Article

Discovery Medicine 2025; 37(201): 2334–2348
https://doi.org/10.24976/Discov.Med.202537201.198

Risk of Bladder Cancer Based on Genetic Background of the Han Chinese Population

Yaqin Fan[1], Qingcheng Gao[2], Guangtao Xu[2,3], Deqing Chen[2,3], Yigang Jin[4], Chenye Tang[4], Chun Sun[4], Dandan Cai[4], Fei Wang[4,*]

[1]Department of Oncology, The First Hospital of Jiaxing, 314000 Jiaxing, Zhejiang, China
[2]Medical College, Jiaxing University, 314000 Jiaxing, Zhejiang, China
[3]Institute of Forensic Science, Medical College, Jiaxing University, 314000 Jiaxing, Zhejiang, China
[4]Department of Urology, The Second Hospital of Jiaxing, 314000 Jiaxing, Zhejiang, China
*Correspondence: fei_wang497@163.com (Fei Wang)
Submitted: 2 July 2025 Revised: 26 August 2025 Accepted: 2 September 2025 Published: 20 October 2025

Background: Bladder cancer is a prevalent malignancy of the urinary system, exhibiting extraordinarily complex pathogenesis involving multidimensional interactions between genetic and environmental factors. Substantial evidence confirms the pivotal role of genetic determinants in bladder carcinogenesis and progression. Although genome-wide association studies (GWAS) have successfully identified multiple genetic variants potentially associated with bladder cancer, the population-specific genetic architecture and its clinical implications in the Han Chinese patients remain to be elucidated.
Methods: This study developed a detection system for single nucleotide polymorphisms (SNPs) using multiplex polymerase chain reaction (PCR), single-base extension (SBE), and capillary electrophoresis technology, aimed at identifying potentially pathogenic variants. The method was applied to analyze 142 samples obtained from the Han Chinese individuals (aged 15–82, with an average age of 58 years) into a case group (n = 71) and a control group (n = 71).
Results: Statistical analysis revealed five SNPs with significant association in the case group. Binary Logistic regression analysis further validated its application for disease risk assessment and prediction. A predictive model integrating four significant SNPs, including rs8102137, rs7747724, rs1258767, and rs2042329, yielded an area under the curve (AUC) of 0.797 for predicting bladder cancer, while multifactor dimensionality reduction (MDR) analysis achieved a balanced accuracy of 0.7543.
Conclusion: This study demonstrates that these SNPs hold significant potential for application in genetic testing to predict bladder cancer risk; however, further research is needed to elucidate their functional mechanisms. Our analysis provides comprehensive insights into the association between bladder cancer-related genetic polymorphisms, hereditary susceptibility, and disease progression. These results establish a theoretical foundation for improving early diagnosis, preventive measures, and personalized treatment strategies in bladder cancer.

Keywords: bladder cancer; genetic susceptibility; single nucleotide polymorphisms; predictive model

Introduction

Bladder cancer is one of the most common malignant tumors in the urinary system worldwide, consistently ranking higher in terms of incidence and mortality, and has become a major health concern that poses a significant threat to human life and well-being [1]. An clinical study has confirmed that the pathogenesis and progression of bladder cancer result from the long-term synergistic interactions between genetic predisposition and environmental factors [2]. Among them, inherited genetic susceptibility plays a key role in the occurrence and progression of bladder cancer [3,4]. Exploring the intrinsic link between bladder cancer and susceptibility genes not only aids in revealing the underlying molecular mechanisms but also provides a crucial theoretical foundation for early and precise diagnosis, the development of effective preventive measures, and the advancement of personalized therapeutic approaches. Such advancements hold great promises for improving patient outcomes and enhancing overall quality of life.

Genetic susceptibility refers to the phenomenon that an individual has a significantly higher risk of developing a particular disease compared to the general population due to specific DNA variations or unique genotypes [5]. Such susceptibility often manifests as single nucleotide polymorphisms (SNPs), copy number variations (CNVs), and other genetic changes distributed across one or more loci. When exposed to the same environmental conditions, individuals carrying these genetic variations exhibit great sensitivity to carcinogens, thereby significantly increasing their risk of tumor development [3].

Genome-wide association studies (GWAS) have significantly advanced the genetic underpinning underlying various complex traits and diseases. In bladder cancer,

Copyright: © 2025 The Author(s). Published by Discovery Medicine. This is an open access article under the CC BY 4.0 license.
Note: *Discovery Medicine* stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

GWAS have identified multiple susceptibility loci, including novel regions such as 6p22.3 and enhanced associations in known regions (e.g., 4p16.3), yielding a total of 24 genome-wide significant independent markers [6]. For example, the 19q12 locus harboring the *CCNE1* gene contains SNPs (rs8102137, rs7257330) associated with aggressive disease, likely through cyclin E overexpression and cell cycle dysregulation [7]; rs8102137 may also mediate gene-diet interactions [8]. Furthermore, the *RBFOX3* variant rs978416 *G*>*A* has been linked to a reduced risk in Chinese populations [9]. Smoking further increases the risk, interacting with variants at 8p22 (*NAT2*), 8q21.13 (*PAG1*), and 9p21.3, while a PRS derived from 24 GWAS markers has been found to effectively stratify lifetime risk [6]. Despite these advances, whether these SNPs directly drive bladder cancer pathogenesis remains controversial. Notably, comprehensive GWAS investigating bladder cancer-associated genetic variants and genetic susceptibility in the Han Chinese population are still lacking. Therefore, the underlying molecular mechanisms linking risk-associated SNPs to disease onset warrant comprehensive investigations. Given the relatively homogenous genetic background of this population, elucidating the hereditary characteristics of bladder cancer would not only improve the precision of genetic risk prediction and clarify its biological relevance but also provide critical theoretical support for optimizing early prevention and tailored intervention strategies.

Based on current research, this study uses GWAS findings to comprehensively investigate the genetic background of bladder cancer among the Han Chinese population. By innovatively employing mini-sequencing technology, this study successfully developed a detection system for bladder cancer susceptibility gene. This system not only efficiently identifies risk factors for bladder cancer but also provides precise diagnostic insights and personalized preventive guidance for the first-degree relatives of patients, thus opening new avenues for early screening and accurate prevention of bladder cancer.

Materials and Methods

This study was approved by the Ethics Committee of the Second Hospital of Jiaxing (Approval No.2025-130). Written informed consent was obtained from all participants before their enrollment in the study, in accordance with the Declaration of Helsinki.

Sample Collection

This study included 142 individuals from the Han Chinese population, and their clinical data were retrospectively analyzed from archived specimens. The samples were divided into a case group (n = 71) and a control group (n = 71). The case group included patients diagnosed with bladder cancer at the Second Hospital of Jiaxing between January 2023 and December 2024. However, the control group consisted of individuals who underwent routine health checkups at the same hospital during the same period. Cases and controls were carefully matched by gender, age, and comorbidities to minimize the confounding factors.

The matching ratio and specific criteria were as follows: (1) an equal number of individuals were maintained between the two groups; (2) the absolute difference in gender distribution did not exceed 10%; (3) the difference in mean age was $\leq$1 year; (4) the difference in the prevalence of underlying comorbidities did not exceed 10%. The corresponding sample characteristics are presented in Table 1.

To examine potential confounding, we conducted univariate logistic regression and multivariate logistic regression analyses, including age, sex, smoking status, and underlying diseases as covariates. These analyses evaluated the independent association of each factor with bladder cancer risk while adjusting for confounders. The findings indicate that, apart from smoking status, which demonstrated a statistically significant difference between the groups, all other variables did not differ statistically ($p > 0.05$). These results support substantial baseline comparability between the two groups. The findings of these analyses are presented in Table 2.

The inclusion criteria for cases were as follows: (1) a confirmed diagnosis of bladder cancer based on pathological examination; (2) availability of complete clinical and follow-up data; (3) absence of other malignant tumors, severe systemic diseases, or immune disorders; and (4) voluntary participation with informed consent. For the control group, participants were selected based on the following criteria: (1) matching with the case group by age and gender (where applicable); (2) no history of bladder cancer or other urinary system malignancies; (3) no severe systemic diseases that could affect study outcomes; and (4) voluntary participation with informed consent.

For the control group, peripheral venous blood samples were collected from volunteers, and genomic DNA was extracted using the E. Z. N. A™ SE Blood DNA Kit (Omega Bio-Tek, Inc., Norcross, GA, USA). For bladder cancer patients, surgically resected bladder cancer tissues were used, with genomic DNA extracted from the tissue samples using the Mag-Blind® Tissue DNA Kit M6223 (Omega Bio-Tek, Inc., Norcross, GA, USA), following the manufacturer's instructions. The DNA was quantified using a BioSpectrometer (22331, Eppendorf AG, Hamburg, Germany) at 260 nm. The DNA samples were then diluted to 10 ng/µL with ultrapure water and stored at –20 °C for further analysis.

Establishment of Mini-sequencing Detection System

This study utilized the ABI PRISM® SNaPshot™ Multiplex kit (Applied Biosystems, Thermo Fisher Scientific, Inc., Waltham, MA, USA), which integrates multiplex polymerase chain reaction (PCR), Single-Base Extension (SBE) technology, and capillary electrophoresis (CE)

DM DISCOVERY MEDICINE 2001

Table 1. Baseline characteristics of the study participants.

Group	Gender	Number	Proportion (%)	Gender ratio (M/F)	Age range	Mean age	Smoking proportion (%)	Patients complicated with other diseases (Hypertension, CHD, DM) (%)
Case (n = 71)	M	69	97.2	34.5:1	15–82	58.23	84.5	60.6
	F	2	2.8					
Control (n = 71)	M	66	93.0	13.2:1	35–81	58.07	70.4	54.9
	F	5	7.0					

Note: M, male; F, female; CHD, coronary heart disease; DM, diabetes mellitus.

Table 2. The univariate and the multivariate logistic regression analyses.

Variables	Univariate logistic regression analysis			Multivariate logistic regression analysis		
	Standard error	Significance	95% CI	Standard error	Significance	95% CI
Gender	0.854	0.261	0.490–13.942	0.917	0.851	0.140–5.073
Age	0.442	0.326	0.273–1.541	0.456	0.324	0.641–3.834
Smoking	**0.419**	**0.048**	**1.009–5.203**	**0.461**	**0.041**	**0.158–0.963**
Underlying diseases	0.340	0.497	0.647–2.455	0.408	0.731	0.517–2.559

Note: Bold indicates variables with statistical significance ($p < 0.05$). CI, confidence interval range.

techniques. The SNaPshot technique is based on the principle of dideoxynucleotide chain termination, where fluorescently labeled dideoxynucleotides (ddNTPs) are specifically incorporated at the 3′ end of extension primers, enabling the precise detection of target nucleotide variations via capillary electrophoresis. By integrating these techniques, we developed a highly sensitive mini-sequencing screening system for SNP analysis. The system allows for simultaneous amplification of multiple target fragments through multiplex PCR, precise identification of base variations using SBE, and accurate fragment resolution via CE. This integration enables efficient, parallel, and reliable SNP detection and genotyping, thereby providing robust technical support for the comprehensive investigation of genetic variations in bladder cancer.

Candidate SNPs

This study systematically conducted screening for genetic variations associated with bladder cancer using genome-wide association study (GWAS) data. An initial screening yielded over 40 candidate SNPs with $p < 1 \times 10^{-6}$ obtained from the GWAS database, guided by the latest cutting-edge research and advances in the field. However, during the experimental phase, several loci were excluded due to technical challenges. Following stringent validation and optimization of genotyping protocols, 19 loci were retained for subsequent analyses to determine whether these SNPs constitute key genetic risk factors for bladder cancer incidence in the Han Chinese population.

Linkage disequilibrium (LD) analysis confirmed that the selected loci were independently distributed without forming LD clusters, thereby effectively minimizing potential confounding effects of locus-locus interaction in subsequent association analysis. This ensured scientific robustness and the reliability of our conclusions. Specific information on each SNP locus is summarized in Table 3.

Multiplex Polymerase Chain Reaction

Multiplex PCR primers were designed using the online tool Primer 3.0 (http://primer3.ut.ee), with parameters set as follows: primer length of 18–25 bp, melting temperature (TM) between 52–62 °C (with a maximum difference of 3 °C between forward and reverse primers), GC content of 40–60%, and expected product length of 100–300 bp. Strict exclusion criteria were set for hairpin structures, self-dimers, and cross-dimers, with a minimum free energy threshold of –4 kcal/mol for dimer formation. Furthermore, primer specificity was verified using the National Center for Biotechnology Information (NCBI) Primer Blast platform (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome), ensuring precise target binding. Additionally, Auto Dimer v1.0 software (Gaithersburg, MA, USA) was used to systematically evaluate the risk of primer dimerization and hairpin formation, effectively reducing non-specific amplification. Given that DNA degradation often occurs in clinical formalin-fixed pathological specimens, a short-fragment amplification strategy (101–285 bp) was applied. The optimized PCR design improve adaptability to degraded DNA samples, maximizing amplification efficiency and accuracy. Primers were synthesized by Sangon Biotech Co., Ltd. (Shanghai, China). The primer sequences used in this study are provided in Table 4.

The PCR amplification experiments were performed using a Mastercycler Nexus gradient thermal cycler (Eppendorf, SE, Hamburg, Germany). Each reaction was con-

Table 3. Information on nineteen SNPs involved in this study.

rs ID	Allele (forward stand)	Involving genes	Location (GRCh38.p14)	*p*-value	Population
rs401681	*C/T*	*CLPTM1L*	Chr5:1321972	5.000×10^{-7}	European
rs2294008	*C/T*	*PSCA*	Chr8:142680513	4.000×10^{-11}	European
rs907611	*G/A*	*LSP1*	Chr11:1852842	4.000×10^{-8}	European
rs10936599	*C/G/T*	*MYNN*	Chr3:169774313	5.000×10^{-9}	European
rs1495741	*G/A*	*NAT2, PSD3*	Chr8:18415371	4.000×10^{-11}	European
rs9642880	*G/A/T*	*CASC8, CASC11*	Chr8:127705823	7.000×10^{-12}	European
rs12216499	*G/C/T*	*C6orf99, RSPH3*	Chr6:158947492	1.000×10^{-6}	European
rs62185668	*C/A/T*	*FAT1P1, RPS11P1*	Chr20:10981287	2.000×10^{-11}	European
rs1014971	*C/G/T*	*CBX6, APOBEC3A*	Chr22:38936618	1.000×10^{-11}	European
rs710521	*T/C/G*	*MIR944, P3H2*	Chr3:189928144	1.000×10^{-11}	European
rs798766	*T/C*	*TACC3*	Chr4:1732512	4.000×10^{-13}	European
rs7238033	*T/C*	*SLC14A1*	Chr18:45737001	9.000×10^{-9}	European
rs11892031	*A/C/T*	*UGT1A10, UGT1A8*	Chr2:233656637	1.000×10^{-7}	European
rs8102137	*T/C*	*C19orf12, CCNE1*	Chr19:29805946	1.000×10^{-11}	European
rs10094872	*A/T*	*CASC8, CASC11*	Chr8:127707639	2.000×10^{-7}	European
rs7747724	*T/C/G*	*CDKAL1*	Chr6:20751084	1.000×10^{-6}	European
rs1258767	*A/C/G*	*FMN1*	Chr15:32834253	7.000×10^{-7}	European
rs5003154	*T/A/C/G*	*PAG1*	Chr8:81074718	1.000×10^{-6}	European
rs2042329	*T/A/C/G*	*CWC27*	Chr5:64771925	5.000×10^{-11}	East Asian

Note: SNPs, single nucleotide polymorphisms.

ducted in a 10 µL reaction system consisting of 5 µL 2× Multiplex PCR mixture (M5 HiPer Multiplex PCR Master Mix (Mei5 Bioservices Co., Ltd., Beijing, China), 1 µL primer mixture (final concentrations details given in Table 4), and 5 ng DNA template. The thermal cycling conditions were set as follows: an initial denaturation at 95 °C for 10 minutes, followed by 30 cycles of denaturation at 94 °C for 20 seconds, annealing at 58 °C for 20 seconds, and extension at 72 °C for 30 seconds. A final extension was performed at 72 °C for 5 minutes to ensure the complete synthesis of amplification products, thereby ensuring the reliability and accuracy of experimental results.

Single-base Extension Reaction

This study targeted SNP sites by designing primers that precisely complement the adjacent bases flanking each of the SNP. To ensure the high-efficiency separation of SBE products during CE, a specific length tag comprising $(CT)_n$ or $(AGCT)_n$ repeat sequences was introduced at the 5′ end of each primer, following the principle of fragment detection length design. This approach ensured that the expected SBE product length fell within the optimal size range of 20–68 bp. The SBE primers were synthesized by Sangon (Shanghai, China) Biotech Co., Ltd. The primer sequences used in SBEs are detailed in Table 4.

During the sample pretreatment and SBE reaction, a dual-enzyme cascade purification strategy was adopted. Shrimp alkaline phosphatase (rSAP, New England BioLabs, Inc., Frankfurt am Main, Germany) was used to specifically hydrolyze residual deoxynucleoside triphosphates (dNTPs), while exonuclease I (Exo I, New England BioLabs, Inc., Frankfurt am Main, Germany) was applied to degrade excess PCR primers. The purification system was prepared in a 10 µL reaction volume, containing 6.6 µL of PCR product, 1 U of rSAP, and 4 U of Exo I. After a 1-hour incubation at 37 °C, enzymatic activity was inactivated by heat at 95 °C for 15 minutes.

The purified PCR products were used for the subsequent SBE reaction. The SBE reaction was performed in 5 µL optimized system, consisting of 2 µL purified PCR product, 2.5 µL SNaPshot reaction premix, and 0.5 µL SBE primer mixture. The SBE reaction followed a three-step cycling protocol: denaturation at 96 °C for 5 seconds, annealing at 50 °C for 10 seconds, and extension at 60 °C for 15 seconds, repeated 35 cycles. However, to prevent residual dideoxynucleoside triphosphates (ddNTPs) from interfering with CE detection, an additional purification step with rSAP was performed. This additional purification involved incubation at 37 °C for 1 hour, followed by enzyme heat-inactivation at 95 °C for 15 minutes, ensuring high detection specificity and accuracy of the outcomes.

Capillary Electrophoresis Separation and Visualization

For visual analysis of SBE products, 1.5 µL of SNaPshot reaction product was mixed with 10 µL of Hi-Di formamide (Applied Biosystems; Thermo Fisher Scientific, Inc., Waltham, MA, USA) and 0.1 µL of GeneScan™ 120 LIZ™ size standard (Applied Biosystems, Thermo Fisher Scientific, Inc., Waltham, MA, USA). The mixture was first denatured at 95 °C for 5 minutes to fully unwind the nucleic acid double-strand helix, then followed by rapid cooling to 0 °C for 3 minutes to stabilize the single-strand structure.

Table 4. Information on PCR and SBE primers used in this study.

rs ID	Primer sequence (5′-3′)	Length of PCR products (bp)	PCR primer concentration	SBE primers	SBE primer concentration
rs401681	F: GTGGGCAGAAAACAAGGTCT	148	0.04	ATCCAGACAACTTCAGAGTC	0.3
	R: TGGGCAGCAGTTACACTGATTA				
rs2294008	F: CCAGCCTCTCCCCATTTGAG	109	0.01	$(GACT)_{1}$CCACAGCCCACCAGTGACCA	0.1
	R: CCAAGCCTGCCATCAACAGG				
rs907611	F: GACCTGATGGCCAGAGCTAC	148	0.04	(CT)GCCATGAAGAGCTGCCTGCG	0.3
	R: CCCTCAGGAAGCGTTAAGCA				
rs10936599	F: GATGAGAACTTACGGTTACTG	167	0.04	$(GACT)_{2}$AGCATGAAACAGAGTAATTG	0.3
	R: ATCACTATGGTACAGTCACA				
rs1495741	F: CAGAGGAGCCTCTCTCAGGA	114	0.04	$(CT)_{3}$AGGATGATTTTCATAATAAT	0.3
	R: TATGGGGCCTCACATGGTCA				
rs9642880	F: CCAGGTTACCCAAAGTACTG	285	0.04	$(CT)_{9}$GGCTGGAGTTAGGAGAACCC	0.3
	R: AATTCCTGAACTCAGGTGAT				
rs12216499	F: ACTTTCCGCTAACTTGGGCA	154	0.04	$(CT)_{11}$ATTGCACAAAGTATTGTTTA	0.3
	R: GGCAAACTTCCTCAGCCCTA				
rs62185668	F: GCTCCCACATAGGGCATCTT	101	0.04	$(GACT)_{5}$AGCAGTCATTATTTGCAGAG	0.1
	R: ACAAGACCACCCAAACCATTT				
rs1014971	F: AGAGGAGGAGCCTTTAGGGG	192	0.04	$(GACT)_{6}$AGTGGGACTACTAGCAGAAG	0.1
	R: AAACTGGTAATGGGAGCGGG				
rs710521	F: GCATTCACTTTTGCCTTAGTGC	150	0.04	$(GACT)_{7}$GTGCTTATTCATTTTTTAAG	0.1
	R: GCTCTGCTCATGCCCAACAT				
rs798766	F: GGAAGGTGGAACGGTCCTTA	125	0.04	$(GACT)_{8}$TAAAGTGCATTTGGTAGGAA	0.1
	R: TGTGTGCCTGGGTTTTCAGT				
rs7238033	F: TTAAGGCATGTGGCCTCCAG	134	0.04	$(CT)_{15}$TTGATTGGAGGGAGTAGACT	0.01
	R: CTATGAGCTATGCCATAAGGAA				
rs11892031	F: CTTTAAGTGTAGGCGCCCAT	270	0.04	$(CT)_{17}$TTGCCACAAAAAGTCCACTG	0.1
	R: TCTCCCACAACAAGTGAATCG				
rs8102137	F: CAGGCCTGGCATCTACACAA	206	0.04	$(GACT)_{9}$ACTACTAGGAGGCTCACTGG	0.1
	R: AGCCAGTTGGGTTTGGACTT				
rs10094872	F: ATCCTGCTCCCTGAGTTTCT	218	0.04	$(CT)_{19}$TACTAGGTGTAAGTTGCCTG	0.1
	R: GGGCAAAGTTAAGTCCACACC				
rs7747724	F: TGCACGTGAAGTGTCAGTGT	159	0.04	$(GACT)_{10}$AGCATGAAACAGAGTAATTG	0.1
	R: CTCTGCAGTCGAAGGTTCTGT				
rs1258767	F: GGTTAGTTGAGGCTGGACCC	273	0.04	$(CT)_{21}$TAAAGTGCATTTGGTAGGAA	0.3
	R: GCCCAAGGGAGACCAAGTTC				
rs5003154	F: TGGAGTTAAGGCACAGGGGA	265	0.04	$(GACT)_{11}$CAGGTGATTCACACTAAATT	0.3
	R: TGGTGCAGATGTGACAGTGG				
rs2042329	F: TGAGCAAAGTGGCCCGATTA	160	0.04	$(GACT)_{12}$TACTAGGTGTAAGTTGCCTG	0.3
	R: CTGTGAGGACAAAATGAGCTGT				

Note: F, forward primer; R, reverse primer; PCR, polymerase chain reaction; SBE, single base extension.

DM

SNP genotyping were conducted on a 3500Dx Genetic Analyzer (Applied Biosystems; Thermo Fisher Scientific, Inc., Waltham, MA, USA) equipped with a 50 cm capillary containing POP-7 gel (Applied Biosystems, Thermo Fisher Scientific, Inc., Waltham, MA, USA), taking advantages of its high-resolution capability for precise SNP identification. Fluorescent signals generated by various labels were analyzed using GeneMapper ID v3.2 software (Applied Biosystems, Thermo Fisher Scientific, Inc., Waltham, MA, USA), which enabled accurate identification and genotyping of both mutant and wild-type SNPs, thereby ensuring the validity and reliability of the detection results.

Statistical Analysis

Data processing and statistical analysis was conducted using Microsoft 365 (Microsoft Corporation, Waltham, MA, USA), SPSS v23.0 (IBM Corp., Armonk, NY, USA), and MDR v3.0.2 software (SourceForge, San Diego, CA, USA). Comparative analyses between the case and control groups were performed to predict the onset of bladder cancer risk. The chi-square test was used to calculate the odds ratio (OR), with $p < 0.05$ considered the threshold for determining statistical significance.

To precisely evaluate the role of SNPs in bladder cancer risk, a predictive model was constructed. SNP loci exhibiting significant differences by chi-square test were incorporated into the model. The ORs, which quantifies the strength of association between variables, were calculated using the cross-tabulated contingency data. To reduce the risk of false positive findings, Bonferroni correction was applied for multiple comparisons.

Binary Logistic regression analysis was then used to identify SNPs suitable for risk assessment and prediction. Model performance was determined by comparing the predicted probabilities with actual disease status, using the area under the curve (AUC) as a quantitative measure of predictive accuracy. Additionally, the multifactor dimensionality reduction (MDR) method was applied not only to validate model robustness but also to investigate potential interaction mechanisms among different SNP loci, providing a more comprehensive understanding of the genetic factors on bladder cancer risk.

Results

Establishment of Mini-sequencing Detection System

The study successfully developed a highly sensitive and specific mini-sequencing detection system by systematically optimizing the multiplex PCR conditions and primer design. The system was designed to screen SNPs closely associated with the susceptibility of bladder cancer. To address the challenges of DNA degradation in formalin-fixed clinical samples, an innovative short fragment amplification strategy (101–285 bp) was employed, significantly enhancing detection efficiency and achieving complete genotyping across all samples.

Genotyping results for nineteen target SNPs in randomly selected samples are illustrated in Fig. 1. The abscissa denotes single-base extension products of varying lengths, with nucleotides color coded as follows: blue (guanine, G), green (adenine, A), black (cytosine, C), and red (thymine, T). For instance, SNP rs401681 exhibits a single black peak (C), corresponding to a homozygous *GG* genotype on the antisense strand. Conversely, SNP rs2042329 shows dual peaks (blue for G and red for T), representing a heterozygous *CA* genotype on the antisense strand. Using this optimized approach, complete genotyping results for all 19 SNPs was obtained, providing a robust dataset to support subsequent genetic association analyses.

Genotype Data of the Han Chinese Population

This study examined nineteen bladder cancer-linked SNPs identified through GWAS, aiming to explore their genetic characteristics in the Han Chinese population. The findings revealed that twelve SNP (rs401681, rs2294008, rs907611, rs10936599, rs1495741, rs9642880, rs12216499, rs710521, rs798766, rs7238033, rs10094872, and rs5003154) did not exhibit polymorphism changes in this cohort. These findings suggest significant genetic differences between the Han Chinese and Central European populations regarding bladder cancer susceptibility, revealing the presence of population-specific genetic regulatory mechanisms.

Seven additional SNPs (rs62185668, rs1014971, rs11892031, rs8102137, rs7747724, rs1258767, and rs2042329) demonstrated significant genetic polymorphism in the Han population (Fig. 2 and Tables 5,6). Further statistical analyses identified that five of these SNPs (rs1014971, rs8102137, rs7747724, rs1258767, and rs2042329) were significantly associated with bladder cancer risk. Genotype frequency comparison revealed the risk genotypes as follows: *CA* for rs1014971, *CT* for rs8102137, *GT* for rs7747724, *GG* for rs1258767, and *CA* for rs2042329. These genotypes were classified as "risk-associated" variants based on their significantly higher frequencies in the case group compared to the control group. The corresponding odds ratios (ORs) for risk-associated genotypes are detailed in Table 7.

To reduce risk of false-positive findings inherent to multiple comparisons, Bonferroni correction was applied. The nominal significance level (α = 0.05) was adjusted by dividing it by the total number of SNPs analyzed, thereby accounting for the multiple testing burden. Specifically, the correction was applied to the five SNPs that showed statistically significant ORs in the initial screening (Table 7). All statistical results were reported with 95% confidence intervals (95% CIs) to quantify the reliability of the estimates. The 95% CI serves both to indicate the precision of the estimates (narrower intervals reflect higher precision, wider intervals indicate greater uncertainty) and to aid in the assessment of statistical significance (particularly for relative

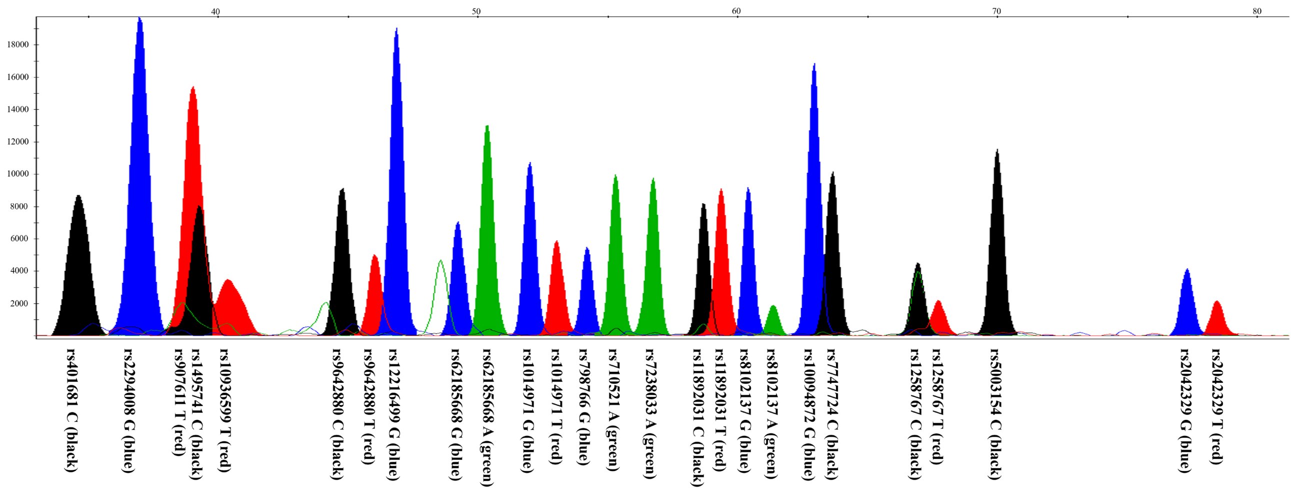

Fig. 1. The genotyping results of nineteen SNPs in random samples. Blue represents guanine (G), green represents adenine (A), black represents cytosine (C) and red represents thymine (T). Abbreviations: SNP, single nucleotide polymorphism.

effect measures such as ORs). These findings provide robust genetic evidence supporting the association between specific genetic variations and bladder cancer risk, while also offering valuable theoretical and empirical insights to inform precise risk prediction, early screening, and personalized treatment strategies.

Using AUC to Predict Disease-risk Alleles

The AUC provides a quantitative measure of the predictive performance of genotypes. To evaluate the contribution of SNPs to bladder cancer susceptibility, five statistically significant SNPs identified by chi-square analysis were incorporated into the prediction model. Of these, four SNPs confirmed through binary Logistic regression analysis were retained as predictors of disease risk (Table 8). The predicted probabilities were compared with the observed disease status, and the area under the receiver operating characteristic (ROC) curve served as a comprehensive indicator of prediction accuracy, including both sensitivity and specificity across different thresholds. An AUC of 0.5 indicates no predictive ability, while an AUC of 1.0 represents perfect discrimination. In this study, the combined prediction model of four SNPs yielded an AUC of 0.797, indicating strong predictive capacity. These results suggest that the four statistically significant SNPs included in the prediction model may serve as reliable genetic risk factors for bladder cancer. The ROC curve is shown in Fig. 3, and AUC values of SNPs are detailed in Table 9.

Using MDR to Evaluate the Interaction Between Risk Alleles

The MDR method was used to evaluate both predictive accuracy and interactions among SNPs. Interaction analysis of these four SNPs demonstrated that two models achieved a 10-fold cross-validation consistency of 10, indicating strong model stability. The cross-validation strategy helps minimize false-positive results arising from increased Type I errors due to random variation in data transformation.

As shown in Table 10, the model consisting of rs8102140, rs7747724, rs1258767, and rs2042329 yielded the highest test balanced accuracy (0.7543), while the balanced accuracy for rs8102137 also reached 0.7042. These results suggest that the four SNPs collectively predict the disease risk with an accuracy exceeding 70%. The interaction cell diagram is shown in Fig. 4, and the SNPs interaction network diagram is illustrated in Fig. 5.

Discussion

Study Rationale and Identification of Bladder Cancer-associated SNPs

Although extensive research has been performed in recent years on the etiology, pathogenesis, and risk factors of bladder cancer, significant gaps remain in effective prevention and treatment strategies [1,10,11]. Meanwhile, GWAS face interpretive dilemmas in clinical translation. The primary limitation of GWAS lies in its ability to reveal correlations between variables rather than establish causal relationships; its clinical application still needs validation through prospective studies that combine identified risk factors [6]. Based on these observations, the present study constructed a mini-sequencing detection system incorporating nineteen bladder cancer-related SNP loci previously reported by GWAS to systematically evaluate their association with bladder cancer susceptibility in the Han Chinese population.

Candidate SNPs were selected through a rigorous two-step process: (i) loci previously identified in retrospective GWAS of bladder cancer, and (ii) validated variants cataloged in authoritative databases. The initial screen-

Table 5. Polymorphic SNPs and the percentage of each genotype.

rs ID	Genotyping	Normal group	Percentage (%)	Case group	Percentage (%)
rs62185668	*CC*	4	6	3	4
	CT	67	94	68	96
rs1014971	*CC*	40	56	14	20
	CA	29	41	55	77
	AA	2	3	2	3
rs11892031	*GG*	3	4	2	3
	GA	68	96	69	97
rs8102137	*CC*	59	83	30	42
	CT	12	17	41	58
rs7747724	*GG*	22	31	4	6
	GT	49	69	67	94
rs1258767	*GG*	61	86	49	69
	GA	10	14	22	31
rs2042329	*CC*	25	35	7	10
	CA	38	54	56	79
	AA	8	11	8	11

Note: SNP, single nucleotide polymorphism.

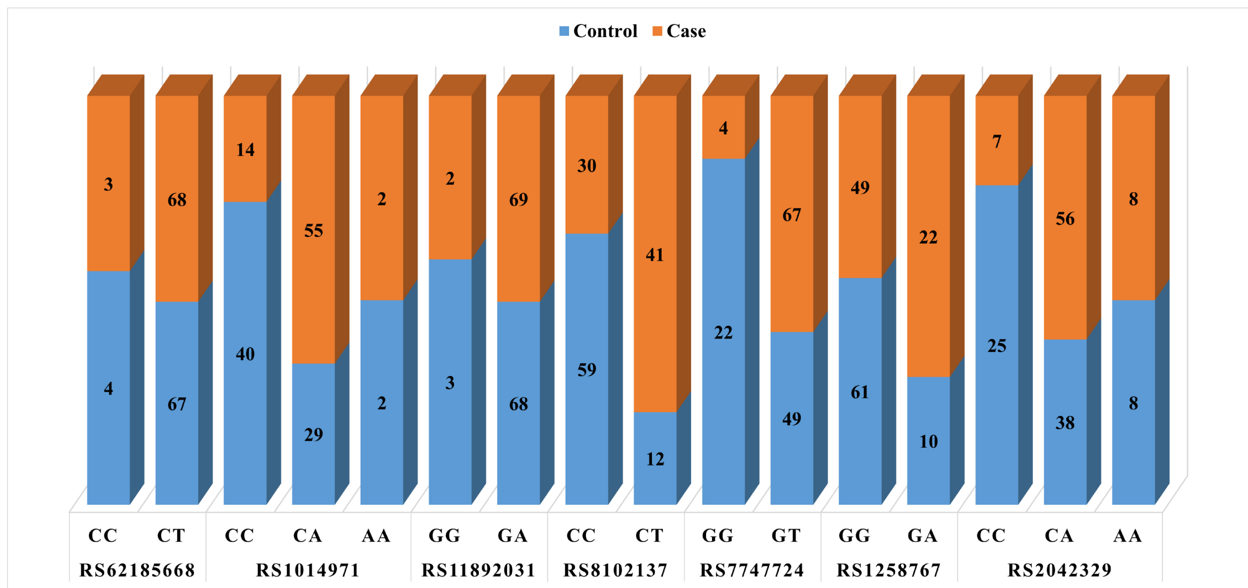

Fig. 2. Percentage of genotyping of seven significant SNPs. Abbreviations: SNP, single nucleotide polymorphism.

ing yielded over 40 candidate loci; however, several loci were excluded due to technical constraints such as challenges in primer design or insufficient specificity, to ensure reliable genotyping. After optimization, nineteen SNPs were retained, all exhibiting stable amplification efficiency and consistent genotyping performance, supported by evidence from prior GWAS findings and databases [12–19]. Analysis of these nineteen SNPs revealed that seven (rs62185668, rs1014971, rs11892031, rs8102137, rs7747724, rs1258767, and rs2042329) displayed polymorphic characteristics in the Han Chinese population, with four SNPs (rs8102137, rs7747724, rs1258767, and rs2042329) exhibiting significant associations with bladder cancer susceptibility.

These four associated SNPs, which have not been extensively reported in the studied population, provide novel insights into the genetic basis of bladder cancer. The risk prediction model constructed based on these four SNPs demonstrated robust discriminative capability, with an AUC of 0.797, indicating its potential as a promising tool for large-scale population risk stratification and potentially opening new avenues for bladder cancer prevention and management. Such a model facilitates targeted screening of high-risk populations, thereby reducing incidence, while simultaneously minimizing unnecessary evaluations in low-risk groups to optimize healthcare resource allocation. Although direct comparisons with existing models were not conducted, this strategy may offer distinct advan-

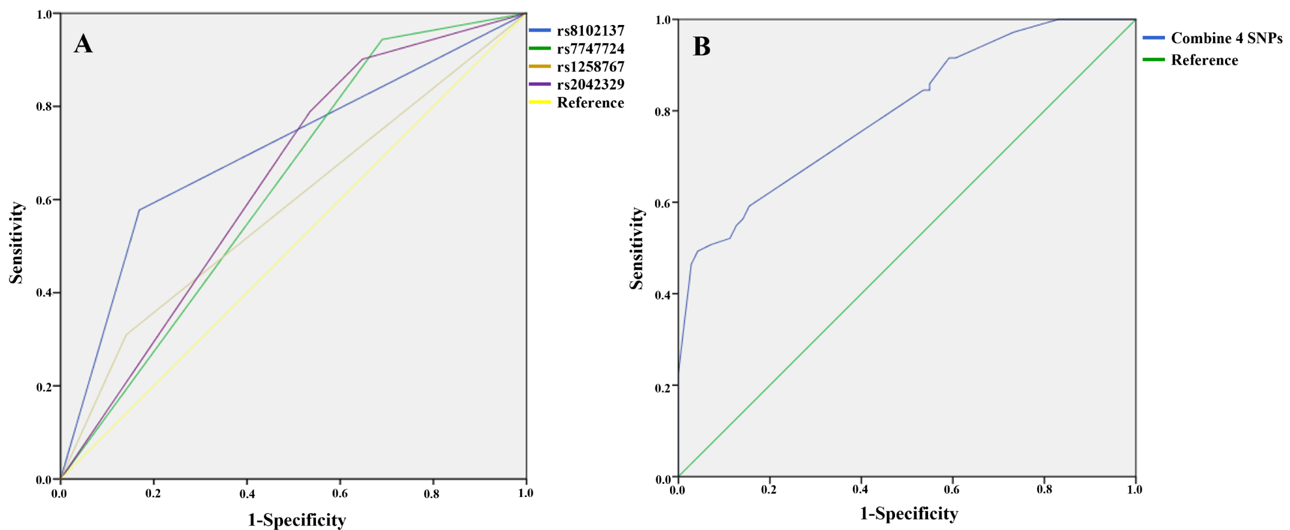

Fig. 3. The ROC curve analysis of the selected SNPs. (A) The results of predicting disease risk with four SNPs as individual risk factors. The AUC of rs8102137 was 0.704; the AUC of rs7747724 was 0.627; the AUC of rs1258767 was 0.585; and the AUC of rs2042329 was 0.641. (B) The results of predicting disease risk combining four SNPs; the AUC of combining four SNPs was 0.797. Abbreviations: ROC, receiver operating characteristic; AUC, area under the curve; SNP, single nucleotide polymorphism.

Table 6. Allele frequencies of the Polymorphic SNPs.

rs ID	Allele	Normal frequency	Case frequency
rs62185668	*C*	0.53	0.52
	T	0.47	0.48
rs1014971	*C*	0.77	0.58
	A	0.23	0.42
rs11892031	*G*	0.52	0.51
	A	0.48	0.49
rs8102137	*C*	0.71	0.92
	T	0.29	0.08
rs7747724	*G*	0.53	0.65
	T	0.47	0.35
rs1258767	*G*	0.85	0.93
	A	0.15	0.07
rs2042329	*C*	0.49	0.62
	A	0.51	0.38

Note: SNP, single nucleotide polymorphism.

tages and serve as a complementary tool to enhance the accuracy of bladder cancer risk assessment.

Population-specific Genetic Heterogeneity in Bladder Cancer Susceptibility

The genetic architecture of complex human traits shows significant heterogeneity across populations [20]. Currently, most GWASs are performed in individuals of European descent [12], raising concerns about the general applicability of these identified SNPs to other populations. This situation underscores the urgent need for GWAS in diverse, non-European cohorts. Our analysis of the Han Chinese population revealed that twelve of the nineteen examined SNPs (rs401681, rs2294008, rs907611, rs10936599, rs1495741, rs9642880, rs12216499, rs710521, rs798766, rs7238033, rs10094872, and rs5003154) demonstrated no polymorphic variation, seven (rs62185668, rs1014971, rs11892031, rs8102137, rs7747724, rs1258767, and rs2042329) were polymorphic, and four (rs8102137, rs7747724, rs1258767, and rs2042329) were substantially associated with bladder cancer risk. These findings provide strong evidence that the biological effects of genetic variations differ significantly across populations, highlighting the population-specific nature of genetic susceptibility. However, the observed link between these variations and bladder cancer phenotypes requires further large-scale studies to clarify their underlying biological mechanisms and clinical significance.

Biological Significance of the Four Associated SNPs

SNPs are the DNA sequence variations caused by the substitution, insertion, or deletion of a single nucleotide at the genomic level. Such variations occur at a certain frequency within populations and often serve as important genetic markers in disease association studies [21,22]. In cancer research, SNPs are strongly and intricately associated with tumor susceptibility. They contribute to tumor occurrence and progression by regulating gene expression, changing protein structure and function, and affecting the activation of key signaling pathways [23,24].

For example, rs8102137, located in the 19q12 region near the cyclin E1 gene (*CCNE1*), has been identified as a crucial bladder cancer susceptibility variant. The risk allele (*C*) functions as a cis-acting methylation quantitative trait locus (mQTL) and contributes to the carcinogenesis of bladder cancer by influencing *CCNE1* splicing, such

Table 7. The chi-square test results for polymorphic SNPs.

rs ID	Allele	p-value	Bonferroni	Risk allele	Odd ratio (OR)	95% CI
rs62185668	$C > T$	0.905	–	C	1.014	0.812–1.265
rs1014971	$C > A$	**0.001**	**<0.0001**	C	**1.313**	**1.113–1.550**
rs11892031	$G > A$	0.905	–	G	1.014	0.810–1.269
rs8102137	$C > T$	**<0.0001**	**<0.0001**	C	**1.287**	**1.146–1.446**
rs7747724	$G > T$	**0.030**	**<0.0001**	G	**1.240**	**1.019–1.509**
rs1258767	$G > A$	**0.024**	**0.016**	G	**1.100**	**1.012–1.196**
rs2042329	$C > A$	**0.032**	**0.008**	C	**1.257**	**1.018–1.552**

Note: Bold indicates SNPs with statistical significance ($p < 0.05$). CI, confidence interval range; SNP, single nucleotide polymorphism.

Table 8. The binary Logistic regression analysis for five SNPs.

Variables	Standard error	Significance	95% CI
rs1014971	1.047	0.316	0.045–2.725
rs8102137	**0.397**	**<0.000**	**3.083–14.645**
rs7747724	**0.575**	**<0.000**	**2.436–23.217**
rs1258767	**0.427**	**0.018**	**1.186–6.323**
rs2042329	**0.318**	**0.010**	**1.218–4.239**

Note: Bold indicates SNPs with statistical significance ($p < 0.05$). CI, confidence interval range; SNP, single nucleotide polymorphism.

as promoting transcripts lacking exon 7 in bladder tissues, and regulating G1-to-S phase transition of the cell cycle [25]. As a bladder cancer susceptibility variant identified by GWAS, this SNP is located 6 kb upstream of *CCNE1* within an LD block containing 47 linked SNPs ($r^2 \geq 0.7$). Furthermore, GWASs have reported that rs8102137 is substantially associated with bladder cancer risk, particularly in aggressive subtypes (high-grade non-muscle-invasive and muscle-invasive bladder cancer), but not in non-aggressive bladder cancer (p = 0.27). In the combined analysis of two GWAS, the per-allele OR was 1.18 (95% CI = 1.09–1.28, $p = 5.6 \times 10^{-5}$), while there was no significant association with non-aggressive bladder cancer (p = 0.27) [7,13]. Moreover, in a comprehensive three-stage GWAS meta-analysis, the association achieved genome-wide significance (per C allele OR = 1.13, 95% CI = 1.09–1.17, $p = 1.7 \times 10^{-11}$), which was validated in independent populations [13].

In our study, we further confirmed the association of rs8102137 with bladder cancer risk in 71 cases and 71 controls. The risk allele C demonstrated an OR of 1.287 (95% CI: 1.146–1.446), consistent with the previously reported trend but with a slightly higher effect than the previously reported 1.13–1.18 [7,13,26]. This discrepancy may be due to the relatively small sample size, differences in ethnic background, geographical variations, or variations in the confounder control. These findings further highlight rs8102137 as a reliable genetic susceptibility marker for bladder cancer.

Functionally, the *CCNE1* gene carrying rs8102137 is involved in several cancer-related pathways, such as the cell cycle and p53 signaling pathways. The risk allele enhances cyclin E overexpression, accelerating the cell cycle progression, leading to genomic instability, which promotes the development of aggressive bladder cancer [7,27]. Additionally, rs8102137 also shows a significant gene-diet interaction with adherence to the long-chain fat guideline in the 2010 Alternate Healthy Eating Index (AHEI-2010). In individuals with the *TT* genotype, high adherence was associated with a reduced risk of bladder cancer (OR = 0.92, 95% CI = 0.87–0.98), suggesting that dietary factors may modify genetic risk through gene-diet interaction [8]. Moreover, this SNP is also part of a strong LD block and is associated with variants such as rs200996365, which regulates *CCNE1* transcription by affecting the binding of the transcription factor KLF5, further supporting its role in the pathogenesis of bladder cancer [26]. Additionally, its association with bladder cancer has been consistently validated across multiple GWAS, with a combined analysis p value reaching 1×10^{-6}, and it frequently appears in four-variant combination models stratified by smoking status. Combined with the results of this study, these findings strongly reinforce rs8102137 as a crucial genetic susceptibility marker for bladder cancer [28,29].

Similarly, rs2042329 is located in the intron of the *CWC27* spliceosome-associated cyclophilin (CWC27) gene on chromosome 5q12.3, exhibiting ethnicity-specific associations in bladder cancer genetic susceptibility. Initially identified in a Chinese cohort of 3406 bladder cancer cases and 4645 controls, this SNP was substantially linked to increased bladder cancer risk, with an OR of 1.40 for the risk allele ($p = 4.61 \times 10^{-11}$). Moreover, the risk allele was also found to upregulate CWC27 mRNA and protein expression in bladder cancer tissues, while functional analyses suggested an oncogenic role through enhanced cell proliferation and inhibition of apoptosis [16]. In contrast, this SNP showed no association with bladder cancer risk in individuals of European descent [16].

In the UK Biobank cohort involving 1534 bladder cancer patients, rs2042329 was associated with an elevated risk of recurrence (OR = 1.26, 95% CI = 1.10–1.48, p = 0.001),

DM

Table 9. The area under the curve results of the predicted diseases for four SNPs.

Inspection result variables	Regional AUC	Standard error (a)	Asymptotic significance (b)	95% CI
Combine 4 SNPs	0.797	0.036	0.000	0.726–0.868
rs8102137	0.704	0.044	0.000	0.617–0.791
rs7747724	0.627	0.047	0.009	0.535–0.719
rs1258767	0.585	0.048	0.082	0.491–0.678
rs2042329	0.641	0.047	0.004	0.550–0.732

Note: $p < 0.05$. AUC, area under the curve; CI, confidence interval range; SNP, single nucleotide polymorphism.

Table 10. Outcomes of the model with multiple SNP interactions in multi-factor dimensionality reduction analysis.

Model	Bal. Acc. training	Bal. Acc. testing	CV consistency	χ^2 (p)
rs8102137	0.7042	0.7042	10/10	25.3174 ($p < 0.0001$)
rs8102138, rs2042329	0.7222	0.6690	9/10	28.9298 ($p < 0.0001$)
rs8102139, rs1258767, rs2042329	0.7363	0.6761	8/10	31.5781 ($p < 0.0001$)
rs8102140, rs7747724, rs1258767, rs2042329	0.7543	0.7113	10/10	38.7273 ($p < 0.0001$)

Note: Bal. Acc., Balance accuracy; CV, cross-validation.

with the strongest log10 Bayes factor (log10(BF) = 1.56) among recurrence-associated SNPs. Although the direction of this association was consistent with the original Chinese study, albeit with a more moderate effect size [30].

In our study, the risk allele *C* of rs2042329 was substantially associated with bladder cancer risk (OR = 1.257, 95% CI: 1.018–1.552). The effect size falls between that observed for recurrence in the UK Biobank cohort (OR = 1.26) and the initial risk association in the original Chinese cohort (OR = 1.40), further supporting its role in bladder cancer pathogenesis. These results suggest that the effect of rs2042329 may vary across populations, influenced by factors such as disease stage and genetic background. Taken together, current evidence reveals that rs2042329 is a valuable genetic marker for both bladder cancer risk and recurrence, warranting further mechanistic studies, particularly in Asian populations.

Furthermore, rs7747724 and rs1258767 are cataloged as bladder cancer-associated SNPs in the Phenotype-Genotype Integrator database of the National Library of Medicine (National Center for Biotechnology Information; https://www.ncbi.nlm.nih.gov/gap/phegeni), yet their association with bladder cancer susceptibility has not been validated to date. Our study identified significant associations of these two SNPs with bladder cancer susceptibility, thereby enriching existing genetic data and providing novel insights into the genetic basis of bladder cancer susceptibility.

Variant rs7747724 is located in an intron of the CDK5 regulatory subunit-associated protein 1 like 1 (*CDKAL1*) gene. The *CDKAL1* gene is localized in the 6p22.3 chromosomal region, which is frequently amplified in bladder cancer, particularly in muscle-invasive bladder cancer (MIBC) and high-grade non-muscle-invasive bladder cancer (NMIBC) [31,32]. Although its expression highly correlates with amplification of this region, shRNA-mediated knockdown of *CDKAL1* exerts no significant effect on cell proliferation in bladder cancer cell lines harboring 6p22.3 amplification, unlike *E2F3*—another gene in the same region that plays a pivotal role in regulating cell proliferation [31,33,34]. Recent studies have revealed the involvement of *CDKAL1* in smoking-related bladder carcinogenesis. For example, the SNP rs9348451 within *CDKAL1* interacts with smoking status ($OR_{interaction}$ = 1.38, $p_{interaction}$ = 1.08×10^{-2}) to increase bladder cancer risk and is associated with *CDKAL1* expression in tumor tissues, with a more pronounced correlation observed in smokers ($p <$ 0.001). Additionally, *CDKAL1* is upregulated in bladder cancer compared to adjacent normal tissues and is further induced by cigarette smoke extracts [35]. Furthermore, *CDKAL1* resides within copy number variation (CNV) regions detected in urine sediment DNA from bladder cancer patients, highlighting its potential as a non-invasive diagnostic marker [36].

In our study, we identified a substantial association between rs7747724 in *CDKAL1* and bladder cancer risk (OR = 1.240, 95% CI: 1.019–1.509), representing the first validation of this locus. This finding provides new genetic evidence for the role of *CDKAL1* in bladder cancer pathogenesis, expands the spectrum of susceptibility loci, and suggests that *CDKAL1* may contribute to bladder carcinogenesis through gene-environment interactions and genomic alterations. The identification of rs7747724 offers a novel target for investigating underlying molecular mechanisms.

Furthermore, rs1258767 is located in the intron of the formin 1 (*FMN1*) gene. *FMN1* belongs to the human FMN1 homology protein family and is homologous to yeast *Bni1p*, which is involved in cell polarity regulation and microtubule organization [37]. As a member of this family, *FMN1* is likely involved in processes such as cell polarity establishment, microtubule organization, cell proliferation, and differentiation. In this study, we report an association between rs1258767 in *FMN1* and bladder cancer risk (OR = 1.100, 95% CI: 1.012–1.196), consistent with the poten-

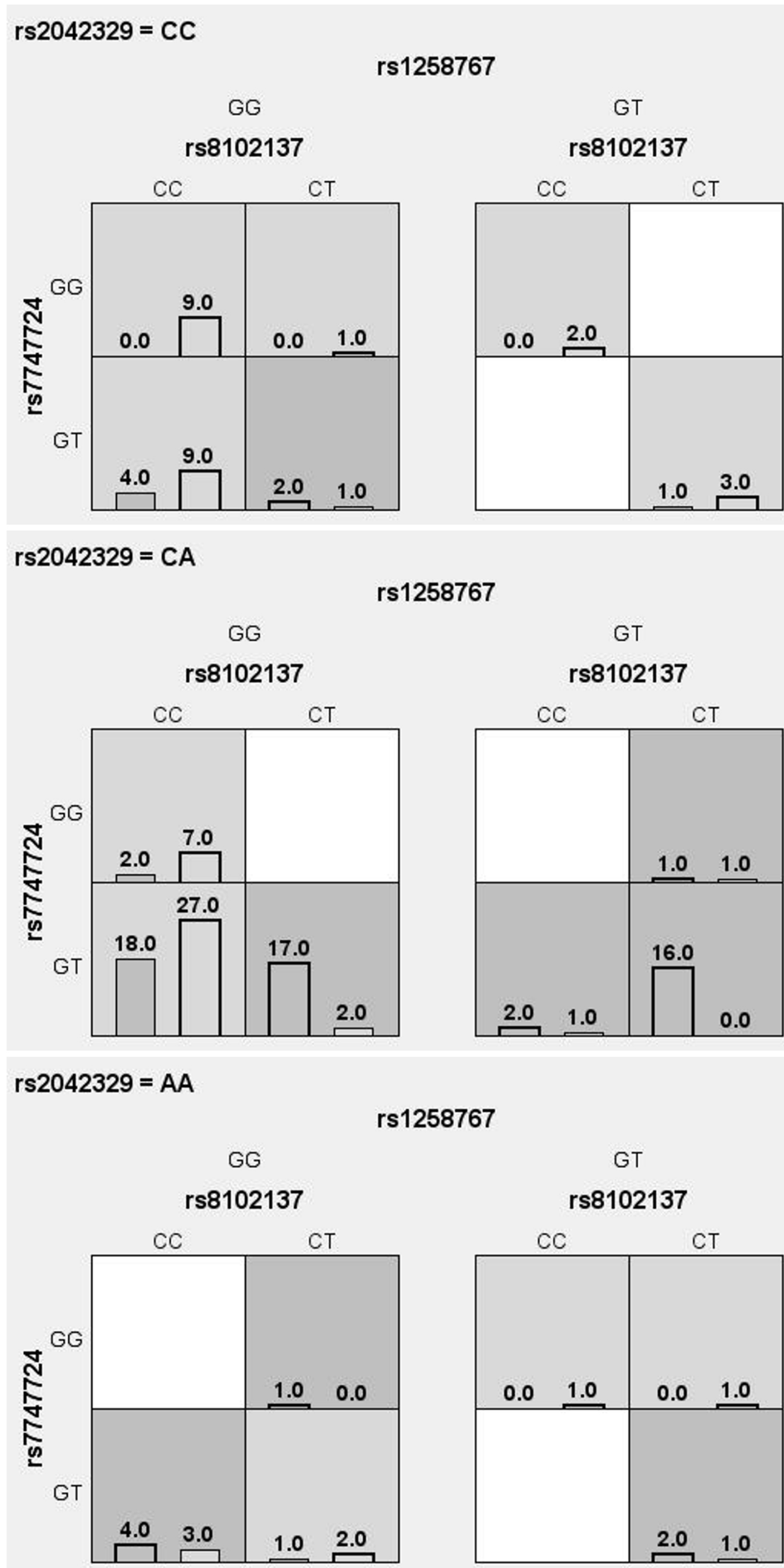

Fig. 4. The MDR analysis of the interaction of four SNPs cell values. Abbreviations: MDR, multi-factor dimensionality reduction; SNP, single nucleotide polymorphism.

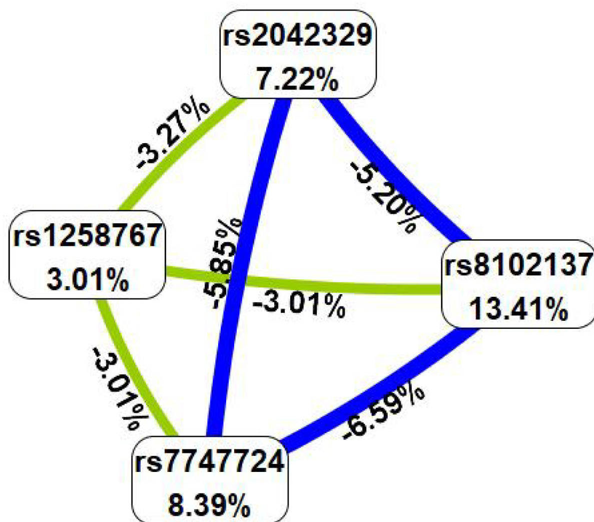

Fig. 5. The MDR analysis of the interaction network of four SNPs. The entropy in a single SNP reflects the primary independent effect. The entropy value marked on the line connecting two SNPs represents the interaction. The blue line indicates high redundancy, while the green line represents reduced redundancy. Abbreviations: MDR, multifactor dimensionality reduction; SNP, single nucleotide polymorphism.

tial regulatory role of *FMN1*. This finding expands the research scope on *FMN1* genetic variants in bladder cancer and provides a novel genetic marker for exploring its functional mechanisms.

Advantages of the Mini-sequencing Assay

In this study, the mini-sequencing detection system combined multiplex PCR, SNaPshot, and CE technologies. Mini-sequencing is based on the principle of dideoxynucleotide chain termination, incorporating fluorescently labeled ddNTPs into the 3′ end of SBE primers to generate fluorescently labeled amplicons. This technology is widely used in DNA laboratories due to its high sensitivity and accuracy. There are multiple technical approaches available for SNP screening, such as Sanger sequencing, whole-genome sequencing (WGS), and whole-exome sequencing (WES) via next-generation sequencing (NGS) technology. Compared with NGS, Sanger sequencing and mini-sequencing methods are generally more accurate, cost-effective, and suitable for degraded DNA. Clinical practices and workflow often involve delays and formalin fixation of specimens, which leads to tissue degradation and reduces the quality of DNA for WGS or WES. Moreover, many diagnostic laboratories lack NGS facilities, limiting their clinical application for common diseases.

In contrast, the methodologies evaluated in our study exhibit greater feasibility for analyzing degraded DNA samples and are readily adaptable to routine molecular diagnostic laboratories, thereby addressing the escalating demand for large-scale screening bladder cancer. Beyond this, mini-sequencing demonstrates substantial advantages for targeted low-to-moderate throughput SNP genotyping, particularly for the rapid validation of specific genetic markers in clinical translation.

Study Limitations and Future Directions

This study is constrained by three notable limitations. First, imaging data were not integrated due to inherent study design constraints. The research dataset, sourced from a prospective clinical registry, primarily comprises clinical variables and genetic markers (including the nineteen SNPs analyzed) but lacks radiological findings. This omission limits the comprehensive interpretability of the results: (i) imaging data could contextualize SNP associations by linking genetic variants to tumor-specific traits, including dimensions, invasiveness, anatomical location, and (ii) integrating radiological parameters could facilitate the development of combined genetic-radiological nomograms, thereby enhancing clinical translational utility.

Second, the sample size was relatively small. A limited cohort size diminishes statistical power to detect weak or moderate effects, particularly for low-frequency variants among the nineteen SNPs, where the paucity of variant carriers hinders robust association testing. It also compromises the precision of effect size estimates and may elevate the risk of overestimating association strength—an established bias in underpowered genetic association studies. Furthermore, the small sample size constrains the analysis of gene-gene and gene-environment interactions, which are pivotal for unraveling the multifactorial etiology of bladder cancer.

Third, external validation was not conducted. The predictive performance of the four associated SNPs has not been corroborated in independent cohorts. Without external replication, the generalizability of the findings remains equivocal, as observed associations may reflect population-specific attributes (e.g., regional demographic nuances, institutional clinical protocols) or stochastic variation.

However, these limitations do not undermine the core findings of the study. Despite these challenges, our study underscores four SNPs that offer novel insights into the genetic basis of bladder cancer. Their application is particularly promising in contexts where imaging data is unavailable, making them accessible genetic predictors for early screening and risk assessment, complementary to radiological examination rather than a replacement.

Future research should prioritize three avenues: (i) incorporating imaging data to explore interactions between SNPs and radiological features for improved risk stratification, (ii) expanding multicenter cohorts to enhance statistical power, capture subgroup effects, and facilitate gene-environment interaction analyses, and (iii) validating the performance of these four associated SNPs in independent populations with long-term follow-up data to appraise their prognostic value, including recurrence, metastasis, and treatment response.

Conclusion

This study developed a mini-sequencing detection system targeting nineteen bladder cancer-related SNP. Statistical analyses revealed four SNP significantly associated with bladder cancer risk. Based on these findings, the study further developed a bladder cancer risk prediction model by integrating these variants, achieving a more precise assessment and prediction of individual susceptibility to bladder cancer. Our findings highlight the crucial role of specific genetic variations in the pathogenesis of bladder cancer and provide both theoretical insights and practical approaches for improving early risk screening, personalized prevention strategies, and precision treatment of bladder cancer.

Availability of Data and Materials

The data and materials in the current study are available from the corresponding author on reasonable request.

Author Contributions

YQF, QCG, and GTX made substantial contributions to the conception or design of the work. DQC and CS made substantial contributions to the analysis of data for the work. YGJ, CYT and DDC acquired the data. YQF wrote the article and performed data analysis. FW interpretation of data for the work and revised the article and gave the final approval of the version to be submitted. All authors contributed to important editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

This study was approved by the Medical Ethics Committee of The Second Hospital of Jiaxing (Approval No.2025-130). Written informed consent was obtained from all participants before their enrollment in the study, in accordance with the Declaration of Helsinki.

Acknowledgment

Not applicable.

Funding

This work was partially supported by fund of Zhejiang Province Health Science and Technology Program (2022KY1256), Jiaxing University Student Research Training (SRT) program (8517241017).

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Dobruch J, Oszczudłowski M. Bladder Cancer: Current Challenges and Future Directions. Medicina (Kaunas, Lithuania). 2021; 57: 749. https://doi.org/10.3390/medicina57080749.
[2] He Q, Wu S, Zhou Y, Liu Y, Xia B, Li W, *et al*. Genetic factors, adherence to healthy lifestyle behaviors, and risk of bladder cancer. BMC Cancer. 2023; 23: 965. https://doi.org/10.1186/s12885-023-11455-4.
[3] Tadesse DA, Rothman N, Xie S, Hurwitz LM, Friesen MC, Baris D, *et al*. Solvent Exposure, Genetic Susceptibility, and Risk of Bladder Cancer. Cancer Prevention Research (Philadelphia, Pa.). 2025; 18: 283–290. https://doi.org/10.1158/1940-6207.CAPR-24-0434.
[4] Siracusano S, Rizzetto R, Porcaro AB. Bladder cancer genomics. Urologia. 2020; 87: 49–56. https://doi.org/10.1177/0391560319899011.
[5] Kourie HR, Succar B, Chouery E, Mehawej C, Ahmadieh N, Zouein J, *et al*. Genetic susceptibility of bladder cancer in the Lebanese population. BMC Medical Genomics. 2022; 15: 217. https://doi.org/10.1186/s12920-022-01372-z.
[6] Koutros S, Kiemeney LA, Pal Choudhury P, Milne RL, Lopez de Maturana E, Ye Y, *et al*. Genome-wide Association Study of Bladder Cancer Reveals New Biological and Translational Insights. European Urology. 2023; 84: 127–137. https://doi.org/10.1016/j.eururo.2023.04.020.
[7] Del Pizzo JJ, Borkowski A, Jacobs SC, Kyprianou N. Loss of cell cycle regulators p27(Kip1) and cyclin E in transitional cell carcinoma of the bladder correlates with tumor grade and patient survival. American Journal of Pathology. 1999; 155: 1129–1136. https://doi.org/10.1016/S0002-9440(10)65216-9.
[8] Leeming RC, Koutros S, Karagas MR, Baris D, Schwenn M, Johnson A, *et al*. Diet quality, common genetic polymorphisms, and bladder cancer risk in a New England population-based study. European Journal of Nutrition. 2022; 61: 3905–3913. https://doi.org/10.1007/s00394-022-02932-w.
[9] Wang X, Guo Z, Zhu H, Xin J, Yuan L, Qin C, *et al*. Genetic variants in splicing factor genes and susceptibility to bladder cancer. Gene. 2022; 809: 146022. https://doi.org/10.1016/j.gene.2021.146022.
[10] Dyrskjøt L, Hansel DE, Efstathiou JA, Knowles MA, Galsky MD, Teoh J, *et al*. Bladder cancer. Nature Reviews. Disease Primers. 2023; 9: 58. https://doi.org/10.1038/s41572-023-00468-9.
[11] Lopez-Beltran A, Cookson MS, Guercio BJ, Cheng L. Advances in diagnosis and treatment of bladder cancer. BMJ (Clinical Research Ed.). 2024; 384: e076743. https://doi.org/10.1136/bmj-2023-076743.
[12] Figueroa JD, Ye Y, Siddiq A, Garcia-Closas M, Chatterjee N, Prokunina-Olsson L, *et al*. Genome-wide association study identifies multiple loci associated with bladder cancer risk. Human Molecular Genetics. 2014; 23: 1387–1398. https://doi.org/10.1093/hmg/ddt519.
[13] Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, *et al*. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nature Genetics. 2010; 42: 978–984. https://doi.org/10.1038/ng.687.
[14] Kiemeney LA, Sulem P, Besenbacher S, Vermeulen SH, Sigurdsson A, Thorleifsson G, *et al*. A sequence variant at 4p16.3 confers susceptibility to urinary bladder cancer. Nature Genetics. 2010; 42: 415–419. https://doi.org/10.1038/ng.558.
[15] Rafnar T, Sulem P, Thorleifsson G, Vermeulen SH, Helgason H, Saemundsdottir J, *et al*. Genome-wide association study yields variants at 20p12.2 that associate with urinary bladder cancer. Human Molecular Genetics. 2014; 23: 5545–5557. https://doi.org/10.1093/hmg/ddu264.

[16] Wang M, Li Z, Chu H, Lv Q, Ye D, Ding Q, *et al*. Genome-Wide Association Study of Bladder Cancer in a Chinese Cohort Reveals a New Susceptibility Locus at 5q12.3. Cancer Research. 2016; 76: 3277–3284. https://doi.org/10.1158/0008-5472.CAN-15-2564.

[17] Kiemeney LA, Thorlacius S, Sulem P, Geller F, Aben KKH, Stacey SN, *et al*. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nature Genetics. 2008; 40: 1307–1312. https://doi.org/10.1038/ng.229.

[18] Rafnar T, Vermeulen SH, Sulem P, Thorleifsson G, Aben KK, Witjes JA, *et al*. European genome-wide association study identifies SLC14A1 as a new urinary bladder cancer susceptibility gene. Human Molecular Genetics. 2011; 20: 4268–4281. https://doi.org/10.1093/hmg/ddr303.

[19] Garcia-Closas M, Ye Y, Rothman N, Figueroa JD, Malats N, Dinney CP, *et al*. A genome-wide association study of bladder cancer identifies a new susceptibility locus within SLC14A1, a urea transporter gene on chromosome 18q12.3. Human Molecular Genetics. 2011; 20: 4282–4289. https://doi.org/10.1093/hmg/ddr342.

[20] Urbanowicz RJ, Andrew AS, Karagas MR, Moore JH. Role of genetic heterogeneity and epistasis in bladder cancer susceptibility and outcome: a learning classifier system approach. Journal of the American Medical Informatics Association: JAMIA. 2013; 20: 603–612. https://doi.org/10.1136/amiajnl-2012-001574.

[21] Xiong J, Yang L, Deng YQ, Yan SY, Gu JM, Li BH, *et al*. The causal association between smoking, alcohol consumption and risk of bladder cancer: A univariable and multivariable Mendelian randomization study. International Journal of Cancer. 2022; 151: 2136–2143. https://doi.org/10.1002/ijc.34228.

[22] Samara M, Papathanassiou M, Farmakioti I, Anagnostou M, Satra M, Mitrakas L, *et al*. Renin-Angiotensin System Single Nucleotide Polymorphisms Are Associated with Bladder Cancer Risk. Current Oncology (Toronto, Ont.). 2021; 28: 4702–4708. https://doi.org/10.3390/curroncol28060396.

[23] Gakis G, Perner S, Stenzl A, Renninger M. The role of single-nucleotide polymorphisms of the 8q24 chromosome region in patients with concomitant bladder and prostate cancer. Scandinavian Journal of Urology. 2022; 56: 126–130. https://doi.org/10.1080/21681805.2022.2049362.

[24] Zhang X, Han C, He J. Research progress of oncogene and tumor suppressor gene in bladder cancer. Panminerva Medica. 2015; 57: 191–200.

[25] Jordahl KM, Phipps AI, Randolph TW, Tinker LF, Nassir R, Hou L, *et al*. Mediation by differential DNA methylation of known associations between single nucleotide polymorphisms and bladder cancer risk. BMC Medical Genetics. 2020; 21: 228. https://doi.org/10.1186/s12881-020-01172-1.

[26] Pattison JM, Posternak V, Cole MD. Transcription Factor KLF5 Binds a Cyclin E1 Polymorphic Intronic Enhancer to Confer Increased Bladder Cancer Risk. Molecular Cancer Research: MCR. 2016; 14: 1078–1086. https://doi.org/10.1158/1541-7786.MCR-16-0123.

[27] de Maturana EL, Rava M, Anumudu C, Sáez O, Alonso D, Malats N. Bladder Cancer Genetic Susceptibility. A Systematic Review. Bladder Cancer (Amsterdam, Netherlands). 2018; 4: 215–226. https://doi.org/10.3233/BLC-170159.

[28] Zhang X, Zhang Y. Bladder Cancer and Genetic Mutations. Cell Biochemistry and Biophysics. 2015; 73: 65–69. https://doi.org/10.1007/s12013-015-0574-z.

[29] Selinski S, Blaszkewicz M, Ickstadt K, Gerullis H, Otto T, Roth E, *et al*. Identification and replication of the interplay of four genetic high-risk variants for urinary bladder cancer. Carcinogenesis. 2017; 38: 1167–1179. https://doi.org/10.1093/carcin/bgx102.

[30] Lipunova N, Wesselius A, Cheng KK, van Schooten FJ, Cazier JB, Bryan RT, *et al*. External Replication of Urinary Bladder Cancer Prognostic Polymorphisms in the UK Biobank. Frontiers in Oncology. 2019; 9: 1082. https://doi.org/10.3389/fonc.2019.01082.

[31] Shen H, Morrison CD, Zhang J, Underwood W, 3rd, Yang N, Frangou C, *et al*. 6p22.3 amplification as a biomarker and potential therapeutic target of advanced stage bladder cancer. Oncotarget. 2013; 4: 2124–2134. https://doi.org/10.18632/oncotarget.1485.

[32] Bellmunt J. Stem-Like Signature Predicting Disease Progression in Early Stage Bladder Cancer. The Role of E2F3 and SOX4. Biomedicines. 2018; 6: 85. https://doi.org/10.3390/biomedicines6030085.

[33] Olsson AY, Feber A, Edwards S, Te Poele R, Giddings I, Merson S, *et al*. Role of E2F3 expression in modulating cellular proliferation rate in human bladder and prostate cancer cells. Oncogene. 2007; 26: 1028–1037. https://doi.org/10.1038/sj.onc.1209854.

[34] Hurst CD, Tomlinson DC, Williams SV, Platt FM, Knowles MA. Inactivation of the Rb pathway and overexpression of both isoforms of E2F3 are obligate events in bladder tumours with 6p22 amplification. Oncogene. 2008; 27: 2716–2727. https://doi.org/10.1038/sj.onc.1210934.

[35] Gao F, Li H, Mao Z, Xiao Y, Du M, Wang S, *et al*. An integrative approach for identification of smoking-related genes involving bladder cancer. Archives of Toxicology. 2023; 97: 177–188. https://doi.org/10.1007/s00204-022-03380-5.

[36] Cai YX, Yang X, Lin S, Xu YW, Zhu SW, Fan DM, *et al*. Low-Coverage Sequencing of Urine Sediment DNA for Detection of Copy Number Aberrations in Bladder Cancer. Cancer Management and Research. 2021; 13: 1943–1953. https://doi.org/10.2147/CMAR.S295675.

[37] Katoh Y, Katoh M. Identification and characterization of CDC50A, CDC50B and CDC50C genes in silico. Oncology Reports. 2004; 12: 939–943.