Polymorphisms at *XPD* and *XRCC1* DNA repair loci and increased risk of oral leukoplakia and cancer among *NAT2* slow acetylators

Mousumi Majumder, Nilabja Sikdar, Saurabh Ghosh and Bidyut Roy*

Human Genetics Unit, Biological Sciences Division, Indian Statistical Institute, Kolkata, India

Polymorphisms at *N*-acetyl transferase 2 locus (*NAT2*) lead to slow, intermediate and rapid acetylation properties of the enzyme. Improper acetylation of heterocyclic and aromatic amines, present in tobacco, might cause DNA adduct formation. Generally, DNA repair enzymes remove these adduct to escape malignancy. But, tobacco users carrying susceptible *NAT2* and DNA repair loci might be at risk of oral leukoplakia and cancer. In this study, 389 controls, 224 leukoplakia and 310 cancer patients were genotyped at 5 polymorphic sites on *NAT2* and 3 polymorphic sites on each of *XRCC1* and *XPD* loci by PCR-RFLP method to determine the risk of the diseases. None of the SNPs on these loci independently could modify the risk of the diseases in overall population but variant genotype (*Gln*/*Gln*) at codon 399 on *XRCC1* and major genotype (*Lys*/*Lys*) at codon 751 on *XPD* were associated with increased risk of leukoplakia and cancer among slow acetylators, respectively (OR = 4.2, 95% CI = 1.2–15.0; OR = 1.6, 95% CI = 1.1–2.3, respectively). Variant genotype (*Asn*/*Asn*) at codon 312 on *XPD* was also associated with increased risk of cancer among rapid and intermediate acetylators (OR = 1.9, 95% CI = 1.2–2.9). Variant *C-G-A* haplotype at *XRCC1* was associated with increased risk of leukoplakia (OR = 1.7, 95% CI = 1.2–2.4) but leukoplakia and cancer in mixed tobacco users (OR = 3.1, 95% CI = 1.4–7.1, OR = 2.4, 95% CI = 1.1–5.4, respectively) among slow acetylators. Although none of the 3 loci could modulate the risk of the diseases independently but 2 loci in combination, working in 2 different biochemical pathways, could do so in these patient populations.
© 2007 Wiley-Liss, Inc.

Key words: tobacco use; oral cancer; leukoplakia, *NAT2*; *XPD*; *XRCC1*; polymorphism

As an early sign of damage to oral mucosa, tobacco smokers and chewers often develop different precancerous lesions, such as leukoplakia, erythroplakia, submucous fibrosis, *etc*., and these lesions are easily accessible to diagnosis. Annual incidence of oral leukoplakia has been reported as 0.2–11.7% in different populations of India[1–3] and about 2–12% of leukoplakia becomes malignant within several years.[2] Since leukoplakia is one of the good predictors of oral cancer so diagnosis and treatment of leukoplakia will be a useful strategy to control oral cancer incidence. Annually about 270,000 cases of oral cancer are reported worldwide but about 82,000 of them are diagnosed in India.[4]

Major procarcinogens present in the tobacco smoke are polycyclic aromatic hydrocarbons, aromatic and heterocyclic amines and nitroso-compounds whereas nitrosamines and aromatic and heterocyclic amines are major components present in smokeless tobacco. Most of the tobacco carcinogens generally undergo bioactivation and inactivation by phase I and phase II enzymes respectively. Human *N*-acetylation transferase 2 (NAT2) is one of the phase II enzymes that participate in the bioconversion of aromatic and heterocyclic amines and variation in NAT2 enzyme activity is defined as polymorphism in *N*-acetylation capacity. This polymorphism arises from variations in DNA sequence resulting in the production of NAT2 proteins with variable enzyme activity or stability.[5] The impact of different acetylation activities of NAT2 enzyme on cancer susceptibility varies among different organs. *O*-acetylation by rapid NAT2 enzyme increased the risk of colon cancer, probably owing to extensive activities of heterocyclic amines (such as 2-amino-1-methyl-6-phenylimidazo [4,5-b] pyridine, *i.e.*, PhIP).[6,7] Again, NAT2 slow acetylation increased the risk of bladder cancer, probably owing to slow acetylation at *N*-position of aromatic amines (such as 4-amino biphenyl), resulting in accumulation in the affected tissue.[8] A few studies, with small sample sizes, also observed positive associations between NAT2 slow acetylation status and risk of head and neck cancer.[9,10] This association was not reproduced in subsequent studies except for a subset of samples.[11,12] Studies on laryngeal cancer also reported inconsistent association between NAT2 acetylation status and risk of cancer.[7,13]

Bulky DNA adducts, like those formed by aryl and heterocyclic amines, are generally repaired by DNA repair enzymes. The *XPD* is a component of the transcription factor TFIIH, which is a multiprotein complex involved in different functions including transcription, nucleotide excision repair (NER), transcription coupled repair, apoptosis and cell cycle regulation. It posses both ssDNA-dependent ATPase and 5′–3′ DNA helicase activities and participate in DNA unwinding during NER and transcription. Several synonymous and nonsynonymous SNPs including those at codons 156 (exon 6 $C > A$, $Arg > Arg$), 312 (exon 10 $G > A$, $Asp > Asn$) and 751 (exon 23 $A > C$, $Lys > Gln$) have been described in *XPD* locus and reported to be associated with tobacco related cancers.[14,15] The *XRCC1* plays an important role in base excision repair (BER) pathway, and interacts with DNA polymerase β, poly ADP-ribose polymerase and DNA ligase III. It also contains a BRCT (BRCA1 COOH terminus) domain, which is characteristic of proteins involved in cell cycle checkpoint functions and this domain can be responsive to DNA damage. Three nonsynonymous polymorphisms at *XRCC1* were detected at codons 194 ($Arg > Trp$, $C > T$), 280 ($Arg > His$, $G > A$), 399 ($Arg > Gln$, $G > A$) and have been associated with presence or absence of cancer risk in breast, stomach, head and neck and lung in different studies.[16–18]

Since both slow and rapid acetylators are susceptible to cancer depending on the carcinogens and tissues, it is highly probable that different bulky DNA adducts are formed in the affected tissues. So, we hypothesize that individuals carrying NAT2 slow or rapid acetylation genotypes and risk genotypes at DNA repair loci will be susceptible to oral cancer and leukoplakia. In a previous case–control study (on 197 leukoplakia and 310 cancer patients and 348 controls) we reported that *XRCC1* variant haplotypes increased the risk of cancer and leukoplakia in mixed tobacco users.[19] In this study, additionally 27 leukoplakia patients and 41 controls were recruited and pooled with the previously collected samples. In the newly pooled samples of 224 leukoplakia and 310 cancer patients and 389 controls, we explored the potential relationships between polymorphisms at *NAT2*, *XRCC1* and *XPD* loci

Grant sponsor: Department of Science and Technology, Government of India.

Nilabja Sikdar's current address is: Genomic Instability Section, Genetics and Molecular Biology Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA.

***Correspondence to:** Human Genetics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India. Fax: +033-2577-3049. E-mail: broy@isical.ac.in

Received 11 July 2006; Accepted after revision 21 November 2006

DOI 10.1002/ijc.22547

Published online 8 February 2007 in Wiley InterScience (www.interscience.wiley.com).

and risk of oral leukoplakia and cancer, independently and in combination of 2 loci.

Materials and methods

Patients, controls and tobacco habit

Unrelated oral leukoplakia ($n = 224$) and cancer ($n = 310$) patients were recruited during 1999–2005 from the R. Ahmed Dental College and Hospital (a primary referral hospital at Kolkata, India) and the department of pathology from the same hospital confirmed the diseases by histopathology. Unrelated individuals ($n = 389$) who came for treatment of dental ailments but without any lesions in oral cavity were recruited from outpatient department of the same hospital as ''controls.'' After getting written consents, all patients and controls were personally interviewed using a questionnaire having information on age, sex, occupation, alcohol consumption, type of tobacco habit, daily tobacco use frequency, duration of habits and economic status. Data related to diagnosis and cellular morphology of the biopsy materials were obtained from the pathology reports.

All subjects in this study reported tobacco habits such as smoking of cigarette and/or *bidi* (a native cigarette-like stick of coarse tobacco hand-rolled in a dry *tembuhurni leaf*) and chewing/dipping of tobacco in different forms.[19] Some patients and controls reported dual habits comprising both smoking and chewing/dipping of tobacco, while the majority had single habit. Lifetime tobacco chewing/dipping exposure was measured in terms of the frequency of chewing/dipping per day multiplied by the duration of habit. This is termed as chewing-year (CY): taking smokeless tobacco once in a day for 1 year = 1 CY and similarly, dose of tobacco smoking was measured as pack-years (PY): 1 packet per day for 1 year = 1 PY (1 pack = 20 cigarettes or 40 *bidies*, since amount of tobacco present in 1 cigarette (750–1000 mg) is similar to that present in 2 *bidies* (850–1050 mg).

Sample collection and processing

About 3–4 ml blood was collected by vein puncture from all patients and controls and stored at −20°C until DNA isolation. Genomic DNA was isolated from whole blood by salt precipitation method.[20] Biopsy materials collected from the patients were used to study histopathology.

Genotyping assays

XRCC1 locus. Most of the genotyping data at 3 polymorphic sites (codons 194 (*Arg* > *Trp*), 280 (*Arg* > *His*) and 399 (*Arg* > *Gln*) on *XRCC1* locus were taken from the previously published paper.[19] But the remaining samples, which were recruited later, were genotyped using same methods.

XPD locus. The *A* > *C* (*Arg* > *Arg*) polymorphism at codon 156 of *XPD* was screened in all samples generating a 644-bp PCR product, digesting with *Tfi*I and electrophoresing in 2% agarose gel.[14] In addition to the *TfiI* site at codon 156, one additional monomorphic *Tfi*I site, producing a 57-bp DNA fragment, served as an internal control for restriction enzyme digestion. Genotypes were determined by banding patterns such as *A*/*A* (474, 113 and 57 bp); *A*/*C* (587, 474, 113 and 57 bp); and *C*/*C* (587, 57 bp).

Polymorphism (*Asp* > *Asn*) at codon 312 of *XPD* was screened digesting the PCR product with *Taq*I for 4 hr at 65°C and electrophoresing in 3% agarose gel. DNA banding patterns were *Asp*/*Asp* = 166, 22 bp, *Asp*/*Asn* = 188, 166, 22 bp and *Asn*/*Asn* = 188 bp.

Polymorphism (*Lys* > *Gln*) at codon 751 of *XPD* was screened digesting the PCR product with *Pst*I and electrophoresing in 3% agarose gel to determine genotypes (*Lys*/*Lys* = 155, 104, 63 bp, *Lys*/*Gln* = 218, 155, 104, 63 bp, *Gln*/*Gln* = 218, 104 bp).[21]

NAT2 locus and acetylation status. Five SNPs at nucleotide positions (np) 341 (T > C), 481 (C > T), 590 (G > A), 803 (A > G) and 857 (G > A) on exon 2 were screened according to the published procedure.[22] Instead of analyzing the genotype data at these 5 polymorphic sites separately, they were expressed as alleles/haplotypes, such as *NAT2*4*, *NAT2*5A*, *NAT2*5B*, *NAT2*5C*, *NAT2*6B*, *NAT2*6C*, *NAT2*7A*, *NAT2*12B*, *NAT2*12C*, *etc.*, depending on the arrangements of nucleotides (*T* > *C*, *C* > *T*, *G* > *A*, *A* > *G*, *G* > *A*) at the above-mentioned 5 polymorphic sites.[8,23] These alleles/haplotypes will provide information of nucleotides present at these polymorphic sites at a glance and allelic combination or genotype will be used to determine the acetylation status of an individual. If an individual had only homozygous wild or variant genotypes at 1, 2, 3, 4 or 5 sites of these 5 polymorphic positions then allelic combination or genotype of that individual could be ascertained easily looking at the genotype data at 5 polymorphic sites of that individual. Again if an individual had only one heterozygous genotype at any one of the above-mentioned 5 polymorphic positions and homozygous wild or variant genotypes at the remaining 4 positions, then also allelic combination or genotype of that individual could be also ascertained easily. In this manner, allelic combinations or genotypes in 50% of the patients and controls were ascertained. But the remaining 50% of the patients and controls had heterozygous genotypes at more than one positions. Allelic combinations or genotypes at *NAT2* locus of these individuals were ascertained with a probabilistic approach such that alleles/haplotypes with high frequencies in overall samples, estimated by HAPLOFREQ software,[24] will also be represented similarly in this bunch of individuals. Then, empirical formula for estimation of frequency of P1P2 genotype in the population is:

$$N(P_1P_2)/(P_1P_2 + Q_1Q_2)$$

where P_1, P_2, Q_1, Q_2 are haplotype/allele frequencies at 2 polymorphic sites and N is the number of individuals having heterozygous genotypes at these 2 positions. Similar methodology was also applied for estimation of allelic combination/genotypes of individuals having heterozygous genotypes at 3, 4 or 5 loci. Although haplotypes and genotypes determined by this estimation method will have some error, but this is the best possible method to know the genotypes of a bunch of individuals who are heterozygous at more than 2 polymorphic sites. Wild type allele (*i.e.*, *NAT2*4*) having wild type nucleotides at all 5 polymorphic sites is known as rapid acetylating allele. Since polymorphisms at 481 and 803 np do not change the acetylation status so variant alleles/haplotypes, having at least one variant nucleotide at one of the remaining 3 polymorphic sites (341, 590 and 857 np), are known as slow acetylating alleles (*e.g.*, *NAT2*5B*, *NAT2*6B*, *NAT2*7A*, *etc*).[8] So, individuals carrying 2 rapid acetylating alleles (such as *NAT2*4*/*NAT2*4* genotype) in the pair of chromosomes are rapid acetylators; carrying 2 slow acetylating alleles (such as *NAT2*5A*/*NAT2*5B* genotype) are slow acetylators; and carrying one slow and one rapid acetylating allele (*i.e.*, *NAT2*4*/*NAT2*5A* genotype) is intermediate acetylators.

Sequencing of PCR products

Few PCR products (8% of total samples) from all loci were resequenced (ABI 3100 Genetic Analyzer; Applied Biosystem, Foster City, CA) to confirm the genotypes determined by PCR-RFLP methods.

Statistical analysis

Age-, sex- and tobacco dose-adjusted risk of oral cancer and leukoplakia was calculated as odds ratios (ORs) with 95% confidence intervals (CIs) for all genotypes in all and stratified patient samples by multiple logistic regression analysis using SPSS statistical package. Chi-square test with Yates' correction, when necessary, was used for comparison of genotype proportions. Frequencies of different alleles/haplotypes at *NAT2*, *XPD* and *XRCC1* loci in all patients and controls were estimated using genotype data at 5, 3 and 3 polymorphic sites, respectively, by the maximum-likelihood method using the expectation maximization algorithm named as HAPLOFREQ.[24] Frequencies of these alleles/haplotypes in cancer, leukoplakia and control populations and subset of the

TABLE I – DEMOGRAPHICS AND TOBACCO EXPOSURES OF PATIENTS AND CONTROLS

Subjects and tobacco habits	Controls [n = 389] (%)	Leukoplakia [n = 224] (%)	p-value (leukoplakia *vs.* control)[1]	Cancer [n = 310] (%)	p value (cancer *vs.* control)[1]
Sex					
Male	302 (78)	196 (87)	0.004	198 (64)	<0.0001
Female	87 (22)	28 (13)		112 (36)	
Age distribution (years)					
Range	25–80	25–75		25–88	
Below 35	55 (14)	29 (13)	NS	14 (5)	<0.001
36–45	110 (28)	81 (36)	0.05	69 (22)	NS
46–55	109 (28)	70 (31)	NS	85 (27)	NS
56–65	78 (20)	34 (15)	NS	102 (33)	<0.001
Above 65	37 (10)	10 (5)	0.03	40 (13)	NS
Age (years)					
Mean ± SD	49 ± 11.9	47 ± 10.3	0.03	55 ± 11.0	< 0.0001
Exclusive smoking habit					
Exclusive smokers	145 (37)	133 (60)	0.001	53 (17)	<0.0001
Lifetime smoking range (PY)	2–90	2–90		2–75	
Mean smoking dose ± SD (PY)	31 ± 18	24 ± 15.5	0.0006	32 ± 13.6	NS
Exclusive smokeless tobacco habit					
Exclusive smokeless tobacco users	169 (44)	32 (14)	<0.0001	176 (57)	0.001
Lifetime smokeless tobacco using range (CY)	12–925	12–420		4–1250	
Mean smokeless tobacco dose ± SD (CY)	183 ± 145	64 ± 174	0.0001	182 ± 162	NS
Mixed habituals					
Smoking as well as Smokeless tobacco using habit	75 (19)	59 (26)	0.05	81 (26)	0.04
Lifetime smoking range (PY)	2–90	2–80		2–120	
Mean smoking dose ± SD (PY)	22 ± 14	26 ± 19.08	NS	25 ± 15.83	NS
Lifetime smokeless tobacco using range (CY)	10–600	10–600		10–640	
Mean smokeless tobacco dose ± SD (CY)	100 ± 91	58 ± 103.18	0.01	106 ± 88	NS

All patients and controls had tobacco exposure. NS, nonsignificant at 5% level. Exclusive smokers and smokeless tobacco users have only smoking and tobacco chewing/dipping habits, respectively. Mixed habituals have smoking as well as smokeless tobacco habits.
[1]Data of patients were compared with those of controls.

samples were compared to know whether any allele or haplotype is associated with the risk of cancer or/and leukoplakia.

Results

It was evident from the interview that most of the each population (>95%) belonged to same ethnic population, *Bengalee*. Most of the patients and controls had low income (family income < $100 USD per month) and this is one of the reasons for which they visited government hospital for treatment. Majority of the male patients and controls had occupations in diverse areas such as agriculture, small industry, car driving, private sector office, small business, *etc*, whereas most of the females were housewives and doing only household jobs. None of the patients and controls was exposed to specific toxic agents except tobacco. Distribution of demographic characteristics and tobacco habits of patient and control populations are summarized in Table I. About 85% of smokers had habits of both cigarettes and *bidis*, so it was not possible to analyze *bidi* and cigarette smokers separately. In patient and control groups, only few (<5%) had occasional alcohol-drinking habit, so, alcohol consumption was also not considered in the analyses. All cancer and leukoplakia patients were incident cases and none of the controls had family history of cancer but ~3% of the oral cancer patients reported that there was death from cancer at different sites in the first or second-degree relatives. Few of these patients were also not sure about the diagnosis of their relatives. So, these patients were not excluded from data analysis.

Leukoplakia mostly affected buccal mucosa and commissure area (65%), buccal mucosa and alveolar sulcus (20%) and other sites including lip, tongue, *etc.* (15%). Most of the patients suffered from ulcerative lesion (61%), followed by homogeneous (36%) and nodular (3%) types of leukoplakia. Cancer affected sites were buccal mucosa and alveolar sulcus (52%), lip (15%), tongue (12%), buccal sulcus (11%) and retromolar area (10%). Histopathologically, all malignancies were diagnosed as squamous cell carcinoma (SCC) and morphologically they were classified as well (65%), moderately (17%) and poorly (18%) differentiated SCC.

Within each of the 3 groups (control, leukoplakia and cancer) we tested for Hardy–Weinberg equilibrium at 11 SNPs on 3 loci (*XRCC1*, *XPD* and *NAT2*). Some significant p-values, uncorrected for multiple comparisons, were noted (p = 0.02 at codon 194 on *XRCC1* among controls; p = 0.05, 0.05 and 0.02 at codons 341, 481 and 803 on *NAT2* among leukoplakia, respectively). However, if Bonferroni's multiple comparison correction (to take into account 11 tests performed within each group) is used, none of these p-values remained significant because the corrected level of significance is 0.0045. This indicates no significant departure from Hardy–Weinberg equilibrium at any SNP locus in any of the 3 groups.

Few genotypes (8% of total samples), at all loci, determined by sequencing method were observed to be identical to those determined by PCR-RFLP methods. This cross checking of genotypes by 2 different methods was done to be sure that genotypes of all samples were correctly determined by PCR-RFLP methods. Neither the heterozygotes nor variant genotypes at 3 polymorphic sites on *XPD* and *XRCC1* loci increased the risks of diseases in these populations (Table II). The *NAT2* genotypes, which were commonly present in all 3 populations, were pooled as rapid, intermediate and slow acetylators (Table III). In the control population, 4, 36 and 60% were rapid, intermediate and slow acetylators, respectively, and similarly 7, 34 and 59% of leukoplakia patients and 4, 32 and 64% of cancer patients were rapid, intermediate and slow acetylators, respectively. None of the *NAT2* genotypes and acetylation status was associated with the risk of leukoplakia and cancer. Few *NAT2* heterozygote and variant genotypes were present only in control or leukoplakia or cancer population. If genotypes of these individuals (n = 65; 40 controls and 15 leukoplakia and 10 cancer patients) were categorized as rapid, intermediate and slow acetylators, and pooled with acetylators in 3 populations, as shown in Table III, to compare, then also, acetylation status did not modulate the risk of leukoplakia and cancer (data not shown).

TABLE II – DISTRIBUTION OF GENOTYPES AT 3 POLYMORPHIC SITES ON *XPD* AND *XRCC1* LOCI AMONG LEUKOPLAKIA, CANCER AND CONTROL POPULATIONS

	Control	Leukoplakia	Adjusted OR, 95% CI (control *vs.* leukoplakia)[1]	Cancer	Adjusted OR, 95% CI (control *vs.* cancer)[1]
Genotypes at XPD					
Codon 156	$n = 388$ (%)	$n = 220$ (%)		$n = 308$ (%)	
C/C	124 (32)	73 (33)	Ref.	88 (29)	Ref.
A/C	191 (49)	103 (46)	0.9 (0.8–1.1)	156 (50)	1.0 (0.9–1.0)
A/A	73 (19)	44 (21)	0.9 (0.9–1.1)	64 (21)	1.0 (0.9–1.0)
Codon 312	$n = 387$ (%)	$n = 224$ (%)		$n = 305$ (%)	
Asp/Asp	205 (53)	117 (52)	Ref.	152 (50)	Ref.
Asp/Asn	146 (38)	89 (40)	0.9 (0.9–1.1)	119 (39)	1.0 (0.9–1.0)
Asn/Asn	36 (9)	18 (8)	0.9 (0.7–1.2)	34 (11)	1.0 (0.9–1.0)
Codon 751	$n = 388$ (%)	$n = 224$ (%)		$n = 309$ (%)	
Lys/Lys	190 (49)	105 (47)	Ref.	158 (51)	Ref.
Gln/Lys	158 (41)	98 (44)	0.9 (0.9–1.1)	125 (41)	1.0 (0.9–2.3)
Gln/Gln	40 (10)	21(9)	0.9 (0.9–1.1)	26 (8)	1.0 (0.9–2.3)
Genotypes at XRCC1					
Codon 194	$n = 387$ (%)	$n = 224$ (%)		$n = 309$ (%)	
Arg/Arg	317 (82)	177 (79)	Ref.	248 (80)	Ref.
Arg/Trp	62 (16)	43 (19)	0.9 (0.9–1.1)	58 (19)	0.9 (0.9–1.0)
Trp/Trp	8 (2)	4 (2)	0.9 (0.8–1.2)	3 (1)	0.9 (0.9–1.0)
Codon 280	$n = 387$ (%)	$n = 220$ (%)		$n = 307$ (%)	
Arg/Arg	297 (77)	160 (73)	Ref.	225 (73)	Ref.
Arg/His	87 (22)	58 (26)	1.0 (0.9–1.0)	79 (26)	1.0 (0.9–1.0)
His/His	3 (1)	2 (1)	1.0 (0.9–1.0)	3 (1)	1.0 (0.9–1.0)
Codon 399	$n = 385$ (%)	$n = 224$ (%)		$n = 309$ (%)	
Arg/Arg	170 (44)	100 (45)	Ref.	134 (43)	Ref.
Arg/Gln	179 (47)	95 (42)	0.8 (0.6–1.3)	143 (46)	0.9 (0.9–1.0)
Gln/Gln	36 (9)	29 (13)	0.9 (0.9–1.0)	32 (11)	0.9 (0.9–1.0)

A/A (Arg/Arg), Asn/Asn, Gln/Gln at XPD; Trp/Trp, His/His and Gln/Gln at *XRCC1* are variant genotypes. Variant allele frequencies are *Arg*: 0.43, *Asn*: 0.28, *Gln*: 0.31; *Trp*: 0.10, *His*: 0.12 and *Gln*: 0.33, respectively. Few patients and controls could not be genotyped after repeated attempts, so total sample sizes (*N*) became different at different codons.

[1]Adjusted for age, sex and tobacco dose.

Polymorphisms at 3 sites on *XPD* locus were not associated with increased risk of cancer in overall population (Table II) but variant genotype (*Asn/Asn*) at codon 312 of *XPD* locus was associated with increased risk of cancer (OR = 1.9, 95% CI = 1.2–2.9) among the pooled samples of intermediate and rapid acetylators (Table IV). Major/wild genotype (*Lys/Lys*) at codon 751 of *XPD* locus was associated with increased risk of cancer (OR = 1.6, 95% CI = 1.1–2.3) among slow acetylators. In the previous[19] as well as present study, variant genotype (*Gln/Gln*) at codon 399 of *XRCC1* locus did not modulate the risk of leukoplakia and cancer in overall population (Table II) but this genotype was also found to be associated with increased risk of leukoplakia (OR = 4.2, 95% CI = 1.2–15.0) among slow acetylators (Table V). Frequency of simultaneous presence of *NAT2* slow and *XRCC1 Gln* alleles in an individual is more prevalent in the leukoplakia patients than controls ($p = 0.05$, legend in Table V).

Since *XPD* locus was genotyped at 3 polymorphic sites so we also estimated frequencies of haplotypes in overall as well as stratified samples such as rapid, intermediate and slow acetylators separately. *XPD* haplotypes containing variation at 1 or 2 nucleotide position/s did not modulate the risk of leukoplakia and cancer in overall and rapid, intermediate (data not shown) and slow acetylators (Table VI). It is interesting to note that frequency of an *XPD* haplotype (A^{156}-G^{312}-A^{751}), containing one variant allele *A* at codon 156, is more than wild haplotype (C^{156}-G^{312}-A^{751}) in patients and controls (37% *vs.* 25% in controls, respectively). Frequency of another *XPD* allele (C^{156}-A^{312}-C^{751}), containing 2 variant alleles at codons 312 and 751, is similar to that of wild haplotype (C^{156}-G^{312}-A^{751}) in patients and controls (25% *vs.* 25% in controls, respectively). Frequencies of *XRCC1* haplotypes were also estimated from genotypes data in overall as well as stratified samples such as rapid, intermediate and slow acetylators, separately. It was observed that variant haplotype (C^{194}-G^{280}-A^{399}) was associated with increased risk of leukoplakia in slow acetylators (OR = 1.7, 95% CI = 1.2–2.4) and leukoplakia and cancer in mixed habituals among slow acetylators (OR = 3.1, 95% CI = 1.4–7.1, OR = 2.4, 95% CI = 1.1–5.4, respectively, Table VII).

Discussion

The controls that were recruited from the same dental hospital had dental ailments such as dental carries and gingivitis, but we do not have any prior knowledge whether polymorphisms in the studied loci had any effects on these dental ailments. So, the individuals were not truly healthy controls. Moreover, male/female distribution, age and number of smokers and chewers were not similar in controls and leukoplakia and cancer patients. It would have been better if we could have recruited 2 separate sets of controls for leukoplakia and cancer patients because males and smokers are more in leukoplakia patients than controls whereas age of the cancer patients is more than that of controls. So, adjustments for age, sex, tobacco doses were done whenever required since there were significant differences in the collected data in relation to age, sex, smoking (PY) and chewing dose (CY) among 3 studied groups (Table I). In India, males use both smoking and smokeless tobacco whereas females use mostly smokeless tobacco. Although smokers and smokeless tobacco users are equally affected by leukoplakia and cancer but it was observed that comparatively more male patients as well as smokers were present in leukoplakia population (Table I). One of the reasons might be that leukoplakia is not life threatening initially, so females (mostly housewives from low income families) neglected the treatment because of several procedural steps required in the hospital. As a result, females who were mostly smokeless tobacco users were less represented in leukoplakia population.

Polymorphisms at 3 polymorphic sites of *XPD* locus did not modulate the risk of cancer or leukoplakia in this population (Table II). Other studies also reported both presence and absence of association between polymorphisms at *XPD* and risk of head and neck/oral cancer in different populations.[14,25,26] Distribution of

TABLE III – DISTRIBUTION OF *NAT2* GENOTYPES AND ACETYLATION STATUS AMONG LEUKOPLAKIA, CANCER AND CONTROL POPULATIONS

Acetylation status	Genotypes	Control (n = 342) (%)	Leukoplakia (n = 207) (%)	OR, 95% CI (control *vs.* leukoplakia)[1]	Cancer (n = 297) (%)	OR, 95% CI (control *vs.* cancer)[1]
Rapid	*4/*4	15 (4)	15 (7)	Ref.	13 (4)	Ref.
	Total	15 (4)	15 (7)		13 (4)	
Intermediate	*4/*5B	40 (12)	24 (12)	0.6 (0.2–1.6)	39 (13)	1.1 (0.4–2.9)
	*4/*5C	5 (1)	4 (2)		3 (1)	
	*4/*6B	56 (16)	25 (12)	0.4 (0.2–1.1)	36 (12)	0.7 (0.3–1.9)
	*4/*7A	11 (3)	10 (5)		6 (2)	
	*5C/*12C	3 (1)	3 (1)		8 (3)	
	*6B/*12B	7 (2)	4 (2)		2 (1)	
	Total	122 (36)	70 (34)	0.6 (0.3–1.3)[2]	94 (32)	0.7 (0.3–1.7)[2]
Slow	*5A/*6B	3 (1)	3 (1)		7 (2)	
	*5B/*5C	5 (1)	5 (2)		8 (3)	
	*5B/*6B	67 (20)	31 (15)	0.4 (0.2–1.2)	59 (20)	1.0 (0.4–2.5)
	*5B/*7A	12 (4)	8 (4)		13 (5)	
	*5C/*6B	13 (4)	6 (3)		8 (3)	
	*5B/*5B	39 (11)	20 (10)	0.5 (0.2–1.4)	30 (10)	0.9 (0.3–2.3)
	*6B/*6B	42 (12)	31 (15)	0.7 (0.3–1.9)	37 (12)	1.0 (0.4–2.6)
	*6B/*7A	16 (5)	15 (7)	0.9 (0.3–2.9)	21 (7)	1.5 (0.5–4.5)
	*6B/*6C	5 (1)	1 (1)		4 (1)	
	*7A/*7A	3 (1)	2 (1)		3 (1)	
	Total	205 (60)	122 (59)	0.8 (0.5–1.2)[2]	190 (64)	1.0 (0.7–1.5)[2]

Other genotypes such as *4/*12A, *4/*12C, *4/*5A, *5A/*5B, *5C/*5C, *5C/*7A, *5A/*11, *5B/*12C, *5C/*12A, *7A/*5F, *7A/*11 and *7A/*12A were also present either in leukoplakia or cancer or control group but with frequency $<1\%$. These genotypes from 65 samples (40 control, 15 leukoplakia and 10 cancer) were not included in this table. When these genotypes were pooled with the respective acetylation genotypes as shown in this table and compared among 3 populations, then there was no significant change in the risk of diseases. Acetylation status of 7 controls, 2 leukoplakia and 3 cancer patients remained undefined because of failure of genotyping at few polymorphic sites and absence of data at SNPs other than our studied SNPs.
[1]Crude.–[2]Age, sex and tobacco dose adjusted.

TABLE IV – DISTRIBUTION OF GENOTYPES AT *XPD* LOCUS AMONG DIFFERENT *NAT2* ACETYLATORS OF LEUKOPLAKIA, CANCER AND CONTROL POPULATIONS

Acetylation status	Genotypes at *XPD*	Control	Leukoplakia	Adjusted OR, 95% CI (control *vs.* leukoplakia)[1]	Cancer	Adjusted OR, 95% CI (control *vs.* cancer)[1]
Slow	Codon 156	n = 219 (%)	n = 134 (%)		n = 193 (%)	
	C/C	70 (32)	43 (32)	Ref.	54 (28)	Ref.
	A/C	111 (51)	65 (49)	1.0 (0.6–1.7)	100 (52)	1.1 (0.8–1.3)
	A/A	38 (17)	26 (19)	1.1 (0.8–1.5)	39 (20)	1.2 (0.9–1.7)
Intermediate + rapid	Codon 312	n = 143 + 20 (%)	n = 73 + 15 (%)		n = 96 + 16 (%)	
	Asp/Asp	80 + 10 (55)	35 + 8 (49)	Ref.	46 + 5 (46)	Ref.
	Asp/Asn	56 + 8 (39)	31 + 6 (42)	1.3 (0.8–2.3)	36 + 9 (40)	1.1 (0.7–1.9)
	Asn/Asn	7 + 2 (6)	7 + 1 (9)	1.3 (0.8–2.8)	14 + 2 (14)	1.9 (1.2–2.9)
Slow	Codon 751	n = 219 (%)	n = 134 (%)		n = 194 (%)	
	Gln/Gln	28 (13)	11 (8)	Ref.	13 (7)	Ref.
	Gln/Lys	93 (42)	57 (43)	1.6 (0.7–3.6)	76 (39)	1.6 (0.8–3.4)
	Lys/Lys	98 (45)	66 (49)	1.3 (0.9–1.9)	105 (54)	1.6 (1.1–2.3)

Genotypes at codon 312 were compared in pooled samples of rapid and intermediate acetylators, since rapid acetylators are few in patients and controls. The numbers of rapid, intermediate and slow acetylators in Table III and this table will not match since genotypes, those were present in all 3 populations were only shown in Table III, but all different genotypes at *NAT2* were considered in this Table. Genotype of one cancer patient remained undefined at codon 156.
[1]Adjusted for age, sex and tobacco dose.

rapid (4%), intermediate (36%) and slow acetylators (60%) in this control population (Table III) were observed to be similar to those (6, 37 and 57%, respectively) in Caucasian population[11] but different from those in South East Asians and Eskimos (10–30% and 5% slow acetylators, respectively).[7] None of the *NAT2* genotypes and acetylation status was associated with increased risk of cancer or leukoplakia in overall population (Table III) although only a few studies on Japanese and Caucasian populations have shown increased risk of head and neck cancer in a subset of slow acetylators.[9–11] A few slow acetylation genotypes also acted as protective or risk factors in laryngeal cancer.[7,13] Independently, *NAT2* acetylation status or *XPD* genotypes were not associated with increased risk of leukoplakia and cancer. So, we hypothesized that combination of NAT2 acetylation status and other genetic factor/s, such as DNA repair loci, may be involved in disease susceptibility. Intermediate carcinogens formed by NAT2 enzyme may be an active ingredient for DNA adducts formation, which should be removed by DNA repair enzymes involved in removal of bulky DNA adducts. Bulky DNA adducts could be repaired by XPD and XRCC1 by NER[14,15] and BER[27] mechanisms, respectively. So, instead of considering the acetylation status alone genotypes/haplotypes at *XPD* and *XRCC1* were also considered, in combination, to determine the risk of the diseases. The major genotype (*Lys/Lys*) at codon 751 of *XPD* locus increased the risk of cancer (OR = 1.6, 95% CI = 1.1–2.3) among slow acetylators (Table IV). Similar observation was also reported in laryngeal[13] and bladder cancers.[28] It has been shown that rate of *N*-acetylation to detoxify the aromatic amines is less in slow acetylators so these carcinogens are accumulated in the affected tissue.[8] The codon 751 polymorphism, which may change the activity of the enzyme, is located in COOH terminal domain of XPD. This region of the protein possesses helicase activity and deletion in this region causes reduced XPD DNA helicase activity.[28] This may account for the observed association

TABLE V – DISTRIBUTION OF *XRCC1* GENOTYPES AT 3 CODONS IN SLOW, INTERMEDIATE AND RAPID ACETYLATORS OF LEUKOPLAKIA, CANCER AND CONTROL POPULATIONS

Genotypes at	Acetylation status at *NAT2*	Control	Leukoplakia	Adjusted OR, 95% CI (control *vs.* leukoplakia)[1]	Cancer	Adjusted OR, 95% CI (control *vs.* cancer)[1]
Codon 194		$n = 69$ (%)	$n = 46$ (%)		$n = 61$ (%)	
(Arg/Trp + Trp/Trp)	Intermediate + rapid	20 + 4 (35)	19 + 3 (48)	Ref.	16 + 6 (36)	Ref.
(Heterozygote + Variant)	Slow	45 (65)	24 (52)	0.7 (0.4–1.3)	39 (64)	0.9 (0.5–1.5)
Codon 280		$n = 89$ (%)	$n = 59$ (%)		$n = 81$ (%)	
(Arg/His + His/His)	Intermediate + rapid	39 + 3 (47)	27 + 2 (49)	Ref.	28 + 4 (39)	Ref.
(Heterozygote + Variant)	Slow	47 (53)	30 (51)	1.0 (0.6–1.9)	49 (61)	1.1 (0.6–1.9)
Codon 399		$n = 36$ (%)	$n = 29$ (%)		$n = 32$ (%)	
Gln/Gln (Variant)	Intermediate + rapid	18 + 1 (53)	5 + 1 (20)	Ref.	14 + 0 (44)	Ref.
	Slow	17 (47)[2]	23 (80)[2]	4.2 (1.2–15.0)	18 (56)	1.4 (0.5–4.1)

Frequencies of variants at codons 194 and 280 were few so heterozygotes, which contain one variant allele, were pooled with variants to compare in 3 populations. Since the rapid acetylators are few in each genotype group so they were pooled with intermediate acetylators as one group to compare in 3 populations. Total numbers of heterozygotes and variants at codons 194 and 280 differ from those present in Table II because of lack of acetylation status data of few individuals.

[1]Adjusted for age, sex, and tobacco doses.–[2]Frequency of slow acetylating allele in leukoplakia patients with *Gln/Gln* genotype is significantly more than the respective controls ($p = 0.05$).

TABLE VI – ESTIMATED FREQUENCIES OF *XPD* HAPLOTYPES IN *NAT2* SLOW ACETYLATORS AND STRATIFIED SAMPLES AMONG THEM PRESENT IN LEUKOPLAKIA, CANCER AND CONTROL POPULATIONS

Slow acetylators and stratified samples among them	Haplotypes		Control (chromosome number) *n* (%)	Leukoplakia (chromosome number) *n* (%)	OR, 95% CI (control *vs.* leukoplakia)[1]	Cancer (chromosome number) *n* (%)	OR, 95% CI (control *vs.* cancer)[1]
Slow acetylators	Wild	*C-G-A*	107 (25)	71 (26)	Ref.	94 (25)	Ref.
	Variants with one nucleotide	*C-A-A*	20 (5)	10 (4)		20 (6)	
		A-G-A	159 (37)	107 (40)	1.0 (0.7–1.5)	163 (43)	1.1 (0.8–1.6)
		C-G-C	16 (3)	9 (3)		3 (1)	
	Variants with two nucleotides	*C-A-C*	107 (25)	60 (23)	0.8 (0.5–1.3)	89 (23)	0.9 (0.6–1.4)
		A-G-C	21 (5)	9 (4)		9 (2)	
		Total[2]	430	266		378	
Exclusive smokers	Wild	*C-G-A*	37 (23)	48 (32)	Ref.	15 (21)	Ref.
	Variants with one nucleotide	*C-A-A*	9 (6)	4 (3)		4 (6)	
		A-G-A	61 (39)	54 (36)	0.7 (0.4–1.2)	32 (46)	1.3 (0.6–2.9)
		C-G-C	5 (3)	7 (4)		1 (2)	
	Variants with two nucleotides	*C-A-C*	39 (25)	32 (21)	0.6 (0.3–1.2)	18 (26)	1.1 (0.5–2.8)
		A-G-C	5 (4)	7 (4)		0	
		Total	156	152		70	
Mixed habituals	Wild	*C-G-A*	16 (20)	17 (21)	Ref.	22 (24)	Ref.
	Variants with one nucleotide	*C-A-A*	5 (6)	2 (3)		4 (5)	
		A-G-A	31 (38)	40 (50)	1.2 (0.5–3.1)	41 (47)	0.9 (0.4–2.3)
		C-G-C	3 (3)	0		2 (3)	
	Variants with two nucleotides	*C-A-C*	20 (25)	18 (22)	0.8 (0.3–2.4)	16 (18)	0.6 (0.2–1.6
		A-G-C	7 (8)	3 (4)		3 (3)	
		Total	82	80		88	
Exclusive smokeless tobacco users	Wild	*C-G-A*	53 (28)	6 (18)	Ref.	56 (25)	Ref.
	Variants with one nucleotide	*C-A-A*	6 (3)	3 (9)		11 (5)	
		A-G-A	66 (35)	13 (38)	1.7 (0.6–5.5)	92 (42)	1.3 (0.8–2.2)
		C-G-C	8 (4)	1 (3)		1 (1)	
	Variants with two nucleotides	*C-A-C*	47 (25)	11 (32)	2.0 (0.6–6.8)	55 (25)	1.1 (0.6–1.9)
		A-G-C	10 (5)	0		5 (2)	
		Total	190	34		220	

Genotypes at codons *156 (C/A)*, *312 (G/A)* and *751 (A/C)* on *XPD* in slow acetylators and stratified samples in 3 populations were used to estimate the haplotype frequencies. Major variant haplotypes were compared in 3 populations.

[1]Crude OR.–[2]Few haplotypes could not be determined because of lack of genotype data.

between risk of cancer and polymorphism at this position among slow acetylators. Additionally, polymorphism at codon 751 of *XPD* (*Lys* > *Gln*) is located at ~50 bp upstream from poly (A) signal and thus may affect XPD protein expression or function. Therefore, individuals with the *Lys/Lys* genotype might have lesser DNA repair capacity, because of less XPD expression/activity than those with the *Gln/Gln* genotype.[14,15] Variant genotype (*Asn/Asn*) at codon 312 of *XPD* locus was also associated with increased risk of cancer among rapid and intermediate acetylators (OR = 1.9, 95% CI = 1.2–2.9) (Table IV). Rapid acetylation might also be involved in the activation of carcinogens, since hydroxylamines and hydroxamic acids of heterocyclic amines can be further activated by *NAT2 via O*-acetylation.[28,29] So, rapid NAT2 acetylation might result in *O*-acetylation of heterocyclic amines, which might lead to DNA adduct formation. As a result, NAT2 rapid and intermediate acetylators with variant genotype (*Asn/Asn*) at codon 312 of *XPD* locus had more risk of cancer (Table IV). This observation is consistent with the report that suggests that *Asn/Asn* genotype at *XPD* locus is associated with low DNA repair capacity.[30]

Since haplotype determines amino acid sequence and, hence, the activity of a protein, so it is important to compare the haplotype frequencies in overall and stratified samples. Unlike reports on breast cancer[31] and basal cell carcinoma patients[32] none of the variant *XPD* haplotypes could modulate the risk of oral leukoplakia and cancer in overall samples and *NAT2* rapid, intermediate (data not shown) and slow acetylators (Table VI). It is interesting to note that frequency of the haplotype (A^{156}-G^{312}-A^{751}) contain-

TABLE VII – ESTIMATED FREQUENCIES OF *XRCC1* HAPLOTYPES IN *NAT2* SLOW ACETYLATORS AND STRATIFIED SAMPLES AMONG THEM PRESENT IN LEUKOPLAKIA, CANCER AND CONTROL POPULATIONS

Slow acetylators and stratified samples among them	Haplotypes		Control (chromosome number) [*n* (%)]	Leukoplakia (chromosome number) [*n* (%)]	OR, 95% CI (control *vs.* leukoplakia)[1]	Cancer (chromosome number) [*N* (%)]	OR, 95% CI (control *vs.* cancer)[1]
Slow acetylators	Wild	C-G-G	219 (52)	102 (39)	Ref	174 (46)	Ref
	Variants with one nucleotide	*C-G-A*	116 (27)	103 (39)	1.7 (1.2–2.4)	118 (31)	1.2 (0.9–1.6)
		C-A-G	50 (12)	32 (12)		51 (14)	
		T-G-G	39 (9)	25 (10)		33 (9)	
		Total[2]	424	262		376	
Exclusive smokers	Wild	C-G-G	77 (50)	60 (41)	Ref	32 (46)	Ref
	Variants with one nucleotide	*C-G-A*	49 (32)	61 (41)	1.5 (0.9–2.4)	19 (27)	0.8 (0.4–1.5)
		C-A-G	14 (9)	12 (8)		12 (17)	
		T-G-G	14 (9)	15 (10)		7 (10)	
		Total	154	148		70	
Mixed Habituals	Wild	C-G-G	55 (67)	28 (35)	Ref	38 (42)	Ref
	Variants with one nucleotide	*C-G-A*	18 (22)	29 (36)	3.1 (1.4–7.1)	31 (34)	2.4 (1.1–5.4)
		C-A-G	6 (7)	14 (18)		11 (12)	
		T-G-G	3 (4)	9 (11)		10 (11)	
		Total	82	80		90	
Exclusive smokeless tobacco users	Wild	C-G-G	88 (47)	16 (47)	Ref	104 (48)	Ref
	Variants with one nucleotide	*C-G-A*	49 (26)	12 (35)	1.5 (0.6–3.5)	70 (32)	1.3 (0.8–2.1)
		C-A-G	29 (15)	6 (18)		26 (12)	
		T-G-G	22 (12)	0 (0)		16 (7)	
		Total	188	34		216	

[1]Crude OR; genotype data at codon *194* (*C/T*), *280* (*G/A*) and *399* (*G/A*) on *XRCC1* among slow acetylators and stratified samples among them in 3 populations were used to estimate the haplotype frequencies.–[2]Few haplotypes could not be determined because of lack of genotype data, so ''*n*'' differs from Table VI. Haplotypes containing more than one variant allele were either absent or less frequent (<2%) and, therefore, were not considered for comparison. The wild (*C-G-G*) and major variant (*C-G-A*) haplotypes were compared among 3 populations.

ing the variant allele *A* at codon 156 was higher compared with the wild haplotype (C^{156}-G^{312}-A^{751}) among slow acetylators of controls (37% and 25%, respectively). Again, frequency of the haplotype (C^{156}-A^{312}-C^{751}), containing variant alleles at both codons 312 and 751, was equal to that of the wild haplotype (C^{156}-G^{312}-A^{751}) among slow acetylators of controls (25% in each case) (Table VI). Similar phenomenon was also observed in overall control population (data not shown). To explain this overrepresentation of variant haplotype, *XPD* genotypes at codons 156, 312 and 751 polymorphic sites among *NAT2* slow acetylators were analyzed for pair-wise linkage disequilibrium using Haploview (http://www.broad.mit.edu/mpg/haploview) and Arlequin Version 2.0 (http://anthro.unige.ch/software/arlequin). It was observed that the allele combinations (A^{156},G^{312}), (G^{312}-A^{751}) and (A^{156}-A^{751}) of *XPD* were in strong positive linkage disequilibrium (r^2 = 0.27, $p < 0.0001$; r^2 = 0.45, $p < 0.0001$ and r^2 = 0.16, $p < 0.0001$, respectively). Because of positive linkage disequilibrium, frequency of variant haplotype (A^{156}-G^{312}-A^{751}) is overrepresented in the samples. Similar observations have also been reported in other studies on populations from Germany,[31] United Kingdom[32] and Poland.[33]

Patients and controls carrying variant *Gln/Gln* genotype at codon 399 of *XRCC1* were further stratified into slow, rapid and intermediate acetylators and it was observed that frequency of slow acetylating allele in leukoplakia patients is significantly more than that in controls ($p = 0.05$, legend in Table V). To compare the frequencies of slow acetylators carrying *Gln/Gln* genotype, rapid and intermediate acetylators (which are few in number in leukoplakia group) were pooled as one group. Then, it was observed that *Gln/Gln* genotype was associated with risk of leukoplakia (OR = 4.2, 95% CI = 1.2–15.0) among slow acetylators (Table V). But we like to mention that sample sizes were few in intermediate and rapid acetylators (5 and 1, respectively) in leukoplakia group so significance of this result lies on the reproducibility of the same result with larger sample sizes. The *G* to *A* transition at *XRCC1* codon 399 results in change from *Arg* to *Gln* in the *XRCC1* BRCT domain that interacts with poly ADP ribose phosphorylase. So, variant amino acid (*i.e.*, *Gln*) at 399 of XRCC1 protein may be less efficient in DNA repair.[34] Variant genotype at codon 399 of *XRCC1* also increased the risk of oral SCC in different populations.[16,18] Since aromatic amines are less detoxified in NAT2 slow acetylators, so these acetylators with less active *XRCC1 Gln/Gln* genotype might become susceptible to leukoplakia. Although NER is mainly involved in the removal of bulky DNA adducts, these adducts are also repaired by BER mechanism. DNA bulky adducts could destabilize the *N*-glycosyl bonds, thus it can induce rapid depurination or depyrimidation of adducted bases for BER mechanism.[27] Haplotype frequencies at *XRCC1* were also estimated from the genotype data at 3 polymorphic sites. In the previous study (controls = 348, leukoplakia = 197 and cancer = 310),[19] we reported that variant haplotypes at *XRCC1* increased the risk of leukoplakia in overall population as well as risk of leukoplakia and cancer in only mixed tobacco users of this population. In this study, we recruited additionally 41 controls and 27 leukoplakia samples and pooled data with the previously collected samples. After addition of these controls and leukoplakia patients, similar results were also observed (data not shown). In this study, frequencies of haplotypes at *XRCC1 locus* were estimated among the slow acetylators of patient and control populations. It was observed that variant haplotype (C^{194}-G^{280}-A^{399}) was associated with increased risk leukoplakia among slow acetylators (OR = 1.7, 95% CI = 1.2–2.4) and leukoplakia and cancer in mixed habituals among slow acetylators (OR = 3.1, 95% CI = 1.4–7.1, OR = 2.4, 95% CI = 1.1–5.4, respectively, Table VII). Mixed tobacco users had habit of both smoking and smokeless tobacco (Table I). Among slow acetylators, mean smoking doses of mixed habituals in control, leukoplakia and cancer populations were similar to or more than those of exclusive smokers in overall control, leukoplakia and cancer populations (data not shown). Additionally, mixed habituals had smokeless tobacco exposure. Among slow acetylators, mean smokeless tobacco doses in mixed habituals of control, leukoplakia and cancer populations were also similar to or less than those of exclusive smokeless tobacco users in overall control, leukoplakia and cancer populations (data not shown). Although it has not been ascertained but it could be assumed that mixed tobacco habituals were exposed to more tobacco carcinogens than exclusive smokers or smokeless tobacco users among the slow acetylators. Since aromatic amines are less detoxified in slow acetylators, so more DNA adducts were formed. As a result, slow acetylators with variant *XRCC1* haplotypes became susceptible to leukoplakia and cancer since variant alleles at *XRCC1* are less effective in DNA

repair.[34] Apart from *XPD* and *XRCC1* loci, polymorphisms at other 2 DNA repair loci (*Thr* > *Met* at codon 241 on *XRCC3* and *Ser* > *Cys* at codon 326 on *OGG1*), which are involved in the repair of DNA double-strand breaks and removal of 8-oxoguanine, respectively, were also analyzed in overall samples and different acetylators. But none of these polymorphisms was associated with the risks of leukoplakia and cancer in overall and stratified samples (data not shown). We had also analyzed polymorphism data at *GSTM3* locus[19] in slow, intermediate and rapid acetylators, but risks of leukoplakia and cancer were not modulated by *GSTM3* polymorphism in different acetylators (data not shown).

Acetylation status did not modulate the risk of oral cancer and leukoplakia in overall samples but variant *XRCC1* genotype at codon 399 (*Gln*/*Gln*, *i.e.*, *A*/*A*) and haplotype (C^{194}-G^{280}-A^{399}), both containing A^{399} allele either in diploid and haploid form, respectively, increased the risk of leukoplakia among slow acetylators. *XRCC1* haplotype (C^{194}-G^{280}-A^{399}) was also associated with increased risk of cancer and leukoplakia in a subset of slow acetylators. Again, wild (*Lys*/*Lys*) and variant (*Asn*/*Asn*) genotypes at *XPD* could also increase the risk of cancer among slow and intermediate and rapid acetylators, respectively. So, it is concluded that combination of 2 loci, working in carcinogen metabolism and DNA repair pathways, played important roles to enhance the risk of oral cancer and leukoplakia. To explain this association at molecular level, expression profiles and biological functions of these variant genotypes and haplotypes should be studied in the similar context. We like to mention few limitations in our study. First, this is a hospital-based case–control study. The controls had dental ailments such as carries and gingivitis, so selection bias may occur and they may not be representative of the general population. Second, although age, sex and tobacco dose were adjusted whenever it required, but second hand smoke exposure, diet and other environmental exposures were not adjusted in our logistic regression models because of incomplete and missing information. Third, although the sample sizes of this case–control study were large, in few observations samples sizes became low after stratification. So, it is necessary to repeat similar study with large case–control samples.

Acknowledgements

Authors are grateful for the cooperation received from all patients and controls. The authors thank Prof. Ranjan Rashmi Paul, presently at Guru Nanak Institute of Dental Sciences and Research, Panihati, Kolkata 700114, for full support during sample collection. We appreciate the suggestions of Prof. P.P. Majumder during statistical analysis. Technical help from Mr. Badal Dey and Mr. Madan Chakraborty during genotyping and sequencing of PCR products is highly appreciated. Lady Tata Memorial Trust, Mumbai, India, supported a SRF fellowship to Miss. Mousumi Majumder.

References

1. World Health Organization. Definition of leukoplakia and related lesions: an aid to studies to precancer. Oral Surg Oral Med Oral Pathol 1978;46:517–39.
2. Gupta PC, Mehta FS, Daftary DK, Pindborg JJ, Bhonsle RB, Jalnawala PN, Sinor PN, Pitkar VK, Murti PR, Irani RR, Shah HT, Kadam PM et al. Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10-year follow-up study of Indian villagers. Comm Dent Oral Epidemiol 1980;8:287–333.
3. Nair UJ, Obe G, Friesen MD, Goldberg MT, Bartsch H. Role of lime in the generation of reactive oxygen species from *betel quid* ingredients. Environ Health Prespect 1997;98:203–5.
4. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002: cancer incidence, mortality and prevalence (Worldwide IARC Cancer Base No. 5, Version 2.0). Lyon: IARC, 2004.
5. Butcher NJ, Boukouvala S, Sim E, Minchin RF. Pharmacogenetics of the arylamine *N*-acetyltransferases. Pharmacogenomics J 2002;2:30–42.
6. Hein DW. Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. Mutat Res 2002;506/507:65–77.
7. Unal M, Tamer L, Akbas Y, Pata YS, Vayisoglu Y, Degirmenci U, Camdeviren H. Genetic polymorphism of *N*-acetyltransferase 2 in the susceptibility to laryngeal squamous cell carcinoma. Head Neck 2005;27:1056–60.
8. Hein DW. *N*-acetyltransferase 2 genetic polymorphism: effects of carcinogen and haplotype on urinary bladder cancer risk. Oncogene 2006;25:1649–58.
9. Jourenkova-Mironova N, Wikman H, Bouchardy C, Mitrunen K, Dayer P, Benhamou S, Hirvonen A. Role of arylamine *N*-acetyltransferase 1 and 2 (NAT1 and NAT2) genotypes in susceptibility to oral/pharyngeal and laryngeal cancers. Pharmacogenetics 1999;9:533–7.
10. Katoh T, Keneko S, Boissy R, Watson M, Ikenmura K, Bell DA. A pilot study testing the association between *N*-acetyltransferases 1 and 2 and risk of oral squamous cell carcinoma in Japanese people. Carcinogenesis 1998;19:1803–07.
11. Chen C, Ricks S, Doody DR, Fitzgibbons ED, Porter PL, Schwartz SM. *N*-acetyltransferase 2 polymorphisms, cigarette smoking and alcohol consumption, and oral squamous cell cancer risk. Carcinogenesis 2001;22:1993–9.
12. Marques CFS, Koifman S, Koifman RJ, Boffetta P, Brennan P, Hatagima A. Influence of CYP1A1, CYP2E1, GSTM3 and NAT2 genetic polymorphisms in oral cancer susceptibility: results from a case-control study in Rio de Jeneiro. Oral Oncol 2006;42:632–7.
13. Gajecka M, Rydzanicz M, Jaskula-Sztul R, Kujawski M, Szyfter W, Szyfter K. CYP1A1, CYP2D6, CYP2E1, NAT2, GSTM1 and GSTT1 polymorphisms or their combinations are associated with the increased risk of the laryngeal squamous cell carcinoma. Mutat Res 2005;574:112–23.
14. Sturgis EM, Zheng R, Li L, Castillo EJ, Eicher SA, Chen M, Strom SS, Spitz MR, Wei Q. XPD/ERCC2 polymorphisms and risk of head and neck cancer. Carcinogenesis 2000;21:2219–23.
15. Park DJ, Stoehlmacher J, Zhang W, Tsao-Wei DD, Groshen S, Lenz H-J. A *Xeroderma Pigmentosum Group D* gene polymorphism predicts clinical outcome to platinum-based chemotherapy in patients with advanced colorectal cancer. Cancer Res 2001;61:8654–8.
16. Sturgis EM, Castillo EJ, Li L, Zheng R, Eicher SA, Clayman GL, Strom SS, Spitz MR, Wei Q. Polymorphisms of DNA repair gene *XRCC1* in squamous cell carcinoma of the head and neck. Carcinogenesis 1999;20:2125–9.
17. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1513–30.
18. Tae K, Lee HS, Park BJ, Park CW, Kim KR, Cho HY, Kim LH, Park BL, Shin HD. Association of DNA repair gene *XRCC1* polymorphisms with head and neck cancer in Korean population. Int J Cancer 2004;111:805–8.
19. Majumder M, Sikdar N, Paul RR, Roy B. Increased risk of oral leukoplakia and cancer among mixed tobacco users carrying *XRCC1* variant haplotypes and cancer among smokers carrying two risk genotypes, one on each of GSTM3 and *XRCC1* (codon 280). Cancer Epidemiol Biomarkers Prev 2005;14:2106–12.
20. Miller SA, Dykes DD, Polesky HK. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
21. Hemminki K, Xu G, Angelini S, Snellman E, Jansen CT, Lambert B, Hou SM. *XPD* exon 10 and 23 polymorphisms and DNA repair in human skin in situ. Carcinogenesis 2001;22:1185–8.
22. Cascorbi I, Brockmoller J, Mrozikiewicz PM, Bauer S, Loddenkemper R, Roots I. Homozygous rapid arylamine *N*-acetyltransferase (NAT2) genotype as susceptibility factor for lung cancer. Cancer Res 1996;56:3961–6.
23. http://www.louisville.edu/medschool/pharmacology/nat.html
24. Majumdar P, Majumder PP. HAPLOFREQ: a computer program for maximum likelihood estimation of haplotype frequencies in a population from genotype data on unrelated individuals via the EM algorithm, Tech. Rep. No. (2000) ISI/AHGU/1, Indian Statistical Institute, Kolkata, India, 2000.
25. Ramachandran S, Ramadas K, Hariharan R, Kumar RR, Pillai MR. Single nucleotide polymorphisms of DNA repair genes *XRCC1* and *XPD* and its molecular mapping in Indian oral cancer. Oral Oncol 2006;42:350–62.
26. Kietthubthew S, Sriplung H, Au WW, Ishida T. Polymorphism in DNA repair genes and oral squamous cell carcinoma in Thailand. Int J Hyg Environ Health 2006;209:21–9.
27. Matullo G, Peluso M, Polidoro S, Guarrera S, Munnia A, Krogh V, Masala G, Berrino F, Panico S, Tumino R, Vineis P, Palli D. Combi-

nation of DNA repair gene single nucleotide polymorphisms and increased levels of DNA adducts in a population-based study. Cancer Epidemiol Biomarkers Prev 2003;12:674–7.

28. Stern MC, Johnson LR, Bell DA, Taylor JA. XPD Codon 751 polymorphism, metabolism genes, smoking, and bladder cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1004–11.
29. Hein DW, Doll MA, Rustan TD, Gray K, Feng Y, Ferguson RJ, Grant DM. Metabolic activation and deactivation of arylamine carcinogens by recombinant human NAT1 and polymorphic NAT2 acetyltransferases. Carcinogenesis 1993;14:1633–8.
30. Spitz MR, Wu X, Wang Y, Wang LE, Shete S, Amos CI, Guo Z, Lei L, Mohrenweiser H, Wei Q. Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. Cancer Res 2001;61:1354–7.
31. Justenhoven C, Hamann U, Pesch B, Harth V, Rabstein S, Baisch C, Vollmert C, Illig T, Ko YD, Bruning T, Brauch H. ERCC2 genotypes and a corresponding haplotype are linked with breast cancer risk in a German population. Cancer Epidemiol Biomarkers Prev 2004;13:2059–64.
32. Lovatt T, Alldersea J, Lear JT, Hoban PR, Ramachandran S, Fryer AA, Smith AG, Strange RC. Polymorphism in the nuclear excision repair gene *ERCC2/XPD*: association between an exon 6-exon 10 haplotype and susceptibility to cutaneous basal cell carcinoma. Hum Mutat 2005;25:353–9.
33. Butkiewicz D, Rusin M, Enewold L, Shields PG, Chorazy M, Harris CC. Genetic polymorphisms in DNA repair genes and risk of lung cancer. Carcinogenesis 2001;22:593–7.
34. Tuimala J, Szekely G, Wikman H, Järventaus HH, Hirvonen A, Gundy S, Norppa H. Genetic polymorphisms of DNA repair and xenobiotic-metabolizing enzymes: effects on levels of sister chromatid exchanges and chromosomal aberrations. Mutat Res 2004;554:319–33.