Article

A Significant Increasing Risk Association between Cigarette Smoking and *XPA* and *XPC* Genes Polymorphisms

Safiah Almushawwah [1], Mikhlid H. Almutairi [2], Abdullah M. Alamri [1] and Abdelhabib Semlali [3,*]

[1]   Biochemistry Department, College of Science, King Saud University, P.O. Box 2455,
      Riyadh 11451, Saudi Arabia
[2]   Zoology Department, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[3]   Groupe de Recherche en Écologie Buccale, Faculté de Médecine Dentaire, Université Laval, 2420 Rue de la
      Terrasse, Local 1758, Québec, QC G1V 0A6, Canada
[*]   Correspondence: abdelhabib.semlali@greb.ulaval.ca; Tel.: +1-418-656-2131 (ext. 400062)

Abstract: Cigarette smoking (CS) is a major cause of various serious diseases due to tobacco chemicals. There is evidence suggesting that CS has been linked with the DNA damage repair system, as it can affect genomic stability, inducing genetic changes in the genes involved in the repair system, specifically the nucleotide excision repair (NER) pathway, affecting the function and/or regulation of these genes. Single nucleotide polymorphism (SNP), along with CS, can affect the work of the NER pathway and, therefore, could lead to different diseases. This study explored the association of four SNPs in both *XPA* and *XPC* genes with CS in the Saudi population. The Taq Man genotyping assay was used for 220 healthy non-smokers (control) and 201 healthy smokers to evaluate four SNPs in the *XPA* gene named rs10817938, rs1800975, rs3176751, and rs3176752 and four SNPs in the *XPC* gene called rs1870134, rs2228000, rs2228001, and rs2607775. In the *XPA* gene, SNP rs3176751 showed a high-risk association with CS-induced diseases with all clinical parameters, including CS duration, CS intensity, gender, and age of smokers. On the other hand, SNP rs1800975 showed a statistically significant low-risk association with all clinical parameters. In addition, rs10817938 showed a high-risk association only with long-term smokers and a low-risk association only with younger smokers. A low-risk association was found in SNP rs3176752 with older smokers. In the *XPC* gene, SNP rs2228001 showed a low-risk association only with female smokers. SNP rs2607775 revealed a statistically significant low-risk association with CS-induced diseases, concerning all parameters, except for male smokers. However, SNP rs2228000 and rs1870134 showed no association with CS. Overall, the study results demonstrated possible significant associations (effector/and protector) between CS and SNPs polymorphisms in DNA repair genes, such as *XPA* and *XPC*, except for rs2228000 and rs1870134 polymorphisms.

Keywords: smoking; nucleotide excision repair; *XPA*; *XPC*; single nucleotide polymorphism; Taq man genotyping

1. Introduction

In the 1930s, epidemiologists used case-control surveys to investigate the relationship between lung cancer and smoking. The first study was published in 1939 by Müller, and this study shows that the incidence of lung cancer was higher in smokers than in non-smokers [1,2]. Over the years, the number of smokers in the world has increased, and many various serious diseases have been attributed to smoking, such as cardiovascular diseases [3], obstruction of the respiratory tract, age-related macular degeneration, and the risk of developing diabetes, which is all 30–40% higher for smokers than nonsmokers. Additionally, many studies showed that there are different types of cancer that are related to CS, such as lung, oral cavity, larynx, tongue, bladder, esophagus, colon, and rectum cancers [4]. According to the World Health Organization, smoking was found to be linked

with more than 6,000,000 million deaths cases every year and is estimated to cause 80% of lung cancer deaths [5]. The incidence of CS in certain regions in the Kingdom of Saudi Arabia is greater than 50% [6]. Therefore, a better understanding of the potential association between CS and relative diseases is critical. Recently, many researchers have demonstrated that CS induces DNA damage and has an impact on the genes that are critical in controlling the cell cycle, such as P53 and KRAS. Mutation in both genes was found to be linked with smoking and various types of cancers [7]. Other genetic analyses of many cancers caused by smoking reveal a number of mutations, particularly in cancer-associated genes, such as GLUT-1, HIF-1 [8]. Furthermore, evidence suggests smoking to be linked with the DNA damage repair system, as it can induce genetic changes involved in the repair system, affecting the function and/or regulation of these genes [9–11]. Damage is detected by a DNA repair system, which is critical for maintaining the integrity and stability of genomic DNA of the cell. Any deficiency in the DNA repair genes causes DNA damage, as well as mutation, leading to diseases [12]. The nucleotide excision repair (NER) is the primary pathway responsible for removing a lesion from bulky DNA adducts caused by chemicals in tobacco smoke. In mammalian cells, over 20 protein factors constitute this critical pathway, including XPA replication protein A (RPA), which is essential for maintaining DNA stability during the repair process, as well as XPC-hHR23B- centrin 2 (CETN2), which is crucial in recognizing DNA damage in the first step of the repair process. In the current study, we will focus on the genes that belong to the XPA and XPC proteins. Furthermore, other proteins are fundamental in the NER pathway, including TFIIH, XPB, and XPD DNA helicases, as well as ERCC1-XPF and XPG [10,13,14]. XPB and XPD helicases are key members of the human TFIIH complex, and they have a critical role in unwinding DNA for the repair of duplex distorting damage by NER [15]. Modifications in NER genes will lead to many diseases. Studies established that a defect in the NER pathway gene causes Xeroderma Pigmentosum (XP) [16]. It was clearly documented that the ERCC-1 gene is closely linked to sensitivity to cisplatin therapy [17].

The xeroderma pigmentosum group C (XPC) is a 940-residue multi-domain protein. The C-terminal portion of XPC (492–940; XPC-C) is important in the interactions with DNA, TFIIH, RAD23B, and CETN2. Biochemical and structural analyses confirmed that the XPC complex is capable of binding specifically to DNA damage sites associated with a relatively bulky distortion of the DNA duplex [18,19]. XPA is an acronym for XP complementation group A (XPA). The recessive genetic disorder of XP in children leads to extreme sensitivity to UV light, which causes skin cancer (Bowden et al., 2015). XPA is involved in the NER pathway and has a critical role in the verification of DNA damage and mobilization of repair proteins. Unlike other proteins in the NER pathway, XPA is unique, as it is required for both GG-NER and TC-NER [20,21]. This research proposal was focused on the genetic variation effects of CS on DNA repair system dysfunction, resulting in health problems.

The specific aim was to investigate the polymorphism variations in genes of the NER pathway caused by CS in Saudi smokers versus non-smoker subjects.

2. Materials and Methods

2.1. Ethics Statement and Samples Collection

Written ethical consent was already reviewed and obtained by the Research Ethics Committee of the College of Applied Medical Sciences at King Saud University (KSU) in Riyadh, Saudi Arabia (Approval Number: CAMS 13/3536). As described in our previous work [6], smokers were divided into two groups based on daily quantity of cigarette consumption ($\geq$10 and <10 cigarettes/day). All volunteer participant in the current study signed written informed consent. Clinical data on smoking history, allergic symptoms and diseases, number of cigarettes smoked daily, and body mass index (BMI) were obtained through a self-completed questionnaire. The control group corresponds to non-smokers, and the former smokers were excluded from the control group. Additionally, as described in our previously study [22], all samples were collected from self-reported healthy smokers and non-smokers (controls) who had signed informed consent forms, confirming their

participation in the present study. We excluded any potential participants who self-reported having symptoms, such as metabolic disorders, inflammatory diseases, autoimmune diseases, cancer, or blood diseases. The blood samples (3 mL) were collected via venipuncture in EDTA-containing tubes. A total of 421 participants were divided into 220 healthy non-smokers (control) and 201 healthy smokers. Baseline characteristics of cases and controls were performed (Age, no. of CS/day $\leq$ 10 sticks per day (moderate) and >10 sticks per day as heavy smokers, and period of smoking). In addition, it was made sure that they had not taken any other drug and that they had no disease. Samples of participants who did not meet the criteria were excluded from the study.

2.2. DNA Extraction

Genomic DNA was extracted from 200 μL of EDTA anticoagulated peripheral blood using Qiagen QIAamp® DNA Mini Kit (Q), as per the manufacturer's instructions. The DNA concentration quantity was determined using a NanoDrop 8000 (Thermo Fisher Scientific, Waltham, MA, USA). The purity of each DNA sample was then determined by calculating the ratio of A260/A280 nm and A260/A230 nm. The samples were considered contamination-free when the ratios were 1.7–2.0. The DNA samples were preserved at −20 °C.

2.3. SNP Selection Taq Man Genotyping Assay

A Taq Man assay was carried out for all 421 samples to examine the polymorphism variation of *XPA* and *XPC* genes. The DNA blood samples were diluted to obtain a final concentration of 10 ng/μL. Four SNPs in the *XPA* gene, named rs10817938, rs1800975, rs3176751, and rs3176752, as well as four SNPs in the *XPC* gene, called rs2607775, rs2228000, rs2228001, and rs1870134, were evaluated by the genotyping assay. The selection of these SNPs was based on the previous review, and they were selected from the NCBI database (http://www.ncbi.nlm.nih.gov/snp, accessed on 1 February 2023). Each SNP was located either in promoter, intron, or exon regions (Tables S1 and S2), which may lead to a change in the regulation of the protein or modification in protein folding or function. As described by our previous works [23–26], a total of 8 μL of the final reaction mix and 2 μL of DNA (10 ng/μL) were distributed in an optical reaction plate. The SNP reaction mix contained 5.3 μL of TaqMan® genotyping master mix (Applied Biosystems, Foster City, CA, USA), 2.5 μL of nuclease-free water, and 0.2 μL of SNP. The negative control (no DNA) was included in each plate. PCR amplification was carried out under the following conditions: a primary denaturation step at 95 °C for 7 min, followed by 40 cycles of 95 °C for 30 s, 60 °C for 1 min, and 72 °C for 30 s. The PCR reaction mixture was terminated for a final extension at 72 °C for 5 min. Quant Studio™ 7 Flex Real-Time PCR System (Applied Biosystems) with an endpoint reading of the genotypes was used to perform the reaction.

2.4. Statistical Analysis

Statistical analysis was carried out using the Statistical Package for Social Sciences (SPSS) version 26.0 (IBM-SPSS, Armonk, NY, USA). Hardy Weinberg equilibrium was used to check the deviation of the computed genotypic and allelic frequencies of each SNP. Genetic comparisons were performed with the aid of the χ2 test and allelic odds ratios (ORs). The chi-square test was used to determine the proportion of cases versus control and according to different groupings (based on duration of smoking, frequency of smoking, gender, and age for the SNPs and alleles). Odds ratio and 95% confidence interval were calculated using the online OR calculator (Medcalc, https://www.medcalc.org/calc/odds_ratio.php, accessed on 1 February 2023). In addition, Fisher's exact test (two-tailed) was applied. Results were expressed as the mean and the standard deviation for age. The proportion of SNPs and alleles were reported as numbers and percentages. An independent sample t-test was performed to compare the mean age between smokes and non-smokers. *p* values of less than 0.05 were considered statistically significant. OR more than one indicates high-risk association, and less than one indicates a low-risk association.

3. Results

3.1. Baseline and Clinical Characteristics of Participants

A total of 220 Saudi healthy non-smokers and 201 healthy smokers were used in the study. Participants were further classified into different groups based on age, gender, period of smoking, and the number of CS per day. Table 1 illustrates the baseline and clinical characteristics of participants measured for all participants. These variable parameters were used to study the association between SNPs in tested genes and the risk of CS causing disease.

Table 1. Demographic characteristics of 421 smokers and non-smokers in the study.

Characteristics	All $n = 421$	Smokers $n = 201$	Non-Smokers $n = 220$	*p* Values
Gender, %				
Male	314 (74.6%)	187 (59.9%)	127 (40.1%)	<0.001
Female	107 (25.4%)	14 (13.1%)	93 (86.9%)	
Age, mean (±SD)	27.8 (7.3)	28.1 (6.1)	27.6 (8.3)	0.413
Smoking frequency per day, %				
≤10 per day	77 (18.3%)	77 (38.3%)		
>10 per day	97 (23.0%)	97 (48.3%)		
Not specified	27 (6.4%)	27 (13.4%)		
Duration of smoking in years, %				
≤5 years	65 (15.4%)	65 (32.3%)		
>5 years	112 (26.6%)	112 (55.7%)		
Not specified	24 (5.7%)	24 (11.9%)		

n: Number, SD: Standard deviation. *p*-Value in bold represents significant result.

3.2. Global Genotyping Analysis of XPA and XPC among Smokers and Non-Smokers

To evaluate the association of SNPs in the *XPA* and *XPC* genes with the effects of CS on induced diseases, we evaluated rs10817938, rs1800975, rs3176751, rs3176752, rs1870134, rs2228000, rs2228001, and rs2607775 variants and CS in 421 participants. The distributions of genotyping and allele frequencies of the smoker and non-smoker groups are summarized in Table 2.

Table 2. General genotype frequencies of four *XPA* and four *XPC* gene polymorphisms in smoker cases and controls.

SNP	Variant	Smoker	Non-Smoker	Chi-Square *p* Values	OR (CI)
XPA gene					
rs10817938	TT	7 (3.7%)	12 (6.0%)		Ref.
	TC	69 (36.7%)	47 (23.6%)	0.071	2.517 (0.923–6.863)
	CC	112 (59.6%)	140 (70.4%)	0.521	1.371 (0.523–3.599)
	TC+CC	181 (96.3%)	187 (94.0%)	0.868	1.025 (0.770–1.363)
	T	83 (22.1%)	71 (18.0%)	Ref.	Ref.
	C	293 (77.9%)	327 (82.0%)	0.141	0.767 (0.538–1.092)
rs1800975	TT	186 (96.4%)	119 (55.1%)		Ref.
	TC	6 (3.1%)	14 (6.5%)	0.009 *	0.274 (0.103–0.733)
	CC	1 (0.5%)	83 (38.4%)	<0.001 *	0.008 (0.001–0.056)
	TC+CC	7 (3.6%)	97 (45.97%)	<0.001 *	0.081 (0.037–0.178)
	T	378 (97.9%)	252 (58.3%)	Ref.	Ref.
	C	8 (2.1%)	180 (41.7%)	<0.001 *	0.029 (0.014–0.061)

Table 2. *Cont.*

SNP	Variant	Smoker	Non-Smoker	Chi-Square *p* Values	OR (CI)
rs3176751	CC	65 (33.3%)	151 (69.0%)		Ref.
	CG	17 (8.7%)	25 (11.4%)	0.188	1.579 (0.799–3.122)
	GG	113 (58.0%)	43 (19.6%)	**<0.001 ***	6.105 (3.869–9.631)
	CG+GG	130 (66.7%)	68 (31.1%)	**<0.001 ***	2.147 (1.511–3.050)
	G	147 (37.7%)	327 (74.7%)	Ref.	Ref.
	C	243 (62.3%)	111 (25.3%)	**<0.001 ***	4.869 (3.618–6.555)
rs3176752	GG	195 (97.0%)	208 (96.0%)		Ref.
	GT	6 (3.0%)	8 (4.0%)	0.0685	0.800 (0.273–2.347)
	TT	0	0		
	GT+TT	6 (2.99%)	8 (3.7%)	0.094	0.806 (0.275–2.363)
	G	396 (98.5%)	424 (98.1%)	Ref.	Ref.
	T	6 (1.5%)	8 (1.9%)	0.687	0.803 (0.276–2.335)

XPC gene

SNP	Variant	Smoker	Non-Smoker	Chi-Square *p* Values	OR (CI)
rs1870134	GG	1 (0.5%)	0		Ref.
	GC	7 (3.7%)	13 (6.0%)	0.319	0.185 (0.007–5.137)
	CC	184 (95.8%)	205 (94.0%)	0.461	0.299 (0.012–7.392)
	GC+CC	191 (99.5%)	218 (100%)	0.970	0.995 (0.756–1.309)
	G	9 (2.3%)	13 (3.0%)	Ref.	Ref.
	C	375 (97.7%)	423 (97.0%)	0.574	1.281 (0.541–3.029)
Rs2228000	CC	3 (1.5%)	2 (0.9%)		Ref.
	CT	45 (23.2%)	49 (22.9%)	0.600	0.612 (0.098–3.834)
	TT	146 (75.3%)	163 (76.2%)	0.575	0.597 (0.098–3.624)
	CT+TT	191 (98.0%)	212 (99.0%)	0.965	0.994 (0.754–1.309)
	C	51 (13.1%)	53 (12.4%)	Ref.	Ref.
	T	337 (86.9%)	375 (87.6%)	0.745	0.934 (0.619–1.409)
rs2228001	AA	64 (34.4%)	72 (34.5%)		Ref.
	AC	82 (44.1%)	95 (45.4%)	0.898	0.971 (0.620–1.519)
	CC	40 (21.5%)	42 (20.1%)	0.805	1.071 (0.619–1.854)
	AA+CC	122 (65.6%)	137 (65.6%)	0.932	0.987 (0.726–1.341)
	A	210 (56.5%)	239 (57.2%)	Ref.	Ref.
	C	162 (43.5%)	179 (42.8%)	0.837	1.030 (0.777–1.366)
rs2607775	GG	79 (41.8%)	111 (51.4%)		Ref.
	GC	72 (38.1%)	86 (39.8%)	0.455	1.176 (0.769–1.801)
	CC	38 (20.1%)	19 (8.8%)	0.001	2.810 (1.509–5.233)
	GC+CC	110 (58.2%)	105 (48.6%)	0.287	1.197 (0.859–1.667)
	G	230 (60.8%)	308 (71.3%)	Ref.	Ref.
	C	148 (39.2%)	124 (28.7%)	**0.002 ***	1.598 (1.192–2.143)

Note: * number of valid cases specified in each of the SNPs, missing data were excluded from the analysis. SNP: Single nucleotide polymorphism, Ref = Reference allele, OR: Odd ratio, CI = Confidence interval. *p*-Values in bold represent significant results.

For SNP rs10817938 (T/C) of *XPA* gene, no association was found with the risk related to smoking induced diseases. In SNP rs1800975 of the *XPA* gene, the genotyping distribution was as follows: 96.4% TT, 3.1% TC, and 0.5% CC in smokers, while it was 55.1% TT, 6.5% TC, and 38.4% CC in non-smokers. The TC, CC, and TC+CC alleles of rs1800975 decreased the risk of developing diseases related to smoking by approximately 72.6%, 99.2%, and 91.9% compared to the TT homozygous allele [TC (OR = 0.274; CI = [0.103–0.733]; *p* = 0.009); CC (OR = 0.008; CI = [0.001–0.056]; *p* < 0.001; and TC+CC (OR = 0.081; CI = [0.037–0.178]; *p* < 0.001]. The phenotypic distribution was 97.9% T and 2.1% C in smokers and

58.3% T and 41.7% C in controls with a protective affect (OR = 0.029; CI = [0.014–0.061]; *p* < 0.001).

SNP rs3176751 exhibited significant differences between smoker and non-smokers with a higher risk of smoking-induced diseases. The genotyping distribution was as the follows: 58% GG and 66.7% CG+GG in smokers, while it was 19.6% GG and 31.1% CG+GG in non-smokers when compared to the CC reference genotype (GG: OR = 6.105, CI = [3.869–9.631], *p* < 0.001; CG+GG = OR = 2.147, CI = [1.511–3.050], *p* < 0.001). The C allele was used as a reference. The G allele was found to be more frequent in smokers (62.3%) and in the non-smoker group (25.3%) compared to the C allele. The G allele showed significant high-risk association with smoking-induced diseases among smokers, as shown in Table 2 (OR = 4.869, CI = [3.618–6.555], *p* < 0.001). Additionally, SNP rs3176752 of the XPA gene did not show any association with CS. The genotyping distribution of GG, GT, and TT variants was estimated to be 97%, 3%, and 0%, respectively, in smokers, and it was 96%, 4%, and 0%, respectively, in the controls.

We have evaluated the association SNPs in the *XPC* gene (rs1870134, rs2228000, rs2228001, and rs2607775) variations and CS in 421 participants. The data for three SNPs, rs1870134, rs2228000 and rs2228001, did not show any association with the increase or decrease in the risk of smoking-induced diseases in the Saudi population. However, SNP rs2607775 exhibited significant differences between smokers and non-smokers, presenting a high risk of smoking-induced diseases in the CC genotype and the C allele. The genotyping distribution was as follows: 20.1% CC in smokers and 8.8% CC in non-smokers; CC: OR = 2.810, CI = [1.509–5.233], *p* = 0.001; the G allele was used as a reference. The C allele was found to be more frequent in smokers (39.2%) and in the non-smoker group (28.7%) compared to the G allele. The C allele showed significant high-risk association with smoking-induced diseases among smokers, as shown in Table 3 (OR = 1.598, CI = [1.192–2.143], *p* = 0.002).

Table 3. Comparison of *XPA* genotypes and allele distributions according to the duration of smoking, frequency of smoking, and gender.

A. Based on the duration of smoking								
			Duration of smoking		Chi-square *p* values		OR (CI)	
SNP	Variant	Control	≤5 years	>5 years	Control vs. <5 years	Control vs. >5 years	Control vs. <5 years	Control vs. >5 years
rs10817938	TT	12 (6.0%)	3 (3.7%)	1 (1.2%)	Ref.	Ref.	Ref.	Ref.
	TC	47 (23.6%)	28 (34.1%)	29 (34.5%)	0.207	0.061	2.383 (0.618–9.182)	7.404 (0.914–59.975)
	CC	140 (70.4%)	51 (62.2%)	54 (64.3%)	0.572	0.146	1.457 (0.395–5.374)	4.629 (0.587–36.463)
	TC+CC	187 (94.0%)	31 (37.8%)	30 (35.7%)	**<0.001 ***	**<0.001 ***	0.402 (0.254–0.637)	0.380 (0.239–0.603)
	T	71 (18.0%)	37 (22.2%)	31 (18.5%)	Ref.	Ref.	Ref.	Ref.
	C	327 (82.0%)	130 (77.8%)	137 (91.5%)	0.235	0.862	0.763 (0.488–1.192)	0.960 (0.602–1.530)
rs1800975	TT	119 (55.1%)	84 (96.6%)	80 (95.2%)	Ref.	Ref.	Ref.	Ref.
	TC	14 (6.5%)	2 (2.3%)	4 (4.8%)	**0.038 ***	0.144	0.202 (0.045–0.914)	0.425 (0.135–1.338)
	CC	83 (38.4%)	1 (1.1%)	0	**0.049 ***	**0.009 ***	0.084 (0.072–0.994)	0.019 (0.001–0.145)
	TC+CC	97 (45.97%)	3 (3.4%)	4 (4.8%)	**<0.001 ***	**<0.001 ***	0.077 (0.024–0.249)	0.106 (0.037–0.297)
	T	252 (58.3%)	170 (97.7%)	164 (97.6%)	Ref.	Ref.	Ref.	Ref.
	C	180 (41.7%)	4 (2.3%)	4 (2.4%)	**<0.001 ***	**<0.001 ***	0.033 (0.012–0.090)	0.034 (0.012–0.093)
rs3176751	CC	151 (69.0%)	25 (28.4%)	29 (34.1%)	Ref.	Ref.	Ref.	Ref.
	CG	25 (11.4%)	7 (8.0%)	9 (10.6%)	0.273	0.152	1.691 (0.661–4.325)	1.874 (0.794–4.427)
	GG	43 (19.6%)	56 (63.6%)	47 (55.3%)	**<0.001 ***	**<0.001 ***	7.866 (4.402–14.057)	5.691 (3.207–10.099)
	CG+GG	68 (31.1%)	63 (71.6%)	56 (65.9%)	**<0.001 ***	**0.001 ***	2.305 (1.510–3.518)	2.121 (1.376–3.273)
	C	327 (74.7%)	57 (32.4%)	67 (39.4%)	Ref.	Ref.	Ref.	Ref.
	G	111 (25.3%)	119 (67.6%)	103 (60.6%)	**<0.001 ***	**<0.001 ***	6.150 (4.197–9.013)	4.529 (3.111–6.593)

Table 3. *Cont.*

SNP	Variant	Control						
rs3176752	GG	208 (96.0%)	88 (100%)	85 (94.4%)	Ref.	Ref.	Ref.	Ref.
	GT	8 (4.0%)	0	5 (5.6%)	0.161	0.544	0.129 (0.007–2.264)	1.426 (0.453–4.487)
	TT	0	0	0	0.694	0.684	2.197 (0.043–111.66)	2.275 (0.044–115.60)
	GT+TT	8 (3.7%)	0	5 (5.6%)	0.169	0.562	0.134 (0.007–2.357)	1.402 (0.446–4.406)
	G	424 (98.1%)	176 (100%)	175 (97.2%)	Ref.	Ref.	Ref.	Ref.
	T	8 (1.9%)	0	5 (2.8%)	0.180	0.472	0.142 (0.008–2.465)	1.514 (0.489–4.693)

B. Based on frequency of smoking

SNP	Variant	Control	Frequency of smoking per day		Chi-square *p* values		OR for cases (CI)	
			≤10	>10	Control vs. ≤10	Control vs. >10	Control vs. ≤10	Control vs. >10
rs10817938	TT	12 (6.0%)	1 (1.5%)	3 (3.2%)	Ref.	Ref.	Ref.	Ref.
	TC	47 (23.6%)	24 (35.3%)	32 (34.0%)	0.090	0.144	6.128 (0.752–49.963)	2.723 (0.711–10.427)
	CC	140 (70.4%)	43 (63.2%)	59 (62.8%)	0.216	0.432	3.686 (0.466–29.164)	1.686 (0.459–6.193)
	TC+CC	187 (94.0%)	67 (98.5%)	91 (96.8%)	0.813	0.868	1.048 (0.708–1.552)	1.030 (0.725–1.463)
	T	71 (18.0%)	26 (19.1%)	41 (21.5%)	Ref.	Ref.	Ref.	Ref.
	C	327 (82.0%)	110 (80.9%)	150 (78.5%)	0.739	0.294	0.919 (0.558–1.512)	0.794 (0.517–1.222)
rs1800975	TT	119 (55.1%)	68 (93.2%)	92 (97.9%)	Ref.	Ref.	Ref.	Ref.
	TC	14 (6.5%)	4 (5.5%)	2 (2.1%)	0.238	**0.028**	0.500 (0.158–1.579)	0.185 (0.041–0.834)
	CC	83 (38.4%)	1 (1.4%)	0	**0.001 ***	**<0.001 ***	0.016 (0.002–0.114)	0.008 (0.001–0.126)
	TC+CC	97 (45.97%)	5 (6.8%)	2 (2.1%)	**0.001 ***	**<0.001 ***	0.152 (0.059–0.389)	0.047 (0.011–0.196)
	T	252 (58.3%)	140 (96.6%)	186 (98.9%)	Ref.	Ref.	Ref.	Ref.
	C	180 (41.7%)	5 (3.4%)	2 (1.1%)	**<0.001 ***	0.738	0.050 (0.020–0.124)	1.400 (0.195–10.032)
rs3176751	CC	151 (69.0%)	21 (28.8%)	31 (32.3%)	Ref.	Ref.	Ref.	Ref.
	CG	25 (11.4%)	6 (8.2%)	10 (10.4%)	0.285	0.115	1.725 (0.634–4.696)	1.948 (0.850–4.463)
	GG	43 (19.6%)	46 (63.0%)	55 (57.3%)	**<0.001 ***	**<0.001 ***	7.692 (4.148–14.262)	6.230 (3.575–10.858)
	CG+GG	68 (31.1%)	52 (71.2%)	65 (97.7%)	**0.001 ***	**0.001 ***	2.294 (1.466–3.590)	2.180 (1.438–3.305)
	C	327 (74.7%)	48 (32.9%)	72 (37.5%)	Ref.	Ref.	Ref.	Ref.
	G	111 (25.3%)	98 (67.1%)	120 (62.5%)	**<0.001 ***	**<0.001 ***	6.015 (4.004–9.035)	4.909 (3.415–7.058)
rs3176752	GG	208 (96.0%)	73 (94.8%)	96 (99.0%)	Ref.	Ref.	Ref.	Ref.
	GT	8 (4.0%)	4 (5.2%)	1 (1.0%)	0.573	0.221	1.425 (0.417–4.872)	0.271 (0.033–2.196)
	TT	0	0	0	nan	nan	nan	nan
	GT+TT	8 (3.7%)	4 (5.2%)	1 (1.0%)	0.589	0.231	1.403 (0.411–4.789)	0.278 (0.034–2.256)
	G	424 (98.1%)	146 (97.3%)	193 (99.5%)	Ref.	Ref.	Ref.	Ref.
	T	8 (1.9%)	4 (2.7%)	1 (0.5%)	0.547	0.225	1.452 (0.431–4.893)	0.275 (0.034–2.211)

C. Based on gender

SNP	Variant	Control	Gender		Chi-square *p* values		OR (CI)	
			Male	Female	Control vs. Male	Control vs. Female	Control vs. Male	Control vs. Female
rs10817938	TT	12 (6.0%)	7 (4.0%)	0	Ref.	Ref.	Ref.	Ref.
	TC	47 (23.6%)	67 (38.5%)	2 (14.3%)	0.081	0.862	2.444 (0.895–6.669)	1.316 (0.059–29.198)
	CC	140 (70.4%)	100 (57.5%)	12 (85.7%)	0.681	0.587	1.225 (0.466–3.220)	2.224 (0.124–39.835)
	TC+CC	187 (94.0%)	167 (96.0%)	14 (100%)	0.887	0.874	1.021 (0.763–1.367)	1.064 (0.491–2.292)
	T	71 (18.0%)	111 (29.4%)	2 (7.1%)	Ref.	Ref.	Ref.	Ref.
	C	327 (82.0%)	267 (70.6%)	26 (92.9%)	0.158	0.164	0.522 (0.372–0.733)	2.823 (0.655–12.166)
rs1800975	TT	119 (55.1%)	175 (97.8%)	11 (78.6%)	Ref.	Ref.	Ref.	Ref.
	TC	14 (6.5%)	3 (1.7%)	3 (21.4%)	**0.003 ***	0.236	0.146 (0.041–0.518)	2.318 (0.577–9.322)
	CC	83 (38.4%)	1 (0.6%)	0	**<0.001 ***	0.056	0.008 (0.001–0.059)	0.006 (0.004–1.071)
	TC+CC	97 (45.97%)	4 (2.2%)	3 (21.4%)	**<0.001 ***	0.253	0.049 (0.018–0.137)	0.477 (0.134–1.698)
	T	252 (58.3%)	353 (98.6%)	25 (89.3%)	Ref.	Ref.	Ref.	Ref.
	C	180 (41.7%)	5 (1.4%)	3 (10.7%)	**<0.001 ***	**0.004 ***	0.019 (0.008–0.018)	0.168 (0.050–0.565)

Table 3. *Cont.*

SNP	Variant				Ref./Ref. columns			
rs3176751	CC	151 (69.0%)	65 (35.9%)	0	Ref.	Ref.	Ref.	Ref.
	CG	25 (11.4%)	17 (9.4%)	0	0.188	**0.018 ***	1.580 (0.799–3.122)	5.941 (0.115–306.217)
	GG	43 (19.6%)	99 (54.7%)	14 (100%)	**<0.001 ***	**0.001 ***	5.349 (3.372–8.482)	101.00 (5.905–1717.54)
	CG+GG	68 (31.1%)	116 (64.1%)	14 (100%)	**0.001 ***	**0.004 ***	2.064 (1.442–2.953)	3.221 (1.463–7.091)
	C	327 (74.7%)	147 (40.6%)	0	Ref.	Ref.	Ref.	Ref.
	G	111 (25.3%)	215 (59.4%)	28 (100%)	**<0.001 ***	**0.001 ***	4.308 (3.190–5.819)	167.42 (10.137–2765-09)
rs3176752	GG	208 (96.0%)	181 (96.8%)	14 (100%)	Ref.	Ref.	Ref.	Ref.
	GT	8 (4.0%)	6 (3.2%)	0	0.787	0.910	0.862 (0.294–2.531)	0.846 (0.047–15.395)
	TT	0	0	0	nan	nan	nan	nan
	GT+TT	8 (3.7%)	6 (3.2%)	0	0.794	0.930	0.866 (0.295–2.542)	0.878 (0.048–15.984)
	G	424 (98.1%)	368 (98.4%)	28 (100%)	Ref.	Ref.	Ref.	Ref.
	T	8 (1.9%)	6 (1.6%)	0	0.789	0.928	0.864 (0.297–2.514)	0.876 (0.049–15.567)

D. Based on age

SNP	Variant	Control	age		Chi-square		OR (CI)	
			≤ 28 years	>28 years	≤28 years vs. control	≤28 years vs. control	≤28 years	>28 years vs. control
rs10817938	TT	12 (6.0%)	12 (5.0%)	7 (4.8%)	Ref.	Ref.	Ref.	Ref.
	TC	47 (23.6%)	62 (26.1%)	52 (35.9%)	0.540	0.215	1.319 (0.544–3.198)	1.897 (0.689–5.219)
	CC	140 (70.4%)	164 (68.9%)	86 (59.3%)	0.709	0.917	1.171 (0.510–2.690)	1.053 (0.399–2.778)
	TC+CC	187 (94.0%)	226 (95.0%)	138 (95.2%)	0.940	0.935	1.011 (0.771–1.324)	1.013 (0.745–1.377)
	T	71 (18.0%)	86 (18.1%)	66 (22.8%)	Ref.	Ref.	Ref.	Ref.
	C	327 (82.0%)	390 (81.9%)	224 (77.2%)	0.930	0.111	0.985 (0.696–1.393)	0.737 (0.506–1.073)
rs1800975	TT	119 (55.1%)	175 (68.6%)	126 (84.0%)	Ref.	Ref.	Ref.	Ref.
	TC	14 (6.5%)	15 (5.9%)	5 (3.3%)	0.417	**0.043 ***	0.729 (0.339–1.565)	0.337 (0.118–0.965)
	CC	83 (38.4%)	65 (25.5%)	19 (12.7%)	**0.002 ***	**<0.001 ***	0.533 (0.357–0.794)	0.216 (0.124–0.378)
	TC+CC	97 (45.97%)	80 (31.4%)	24 (16.0%)	**0.043 ***	**<0.001 ***	0.699 (0.493–0.989)	0.356 (0.218–0.583)
	T	252 (58.3%)	365 (71.6%)	257 (85.7)	Ref.	Ref.	Ref.	Ref.
	C	180 (41.7%)	145 (28.4%)	43 (14.3%)	**<0.001 ***	**<0.001 ***	0.556 (0.424–0.729)	0.234 (0.161–0.341)
rs3176751	CC	151 (69.0%)	138 (53.5%)	75 (49.3%)	Ref.	Ref.	Ref.	Ref.
	CG	25 (11.4%)	29 (11.2%)	12 (7.9%)	0.422	0.928	1.269 (0.709–2.273)	0.966 (0.460–2.029)
	GG	43 (19.6%)	91 (35.3%)	65 (42.8%)	**<0.001 ***	**<0.001 ***	2.316 (1.507–3.229)	2.316 (1.507–3.559)
	CG+GG	68 (31.1%)	120 (46.5%)	77 (50.7%)	**0.023 ***	**0.013 ***	1.498 (1.058–2.121)	1.632 (1.109–2.401)
	C	327 (74.7%)	305 (59.1%)	162 (53.3%)	Ref.	Ref.	Ref.	Ref.
	G	111 (25.3%)	211 (40.9%)	142 (46.7%)	**<0.001 ***	**<0.001 ***	2.038 (1.544–2.691)	2.582 (1.891–3.527)
rs3176752	GG	208 (96.0%)	250 (96.9%)	149 (96.1%)	Ref.	Ref.	Ref.	Ref.
	GT	8 (4.0%)	8 (3.1%)	6 (3.9%)	0.718	0.936	0.832 (0.307–2.255)	1.047 (0.356–3.081)
	TT	0	0	0	nan	nan	nan	nan
	GT+TT	8 (3.7%)	8 (3.1%)	6 (3.9%)	0.727	0.936	0.837 (0.309–2.268)	1.045 (0.355–3.073)
	G	424 (98.1%)	508 (98.4%)	304 (98.1%)	Ref.	Ref.	Ref.	Ref.
	T	8 (1.9%)	8 (1.6%)	6 (1.9%)	0.720	0.934	0.835 (0.311–2.243)	1.046 (0.359–3.046)

* = $p < 0.05$, SNP: Single nucleotide polymorphism, Ref = Reference allele, OR: Odd ratio, CI = Confidence interval. *p*-Values in bold represent significant results.

3.3. The Association between SNPs in XPA, XPC, CS Duration, Daily CS Average, Gender, and Age

The smokers and non-smokers were divided into different groups based on CS duration, daily CS average gender, and age. The associations of the four SNPs in XPA and the four SNPs in *XPC* genes with clinical characteristics were evaluated. Tables 3 and 4 compare genotyping and allele frequencies for each SNP in *XPA* and *XPC* genes based on different clinical characteristics.

Table 4. Comparison of *XPC* genotypes and alleles distribution according to duration of smoking, frequency of smoking, and gender.

A. Based on duration of smoking

SNP	Variant	Control	Duration of smoking		Chi-square *p* values		OR (CI)	
			≤5 years	>5 years	Control vs. <5 years	Control vs. >5 years	Control vs. <5 years	Control vs. >5 years
rs1870134	GG	0	1 (1.1%)	0	Ref.	Ref.	Ref.	Ref.
	GC	13 (6.0%)	4 (4.5%)	1 (1.2%)	0.202	0.313	0.111 (0.003–3.243)	0.111 (0.002–7.927)
	CC	205 (94.0%)	83 (94.3%)	82 (98.8%)	0.222	0.649	0.135 (0.006–3.359)	0.402 (0.008–20.403)
	GC+CC	218 (100%)	87 (98.8%)	83 (100%)	0.949	1.000	0.989 (0.696–1.404)	1.000 (0.699–1.430)
	G	13 (3.0%)	6 (3.4%)	1 (0.6%)	Ref.	Ref.	Ref.	Ref.
	C	423 (97.0%)	170 (96.6%)	165 (99.4%)	0.783	0.119	0.871 (0.326–2.329)	5.071 (0.658–39.084)
Rs2228000	CC	2 (0.9%)	0	3 (3.5%)	Ref.	Ref.	Ref.	Ref.
	CT	49 (22.9%)	18 (20.7%)	24 (28.2%)	0.691	0.237	1.868 (0.085–40.783)	0.326 (0.051–2.086)
	TT	163 (76.2%)	69 (79.3%)	58 (68.2%)	0.628	0.120	2.125 (0.100–44.849)	0.237 (0.038–1.456)
	CT+TT	212 (99.0%)	87 (100%)	82 (96.5%)	0.831	0.787	0.962 (0.675–1.371)	1.052 (0.728–1.519)
	C	53 (12.4%)	18 (10.3%)	30 (17.6%)	Ref.	Ref.	Ref.	Ref.
	T	375 (87.6%)	156 (89.7%)	140 (82.4%)	0.482	0.095	1.225 (0.695–2.158)	0.660 (0.405–1.075)
rs2228001	AA	72 (34.5%)	25 (29.4%)	31 (37.8%)	Ref.	Ref.	Ref.	Ref.
	AC	95 (45.4%)	43 (50.6%)	31 (37.8%)	0.371	0.353	1.304 (0.729–2.329)	0.758 (0.422–1.360)
	CC	42 (20.1%)	17 (20.0%)	20 (24.4%)	0.678	0.771	1.166 (0.565–2.405)	1.106 (0.561–2.181)
	AA+CC	137 (65.6%)	42 (49.4%)	51 (62.2%)	0.195	0.802	0.754 (0.491–1.156)	0.949 (0.629–1.431)
	A	239 (57.2%)	93 (52.5%)	93 (56.7%)	Ref.	Ref.	Ref.	Ref.
	C	179 (42.8%)	77 (47.5%)	71 (43.3%)	0.584	0.918	1.106 (0.772–1.583)	1.019 (0.708–1.468)
rs2607775	GG	111 (51.4%)	32 (37.6%)	36 (43.9%)	Ref.	Ref.	Ref.	Ref.
	GC	86 (39.8%)	32 (37.6%)	29 (35.4%)	0.376	0.892	1.291 (0.734–2.271)	1.040 (0.591–1.828)
	CC	19 (8.8%)	21 (24.7%)	17 (20.7%)	**0.001 ***	**0.008 ***	3.834 (1.839–7.993)	2.759 (1.297–5.868)
	GC+CC	105 (48.6%)	53 (62.4%)	46 (56.1%)	0.855	0.514	1.038 (0.693–1.556)	1.154 (0.751–1.774)
	G	308 (71.3%)	96 (56.5%)	101 (61.6%)	Ref.	Ref.	Ref.	Ref.
	C	124 (28.7%)	74 (43.5%)	63 (38.4%)	**0.001 ***	**0.023 ***	1.915 (1.325–2.766)	1.549 (1.062–2.260)

B. Based on frequency of smoking

SNP	Variant	Control	Frequency of smoking per day		Chi-square *p* values		OR (CI)	
			≤10	>10	Control vs. ≤10	Control vs. >10	Control vs. ≤10	Control vs. >10
rs1870134	GG	0	0	1 (1.0%)	Ref.	Ref.	Ref.	Ref.
	GC	13 (6.0%)	6 (8.5%)	1 (1.0%)	0.722	0.074	0.482 (0.009–27.092)	0.037 (0.001–1.380)
	CC	205 (94.0%)	65 (91.5%)	94 (97.9%)	0.569	0.252	0.319 (0.006–16.223)	0.153 (0.006–3.798)
	GC+CC	218 (100%)	71 (100%)	95 (99.0%)	1.00	0.952	1.00 (0.685–1.461)	0.989 (0.704–1.390)
	G	13 (3.0%)	6 (4.2%)	3 (1.6%)	Ref.	Ref.	Ref.	Ref.
	C	423 (97.0%)	136 (95.8%)	189 (98.4%)	0.473	0.307	0.697 (0.260–1.868)	1.936 (0.545–6.874)
Rs2228000	CC	2 (0.9%)	1 (1.4%)	2 (2.1%)	Ref.	Ref.	Ref.	Ref.
	CT	49 (22.9%)	22 (30.1%)	19 (20.0%)	0.932	0.360	0.898 (0.077–10.433)	0.388 (0.051–2.953)
	TT	163 (76.2%)	50 (68.5%)	74 (77.9%)	0.693	0.434	0.614 (0.055–6.908)	0.454 (0.063–3.285)
	CT+TT	212 (99.0%)	72 (98.6%)	93 (97.9%)	0.982	0.946	0.996 (0.683–1.451)	0.988 (0.701–1.393)
	C	53 (12.4%)	24 (16.4%)	23 (12.4%)	Ref.	Ref.	Ref.	Ref.
	T	375 (87.6%)	122 (83.6%)	163 (87.6%)	0.216	0.995	0.718 (0.426–1.213)	1.002 (0.594–1.689)
rs2228001	AA	72 (34.5%)	20 (27.8%)	33 (36.3%)	Ref.	Ref.	Ref.	Ref.
	AC	95 (45.4%)	35 (48.6%)	39 (42.9%)	0.379	0.698	1.326 (0.707–2.488)	0.896 (0.514–1.561)
	CC	42 (20.1%)	17 (23.6%)	19 (20.9%)	0.325	0.970	1.457 (0.688–3.086)	0.987 (0.499–1.949)
	AA+CC	137 (65.6%)	37 (51.4%)	52 (57.1%)	0.290	0.505	0.784 (0.499–1.231)	0.872 (0.583–1.305)
	A	239 (57.2%)	75 (52.1%)	105 (57.7%)	Ref.	Ref.	Ref.	Ref.
	C	179 (42.8%)	69 (47.9%)	77 (52.3%)	**<0.001 ***	0.954	2.569 (1.643–4.019)	1.011 (0.707–1.443)
rs2607775	GG	111 (51.4%)	23 (31.9%)	43 (47.3%)	Ref.	Ref.	Ref.	Ref.
	GC	86 (39.8%)	27 (37.5%)	33 (36.3%)	0.191	0.972	0.515 (0.812–2.826)	0.991 (0.581–1.689)
	CC	19 (8.8%)	22 (30.6%)	15 (16.5%)	**<0.001 ***	0.067	5.588 (2.612–11.956)	2.038 (0.950–4.371)
	GC+CC	105 (48.6%)	50 (69.4%)	76 (83.5%)	0.104	**0.006 ***	1.429 (0.929–2.195)	1.718 (1.171–2.521)
	G	308 (71.3%)	73 (50.7%)	119 (65.4%)	Ref.	Ref.	Ref.	Ref.
	C	124 (28.7%)	71 (49.3%)	63 (34.6%)	**<0.001 ***	0.147	2.416 (1.639–3.559)	1.315 (0.909–1.903)

Table 4. *Cont.*

C. Based on gender

SNP	Variant	Control	Gender		Chi-square *p* values		OR (CI)	
			Male	Female	Control vs. Male	Control vs. Female	Control vs. Male	Control vs. Female
rs1870134	GG	0	1 (0.6%)	0	Ref.	Ref.	Ref.	Ref.
	GC	13 (6.0%)	6 (3.4%)	1 (7.1%)	0.282	0.313	0.161 (0.006–4.505)	0.111 (0.002–7.927)
	CC	205 (94.0%)	171 (96.1%)	13 (92.9%)	0.434	0.178	0.278 (0.012–6.873)	0.066 (0.001–3.441)
	GC+CC	218 (100%)	177 (99.4%)	14 (100%)	0.969	1.00	0.994 (0.751–1.316)	1.00 (0.466–2.147)
	G	13 (3.0%)	8 (2.2%)	1 (3.4%)	Ref.	Ref.	Ref.	Ref.
	C	423 (97.0%)	348 (97.8%)	27 (96.6%)	0.524	0.859	1.337 (0.548–3.262)	0.829 (0.105–6.582)
Rs2228000	CC	2 (0.9%)	3 (1.7%)	0	Ref.	Ref.	Ref.	Ref.
	CT	49 (22.9%)	40 (22.2%)	5 (35.7%)	0.516	0.716	0.544 (0.087–3.412)	0.556 (0.024–13.116)
	TT	163 (76.2%)	137 (76.1%)	9 (64.3%)	0.529	0.435	0.560 (0.092–3.402)	0.291 (0.013–6.488)
	CT+TT	212 (99.0%)	177 (98.3%)	14 (100%)	0.959	0.981	0.993 (0.749–1.315)	1.009 (0.469–2.168)
	C	53 (12.4%)	46 (12.8%)	5 (17.9%)	Ref.	Ref.	Ref.	Ref.
	T	375 (87.6%)	314 (87.2%)	23 (82.1%)	0.868	0.403	0.965 (0.632–1.472)	0.650 (0.237–1.783)
rs2228001	AA	72 (34.5%)	64 (37.2%)	0	Ref.	Ref.	Ref.	Ref.
	AC	95 (45.4%)	73 (42.4%)	9 (64.3%)	0.527	0.068	0.862 (0.544–1.366)	14.381 (0.823–251.404)
	CC	42 (20.1%)	35 (20.3%)	5 (35.7%)	0.986	**0.044 ***	1.005 (0.566–1.758)	20.093 (1.081–373-437)
	AA+CC	137 (65.6%)	99 (57.6%)	5 (35.7%)	0.437	0.254	0.878 (0.633–1.219)	0.545 (0.192–1.547)
	A	239 (57.2%)	163 (53.3%)	9 (32.1%)	Ref.	Ref.	Ref.	Ref.
	C	179 (42.8%)	143 (46.7%)	19 (67.9%)	0.296	**0.013 ***	1.171 (0.871–1.579)	2.819 (1.246–6.377)
rs2607775	GG	111 (51.4%)	78 (44.6%)	1 (7.1%)	Ref.	Ref.	Ref.	Ref.
	GC	86 (39.8%)	69 (39.4%)	3 (21.4%)	0.545	0.245	1.142 (0.743–1.754)	3.872 (0.396–37.881)
	CC	19 (8.8%)	28 (16.0%)	10 (71.4%)	**0.026 ***	**0.001 ***	2.097 (1.094–4.019)	58.421 (7.065–483.07)
	GC+CC	105 (48.6%)	106 (60.6%)	13 (92.8%)	0.198	0.108	1.246 (0.891–1.743)	1.910 (0.867–4.209)
	G	308 (71.3%)	225 (64.3%)	5 (17.9%)	Ref.	Ref.	Ref.	Ref.
	C	124 (28.7%)	125 (35.7%)	23 (82.1%)	**0.037 ***	**<0.001 ***	1.379 (1.020–1.867)	11.426 (4.284–30.729)

D. Based on age

SNP	Variant	Control	age		Chi-square		OR (CI)	
			≤ 28 years	>28 years	≤ 28 years vs. control	>28 years vs. control	≤ 28 years vs. control	>28 years vs. control
rs1870134	GG	0	0	1 (0.7%)	Ref.	Ref.	Ref.	Ref.
	GC	13 (6.0%)	12 (4.7%)	8 (5.3%)	0.969	0.356	0.926 (0.017–50.291)	0.209 (0.008–5.769)
	CC	205 (94.0%)	243 (95.3%)	142 (94.0%)	0.933	0.371	1.185 (0.023–59.983)	0.231 (0.009–5.715)
	GC+CC	218 (100%)	255 (100%)	150 (99.4%)	1.00	0.965	1.0 (0.774–1.291)	0.993 (0.741–1.333)
	G	13 (3.0%)	12 (2.4%)	10 (3.3%)	Ref.	Ref.	Ref.	Ref.
	C	423 (97.0%)	498 (97.6%)	292 (96.7%)	0.549	0.800	1.275 (0.576–2.825)	0.897 (0.388–2.074)
Rs2228000	CC	2 (0.9%)	5 (2.0%)	0	Ref.	Ref.	Ref.	Ref.
	CT	49 (22.9%)	56 (22.0%)	37 (24.7%)	0.362	0.395	0.457 (0.085–2.463)	3.788 (0.177–81.266)
	TT	163 (76.2%)	193 (76.0%)	113 (75.3%)	0.376	0.423	0.474 (0.091–2.474)	3.471 (0.165–72.987)
	CT+TT	212 (99.0%)	249 (98.0%)	150 (100%)	0.937	0.950	0.989 (0.764–1.281)	1.009 (0.751–1.356)
	C	53 (12.4%)	66 (13.0%)	37 (12.3%)	Ref.	Ref.	Ref.	Ref.
	T	375 (87.6%)	442 (87.0%)	263 (87.7%)	0.781	0.984	0947 (0.643–1.394)	1.005 (0.645–1.573)
rs2228001	AA	72 (34.5%)	91 (37.0%)	42 (29.0%)	Ref.	Ref.	Ref.	Ref.
	AC	95 (45.4%)	107 (43.5%)	69 (47.6%)	0.586	0.381	0.891 (0.589–1.349)	1.245 (0.762–2.034)
	CC	42 (20.1%)	48 (19.5%)	34 (23.4%)	0.703	0.277	0.904 (0.539–1.516)	1.388 (0.769–2.506)
	AA+CC	137 (65.6%)	139 (56.5%)	76 (52.4%)	0.331	0.212	0.862 (0.639–1.163)	0.799 (0.5562–1.136)
	A	239 (57.2%)	289 (58.7%)	153 (52.8%)	Ref.	Ref.	Ref.	Ref.
	C	179 (42.8%)	203 (41.3%)	137 (47.2%)	0.634	0.245	0.938 (0.720–1.221)	1.196 (0.885–1.616)
rs2607775	GG	111 (51.4%)	130 (51.4%)	57 (38.5%)	Ref.	Ref.	Ref.	Ref.
	GC	86 (39.8%)	87 (34.4%)	70 (47.3%)	0.463	**0.044 ***	0.864 (0.584–1.277)	1.585 (1.012–2.486)
	CC	19 (8.8%)	36 (14.2%)	21 (14.2%)	0.123	**0.031 ***	1.618 (0.878–2.979)	2.152 (1.071–4.325)
	GC+CC	105 (48.6%)	123 (48.6%)	91 (62.3%)	0.999	0.188	1.001 (0.728–1.374)	1.265 (0.891–1.795)
	G	308 (71.3%)	347 (68.6%)	184 (62.2%)	Ref.	Ref.	Ref.	Ref.
	C	124 (28.7%)	159 (31.4%)	112 (37.8%)	0.366	**0.010 ***	1.138 (0.859–1.507)	1.512 (1.104–2.069)

* = *p* < 0.05, SNP: Single nucleotide polymorphism, Ref = Reference allele, OR: Odd ratio, CI = Confidence interval. *p*-Values in bold represent significant results.

3.3.1. Estimation for SNPs XPA (rs10817938, rs1800975, rs3176751, and rs3176752)

To evaluate the relationships between different *XPA* rs10817938 genotypes and CS in the control and smokers, we distributed the study based on smoking duration (short-term smokers ≤ five years and long-term smokers > five years), frequency of smoking (≤10 times and >10 times), gender (males and females), and the average age of smokers (≤28 years and >28 years). The results of genotype and allele distributions of the rs10817938 variant in smokers and controls, with its different clinical characteristics, are described in Table 3A–D. *XPA* rs10817938 showed an association only with the duration of smoking (smoking for ≤five years and >five years) in the TC+CC genotype. The distribution of genotyping frequency was 37.8% and 35.7%. In short-term and long-term smokers, only the TC+CC genotype showed low-risk association (TC+CC: OR = 0.402, CI = [0.254–0.637], *p* < 0.001 for short -term; while, in long-term-smoking: OR = 0.380, CI = [0.239–0.603], *p* < 0.001).

SNP rs1800975 has a significant association with all clinical parameters. With regards to smoking duration (more/less five years), the genotyping and allele distribution were analyzed in Table 3A. In a period of smoking of more than five years, the results showed highly significant low-risk association in genotype TC compared to the TT genotype. The heterozygous TC genotype in short-term smoking exhibits 0.202-fold low-risk effects in smokers (OR = 0.202; CI = [0.045–0.914]; *p* = 0.038). This went along with the CC genotype, which posed a 0.08-fold low-risk effect in short-term smokers (OR= 0.084; CI = [0.072–0.994]; *p* < 0.049) and long term-smoking (OR = 0.019; CI = [0.001–0.145]. Allele C was shown to be significant with regards to protection from the effects of smoking (OR = 0.034; CI = [0.012–0.093]; *p* < 0.001), (OR = 0.033; CI = [0.012–0.090]; in long term smoking and short-term smoking, respectively). The TC+CC genotype showed a reduced risk of association by 92% and 89.4% in both short- and long-term smoking (OR = 0.077; CI = [0.024–0.249]; *p* < 0.001) (OR = 0.106; CI = [0.037–0.297]; *p* < 0.001). The CC genotype presented a significant correlation with protection effects of moderate smokers (OR = 0.016; CI = [0.002–0.14]; *p* < 0.001). The *C. allele* posed 0.050 low-risk effects with moderate smokers (OR = 0.050, CI = [0.020–0.124] and *p* < 0.001). Additionally, TC+CC showed low risk association with moderate smoking (OR = 0.125; CI = [0.059–0.389]; *p* = 0.001). Heavy smokers showed significant low-risk association with the TC, CC, and TC+CC genotypes (OR = 0.0185 TC, 0.008 CC and 0.047 TC+CC); (CI = [0.041–0.834], *p* = 0.028); (CI = [0.001–0.126], *p* < 0.001); (CI = [0.011–0.0196], *p* < 0.001). In Table 3C, males showed significant low-risk association with smokers. The CC variant, which is homozygous in males, exhibited significant low-risk association with smokers (OR = 0.008; CI = [0.001–0.059]; *p* < 0.001), along with the *C. allele* (OR = 0.019; CI = [0.008–0.018]; *p* < 0.001). Additionally, the *C. Allele* presented 0.168-fold protective effects in female smokers (OR = 0.0168; CI = CI = [0.050–0.565]; *p* < 0.004). The SNPs showed low-risk association with age for both those under and above 28 years, as shown in Table 3D. The heterozygous TC presented a significant protective correlation with older smokers (OR = 0.3372; CI = [0.11809–1.965012]; *p* = 0.043). The CC homozygous variant genotype has a low-risk relationship with smokers in subjects older than 28 years (OR = 0.216; CI = [0.124–0.378]; *p* < 0.001). Additionally, the *C. allele* showed a significant low-risk association with smokers older than 28 (OR = 0.234; CI = [0.161–0.341; *p* < 0.001). In contrast, the TC variant of smokers under 28 years showed no significant association. The CC variant (homozygous) of smokers under 28 years also showed a low-risk relationship with smokers (OR = 0.533010; CI = [0.357–0.079]; *p* < 0.002). Furthermore, there a low-risk association was observed in the *C. allele* (OR = 0.0556; CI = [0.424–0.7290]; *p* < 0.001).

The rs3176751 SNPs showed significant high-risk association with all clinical parameters tested in this study. The genotyping and allele distribution were shown in Table 3A–D. For example, the allele frequencies analysis of the GG genotype and the *G. allele* showed that short-term smokers (≤ five years) and long-term smokers (>five years) revealed high-risk significant association when compared to non-smokers (for short-term smokers: (OR = 7.866 GG and 6.150 G), (CI = [(4.4012–14.056)], *p* < 0.001), (CI = [4.197–9.013], *p* < 0.001), respectively; for short-term smokers: (OR = 5.691 GG and 4.529 G), (CI = [3.207–10.099], *p* < 0.001); (CI = [3.111–6.593], *p* < 0.001), respectively). The moderate and heavy smokers both dis-

played a significant high-risk relationship when compared to non-smokers (Table 3B). The GG genotype presented a significant high-risk association of moderate smokers (OR = 7.692; CI = [4.148–14.262]; $p < 0.001$). The *G. allele* presented 6.015 high-risk effects with moderate smokers (OR = 6.015, CI = [4.004–9.035] and $p < 0.001$). Heavy smokers showed a significant high-risk association with the GG genotype and the *G. allele* ((OR = 6.230 GG, 4.909 G); (CI = [3.575–10.858] $p < 0.001$); (CI = [3.415–7.058], $p < 0.001$), respectively). The combination of CG+GG has a significant high-risk association in both moderate and heavy smokers ($p < 0.001$ (OR = 2.294, 2.180; CI = [1.466–3.590], CI = [1.438–3.305], respectively). In addition, there was significant high-risk association for both genders, male and female, as well as the age of subjects for the GG genotype and the *G. allele,* as shown in Table 3C, D. The GG genotype presented a significant high-risk association of male smokers (OR = 5.349; CI = [3.372–8.482]; $p < 0.001$). The *G. allele* presented 4.308 high-risk effects with male smokers (OR = 4.308, CI = [3.190–5.819] and $p < 0.001$). Female smokers showed significant high-risk association with the GG genotype and the *G. allele* ((OR = 101 GG, 167.42 G); (CI = [5.905–1717.54, $p = 0.001$); (CI = [10.137–276509], $p = 0.001$), respectively). The SNPs showed a high-risk association with age for both those under and those above 28 years, as shown in Table 3D. The GG variant (homozygous) in smokers under 28 years showed a high-risk relationship (OR = 2.316; CI = [1.507–3.229]; $p < 0.001$). Furthermore, a high-risk association was observed in the *G. allele* (OR = 2.038; CI = [1.544–2.691]; $p < 0.001$). Furthermore, the GG and G variants in smokers above 28 years showed significant association. The GG homozygous variant genotype has a high-risk relationship with smokers in subjects older than 28 years (OR = 3.0434; CI = [1.8937–4.8912]; $p < 0.001$). Additionally, the *G. allele* showed a significant high-risk association with smokers older than 28 years (OR = 2.582; CI = [1.891–3.527]; $p < 0.001$).

SNP rs3176752 showed no significant association with clinical parameters, including CS duration, daily CS average, gender, and younger smokers, as shown in Table 3A–D.

3.3.2. Estimation for SNPs XPC (rs1870134, rs2228000, rs2228001, and rs2607775)

The genotyping and allele distributions of SNPs *XPA* rs1870134 and rs2228000 were estimated in order to investigate the link between clinical parameters and the risk of smoking-induced diseases. These SNPs showed no significant association with all clinical parameters. The comparison of alleles and genotyping frequencies between subjects with the four clinical characteristics did not show any correlation because the p value is not statistically significant.

The analysis result in SNP rs2228001 does not show a significant association with the risk of smoking causing disease, considering genotyping and allele frequencies and statistical values, except for the female gender. The CC genotype and the C allele have a high-risk association with smoking-induced diseases ((OR = 18.765 CC, 2.810 C); CI = [1.012–347.841], [1.246–6.377], $p = 0.044$, 0.013, respectively).

SNP 2607775CC presented a high-risk association with all clinical parameters. Based on the C allele and the CC genotype frequencies comparison, the SNP between subjects with short-term smokers and long-term smokers compared to non-smokers showed significant high-risk association with smoking duration (Table 4A). Similarly, the CC and the *C. allele* genotype displayed a significant high-risk relationship with moderate smokers when compared to non-smokers (see Table 4B). Additionally, the SNP showed significant high-risk association with both male and female smokers (Table 4C). Lastly, there was a significant high-risk association for age above 28 between smokers and non-smokers (Table 4D).

3.4. Observed and Expected Counts

The null hypothesis is that there is no difference in the genotypes or alleles and the results in the equilibrium. So, if the p value is <0.05, we will reject the null hypothesis, and the genotype/alleles will be in disequilibrium (Table 5).

Table 5. Hardy-Weinberg Equilibrium Test and for all genotypes and alleles in XPA and XPC genes.

		Nonsmokers		Hardy-Weinberg Chi-Square Equilibrium Test	Smokers		Hardy-Weinberg Chi-Square Equilibrium Test	Interpretation
		Observed Expected	Expected		Observed	Expected		
rs10817938	TT	12	9	5.959	7	9.2	5.435	
	TC	47	59.6	$p = 0.051$	69	56.4	$p = 0.057$	Equilibrium
	CC	140	129.6		112	122.4		
rs1800975	TT	119	161.1	5.865	186	143.9	5.911	
	TC	14	10.6	$p = 0.053$	6	9.4	$p = 0.052$	Equilibrium
	CC	83	44.4		1	39.6		
rs3176751	CC	151	114.3	1.812	65	101.7	1.806	
	CG	25	22.2	$p = 0.404$	17	19.8	$p = 0.405$	Equilibrium
	GG	43	82.5		113	73.5		
rs3176752	GG	208	208.7	13.312	195	194.3	13.517	disEquilibrium
	GT	8	7.3	$p < 0.001$ *	6	6.7	$p < 0.001$ *	
rs1870134	GG	0	5	19.081	1	5	17.438	
	GC	13	10.6	$p < 0.001$ *	7	9.4	$p < 0.001$ *	disEquilibrium
	CC	205	205.8		184	182.2		
Rs2228000	CC	2	2.6	26.323	3	2.4	25.831	
	CT	49	49.3	$p < 0.001$ *	45	44.7	$p < 0.001$ *	disEquilibrium
	TT	163	162.1		146	146.9		
rs2228001	AA	72	72	0.318	64	64	0.317	
	AC	95	93.7	$p = 0.853$	82	83.3	$p = 0.853$	Equilibrium
	CC	42	43.4		40	38.6		
rs2607775	GG	111	101.3	0.933	79	88.7	0.933	
	GC	86	84.3	$p = 0.627$	72	73.7	$p = 0.627$	Equilibrium
	CC	19	30.4		38	26.6		

* = $p < 0.05$. *p*-Values in bold represent significant results

4. Discussion

CS is a major cause of various serious diseases, and several studies have shown that the cause of these diseases is due to tobacco chemicals [27]. For decades, scientific studies have not considered the damaging effects of CS on the human body, such as the lungs, respiratory system, and oral cavity. Nevertheless, there is evidence that CS promotes inflammation in the oral cavity and assists the development of gingival and periodontal disease by stimulating the secretion of inflammatory cytokines [28]. These findings illustrate the possibility of a powerful and effective approach to learning more about the effects of CS on oral mucosae. Previous studies showed that there are different types of cancers, and diseases are related to CS. Indeed, CS could affect many cellular pathways, such as KRAS and P53, causing certain defects, and it may increase the risk of cancer and disease [29]. For instance, it has been found that CS induces DNA damage and has an impact on the genes controlling the cell cycle and DNA replication and repair. Mutations or polymorphisms in these genes were found to be linked with smoking and various type of cancer [4,7,30]. Furthermore, it was reported that HIF-1α expression in human NSCLC cells was induced by exposure to Nicotine in s CS [31]. HIF was associated with GLUT-1 in hypoxia conditions. Elevated GLUT1 expression has been associated with many cancers, including hepatic, pancreatic, renal, breast, ovarian, brain, esophageal, lung, cutaneous, endometrial, colorectal, and cervical [32].

The risks of CS were underestimated, since those risks were assumed to be well linked, with machine-determined CS yields. However, the changes in people's smoking behaviors were not accounted for in machine-determined smoke yields [33]. Studies estimating exposure biomarkers would have presented more accurate assessments of risk, as these biomarkers will be a result of smoking behavior, as well the characteristics of the cigarette itself [34]. To better understand the association between CS and relative diseases, it is, therefore, critical to evaluate how polymorphisms function in related genes. For example, *XPA* and *XPC* genes may contribute to the disease because of CS. Due to the serious health

condition caused by smoking in the Saudi population; this study could indicate the early effects of the smoking-induced disease because of genetic variations to several affected genes following the CS exposure. This shows the importance of evaluating the effects of CS on causing disease or cancer by looking into associations with a genetic variation on the number of genes in different pathways.

The present study aimed to investigate the *XPA* and *XPC* gene polymorphisms' variations in the NER pathway, an important part of the DNA repair system damage caused by CS in smokers versus non-smokers of the Saudi population, to detect a genetic marker that could help predict disease, thus reducing the risks caused by CS among healthy individuals. In this work, four SNPs were selected (rs10817938, rs1800975, rs3176751, and rs3176752) and distributed in different regions of *XPA* gene, and four SNPs (rs1870134, rs2228000, rs2228001, and rs2607775) were distributed in different regions of the *XPC* gene. There have been many studies suggesting that *XPA* and *XPC* polymorphisms had a significant effect on the risk of cancer and disease, and they could be a biomarker [35,36].

Furthermore, the associations were validated between *XPA* and *XPC* SNPs and clinical characteristics, including CS duration, daily CS average, gender, and age. The *XPA* and *XPC* SNPs appeared to be significantly affected by CS, resulting in genetic changes in the DNA repair system gene. Given that, the *XPA* polymorphisms are related to the risk of many types of cancers [21,37]. In this study, SNP rs10817938 results showed a significant low-risk association only with the duration of smoking. SNPs rs1800975 and rs3176751 results showed significant low-risk and high-risk associations, respectively, with regard to all clinical parameters. However, rs3176752 showed no significant association in all parameters. The results of *XPA* rs3176751 polymorphisms might increase the influence of these clinical parameters regarding disease caused by CS. It is considered that CS contains chemical carcinogens that are known to produce genetic mutations that may not be repaired by the NER pathway because this mutation may not be recognized by *XPA* with the rs3176751 mutant gene type. The study results of SNP 10817938 for the Saudi population are inconsistent with the study that confirmed the association of *XPA* polymorphisms with oral squamous cell carcinoma (OSCC) risk in the Han Chinese population. The results demonstrated a significant high-risk association between CS and CC homozygous genotype in rs10817938 of OSCC, $p < 0.01$, OR = 3.60 [38]. For rs1800975, the results of this study are similar to a prior study, demonstrating that this variant was associated with a significantly reduced risk of lung cancer [39]. Although the percentage of female smokers used in this study was very low (13.1%) compared to the male smokers (59.9%), it was intriguing to find a significant high-risk difference in the genotypic and the allelic distribution of *XPA* rs3176751 in female smokers, suggesting a possible interference of CS in disease development among women, as reported previously for innate immune genes in acute respiratory distress syndrome [40] and human papillomavirus [41]. The results of the study indicate that gender may have a significant role in the association between the *XPA* rs3176751 polymorphism and the cancer risk or other diseases.

A comparison between smokers with clinical characteristics to controls revealed that there were no associations observed between SNPs rs1870134 and rs2228000 in the *XPC* gene and smokers. These results do not match the previous study, which showed that *XPC* rs1870134 was verified to be correlated with a decreased risk of hepatocellular and prostate cancers [42,43]. Additionally, SNP rs2228000 CT/TT genotype revealed a protective effect of gastric cancer only significant among subjects older than 58 years in a Southern Chinese population [44]. However, there were significant high-risk associations between the rs2228001 polymorphism and female smokers. SNP rs2607775 showed significant high-risk association with all clinical parameters.

Finally, this work offers various strengths and benefits. One of its strengths is that this study determined polymorphisms in the *XPA* and *XPC* genes in three categories of the SNP site, including 3′ UTR, 5′ UTR, and exon variants. Second, all samples were obtained from the same region of Riyadh and not from different regions of Saudi Arabia, and they were carefully monitored and stored according to protocol. However, due to the social traditions

of our community, this study was limited with regards to the adequacy of samples from female smokers.

5. Conclusions

Overall, the present study results demonstrated possible significant associations between CS and SNPs polymorphisms in DNA repair genes, such as *XPA* and *XPC*, and these effects of polymorphism can be a key factor in the development of CS-induced disease. The exact mechanism of how smoking influences genetic changes that cause cancer or disease remains unclear. Therefore, future studies are required to investigate the expressions of *XPA* and *XPC* gene and the link between polymorphisms and the rate of CS. Additionally, we suggest examining the oxidatively generated guanine lesion 8-oxoguanin to evaluate the oxidative stress DNA damage that occurs by CS. A further investigation comparing our results with other previously studied populations involving different ethnicities and CS habits may help define the effects of CS on different genes involved in DNA repair systems. The finding of identified SNPs polymorphisms associated with the CS induced disease could be used as biomarkers.

Abbreviations

NER	Nucleotide excision repair
SNP	Single nucleotide polymorphism
XPA	Xeroderma pigmentosum group A
XPC	Xeroderma pigmentosum group C
KRAS	Gene Kirsten rat sarcoma viral oncogene homolog
TFIIH	Transcription factor IIH
hHR23B	Human RAD23 homolog B
µg	Microgram
ng	Nanogram
CIs	confidence intervals
µL	Microliter
OR	Odds ratio
p-Value	Probability value
SPSS	Statistical Package for the Social Sciences

References

1. Morabia, A. Quality, originality, and significance of the 1939 "Tobacco consumption and lung carcinoma" article by Mueller, including translation of a section of the paper. *Prev. Med.* **2012**, *55*, 171–177. [CrossRef]
2. Proctor, R.N. The history of the discovery of the cigarette-lung cancer link: Evidentiary traditions, corporate denial, global toll. *Tob. Control* **2012**, *21*, 87. [CrossRef]
3. Martin-Timon, I.; Sevillano-Collantes, C.; Segura-Galindo, A.; del Canizo-Gomez, F.J. Type 2 diabetes and cardiovascular disease: Have all risk factors the same strength? *World J. Diabetes* **2014**, *5*, 444–470. [CrossRef]
4. Centers For Disease Control and Prevention (US); NCfCdPahPU; Office on Smoking and Health (US). *How Tabacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General*; Centers for Disease Control and Prevention (US): Atlanta, GA, USA, 2010.
5. Society, A.C. *Cancer Facts & Figures 2016*; American Cancer Society: Atlanta, GA, USA, 2016; pp. 1–9.
6. Semlali, A.; Almutairi, M.; Azzi, A.; Parine, N.R.; AlAmri, A.; Alsulami, S.; Rouabhia, M. TSLP and TSLP receptors variants are associated with smoking. *Mol. Genet. Genom. Med.* **2019**, *7*, e842. [CrossRef]
7. Gibbons, D.L.; Byers, L.A.; Kurie, J.M. Smoking, p53 mutation, and lung cancer. *Mol. Cancer Res.* **2014**, *12*, 3–13. [CrossRef] [PubMed]
8. Yang, H.; Zhong, J.T.; Zhou, S.H.; Han, H.M. Roles of GLUT-1 and HK-II expression in the biological behavior of head and neck cancer. *Oncotarget* **2019**, *10*, 3066–3083. [CrossRef]
9. Popanda, O.; Schattenberg, T.; Phong, C.T.; Butkiewicz, D.; Risch, A.; Edler, L.; Schmezer, P. Specific combinations of DNA repair gene variants and increased risk for non-small cell lung cancer. *Carcinogenesis* **2004**, *25*, 2433–2441. [CrossRef] [PubMed]
10. Scharer, O.D. Nucleotide Excision Repair in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*. [CrossRef] [PubMed]
11. Stern, M.C.; Lin, J.; Figueroa, J.D.; Kelsey, K.T.; Kiltie, A.E.; Yuan, J.M.; Matullo, G.; Fletcher, T. Polymorphisms in DNA Repair Genes, Smoking, and Bladder Cancer Risk: Findings from the International Consortium of Bladder Cancer. *Cancer Res.* **2009**, *69*, 6857–6864. [CrossRef]
12. Chatterjee, N.; Walker, G.C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **2017**, *58*, 235–263. [CrossRef]
13. Lopez-Cima, M.F.; Gonzalez-Arriaga, P.; Garcia-Castro, L.; Pascual, T.; Marron, M.G.; Puente, X.S.; Tardón, A. Polymorphisms in XPC, XPD, XRCC1, and XRCC3 DNA repair genes and lung cancer risk in a population of Northern Spain. *BMC Cancer* **2007**, *7*, 162. [CrossRef]
14. You, J.S.; Wang, M.; Lee, S.H. Biochemical analysis of the damage recognition process in nucleotide excision repair. *J. Biol. Chem.* **2003**, *278*, 7476–7485. [CrossRef] [PubMed]
15. Fuss, J.O.; Tainer, J.A. XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. *DNA Repair* **2011**, *10*, 697–713. [CrossRef]
16. Manandhar, M.; Boulware, K.S.; Wood, R.D. The ERCC1 and ERCC4 (XPF) genes and gene products. *Gene* **2015**, *569*, 153–161. [CrossRef]
17. Usanova, S.; Piee-Staffa, A.; Sied, U.; Thomale, J.; Schneider, A.; Kaina, B.; Köberle, B. Cisplatin sensitivity of testis tumour cells is due to deficiency in interstrand-crosslink repair and low ERCC1-XPF expression. *Mol. Cancer* **2010**, *9*, 248. [CrossRef]
18. Bunick, C.G.; Miller, M.R.; Fuller, B.E.; Fanning, E.; Chazin, W.J. Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein. *Biochemistry* **2006**, *45*, 14965–14979. [CrossRef] [PubMed]
19. Zhou, H.; Saliba, J.; Sandusky, G.E.; Sears, C.R. XPC protects against smoking- and carcinogen-induced lung adenocarcinoma. *Carcinogenesis* **2019**, *40*, 403–411. [CrossRef]
20. Asahina, H.; Kuraoka, I.; Shirakawa, M.; Morita, E.H.; Miura, N.; Miyamoto, I.; Tanaka, K. The Xpa Protein Is a Zinc Metalloprotein with an Ability to Recognize Various Kinds of DNA-Damage. *Mutat. Res.-DNA Repair* **1994**, *315*, 229–237. [CrossRef]
21. Yang, Z.G.; Liu, Y.; Mao, L.Y.; Zhang, J.T.; Zou, Y. Dimerization of human XPA and formation of XPA2-RPA protein complex. *Biochemistry* **2002**, *41*, 13012–13020. [CrossRef]
22. Almutairi, M.H.; Al-Numair, N.S.; Parine, N.R.; Almutairi, B.O.; Alrefaei, A.F.; Rouabhia, M.; Semlali, A. The protective effects of the methylenetetrahydrofolate reductase rs1801131 variant among Saudi smokers. *Saudi J. Biol. Sci.* **2021**, *28*, 3972–3980. [CrossRef]
23. Alhadheq, A.M.; Purusottapatnam Shaik, J.; Alamri, A.; Aljebreen, A.M.; Alharbi, O.; Almadi, M.A.; Alhadeq, F.; Azzam, N.A. The Effect of Poly(ADP-ribose) Polymerase-1 Gene 3′Untranslated Region Polymorphism in Colorectal Cancer Risk among Saudi Cohort. *Dis. Markers* **2016**, *2016*, 8289293. [CrossRef]
24. Semlali, A.; Reddy Parine, N.; Arafah, M.; Mansour, L.; Azzi, A.; Al Shahrani, O.; Alanazi, M.S. Expression and Polymorphism of Toll-Like Receptor 4 and Effect on NF-kappaB Mediated Inflammation in Colon Cancer Patients. *PLoS ONE* **2016**, *11*, e0146333. [CrossRef]
25. Semlali, A.; Almutairi, M.H.; Alamri, A.; Reddy Parine, N.; Arafah, M.; Almadi, M.A.; Aljebreen, A.M.; Alharbi, O. Expression and Polymorphism of TSLP/TSLP Receptors as Potential Diagnostic Markers of Colorectal Cancer Progression. *Genes* **2021**, *12*, 1386. [CrossRef] [PubMed]
26. Almutairi, M.; Almutairi, B.; Almutairi, M.; Parine, N.R.; Alrefaei, A.; Alanazi, M.; Semlali, A. Human beta-defensin-1 rs2738047 polymorphism is associated with shisha smoking risk among Saudi population. *Environ. Sci. Pollut. Res. Int.* **2021**, *28*, 42916–42933. [CrossRef] [PubMed]

27. Saha, S.P.; Bhalla, D.K.; Whayne, T.F.; Gairola, C.G. Cigarette smoke and adverse health effects: An overview of research trends and future needs. *Int. J. Angiol.* **2007**, *16*, 77–83. [CrossRef] [PubMed]

28. Semlali, A.; Witoled, C.; Alanazi, M.; Rouabhia, M. Whole Cigarette Smoke Increased the Expression of TLRs, HBDs, and Proinflammory Cytokines by Human Gingival Epithelial Cells through Different Signaling Pathways. *PLoS ONE* **2012**, *7*. [CrossRef]

29. Pezzuto, A.; Citarella, F.; Croghan, I.; Tonini, G. The effects of cigarette smoking extracts on cell cycle and tumor spread: Novel evidence. *Future Sci. OA* **2019**, *5*, FSO394. [CrossRef] [PubMed]

30. Phillips, D.H.; Venitt, S. DNA and protein adducts in human tissues resulting from exposure to tobacco smoke. *I.J. Cancer* **2021**, *131*, 2733–2753. [CrossRef]

31. Guo, L.; Li, L.; Wang, W.; Pan, Z.; Zhou, Q.; Wu, Z. Mitochondrial reactive oxygen species mediates nicotine-induced hypoxia-inducible factor-1α expression in human non-small cell lung cancer cells. *Biochim. Biophys. Acta BBA—Mol. Basis Dis.* **2012**, *1822*, 852–861. [CrossRef]

32. Pezzuto, A.; D'Ascanio, M.; Ricci, A.; Pagliuca, A.; Carico, E. Expression and role of p16 and GLUT1 in malignant diseases and lung cancer: A review. *Thoracic Cancer* **2020**, *11*, 3060–3070. [CrossRef]

33. Hatsukami, D.K.; Hecht, S.S.; Hennrikus, D.J.; Joseph, A.M.; Pentel, P.R. Biomarkers of tobacco exposure or harm: Application to clinical and epidemiological studies. *Nicotine Tob. Res.* **2003**, *5*, 387–396. [CrossRef]

34. Chang, C.M.; Edwards, S.H.; Arab, A.; Del Valle-Pinero, A.Y.; Yang, L.; Hatsukami, D.K. Biomarkers of Tobacco Exposure: Summary of an FDA-Sponsored Public Workshop. *Cancer Epidemiol. Biomark. Prev.* **2017**, *26*, 291–302. [CrossRef] [PubMed]

35. Gowda, S.; Kumaran, M.; Suma, B.S.; Rao, M.R. Single nucleotide polymorphism analysis of the nucleotide excision repair genes XPC, XPA, and XPG in the Indian population. *Hum. Biol.* **2007**, *79*, 545–562. [CrossRef] [PubMed]

36. Vogel, U.; Overvad, K.; Wallin, H.; Tjonneland, A.; Nexo, B.A.; Raaschou-Nielsen, O. Combinations of polymorphisms in XPD, XPC and XPA in relation to risk of lung cancer. *Cancer Lett.* **2005**, *222*, 67–74. [CrossRef]

37. Zhang, Y.; Guo, Q.; Yin, X.; Zhu, X.; Zhao, L.; Zhang, Z.; Li, X. Association of XPA polymorphism with breast cancer risk: A meta-analysis. *Medicine* **2018**, *97*, e11276. [CrossRef]

38. Gao, C.H.; Wang, J.Z.; Li, C.; Zhang, W.; Liu, G.X. A Functional Polymorphism (rs10817938) in the XPA Promoter Region Is Associated with Poor Prognosis of Oral Squamous Cell Carcinoma in a Chinese Han Population. *PLoS ONE* **2016**, *11*. [CrossRef] [PubMed]

39. Liu, X.; Lin, Q.Y.; Fu, C.P.; Liu, C.Y.; Zhu, F.; Liu, Z.L.; Jiang, L. Association between XPA gene rs1800975 polymorphism and susceptibility to lung cancer: A meta-analysis. *Clin. Respir. J.* **2018**, *12*, 448–458. [CrossRef]

40. Feng, Q.; Liu, N.; Song, S.; Ma, Y. Relationship between beta-defensin-1 gene polymorphism and susceptibility and prognosis of acute respiratory distress syndrome. *Medicine* **2019**, *98*, e14131. [CrossRef]

41. Segat, L.; Zupin, L.; Moura, R.R.; Coelho, A.V.C.; Chagas, B.S.; de Freitas, A.C.; Crovella, S. DEFB1 polymorphisms are involved in susceptibility to human papillomavirus infection in Brazilian gynaecological patients. *Memórias Do Inst. Oswaldo Cruz* **2014**, *109*, 918–922. [CrossRef]

42. Wang, B.G.; Xu, Q.; Yang, H.W.; Sun, L.P.; Yuan, Y. The association of six polymorphisms of five genes involved in three steps of nucleotide excision repair pathways with hepatocellular cancer risk. *Oncotarget* **2016**, *7*, 20357–20367. [CrossRef]

43. Wang, M.Y.; Li, Q.X.; Gu, C.Y.; Zhu, Y.; Yang, Y.J.; Wang, J.C.; Wei, Q. Polymorphisms in nucleotide excision repair genes and risk of primary prostate cancer in Chinese Han populations. *Oncotarget* **2017**, *8*, 24362–24371. [CrossRef] [PubMed]

44. Hua, R.X.; Zhuo, Z.J.; Shen, G.P.; Zhu, J.; Zhang, S.D.; Xue, W.Q.; Li, X.-Z.; Zhang, P.-F. Polymorphisms in the XPC gene and gastric cancer susceptibility in a Southern Chinese population. *OncoTargets Ther.* **2016**, *9*, 5513–5519. [CrossRef]