reason, *HNF1B* was not included in the extended test series; efforts of optimizing amplification conditions (and, if necessary, redesigning primers for this locus) are ongoing and will be addressed separately. For *CASC17*, a plateau and a curve meeting the pilot criterion were observed in 5 out of 8 reactions. Of the 6 reactions subjected to electrophoretic analysis, 4 showed no additional bands. The *Tm* ranged from 79-82 °C, and the median |ΔCq| was 10.5 cycles. For *CASC8*, a plateau was reached in 7 of 8 reactions, and the qualitative curve criterion was met in 6 of 8 reactions. However, only 2 of 6 reactions evaluated via electrophoresis were free of additional bands. The *Tm* ranged from 76-80 °C, and the median |ΔCq| was 3.5 cycles. For *CCAT2*, all 6 pilot reactions reached a plateau and met the established criterion for the curve shape; no additional bands were observed in any of the 6 electropherograms. The *Tm* ranged from 77-79 °C. The median |ΔCq| was 1.0 cycle, with a range of 1-3 cycles (*Table* 3; *Figures* 2 and 3).

Between-method comparison

Exact concordance between PCR-RFLP and allele-specific qPCR was observed for *HNF1B* in 5 out of 22 comparisons (22.7%; 95% CI 7.8-45.4; κ=-0.199), for *CASC17* in 1 out of 20 (5.0%; 95% CI 0.1-24.9; κ=-0.310), for *CASC8* in 22 out of 22 (100%; 95% CI 84.6-100; κ=1.000), and for *CCAT2* in 16 out of 21 (76.2%; 95% CI 52.8-91.8; κ=0.618). According to the Landis-Koch scale [11], the level of concordance for *HNF1B* and *CASC17* is poor, for *CCAT2* it is substantial, and for *CASC8* it is almost perfect. While these metrics reflect inter-method consistency, they do not constitute an assessment of the diagnostic accuracy of the new assay (*Table* 4).

Extended laboratory validation

In the extended series for *CASC17*, a final genotype assignment was obtained in 154 out of 154 analyses (100%; 95% CI 97.63–100); for *CASC8*, the result was also 154 out of 154 (100%; 95% CI 97.63-100); and for *CCAT2*, it was 153 out of 154 (99.35%; 95% CI 96.44-99.98). In the laboratory dataset, one sample lacked a final *CCAT2* assignment; based on the available records, it was not possible to definitively reconstruct the technical reason for this case.

Table 3. Summary characteristics of pilot reactions

Locus	Reaction records, *n*	Reached plateau, *n/N* (%)	Curve meeting pilot criteria, *n/N* (%)	*Tm*, °C	Without additional bands, *n/N* (%)	|ΔCq|, median (range)
HNF1B	4	0/4 (0.0)	0/4 (0.0)	82-84	0/4 (0.0)	4.5 (1-8)
CASC17	8	5/8 (62.5)	5/8 (62.5)	79-82	4/6 (66.7)	10.5 (1-11)
CASC8	8	7/8 (87.5)	6/8 (75.0)	76-80	2/6 (33.3)	3.5 (2-7)
CCAT2	6	6/6 (100)	6/6 (100)	77079	6/6 (100)	1.0 (1-3)

n is the number of reactions with the observed characteristic; *N* is the number of reactions for which the corresponding assessment was available.

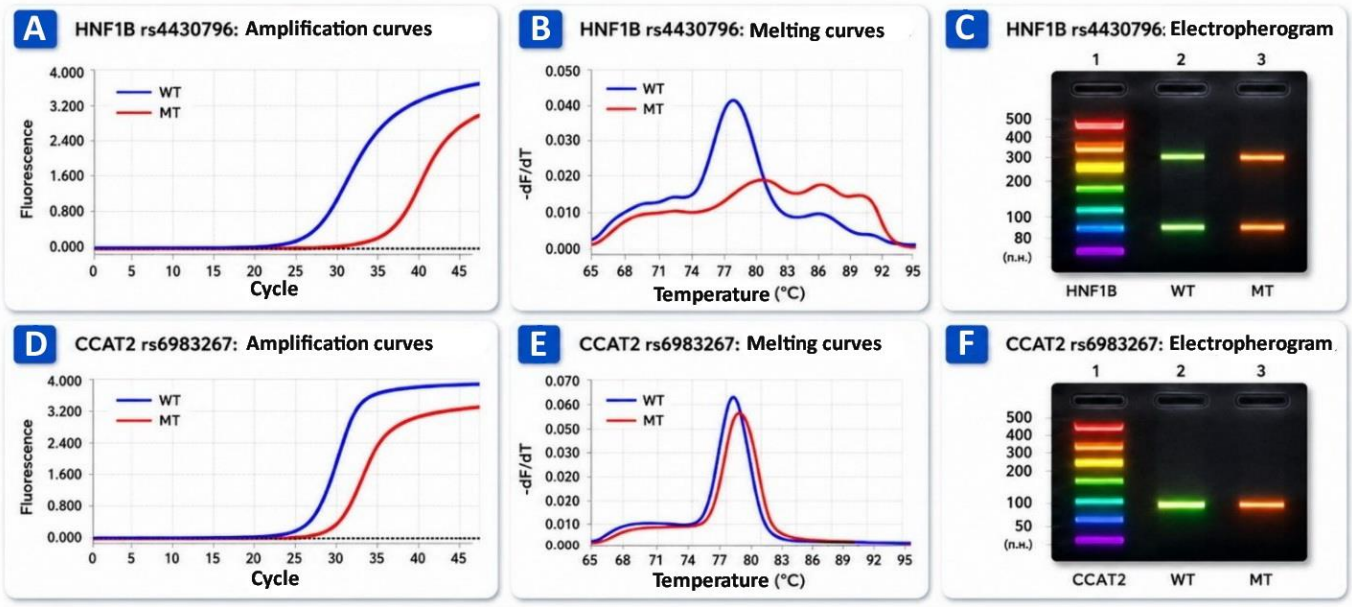

Figure 2. Representative results of a pilot run of allele-specific real-time polymerase chain reaction for *HNF1B* rs4430796 and *CCAT2* rs6983267. A, D – amplification curves; B, E – melting curves; C, F – electropherograms. WT (wild type) and MT (mutant type) – technical designations of primer mixes for preferential amplification of the reference and alternative alleles, respectively.

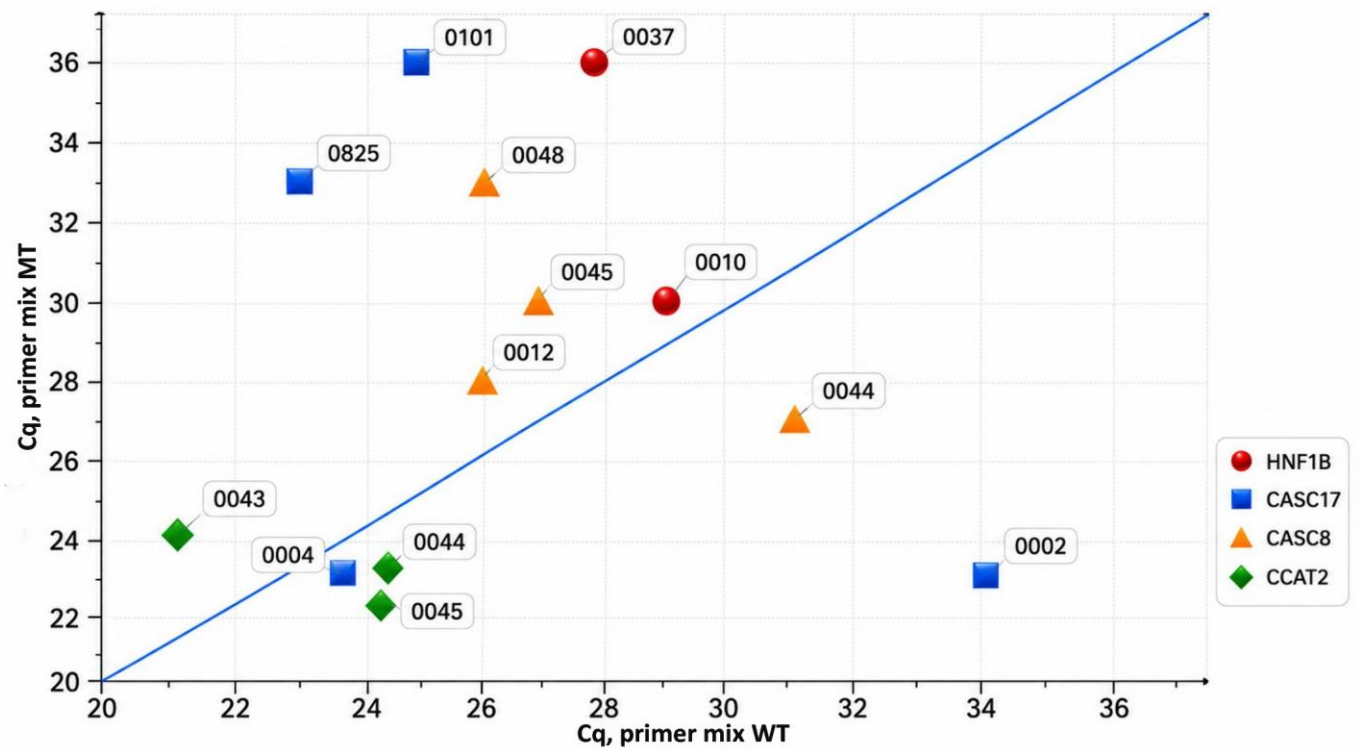

Figure 3. Paired Cq values in reactions with WT and MT primer mixes in the course of the pilot test. The diagonal line corresponds to equal Cq values in two parallel allele-specific reactions. WT (wild type) and MT (mutant type) – technical designations of primer mixes for preferential amplification of the reference and alternative alleles, respectively.

Table 4. Exact concordance between the results of polymerase chain reaction (PCR) with subsequent restriction fragment length polymorphism (RFLP) analysis and real-time allele-specific PCR

Locus	Comparisons, *n*	Matches, *n*	Concordance, %	95% CI; Cohen's κ
HNF1B	22	5	22.7	7.8-45.4; -0.199
CASC17	20	1	5.0	0.1-24.9; -0.310
CASC8	22	22	100.0	84.6-100.0; 1.000
CCAT2	21	16	76.2	52.8-91.8; 0.618

PCR-RFLP was treated as a comparator method rather than a reference method; hence, sensitivity and specificity values relative to PCR-RFLP were not calculated.

A complete three-locus profile (*CASC17*+*CASC8*+*CCAT2*) was created for 153 out of 154 samples, i.e., 99.35% (two-sided 95% CI 96.44-99.98). When testing H0: $\pi \leq 0.95$ against the H1: $\pi > 0.95$, a one-sided exact binomial test (x=153, n=154) yielded p=0.0034; the one-sided lower 95% confidence bound was 96.96%. Thus, at the level of an individual sample, the operational performance of the three scaled components statistically exceeded the established working threshold of 95%. At the level of individual loci, we obtained 461 final genotype assignments out of 462 potential ones (99.78%; two-sided 95% CI 98.80-99.99); this cumulative estimate was used exclusively as a descriptive characteristic and did not represent the basis for a formal statistical inference. The results are presented in *Table* 5 and *Figure* 4.

Testing for Hardy-Weinberg equilibrium

The genotype distributions for *CASC17* (AA 20, AC 70, CC 64; n=154) and *CCAT2* (GG 34, GT 77, TT 42; n=153) did not deviate from those expected in case of Hardy–Weinberg equilibrium: the asymptotic chi-squared test (with one degree of freedom) yielded χ^2=0.016 (p=0.900) and χ^2=0.013 (p=0.908), respectively, while the exact test yielded p=0.863 and p=1.000. For *CASC8* (GG 92, GT 43, TT 19; n=154), we observed a statistically significant deficit of heterozygotes, with 43 observed vs. 59.7 expected under Hardy-Weinberg equilibrium (χ^2=12.05, p<0.001; exact test p=0.001).

Discussion

The development of four components using a unified allele-specific qPCR format revealed pronounced differences in their analytical performance. Despite employing the same reaction setup, the *HNF1B*, *CASC17*, *CASC8*, and *CCAT2* assays differed in terms of their amplification curve shapes, |ΔCq| values, and the electrophoretic profiles of their products. Such locus dependency is predictable in allele-specific amplification, since the efficacy of allelic discrimination is provided by the nucleotide context surrounding the investigated position, the interaction between the 3′ end of the primer and the template, the annealing temperature, and the likelihood of generating by-products. [9]. These findings support the initial working hypothesis that a unified reaction architecture does not eliminate the need for locus-specific tuning of interpretation criteria.

Table 5. Results of extended laboratory validation

Parameter	Final assignments, *n/N*	Proportion, %	95% CI	Genotype distribution
CASC17	154/154	100.00	97.63-100.00	AA 20; AC 70; CC 64
CASC8	154/154	100.00	97.63-100.00	GG 92; GT 43; TT 19
CCAT2	153/154	99.35	96.44-99.98	GG 34; GT 77; TT 42
Total, locus-level	461/462	99.78	98.80-99.99	–
Full profile	153/154	99.35	96.44-99.98	–

CI, exact 95% Clopper–Pearson confidence interval. The row *'Total, locus-level'* represents a descriptive pooled estimate and was not used for formal statistical testing due to repeated measurements of three loci within the same samples. The row *'Full profile'* refers to the presence of combined findings for *CASC17*, *CASC8*, and *CCAT2*.

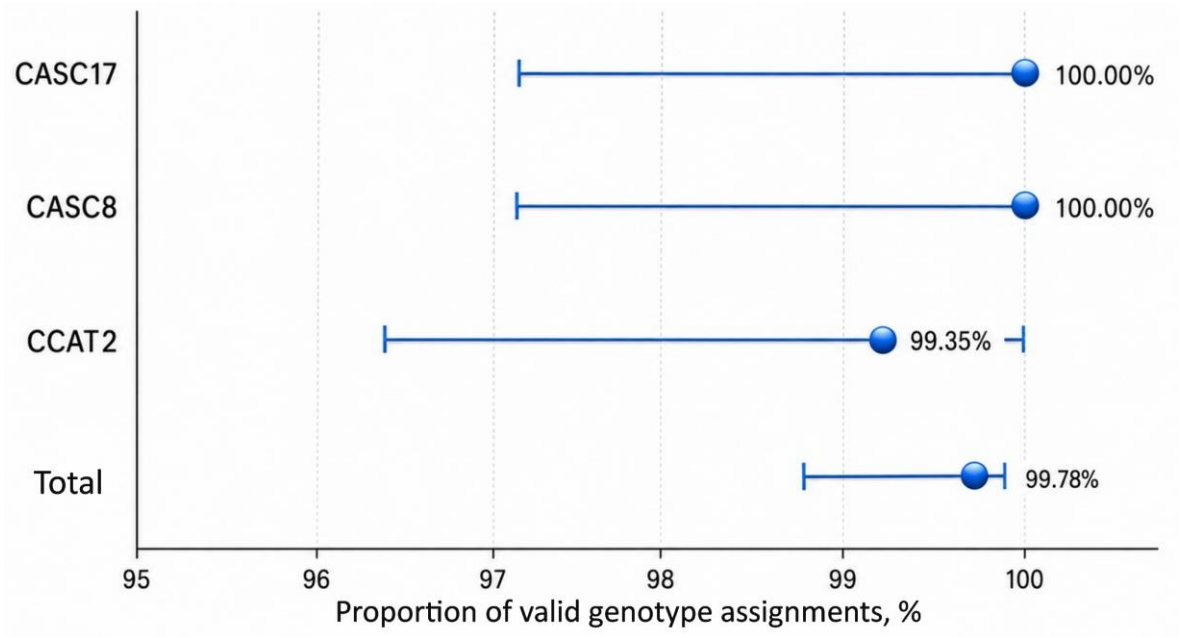

Figure 4. Proportion of valid genotype assignments in the course of extended laboratory validation of the test system. Data points represent the observed proportion of valid results; horizontal bars indicate exact two-sided 95% Clopper–Pearson confidence intervals. For *CASC17*, 154 valid results were obtained out of 154; for *CASC8*, 154 out of 154; and for *CCAT2*, 153 out of 154—yielding a total of 461 valid results out of 462 locus-level analyses.

During the pilot phase, the most pronounced technical limitations were observed for *HNF1B*. None of the four reaction setups yielded a robust amplification curve reaching a plateau; furthermore, additional electrophoretic products were detected across all tested conditions. Consequently, the *HNF1B* component was excluded from the extended test series. While this result reflects the specific primer configuration and reaction conditions used in this study, it does doubt the validity of rs4430796 per se as a genetic marker. Nonspecific amplification remains the primary unresolved analytical issue for this component.

An important observation was made in the course of the control series formation. Among 12 samples preliminarily classified as heterozygous via PCR-RFLP, Sanger sequencing confirmed the initial result in 10 cases. For instance, one *HNF1B* sample was reclassified from CT to CC, and one *CASC17* sample from AC to AA. Consequently, the PCR-RFLP results in this study are treated as comparable data rather than as a reference standard. This fact is a crucial point when interpreting the low between-method concordance observed for *HNF1B* and *CASC17*: the discrepancy makes it impossible to determine what proportion of the mismatches stems from the original (archival) method vs. the new assay. The negative Cohen's κ values for these loci indicate the absence of concordance in terms of what would be expected by chance given the distribution of categories [11], even though the latter do not reveal the underlying cause of the discrepancies.

For *CASC17*, the pilot amplification profile was better than that of *HNF1B*: 5 out of 8 reactions met the established criteria for curve shape, and no extra bands were observed in 4 out of the 6 runs analyzed via electrophoresis. Furthermore, *CASC17* exhibited the highest median |ΔCq| value (10.5 cycles). Yet, the results matched the archival PCR-RFLP data in only 1 out of 20 comparisons, yielding a Cohen's κ of -0.31. Such combination of a marked difference between the two allele-specific reactions alongside extremely low between-method concordance demonstrates that the |ΔCq| value alone cannot serve as a marker of genotyping accuracy. For *CASC17*, the chief uncertainty remains the analytical accuracy of the final genotype assignment.

A different pattern was observed for *CASC8*. All 22 available results matched those from the archival PCR-RFLP analysis, with a Cohen's κ of 1.000, indicating full beween-method concordance. At the same time, only 2 out of the 6 reactions evaluated via electrophoresis were free of additional bands. When SYBR Green is used, nonspecific double-stranded DNA products can also contribute to the fluorescent signal [9]. Consequently, the concordance between two

© Fomkin R.N., Pylaev T.E., Popkov V.M., Mezirov G.G., Fomkina O.A., Bucharskaya A.B., Polukonova N.V., 2026

genotyping methods and the specificity of the amplification process per se represent distinct properties of an analytical system: a favorable outcome for one does not eliminate the need to evaluate the other.

An additional feature observed for *CASC8* was a deviation of the genotype distribution from Hardy–Weinberg equilibrium. In the extended dataset, we detected 92 GG, 43 GT, and 19 TT genotypes; compared to an expected 59.7 heterozygotes, only 43 were observed, yielding an exact test *p*-value of 0.0008. While deviation from Hardy-Weinberg equilibrium can be associated with genotyping errors, it is not per se a definitive marker of a technical error [12]. In our study, interpretation is further complicated by the sample structure, since the overall analysis included both PCa patients and participants in a comparison group. Variants in the 8q24 region are associated with PCa risk [5], and both allele frequencies and effect sizes for these variants are known to vary across populations [5, 6]. Therefore, the observed deficit of heterozygotes could have several explanations; consequently, our finding alone is insufficient to identify the underlying cause. The significance of such deficit in this study lies primarily in providing an additional rationale for a more thorough validation of *CASC8* via an independent method.

Contrariwise, for *CASC17*, the genotype distribution (AA 20, AC 70, and CC 64) is consistent with Hardy-Weinberg equilibrium, with an exact *p*-value of 0.863. However, the absence of a statistically significant deviation does not constitute proof of analytical accuracy. This is particularly true in the case of *CASC17*, where a favorable Hardy-Weinberg equilibrium test result coincides with extremely low between-method concordance. Consequently, in our study, the Hardy–Weinberg equilibrium test cannot replace direct validation via sequencing [12].

The most consistent pilot performance characteristics were discovered for *CCAT2*. All 6 reactions met the established criteria for amplification curves; no extra bands were present in any of the 6 electropherograms, and *Tm* fell within the 77-79 °C range. The genotype distribution (GG 34, GT 77, and TT 42) also complied with Hardy–Weinberg equilibrium expectations, yielding an exact *p*-value of 1.0. However, the median |ΔCq| value was just 1 cycle; therefore, even with this highly favorable set of pilot characteristics, it remains necessary to identify the ΔCq distribution separately for reference homozygotes, heterozygotes, and alternative homozygotes that have been confirmed by sequencing [12].

Extended laboratory validation demonstrated that the 3 components that advanced to serial testing exhibited high operational performance. A final genotype assignment was successfully obtained in 154 out of 154 *CASC17* assays, 154 out of 154 *CASC8* assays, and 153 out of 154 *CCAT2* assays. A complete three-locus profile was generated for 153 out of 154 samples (99.35%; 95% two-sided CI: 96.44-99.98). A one-sided exact binomial test confirmed that this value exceeded the established threshold of 95% ($p<0.001$). Thus, the primary technical endpoint of the study was achieved for the three scaled components.

It is important to distinguish between operational performance and analytical accuracy. Obtaining a final genotype assignment does not necessarily mean that the assigned genotype corresponds to the true state of the locus; therefore, 99.35% value reflects the ability of a laboratory algorithm to produce an interpretable result, but it does not refer to the sensitivity, specificity, or analytical accuracy of the test system. Similarly, 461 final genotype assignments out of 462 potential locus-level assessments (99.78%) have purely descriptive meaning. Since the three loci were analyzed using the same 154 DNA samples, these 462 observations are statistically dependent; consequently, the formal hypothesis testing was appropriately conducted at the level of the individual sample.

The validation via sequencing is also limited in scope. The control set was initially formed of samples identified as heterozygotes by PCR-RFLP. However, Sanger sequencing revealed that two of these had been misclassified and were, in fact, homozygotes. Ultimately, it was not possible to confirm all three potential genotypes by sequencing for every locus. This was particularly the case for *CASC8* and *CCAT2*, where all control samples subjected to Sanger sequencing turned out to be heterozygous. For this reason, an accurate assessment of analytical accuracy and the establishment of locus-specific ΔCq intervals must be based on a series that includes confirmed reference homozygotes, heterozygotes, and alternative homozygotes.

Conclusion

We developed reagent components for a four-locus, allele-specific qPCR-RFLP prototype for the following variants: *HNF1B* rs4430796, *CASC17* rs1859962, *CASC8* rs1447295 and *CCAT2* rs6983267. Sanger sequencing of a control series revealed two mismatches with archival PCR-RFLP, thereby confirming the necessity of an independent reference method for future assessments of analytical accuracy.

We discovered that pilot performance characteristics are locus-dependent. At this stage, the *HNF1B* assay did not meet the criteria for advancement to extended validation; its optimization is ongoing, and the results of that study will be reported separately. For *CASC17*, further sequencing-based verification is planned, given the extremely low concordance between methods. For *CASC8*, the future focus will be on reducing nonspecific amplification and investigating why the observed genotype distribution deviates from Hardy–Weinberg equilibrium. For *CCAT2*, the boundaries of allelic discrimination across the three genotype classes will be formalized. In an extended test series, we obtained a complete three-locus profile for 153 out of 154 samples (99.35%; 95% CI 96.44-99.98; $p=0.0034$ relative to a 95% benchmark value), thereby confirming the high operational efficacy of these components.

Conflict of interest

R.N. Fomkin, T.E. Pylaev, and A.B. Bucharskaya are among the authors of the Russian Federation patent RU 2827104 C1, which pertains to the combination of genetic markers described in this article. An additional patent application is in preparation for the same combination. No other conflicts of interest are declared.

Author contributions

All authors contributed equally to the preparation of this publication.

Informed consent

The patients signed a voluntary informed consent form for the publication of medical information.

Funding

The study was conducted as part of a Government Procurement from the Russian Federation Ministry of Healthcare, *Development of a Real-Time PCR Test System for Genotyping Prostate Cancer Risk Variants: rs4430796 (HNF1B), rs1859962 (CASC17), rs1447295 (CASC8), and rs6983267 (CCAT2)* (registration number NIOKTR 125030703254-0); R&DT registration number 125030703254-0.

Compliance with ethical principles

The study protocol was approved by the local ethics committee. Approval and the implementation of the study protocol were carried out in accordance with the principles of the Declaration of Helsinki.

References

1. Kaprin AD, Starinskiy VV, Shakhzadova AO, eds. Malignant Neoplasms in Russia in 2024: Morbidity and Mortality. Moscow: Hertzen Moscow Research Institute for Oncology, 2025; 178 p. (In Russ.).
2. Conti DV, Darst BF, Moss LC, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. Nat Genet. 2021; 53(1): 65-75. https://doi.org/10.1038/s41588-020-00748-0.
3. Tong Y, Qu Y, Li S, et al. Cumulative evidence for relationships between multiple variants of HNF1B and the risk of prostate and endometrial cancers. *BMC Med Genet* 2018; 19(1): 128. https://doi.org/10.1186/s12881-018-0640-7.
4. Ren F, Zhang P, Ma Z, et al. Association of 17q24 rs1859962 gene polymorphism with prostate cancer risk: A systematic review and meta-analysis. *Medicine (Baltimore)* 2020; 99(3): e18398. https://doi.org/10.1097/MD.0000000000018398.
5. Tong Y, Yu T, Li S, et al. Cumulative evidence for relationships between 8q24 variants and prostate cancer. *Front Physiol* 2018; 9: 915. https://doi.org/10.3389/fphys.2018.00915.
6. Silva-Ramirez B, Macías-González EJ, Frausto-Valdes OS, et al. Association between genetic variations at 8q24 and prostate cancer risk in Mexican men. *Prostate Cancer Prostatic Dis* 2022; 25(3): 507-512. https://doi.org/10.1038/s41391-021-00461-x.
7. Polukonova NV, Fomkin RN, Pylaev TE, et al. Genetic polymorphisms of HNF1B, CASC17, CASC8 and CCAT2 genes associated with prostate cancer risk in urologic patients. *Urologiia* 2025; (5): 44-52. (In Russ.) https://doi.org/10.18565/urology.2025.5.44-52.
8. Mylnikov AM, Demin AG, Polukonova NV, et al. Method for predicting risk of prostate cancer development. Patent RU 2827104 C1, 23 Sep 2024. Application No. 2023134769 of 22 Dec 2023. (In Russ.). https://patents.google.com/patent/RU2827104C1/ru.
9. Dhas DBB, Ashmi AH, Bhat BV, et al. Modified low-cost SNP genotyping technique using cycle threshold and melting temperature values in allele specific real-time PCR. *Indian J Med Res* 2015; 142(5): 555-562. https://doi.org/10.4103/0971-5916.171282.
10. Bustin SA, Ruijter JM, van den Hoff MJB, et al. MIQE 2.0: Revision of the minimum information for publication of quantitative real-time PCR experiments guidelines. *Clin Chem* 2025; 71(6): 634-651. https://doi.org/10.1093/clinchem/hvaf043.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1): 159-174. https://pubmed.ncbi.nlm.nih.gov/843571.
12. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005; 76(5): 887-893. https://doi.org/10.1086/429864.

Authors:

Roman N. Fomkin – PhD (Medicine), Assistant Professor, Department of Urology, Senior Researcher, Research Institute of Fundamental and Clinical Uronephrology, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0000-0001-6895-6445;

Timofey E. Pylaev – PhD (Biology), Director of the Research, Production and Educational Center for Molecular Genetics and Cell Engineering, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0000-0002-2701-3333;

Vladimir M. Popkov – DSc (Medicine), Professor, Chair of the Department of Urology, Director of the Research Institute of Fundamental and Clinical Uronephrology, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0000-0003-2876-9607;

Georgy G. Mezirov – Graduate Student, Department of Urology, Junior Researcher, Research Institute of Fundamental and Clinical Uronephrology, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0009-0008-9372-621X;

Olga A. Fomkina – DSc (Medicine), Associate Professor, Department of Human Anatomy, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0000-0002-1516-0504;

Alla B. Bucharskaya – DSc (Biology), Head of the Shared Core Facility for Experimental Oncology, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0000-0003-0503-6486;

Natalya V. Polukonova – DSc (Biology), Professor, Head of the Laboratory of Cell Engineering, Research Institute of Fundamental and Clinical Uronephrology, Razumovsky State Medical University of Saratov, Saratov, Russia. https://orcid.org/0000-0001-9228-6808.