

TWO CO-INHERITED SNPS OF THE TELOMERASE REVERSE TRANSCRIPTASE (TERT) GENE ARE ASSOCIATED WITH IRAQI PATIENTS WITH LUNG CANCER

DVA KO-NASLEĐENA SNPS GENA TELOMERAZA REVERZNE TRANŠKRIPTAZE (TERT) KOJA SE DOVODE U VEZU SA PACIJENTIMA OD RAKA PLUĆA U IRAKU

Zahraa K. Lawi^{1,2}, Ahmed H. Alkhammas³, Malek Elerouri⁴,
Ibtissem Ben Amara⁵, Mohammed Baqur S. Al-Shuhaib^{3,6*}

¹Department of Biology, College of Science, University of Kufa, Najaf, 54001, Iraq

²Department of Biology, Faculty of Science, University of Sfax, PB 261, 3000 SFAX, Tunisia

³Department of Animal Production, College of Agriculture, Al-Qasim Green University, Al-Qasim, Babil 51001, Iraq

⁴Department of Enzyme Engineering and Microbiology, National Engineering School in Sfax, University of Sfax, B.P. 1173, 3038 Sfax, Tunisia

⁵Laboratory of Medicinal and Environment Chemistry, Higher Institute of Biotechnology, University of SFAX, PB 261, 3000 SFAX, Tunisia

⁶College of Medicine, University of Warith Al-Anbiyaa, Karbala, 56001, Iraq

Summary

Background: The telomerase reverse transcriptase (TERT) gene is essential polymorphic loci linked to most malignant tumors. This study assessed the association between the *TERT* gene and non-small cell lung carcinoma (NSCLC) in Iraq.

Methods: Genomic DNA samples were extracted from a total of 200 samples of blood. Four specific PCR fragments were designed to amplify four high-frequency rs2735940, rs2736098, rs2736100, and rs10069690 SNPs within the *TERT* gene. Single-strand conformation polymorphism (SSCP) followed by sequencing reactions were used for genotyping and validating the amplified fragments.

Results: Individuals with the genotype rs2735940: A/G were at a significantly greater risk of developing NSCLC ($P=0.0299$; OD 2.3158; CI_{95%} 1.0853 to 4.9414). Individuals with the genotype rs2736098: C/T were also significantly associated with the increased likelihood of developing NSCLC ($P=0.0363$; OD 2.1583; CI_{95%} 1.0503 to 4.4351). Linkage disequilibrium analysis showed that both SNPs showed a very high level of patient coinherence. The LD plot showed that allele T of rs2736098 had

Kratak sadržaj

Uvod: Gen za reverznu transkriptazu telomeraze (TERT) je esencijalni polimorfni lokus povezan sa većinom malignih tumora. Ova studija je procenila vezu između TERT gena i karcinoma pluća ne-malih ćelija (NSCLC) u Iraku.

Metode: Uzorci genomske DNK su ekstrahovani iz ukupno 200 uzoraka krvi. Četiri specifična PCR fragmenta su dizajnirana da amplifikuju četiri visokofrekventna rs2735940, rs2736098, rs2736100 i rs10069690 SNP u okviru TERT gena. Za genotipizaciju i validaciju amplifikovanih fragmenata korišćen je jednonančani konformacioni polimorfizam (SSCP) praćen reakcijama sekvenciranja.

Rezultati: Pacijenti sa genotipom rs2735940: A/G su bili u značajno većem riziku od razvoja NSCLC ($P=0,0299$; OD 2,3158; CI_{95%} 1,0853 do 4,9414). Pojedinci sa genotipom rs2736098: C/T su takođe značajno povezani sa povećanom verovatnoćom razvoja NSCLC ($P=0,0363$; OD 2,1583; CI_{95%} 1,0503 do 4,4351). Analiza neravnoteže veze je pokazala da su oba SNP-a pokazala veoma visok nivo konasleđa pacijenata. LD dijagram je pokazao da je alel T od rs2736098 sarađivao sa alelom G od rs2735940 da generiše TG haplotip kod pacijenata.

Address for correspondence:

Mohammed Baqur S. Al-Shuhaib
Department of Animal Production, College of Agriculture
Al-Qasim Green University, Al-Qasim-8, Babil 51001, Iraq
Phone: +9647707115693
e-mail: mohammed79@agre.uoqasim.edu.iq

collaborated with allele G of rs2735940 to generate TG haplotype in patients. According to our findings, both *TERT*-rs2735940: A/G and *TERT*-rs2736098: C/T SNPs were found to be significant associations with the elevated risk of NSCLC. Both SNPs showed the highest values of co-inheritance in patients. This co-inheritance is mainly represented by alleles rs2735940: A and rs2736098: C. Both pathogenic T and G alleles have generated TG haplotype that is only available in patients' samples.

Conclusion: This study suggests employing the haplotype TG as a promising biomarker for the early diagnosis of NSCLC. These findings need further validation by large-scale investigation with a larger size of samples in the study population.

Keywords: *TERT*, polymorphism, NSCLC, rs2736098, rs10069690

Introduction

Lung cancer is one of the most severe public health challenges and a major contributor to cancer-related morbidities and fatalities worldwide. Lung cancer cases are being reported more frequently in the Middle East regions (1). Recent years have shown a steady increase in the clinical risk factors for lung cancer and its associated consequences (2). It has been demonstrated that traditional methods of diagnosing lung cancers are less effective due to several issues related to higher mortality rates reported in developing malignant lung tumors in their advanced stages (3). Non-small cell lung cancer (NSCLC) is one of the most critical types of respiratory tract tumours, making up 80% of lung carcinogenesis worldwide. Unfortunately, it has been reported that surgical treatment may not be possible if NSCLC develops in severe stages III and IV (4). As a result, NSCLC is substantially more challenging to detect and cure early than other types of cancer (5). This consequence is linked to people with lung cancer, typically detected in advanced stages, having a poor prognosis (6). Given the sharp decline in lung cancer survival rates from early to advanced stages, more focus should be given to this issue regarding its early diagnosis (7). Because of this, it is critical to detect this type of cancer early by integrating molecular research with traditional clinical evaluations (8, 9). Recently, the onset and progression of lung cancer have been linked to various extents by a wide range of genetic variants (10, 11). The telomerase reverse transcriptase (*TERT*) gene is one of the essential polymorphic loci linked to various human cancers (12). The *TERT* gene has 16 exons and is positioned in chromosome 5 (5p15.33). The *TERT* gene encodes the catalytic component of telomerase, a ribonucleoprotein that plays a pivotal role in cancer initiation via telomere-dependent or telomere-independent pathways (13). The methods for maintaining telomeres entail intricate cellular adjustments brought on by *TERT*, such as *TERT* structural variants (14), *TERT* gene amplifications (15), *TERT* epigenetic changes (16), and alternative

lengthening of telomere (17), and *TERT* promoter mutations (18). Any alteration of these mechanisms is possibly associated with numerous types of carcinogenesis at a single-cell level (14). Telomere length is influenced by *TERT*-based telomerase activity, which can also act as an informative biomarker in the early diagnosis and prognosis of many malignancies (19). *TERT* aids telomerase in replenishing telomere sequences by extending the extreme ends of telomere strands to enable other polymerases to synthesize the complementary strand (20). It has been established that elevated *TERT* expression leads to telomerase activity restoration (21), implying that the transcriptional control of *TERT* plays a remarkable role in carcinogenesis (22). Due to its significance in numerous critical actions throughout the cells, *TERT* polymorphism may be connected to the early onset and progression of malignancies. It has been shown that the *TERT* gene is linked to a diverse spectrum of malignancies, such as pancreatic cancer (23), prostate cancer (24), gastric cancer (25), thyroid cancer (26), esophageal cancer (27), and bladder cancer (28), melanoma (29). However, controversial data have been found on the possible association of the *TERT* gene with lung cancers (4). This is due to the poor association with squamous cell carcinoma and NSCLC (30,31). But later on, many sources of data have shown that there is a strong connection between *TERT* variants and the risk of NSCLC (32–34). In Iraq, the prevalence of NSCLC has steadily increased. This is due to exposure of Iraqi people the several disastrous wars and the deterioration of health care infrastructures (35). However, the significance of several *TERT* SNPs in the development of NSCLC have not been identified in this population. Taking these data into consideration, this research aim was to investigate the possible association between four *TERT* high-frequency SNPs rs2735940, rs2736098, rs2736100, and rs10069690 with increased risk of NSCLC in cancer subjects in Iraq.

Prema našim nalazima, utvrđeno je da su i *TERT*-rs2735940: A/G i *TERT*-rs2736098: C/T SNP u značajnoj vezi sa povećanim rizikom od NSCLC. Oba SNP-a su pokazala najveće vrednosti ko-nasleđivanja kod pacijenata. Ovo ko-nasleđivanje je uglavnom predstavljeno alelima rs2735940: A i rs2736098: C. I patogeni T i G aleli su generisali TG haplotip koji je dostupan samo u uzorcima pacijenata.

Zaključak: Ova studija predlaže korišćenje haplotipa TG kao biomarkera koji obećava u smislu rane dijagnoze NSCLC. Potrebna je dalja validacija ovih nalaza opsežnim istraživanjem na većoj populaciji.

Ključne reči: *TERT*, polimorfizam, NSCLC, rs2736098, rs10069690

Materials and Methods

Controls and Subjects

The experiments performed in the study were conducted following the Helsinki Declaration for experiments involving people, and the biochemical research involving human subjects was approved by the Institutional Review Board (IRB) in the University of Kufa (IECIH/UOK 088/2020; CAAE 08802212). Signed written informed consent was obtained from all participants before being involved in the study. In this case-control study, a total of 200 individuals, of which 100 patients with NSCLC and 100 controls, were included. The details of the collected samples were described in (36). Briefly, the samples of patients were collected from the Middle-Euphrates Cancer Center (located in Najaf governorate) and Merjan Cancer Center (located in Babil governorate). The involved patients ranged in age from 24 to 80 (average age: 52.2). Each had previously received a diagnosis of NSCLC from trained personnel to diagnose malignan-

cies at the facilities mentioned above in the study area. In this investigation, 100 healthy volunteers with no prior history of lung cancer and ages ranging from 18 to 78 made up the control population (mean age: 44.4). The samples that were screened all belonged to citizens of Iraq, and they were all examined from February to December of 2021.

Genomic DNA extraction

Genomic DNA samples were extracted from the peripheral blood samples using a Blood/Cell DNA Mini Kit (Cat. No. GB100, Geneaid Co., Taipei, Taiwan). A Nanodrop spectrophotometric technique was used to confirm the purity of the extracted genomic DNA (Biodrop Co., UK). The isolated genomic DNA's integrity was assessed by agarose gel electrophoresis following the recommended standard instructions (37).

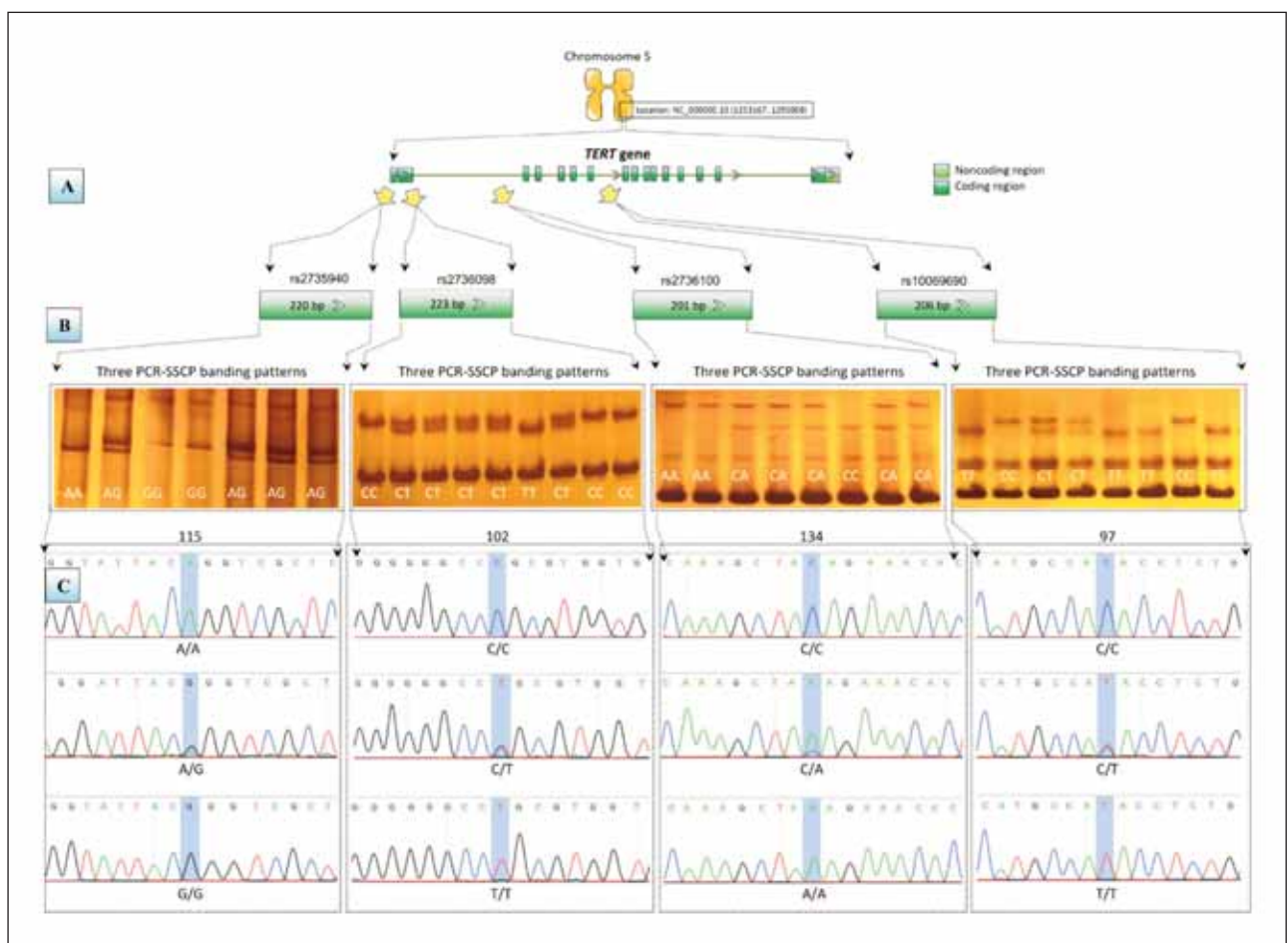


Figure 1 A schematic diagram of TERT genotyping using the PCR-SSCP-sequencing method. A) – PCR designing of four PCR-specific primers for the amplification of 220 bp, 223 bp, 201 bp, and 206 bp to flank rs2735940, rs2736098, rs2736100, and rs10069690, respectively. B – PCR-SSCP genotyping, in which all targeted SNPs showed three polymorphic patterns of nucleic acid variations. C – Sequencing reactions of the targeted loci as positioned in the amplified PCR fragments.

PCR

Four sets of PCR-specific primers were designed using NCBI primer BLAST tool to amplify four distinct PCR fragments within the *TERT* gene (38). In PCR designing, four high-frequency SNPs (rs2735940, rs2736098, rs2736100, and rs10069690) were targeted within total lengths of 220 bp, 223 bp, 201 bp, and 206 bp, respectively (Figure 1a). The specified features of PCR products that are required to provide the best resolutions in PCR-SSCP methods were met by the PCR amplicons when they were prepared with these lengths (39). The details of the PCR oligonucleotide sequences are displayed in Table I. The experiments of PCR were conducted by a lyophilized PCR AccuPower PreMix (Cat# K-2012, Bioneer Co., Korea), with a final volume of 20 μ L for each amplified fragment. After performing PCR experiments, it was confirmed that PCR products were in the expected sizes by electrophoresis on agarose gels.

Genotyping analysis

The genotyping experiments were conducted using PCR-single strand conformation polymorphism (SSCP). This sensitive and inexpensive genotyping approach can discriminate between PCR products that differ in only one nitrogen base using denaturation, chilling, and electrophoresis on polyacrylamide gels (40). Briefly, PCR amplicons were denatured at 94 for 7 to 8 min and then immediately frozen in ice for at least 10 min. Then, PCR products were subsequently loaded in 8% polyacrylamide gels until samples reached the bottom of the gels. Gels were stained with silver nitrate following the recommended procedure (41). Each detected SSCP band-

ing pattern was subjected to Sanger dideoxy-sequencing laboratories to confirm the obtained electrophoretic genotypes following the recommended instructions (Macrogen Inc., South Korea). The DNA chromatogram of each genotype was visualized using SnapGene Viewer ver. 4.0.4. (Insightful Science, Canada). The alignment of the observed variation with the referring DNA sequences was conducted using BioEdit suit, version 7.1 (DNASTAR; Madison, USA).

Statistical and functional analysis

Using the MedCalc online server, the odds ratios (ORs) and associated 95% confidence intervals (CIs) were computed to examine the genotype variations for the targeted SNPs between the patients and controls (42). Gen-Calculator software (www.gen-calculator.pl) was utilized to evaluate the likelihood that both the cases and the controls would deviate from the Hardy-Weinberg equilibrium (HWE). The linkage disequilibrium (LD) plots of the identified SNPs were constructed by Haploview software ver. 4.2 to assess the likely prevalence of haplotypes' co-inheritance in both controls and patients (43). The p-values below 0.001 were regarded as significant. After retrieving the positions of investigated SNPs from the reference genomic sequence, the possible connections between their haplotypes were assessed.

Results

Genotyping analysis

After reviewing the SNPs database for the *TERT* gene according to their frequency in the dbSNP web server, four SNPs with the highest fre-

Table I The specific PCR primers designed for the amplification of the *TERT* gene. PCR primers were designed based on the GenBank acc. no. NC_000005.10.

| Targeted SNP | SNP locus in the amplicon | Primer's direction | Primer's code | Sequence (5-3) | Length (bp) | PCR conditions |
|--------------|---------------------------|--------------------|---------------|-----------------------|-------------|----------------|
| rs2735940 | 115 | Forward | TERT-1F | CTGCCGGAGGAAATTGCTTT | 220 | 60.1 °C |
| | | Reverse | TERT-1R | CAAATTTTCCAAACCGCCCT | | |
| rs2736098 | 102 | Forward | TERT-2F | TCCTTGTCGCCTGAGGAGTA | 223 | 61.0 °C |
| | | Reverse | TERT-2R | CGAGTGACCGTGGTTTCTGT | | |
| rs2736100 | 134 | Forward | TERT-3F | AAAGCTGGGGTTTACCAGCAA | 201 | 59.9 °C |
| | | Reverse | TERT-3R | TCTCAGGCATCTTGACACCC | | |
| rs10069690 | 97 | Forward | TERT-4F | CCAGACCCGGGACCTAGAA | 206 | 60.1 °C |
| | | Reverse | TERT-4R | CTGAGGATCCTGGACCTTGC | | |

Table II Hardy-Weinberg equilibrium (HWE) for rs2735940, rs2736098, rs2736100, and rs10069690 SNPs of the TERT gene in patients and control groups.

| SNP | Patients | | Controls | |
|-------------------------|----------|----------|----------|----------|
| | Observed | Expected | Observed | Expected |
| rs2735940 (A115G) | | | | |
| AA | 68 | 65.61 | 78 | 75.31 |
| AG | 26 | 30.78 | 21 | 26.38 |
| GG | 06 | 03.61 | 05 | 02.31 |
| Chi-square value | 2.41167 | | 4.32537 | |
| P-value | 0.29944 | | 0.11502 | |
| rs2736098 (C102T) | | | | |
| CC | 66 | 62.41 | 82 | 79.21 |
| CT | 26 | 33.18 | 14 | 19.58 |
| TT | 08 | 4.41 | 04 | 1.21 |
| Chi-square value | 4.6827 | | 8.12163 | |
| P-value | 0.0962 | | 0.01723 | |
| Yate's chi-square value | | | 5.71817 | |
| Yate's P-value | | | 0.05732 | |
| rs2736100 (C134A) | | | | |
| CC | 54 | 50 | 59 | 55.8 |
| CA | 32 | 40 | 31 | 38 |
| AA | 12 | 8 | 10 | 6.5 |
| Chi-square value | 3.92 | | 3.3894 | |
| P-value | 0.14086 | | 0.18365 | |
| rs10069690 (C97T) | | | | |
| CC | 70 | 67.24 | 85 | 82.81 |
| CT | 24 | 29.52 | 12 | 16.38 |
| TT | 6 | 3.24 | 3 | 0.81 |
| Chi-square value | 3.4966 | | 7.15024 | |
| P-value | 0.17407 | | 0.02801 | |
| Yate's chi-square value | | | 4.47961 | |
| Yate's P-value | | | 0.10648 | |

The P-value with statistical significance is in bold. CL – confidence interval.

quency were selected respectively. These four SNPs are rs2735940, rs2736098, rs2736100, and rs10069690, which were chosen to be screened in four distinct positions of the *TERT* gene. Three different banding patterns were found for each selected SNP, demonstrating that the rs2735940, rs2736098, rs2736100, and rs10069690 SNPs each have three genotypes (Figure 1b). The conducted sequencing experiments verified the three expected genotypes for all investigated high-frequency SNPs. The electropherograms of the four investigated SNPs demonstrated these three geno-

types (rs2735940: AA: AG: GG, rs2736098CC: CT: TT, rs2736100: CC: CA: AA, and rs10069690: CC: CT: TT) (Figure 1c).

To determine if the study population was in HWE, the genetic diversity of the identified polymorphic SNPs was performed. According to the chi-square values with Yate's corrections, the polymorphisms of all four detected polymorphic SNPs were confirmed to be compatible with HWE in both control and NSCLC groups at the significance level of 0.05 (Table II).

Association analysis

The nucleic acid differences of four investigated SNPs were compared between controls and patients to determine whether there may be an association between the *TERT* gene and NSCLC (Table III).

For rs2735940 (G/A) SNP, It was observed that the heterozygous AG genotype was significantly more frequent in patients than in controls. Thus, individuals with the heterozygous AG genotype showed a significantly higher risk of developing NSCLC with a *P*-value of 0.0299 (OD 2.3158, CI_{95%} 1.0853–4.9414). This data indicated that individuals with allele G had a considerably higher risk of developing NSCLC with a

P-value of 0.0097 (OR 2.2195; CI_{95%} 1.2134–4.0599). For rs2736098 (C/T) SNP, noticeable differences were identified in the distributions of the heterozygous CT genotype between the patients and controls. Individuals with the CT genotype exhibited a significantly higher risk of developing NSCLC with a *P*-value of 0.0363 (OD 2.1583, CI_{95%} 1.0503–4.4351). This data also showed that individuals with allele T exhibited a significantly higher risk of developing NSCLC with a *P*-value of 0.0072 (OR 2.1507; CI_{95%} 1.2303–3.7598). Regarding the other two SNPs under investigation (rs2736100 and rs10069690), no significant associations were identified in their alleles and genotypes distribution between the patients and controls. rs2736100 (CA)

Table III Association of rs2735940, rs2736098, rs2736100, and rs10069690 SNPs with the risk of NSCLC.

| SNP | Patients (n = 100) | Controls (n = 100) | Odds ratio (95% CI) | P-value |
|-------------------|--------------------|--------------------|---------------------|---------|
| rs2735940 (A115G) | | | | |
| AA | 70 (70%) | 85 (85%) | Reference value | |
| AG | 24 (24%) | 12 (12%) | 2.3158 | 0.0299 |
| GG | 06 (06%) | 03 (03%) | 2.0638 | 0.3154 |
| A allele | 164 (82%) | 182 (91%) | Reference value | |
| G allele | 036 (18%) | 018 (09%) | 2.2195 | 0.0097 |
| rs2736098 (C102T) | | | | |
| CC | 66 (66%) | 82 (82%) | Reference value | |
| CT | 26 (26%) | 14 (14%) | 2.1583 | 0.0363 |
| TT | 08 (08%) | 04 (04%) | 2.0870 | 0.2425 |
| C allele | 158 (79%) | 178 (89%) | Reference value | |
| T allele | 042 (21%) | 022 (11%) | 2.1507 | 0.0072 |
| rs2736100 (C134A) | | | | |
| CC | 54 (54%) | 59 (59%) | Reference value | |
| CA | 32 (32%) | 31 (31%) | 1.0474 | 0.8790 |
| AA | 14 (14%) | 10 (10%) | 1.4651 | 0.3861 |
| C allele | 140 (70%) | 149 (74.5%) | Reference value | |
| A allele | 60 (30%) | 051 (25.5%) | 1.2521 | 0.3153 |
| rs10069690 (C97T) | | | | |
| CC | 68 (68%) | 78 (78%) | Reference value | |
| CT | 26 (26%) | 21 (21%) | 1.3218 | 0.4051 |
| TT | 06 (06%) | 05 (05%) | 1.2128 | 0.7567 |
| C allele | 162 (81%) | 177 (88.5%) | Reference value | |
| T allele | 038 (19%) | 031 (15.5%) | 1.2788 | 0.3549 |

The *P*-values with statistical significance are shown in bold. CL – confidence interval.

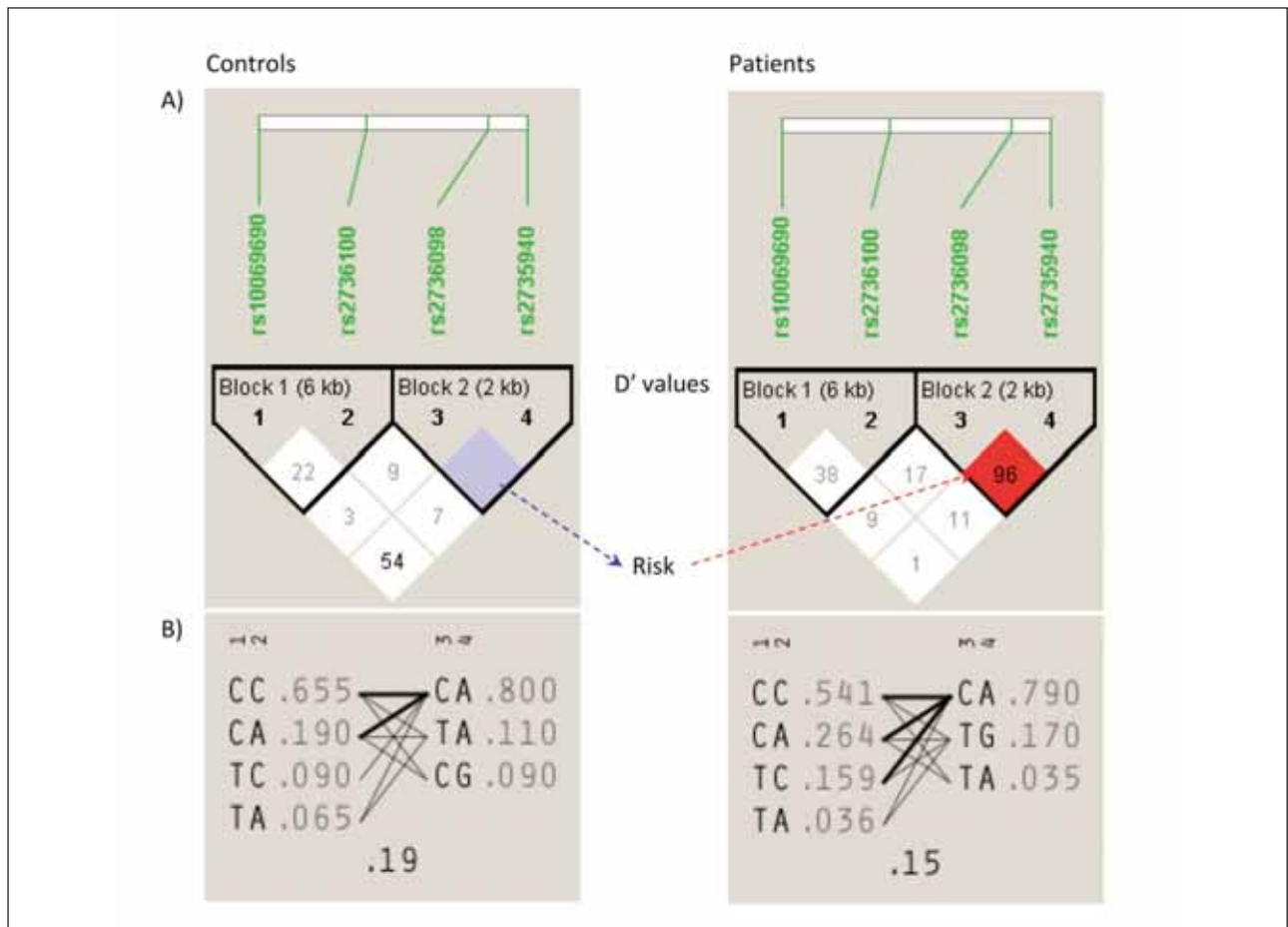


Figure 2 Haplotype block map for four investigated SNPs in the TERT gene. Standard color schemes are used to display LD values. Bright red colour indicates very strong LD (LOD > 2, $D' = 1$), pink colour indicates less strong LD (LOD > 2, $D' < 1$), and blue colour indicates partial linkage (LOD < 2, $D' = 1$), and white colour indicates complete recombination (LOD < 2, $D' < 1$).

exerted no obvious preferences for any genotype or allele to be in patients or controls since no significant association was detected in the distribution of this SNP between both involved groups (P -value 0.8790 and 0.3861 for CA and AA genotypes, respectively). The same observation also applied to the distribution of rs10069690 (CT). Though the heterozygous genotype CT showed a higher frequency in patients (26%) than the controls (21%), The p -value analysis reflected statistically insignificant differences between the patients and controls for this SNP (P -value 0.4051).

LD analysis

The observed haplotypes were evaluated, and the values of linkage disequilibrium were determined between the four polymorphic loci in the current population to assess the co-inheritance potential among them. The observed D' (1.0) and LOD (0.96) values showed the presence of partial linkage between two polymorphic SNPs in the controls, namely rs2735940 and rs2736098. Whereas the other val-

ues of the other two SNPs showed no evidence of co-inheritance (Figure 2a). In the patients' population, LD plot analysis showed that the rs2735940 and rs2736098 were highly co-inherited (D' 0.96, LOD 22.6) (Table IV).

The haplotype analysis further details the LD plot among the investigated polymorphic loci. This high value entailed a high genetic collaboration between rs2735940 and rs2736098 in the patients' group. Based on their positions occupied within the genomic sequences of the TERT gene, these four SNPs were separated accordingly into two adjacent blocks. The 6 kb distance block-1 was created by combining rs2736100 and rs10069690. Whereas block-2 of 2 kb distance was created between rs2735940 and rs2736098 SNPs. In controls, the highest ratio of haplotypes' distributions between both blocks was represented by CC: CA and CA: CA, with a multi-allelic D' of 0.19. Whereas the highest ratio of haplotypes' distributions between both involved blocks were CC: CA, CA: CA, and TC: CA, with a multi-allelic D' of 0.15. LD plot showed that

Table IV The computed linkage values four investigated four SNPs in the *TERT* gene as determined by LD plot analysis.

| | L1 | L2 | D' | LOD | r ² | CIlow | CIhi | Dist | T-int |
|--------------|------------|-----------|-------|------|----------------|-------|------|-------|-------|
| For controls | rs10069690 | rs2736100 | 0.22 | 0.75 | 0.026 | 0.04 | 0.43 | 6726 | 1.01 |
| | rs10069690 | rs2736098 | 0.039 | 0.03 | 0.001 | -0.01 | 0.26 | 14296 | – |
| | rs10069690 | rs2735940 | 0.54 | 0.23 | 0.005 | 0.05 | 0.88 | 16696 | – |
| | rs2736100 | rs2736098 | 0.096 | 0.08 | 0.003 | 0.0 | 0.39 | 7570 | 0.4 |
| | rs2736100 | rs2735940 | 0.073 | 0.06 | 0.002 | 0.0 | 0.36 | 9970 | – |
| | rs2736098 | rs2735940 | 1.0 | 0.96 | 0.012 | 0.14 | 0.99 | 2400 | 1.25 |
| For patients | rs10069690 | rs2736100 | 0.382 | 0.5 | 0.015 | 0.05 | 0.68 | 6726 | 0.52 |
| | rs10069690 | rs2736098 | 0.095 | 0.02 | 0.001 | 0.01 | 0.61 | 14296 | – |
| | rs10069690 | rs2735940 | 0.011 | 0.0 | 0.0 | -0.01 | 0.22 | 16696 | – |
| | rs2736100 | rs2736098 | 0.175 | 0.57 | 0.018 | 0.02 | 0.36 | 7570 | 0.78 |
| | rs2736100 | rs2735940 | 0.112 | 0.19 | 0.006 | 0.0 | 0.33 | 9970 | – |
| | rs2736098 | rs2735940 | 0.963 | 22.6 | 0.763 | 0.85 | 1.0 | 2400 | 22.79 |

the observed CA haplotype represented the highest distribution of 0.8 to 0.79 in controls and patients, respectively (Figure 2b). This haplotype was generated by the collaboration of both rs2735940 (AG) and rs2736098 (CT) SNPs in the block-2 in both populations. Interestingly, it was found that the collaboration between the pathogenic T allele in rs2736098 SNP with the pathogenic G allele in rs2735940 SNP created the TG haplotype. This haplotype is patient-specific and not found in the control population.

Discussion

One of the main reasons for the high mortality rates associated with lung cancer in many parts of the world is related to its inadequate early detection (44). Thus, it is critical to develop useful genetic biomarkers for the early diagnosis of lung cancer (45). The *TERT* gene is one of these intriguing factors that has been shown to have a remarkable connection with the initiation and development of many cancers. Many *TERT* biomarkers have been discovered to be risk factors for NSCLC. In various populations affected by the steady increase in these cases, it is unclear if many certain alleles of the *TERT* gene pose a risk for NSCLC individuals. This study described the connection between *TERT* polymorphism and NSCLC in the Iraqi population by genotyping four high-frequency SNPs positioned in four regions within the *TERT* gene.

Out of four investigated SNPs, the polymorphisms of rs2736100 and rs10069690 SNPs did not show any significant association with NSCLC. This finding implies no evidence of a connection between these variants and lung cancer in the investigated population. However, our results do not align with

other research that indicated the significant association of rs2736100 and rs10069690 SNPs with lung cancer in Chinese and Korean populations (34, 46–51). This may be due to the lower number of investigated samples that represent the main limitation of this case-control study. However, genotyping 100 patients with NSCLC is tedious due to many logistic issues related to the strict identification of this type of cancer, and eliminating other types of lung cancers has reduced the sample size.

In contrast to rs2736100 and rs10069690 SNPs, this study found that both rs2735940 and rs2736098 SNPs showed significant associations with the development of NSCLC in the investigated Iraqi subjects. This association may indicate the potential roles of these SNPs in the progression of NSCLC. In agreement with our results, many reports have shown a significant association between the polymorphism of rs2735940 SNP and the development of lung cancers in several populations, including Chinese (52), Koreans (53), Icelanders (54), and African-Americans (55). In contrast, this SNP is also reported to be associated with a decreased risk of lung cancer in American women (56). However, the inclusion of one gender in the case-control investigation may change the genotype-phenotype association. In addition to lung cancer, rs2735940 SNP has also been associated with multiple malignancies, such as renal cell carcinoma (57), head and neck cancer (31), and gastric cancer (25). Furthermore, rs2735940 SNP showed variable distributions in many populations, which refers to variable biological diversity. According to dbSNP, the rs2735940 SNP allele G is deposited with high frequencies in Latin Americans (G=0.54 to 0.63), East Asians (G=0.53), Asians (G=0.52), Africans (G=0.49), and Europeans (G=0.48).

Whereas it is deposited in lower frequencies in South Asians ($G=0.32$). However, the population of the Middle East has the lowest frequencies of the allele G, with only 0.26 in the Qatari population.

Although rs2736098 SNP is less known than rs2735940 SNP, it is also linked to the development of lung cancer (32, 53). Moreover, it is also linked with the increased risks of other cancers (58), such as nasopharyngeal carcinoma (59), hepatocellular carcinoma (60), cervical carcinoma (61), and breast cancer (62). The rs2736098 SNP allele T is deposited in dbSNP with lower frequencies than in rs2735940 SNP. This can easily be shown in the percentage of T allele that deposited in 0.45, 0.41, 0.38, 0.27, 0.22-0.19, and 0.11 in South Asians, Asians, East Asians, Europeans, Latin Americans, and Africans, respectively. Noteworthy, the A allele was deposited in the Qatari population at the same frequency in which the G allele of rs2735940 SNP was deposited (0.26). This point may support our findings of the presence of strong collaboration between both polymorphic loci in the studied area of the Middle East.

Both the rs2736098 and rs2735940 SNPs have been deposited in the ClinVar database because of their significance in developing many different forms of tumors (63). Due to the synonymous effect of rs2735940 SNP, it seems that this variant exhibits its effect by modulating splicing patterns (64, 65), regulating the velocity of mRNA translation (66), or influencing the protein kinetics (67). Whereas the effect of the 5'-UTR rs2736098 SNP differs based on its different position upstream of the open reading frame. It was reported that 5'-UTR SNPs effects are mainly concerned with altering mRNA translation and decay (68). Whatever the mechanism through which each SNP impacts the TERT, both rs2736098 and rs2735940 SNPs showed a high collaboration in the progression of NSCLC. This can be seen in the heterozygous genotypes of rs2736098: CT and rs2735940: AG as both showed a remarkable tendency to be co-inherited with each other the study population. Further analysis of haplotype distributions revealed that both SNPs are mainly co-inherited by the CA haplotype. In the patient population, the CA haplotype has a greater contribution with further genetic formulations than that found in the control group. This haplotype was generated by the contribution of the C and A alleles from rs2736098 and rs2735940 SNPs, respectively. Most importantly, the pathogenic T allele of rs2735940 and the pathogenic G allele of rs2736098 SNPs have collaborated to generate the TG haplotype, a distinct haplotype for patients. Due to the absence of TG haplotype in control, it is rather important to figure out its severity in association with NSCLC in the studied population. This sort of cooperation may be attributed to the close positioning of both SNPs in the TERT gene sequences. To the best of our knowledge, no previous report has claimed this sort of co-inheritance between rs2736098 and rs2735940 SNPs in the TERT gene. However, more pathological

connections between TERT gene and NSCLC are need to be clarified the get further details on the mechanisms of this con-inheritance.

In conclusion, the genotyping experiments conducted on four variants within the TERT gene indicated the presence of a significant association between the polymorphisms of rs2735940 and rs2736098 SNPs and NSCLC in the studied Iraqi subjects. Individuals with rs2735940: AG and rs2736098: CT genotypes exhibited higher risks of NSCLC. This research demonstrated that both AG and CT genotypes collaborated to generate a high level of co-inheritance represented by the haplotype TG. This haplotype is specific for patients and may be associated with the development of NSCLC. Accordingly, the haplotype TG can serve as a promising biomarker for the early diagnosis of NSCLC. However, large-scale screening experiments are suggested to provide another layer of confirmation of the current findings.

Funding

This work was part of the Ph.D. thesis of ZKA prepared at the Higher Institute of Biotechnology of Sfax, 3000 Sfax University, Tunisia (IQ200476). The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability statement

The data supporting this study's findings are available on request from the corresponding author. The data are not publicly available for privacy and ethical restrictions, as stipulated by the University of Kufa Institutional Review Board.

Authorship contribution statement

ZKL developed the work and conducted genotyping experiments. TMA supervised the research. AHA and ME participated in the genotyping experiments. MBSA contributed to the supervision of the work, analyzed the data, and wrote the manuscript. All authors approved the final version of the work.

Acknowledgments: The authors thank the Middle-Euphrates Cancer Center (MECC), Najaf province, and Merjan Cancer Center (MCC), Babil province, and the postgraduate laboratory unit, College of Medicine, University of Kufa for granting the samples and providing the informative scientific help to accompany the preliminary data preparations.

Conflict of interest statement

All the authors declare that they have no conflict of interest in this work.

References

- Jazieh AR, Algaiz G, Errihani H, Elghissassi I, Mula-Hussain L, Bawazir AA, et al. Lung cancer in the Middle East and North Africa region. *J Thorac Oncol* 2019; 14(11): 1884–91.
- Hussain AMA, Lafta RK. Cancer Trends in Iraq 2000–2016. *Oman Med J* 2021; 36(1): e219.
- Wong M, Lao XQ, Ho K-F, Goggins WB, Tse SLA. Incidence and mortality of lung cancer: global trends and association with socioeconomic status. *Sci Rep* 2017; 7(1): 1–9.
- Zhao X, Wang S, Wu J, Li X, Wang X, Gao Z, et al. Association of TERT polymorphisms with clinical outcome of non-small cell lung cancer patients. *PLoS One* 2015; 10(5): e0129232.
- Gao H, Niu Y, Li M, Fang S, Guo L. Identification of DJ 1 as a contributor to multidrug resistance in human small cell lung cancer using proteomic analysis. *Int J Exp Pathol* 2017; 98(2): 67–74.
- Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* 1999; 354(9173): 99–105.
- Diaz-Lagares A, Mendez-Gonzalez J, Hervas D, Saigi M, Pajares MJ, Garcia D, et al. A novel epigenetic signature for early diagnosis in lung cancer. *Clin Cancer Res* 2016; 22(13): 3361–71.
- Hamann HA, Ver Hoeve ES, Carter-Harris L, Studts JL, Ostroff JS. Multilevel opportunities to address lung cancer stigma across the cancer control continuum. *J Thorac Oncol* 2018; 13(8): 1062–75.
- El-Baz A, Gimel'farb G, Falk R, El-Ghar A, Rainey S, Heredia D, et al. Toward early diagnosis of lung cancer. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2009; p. 682–9.
- Liao Y, Yin G, Wang X, Zhong P, Fan X, Huang C. Identification of candidate genes associated with the pathogenesis of small cell lung cancer via integrated bioinformatics analysis. *Oncol Lett* 2019; 18(4): 3723–33.
- Zawadzka I, Jele A, Pietrzak J, ebrowska-Nawrocka M, Michalska K, Szmajda-Krygier D, et al. The impact of ABCB1 gene polymorphism and its expression on non-small-cell lung cancer development, progression and therapy—preliminary report. *Sci Rep* 2020; 10(1): 1–10.
- Zhang F, Cheng D, Wang S, Zhu J. Human specific regulation of the telomerase reverse transcriptase gene. *Genes (Basel)* 2016; 7(7): 30.
- Dratwa M, Wysoczka B, Łacina P, Kubik T, Bogunia-Kubik K. TERT—Regulation and Roles in Cancer Formation. *Front Immunol* 2020; 11: 589929.
- Barthel FP, Wei W, Tang M, Martinez-Ledesma E, Hu X, Amin SB, et al. Systematic analysis of telomere length and somatic alterations in 31 cancer types. *Nat Genet* 2017; 49(3): 349–57.
- Leão R, Apolónio JD, Lee D, Figueiredo A, Tabori U, Castelo-Branco P. Mechanisms of human telomerase reverse transcriptase (hTERT) regulation: clinical impacts in cancer. *J Biomed Sci* 2018; 25(1): 1–12.
- Dogan F, Forsyth NR. Telomerase regulation: a role for epigenetics. *Cancers (Basel)* 2021; 13(6): 1213.
- Ferreira MSV, Sørensen MD, Pusch S, Beier D, Bouillon A-S, Kristensen BW, et al. Alternative lengthening of telomeres is the major telomere maintenance mechanism in astrocytoma with isocitrate dehydrogenase 1 mutation. *J Neurooncol* 2020; 147(1): 1–14.
- Heidenreich B, Kumar R. TERT promoter mutations in telomere biology. *Mutat Res Mutat Res* 2017; 771: 15–31.
- Xu Y, Goldkorn A. Telomere and telomerase therapeutics in cancer. *Genes (Basel)*. 2016; 7(6): 22.
- Autexier C, Lue NF. The structure and function of telomerase reverse transcriptase. *Annu Rev Biochem* 2006; 75: 493–517.
- Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, Lingner J, et al. Telomerase catalytic subunit homologs from fission yeast and human. *Science* (80-) 1997; 277(5328): 955–9.
- Colebatch AJ, Dobrovic A, Cooper WA. TERT gene: its function and dysregulation in cancer. *J Clin Pathol* 2019; 72(4): 281–4.
- Campa D, Rizzato C, Stolzenberg Solomon R, Pacetti P, Vodicka P, Cleary SP, et al. TERT gene harbors multiple variants associated with pancreatic cancer susceptibility. *Int J cancer* 2015; 137(9): 2175–83.
- Stoehr R, Taubert H, Zinnall U, Giedl J, Gaisa NT, Burger M, et al. Frequency of TERT promoter mutations in prostate cancer. *Pathobiology* 2015; 82(2): 53–7.
- Bayram S, Ülger Y, Sümbül AT, Kaya BY, Genç A, Rencüzo ullari E, et al. Polymorphisms in human telomerase reverse transcriptase (hTERT) gene and susceptibility to gastric cancer in a Turkish population: Hospital-based case–control study. *Gene* 2016; 585(1): 84–92.
- Liu R, Xing M. TERT promoter mutations in thyroid cancer. *Endocr Relat Cancer* 2016; 23(3): R143–55.
- Wu Y, Yan M, Li J, Li J, Chen Z, Chen P, et al. Genetic polymorphisms in TERT are associated with increased risk of esophageal cancer. *Oncotarget* 2017; 8(6): 10523.
- Wan S, Liu X, Hua W, Xi M, Zhou Y, Wan Y. The role of telomerase reverse transcriptase (TERT) promoter mutations in prognosis in bladder cancer. *Bioengineered* 2021; 12(1): 1495–504.
- Llorca-Cardena MJ, Peña-Chilet M, Mayor M, Gomez-Fernandez C, Casado B, Martin-Gonzalez M, et al. Long telomere length and a TERT-CLPTM1 locus polymorphism association with melanoma risk. *Eur J Cancer* 2014; 50(18): 3168–77.
- Zhu CQ, Cutz JC, Liu N, Lau D, Shepherd FA, Squire JA, et al. Amplification of telomerase (hTERT) gene is a poor prognostic marker in non-small-cell lung cancer. *Br J Cancer* 2006; 94(10): 1452–9.

31. Liu Z, Ma H, Wei S, Li G, Sturgis EM, Wei Q. Telomere Length and TERT Functional Polymorphisms Are Not Associated with Risk of Squamous Cell Carcinoma of the Head and Neck. *Telomere Length, TERT Genetic Variations, and Head and Neck Cancer Risk. Cancer Epidemiol Biomarkers Prev* 2011; 20(12): 2642–5.
32. Wu H, Qiao N, Wang Y, Jiang M, Wang S, Wang C, et al. Association between the telomerase reverse transcriptase (TERT) rs2736098 polymorphism and cancer risk: evidence from a case-control study of non-small-cell lung cancer and a meta-analysis. *PLoS One* 2013; 8(11): e76372.
33. Yoon K-A, Park JH, Han J, Park S, Lee GK, Han J-Y, et al. A genome-wide association study reveals susceptibility variants for non-small cell lung cancer in the Korean population. *Hum Mol Genet* 2010; 19(24): 4948–54.
34. Xing Y, Liu F, Li J, Lin J, Zhu G, Li M, et al. Case-Control Study on Impact of the Telomerase Reverse Transcriptase Gene Polymorphism and Additional Single Nucleotide Polymorphism (SNP)–SNP Interaction on Non Small Cell Lung Cancers Risk in Chinese Han Population. *J Clin Lab Anal* 2016; 30(6): 1071–7.
35. Lawi ZK, Al-Shuhaib MBS, Amara I Ben. The rs1801280 SNP is associated with non-small cell lung carcinoma by exhibiting a highly deleterious effect on N-acetyltransferase 2. *J Cancer Res Clin Oncol* 2022; 1–11.
36. Lawi ZK, Al-Shuhaib MBS, Amara I Ben, Alkhammas AH. Two missense variants of the epidermal growth factor receptor gene are associated with non small cell lung carcinoma in the subjects from Iraq. *Mol Biol Rep* 2022; 1–9.
37. Al-Shuhaib MBS. A minimum requirements method to isolate large quantities of highly purified DNA from one drop of poultry blood. *J Genet* 2018; 97: e87–e94.
38. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 2012; 13(1): 134.
39. Badi MA, Al-Shuhaib MBS, Aljoubouri TRS, Al-Thuwaini TM, Dawud HH, Hussein TH, et al. Rapid and optimized protocol for efficient PCR-SSCP genotyping for wide ranges of species. *Biologia (Bratisl)* 2021; 76(8): 2413–20.
40. Hashim HO, Al-Shuhaib MBS. Exploring the potential and limitations of PCR-RFLP and PCR-SSCP for SNP detection: A review. *J Appl Biotechnol Reports* 2019; 6(4): 137–44.
41. Byun SO, Fang Q, Zhou H, Hickford JGH. An effective method for silver-staining DNA in large numbers of polyacrylamide gels. *Anal Biochem* 2009; 385(1): 174–5.
42. Schoonjans F, Zalata A, Depuydt CE, Comhaire FH. MedCalc: a new computer program for medical statistics. *Comput Methods Programs Biomed* 1995; 48(3): 257–62.
43. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21(2): 263–5.
44. Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. *Ann Glob Heal.* 2019; 85(1).
45. Alsubaie LM, Alsuwat HS, Almandil NB, AlSulaiman A, AbdulAzeez S, Borgio JF. Risk Y-haplotypes and pathogenic variants of Arab-ancestry boys with autism by an exome-wide association study. *Mol Biol Rep* 2020; 47(10): 7623–32.
46. Ye G, Tan N, Meng C, Li J, Jing L, Yan M, et al. Genetic variations in TERC and TERT genes are associated with lung cancer risk in a Chinese Han population. *Oncotarget* 2017; 8(66): 110145.
47. Gao L, Thakur A, Liang Y, Zhang S, Wang T, Chen T, et al. Polymorphisms in the TERT gene are associated with lung cancer risk in the Chinese Han population. *Eur J Cancer Prev* 2014; 23(6): 497–501.
48. Lan Q, Cawthon R, Gao Y, Hu W, Hosgood III HD, Barone-Adesi F, et al. Longer telomere length in peripheral white blood cells is associated with risk of lung cancer and the rs2736100 (CLPTM1L-TERT) polymorphism in a prospective cohort study among women in China. *PLoS One.* 2013; 8(3): e59230.
49. Wang H-M, Zhang X-Y, Jin B. TERT genetic polymorphism rs2736100 was associated with lung cancer: a meta-analysis based on 14,492 subjects. *Genet Test Mol Biomarkers* 2013; 17(12): 937–41.
50. Nie W, Zang Y, Chen J, Xiu Q. TERT rs2736100 polymorphism contributes to lung cancer risk: a meta-analysis including 49,869 cases and 73,464 controls. *Tumor Biol* 2014; 35(6): 5569–74.
51. Bae EY, Lee SY, Kang BK, Lee EJ, Choi YY, KANG H, et al. Replication of results of genome wide association studies on lung cancer susceptibility loci in a Korean population. *Respirology* 2012; 17(4): 699–706.
52. Hosgood III HD, Menashe I, Shen M, Yeager M, Yuenger J, Rajaraman P, et al. Pathway-based evaluation of 380 candidate genes and lung cancer susceptibility suggests the importance of the cell cycle pathway. *Carcinogenesis* 2008; 29(10): 1938–43.
53. Choi JE, Kang H-G, Jang JS, Choi YY, Kim MJ, Kim JS, et al. Polymorphisms in Telomere Maintenance Genes and Risk of Lung Cancer. *Telomere Maintenance Gene Polymorphisms in Lung Cancer. Cancer Epidemiol Biomarkers Prev* 2009; 18(10): 2773–81.
54. Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet.* 2009; 41(2): 221–7.
55. Walsh KM, Gorlov IP, Hansen HM, Wu X, Spitz MR, Zhang H, et al. Fine-mapping of the 5p15. 33, 6p22. 1-p21. 31, and 15q25. 1 Regions Identifies Functional and Histology-Specific Lung Cancer Susceptibility Loci in African-Americans. *Fine-mapping Lung Cancer Risk Regions in African-Americans. Cancer Epidemiol Biomarkers Prev* 2013; 22(2): 251–60.
56. Van Dyke AL, Cote ML, Wenzlaff AS, Abrams J, Land S, Iyer P, et al. Chromosome 5p region SNPs are associated with risk of NSCLC among women. *J Cancer Epidemiol* 2009; 2009.
57. Morais M, Dias F, Resende T, Nogueira I, Oliveira J, Maurício J, et al. Leukocyte telomere length and hTERT genetic polymorphism rs2735940 influence the renal

- cell carcinoma clinical outcome. *Futur Oncol* 2020; 16(18): 1245–55.
58. Zhou M, Jiang B, Xiong M, Zhu X. Association between TERT rs2736098 polymorphisms and cancer risk—a meta-analysis. *Front Physiol* 2018; 9: 377.
59. Fachiroh J, Sangrajrang S, Johansson M, Renard H, Gaborieau V, Chabrier A, et al. Tobacco consumption and genetic susceptibility to nasopharyngeal carcinoma (NPC) in Thailand. *Cancer Causes Control* 2012; 23(12): 1995–2002.
60. Yuan X, Cheng G, Yu J, Zheng S, Sun C, Sun Q, et al. The TERT promoter mutation incidence is modified by germline TERT rs2736098 and rs2736100 polymorphisms in hepatocellular carcinoma. *Oncotarget* 2017; 8(14): 23120.
61. Wang S, Wu J, Hu L, Ding C, Kan Y, Shen Y, et al. Common genetic variants in TERT contribute to risk of cervical cancer in a Chinese population. *Mol Carcinog* 2012; 51(S1): E118–22.
62. Hashemi M, Amininia S, Ebrahimi M, Hashemi SM, Taheri M, Ghavami S. Association between hTERT polymorphisms and the risk of breast cancer in a sample of Southeast Iranian population. *BMC Res Notes* 2014; 7(1): 1–8.
63. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* 2014; 42(D1): D980–5.
64. Logette E, Wotawa A, Solier S, Desoche L, Solary E, Corcos L. The human caspase-2 gene: alternative promoters, pre-mRNA splicing and AUG usage direct isoform-specific expression. *Oncogene* 2003; 22(6): 935–46.
65. Druillennec S, Dorard C, Eychene A. Alternative splicing in oncogenic kinases: from physiological functions to cancer. *J Nucleic Acids* 2012; 2012.
66. Kirchner S, Cai Z, Rauscher R, Kastelic N, Anding M, Czech A, et al. Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol* 2017; 15(5): e2000779.
67. Sauna ZE, Kimchi-Sarfaty C, Ambudkar S V, Gottesman MM. Silent polymorphisms speak: how they affect pharmacogenomics and the treatment of cancer. *Cancer Res* 2007; 67(20): 9609–12.
68. Martin KC, Ephrussi A. mRNA localization: gene expression in the spatial dimension. *Cell* 2009; 136(4): 719–30.

Received: January 15, 2023

Accepted: May 19, 2023