Mol Biol Rep (2012) 39:9639–9647
DOI 10.1007/s11033-012-1829-z

An interethnic variability and a functional prediction of DNA repair gene polymorphisms: the example of *XRCC3* (p.Thr241>Met) and *XPD* (p.Lys751>Gln) in a healthy Tunisian population

Ghada Ben Salah · Nourhene Fendri-Kriaa · Hassen Kamoun · Fakhri Kallabi · Emna Mkaouar-Rebai · Amine Fourati · Hammadi Ayadi · Faiza Fakhfakh

Received: 12 January 2012 / Accepted: 10 June 2012 / Published online: 28 June 2012

Abstract Genetic polymorphisms in DNA repair genes might influence the repair activities of the enzymes predisposing individuals to cancer risk. Owing to the presence of these genetic variants, interethnic differences in DNA repair capacity have been observed in various populations. The present study was undertaken to determine the allele and genotype frequencies of two common non-synonymous SNPs, *XRCC3* p.Thr241>Met (C > T, rs861539) and *XPD* p.Lys751>Gln (T > G, rs13181) in a healthy Tunisian population and to compare them with HapMap (http://www.hapmap.org/) populations. Also, we predicted their eventual functional effect based on bioinformatics tools. The genotypes of 154 healthy and unrelated individuals were determined by PCR–RFLP procedure. Our findings showed a close relatedness with Caucasians from European ancestry which might be explained by the strategic geographic location of Tunisia in the Mediterranean, thus allowing exchanges with Europeans countries. The in silico predictions showed that p.Thr241>Met substitution in *XRCC3* protein was predicted as possibly damaging, indicating that it is likely to have functional consequences as well. To the best of our knowledge, this is the first study in this regard in Tunisia. So, these data could provide baseline database and help us to explore the relationship of *XRCC3* and *XPD* polymorphisms with both cancer risk and DNA repair variability in our population.

Keywords DNA repair · Polymorphism · Genotyping · Ethnic variability · Functional prediction · Tunisian population

G. Ben Salah (✉) · N. Fendri-Kriaa · H. Kamoun · F. Kallabi · E. Mkaouar-Rebai · F. Fakhfakh
Laboratory of Human Molecular Genetics, Faculty of Medicine, University of Sfax, Av. Magida Boulila, 3029 Sfax, Tunisia
e-mail: ghada.bsalah@gmail.com

A. Fourati
Research Unit, Chemical Tunisian Group, Sfax, Tunisia

H. Ayadi
Centre of Biotechnology, University of Sfax, Sfax, Tunisia

Introduction

The DNA repair and the maintenance of genome stability are crucial to cellular and organism functions; defects in these processes have been implicated in cancer risk and variability in DNA repair. There are about 150 known DNA repair genes in humans to repair different types of DNA damage involving distinct pathways [1, 2].

X-ray repair cross-complementing group 3 (*XRCC3*), one of the DNA repair genes, encodes for a protein participating in homologous recombination repair (HRR) of DNA double-strand breaks (DSB) and cross-links [3–6]. It is a member of an emerging family of Rad-51-related proteins that may take part in HRR to maintain chromosome stability and to repair DNA damage [3–6]. *XRCC3*-deficient cells were found to be unable to form Rad51 foci after radiation damage and demonstrated genetic instability and increased sensitivity to DNA damaging agents [6, 7].

A single nucleotide polymorphism (SNP) has been reported in *XRCC3* gene at exon 7 (C > T, rs861539); it causes the substitution of Threonine to Methionine at codon 241 (p.Thr241>Met) [8]. It is the most frequent polymorphism in *XRCC3*. Carriers of *XRCC*3 241Met variant allele had high DNA adduct levels in lymphocyte DNA compared to homozygous 241Thr wild-type allele carriers, indicating that this polymorphism was associated

with low DNA repair capacity [9, 10]. Therefore, *XRCC3* has been of considerable interest as a candidate susceptibility gene for cancer. A positive association between the 241Met variant allele and cancer risk has been reported in several studies [10–17].

The xeroderma pigmentosum complementation group D (*XPD*), also known as *ERCC2* gene encodes for a DNA-dependent ATPase/helicase [18]. *XPD* or *ERCC2* has two functions: nucleotide excision repair (NER), and basal transcription as part of human transcriptional initiation factor complex, TFIIH [19, 20]. *XPD* is also essential for the viability of cells [21].

Several single-nucleotide polymorphisms have been identified in the *XPD* locus. Among them, a C > A (rs13181) in codon 751 of exon 23, produces a Lysine to Glutamine substitution (p.Lys751>Gln) [8]. The mutant phenotype has shown to be associated with lower DNA repair capacity by altering the functional properties of DNA repair enzymes [9, 21–23]. To date, epidemiological studies have investigated the relationship between the p.Lys751>Gln polymorphisms and predisposition to cancer. However, the results remain conflicting rather than conclusive. A number of case–control studies have shown the association of the 751Gln variant with cancer [24–29]. However, no overall association has been seen in Skjelbred report [30] and recently in Zhang meta-analysis [31].

Such associations between *XPD* and *XRCC3* polymorphisms and cancers address an increased need to elucidate their effects on cancer susceptibility and adverse health outcomes in Tunisian population. However, it has been reported in HapMap[1] and NCBI dbSNP[2] databases that the prevalence of the repair gene polymorphisms are not randomly distributed throughout the human population, but follow diverse ethnic and/or geographic-specific patterns. These findings highlight the necessity to establish a specific genotype profile for our population in order to provide their basic prevalence to evaluate their significance in risk assessment in cancer and other phenotypes.

Therefore, the present study aimed to determine the allele and genotype frequencies of *XRCC3* p.Thr241>Met and *XPD* p.Lys751>Gln polymorphisms in a healthy Tunisian population and to compare them with those observed in various ethnic groups from HapMap (see footnote 1) populations. Also, we predicted their eventual functional effect based on bioinformatics tools, in order to distinguish between tolerated and deleterious amino acid substitutions for further clinical and genetic studies as well.

[1] http://www.hapmap.org/.

[2] http://www.ncbi.nlm.nih.gov/projects/SNP/.

Materials and methods

Studied population

In the present study, 154 healthy unrelated individuals (56 women and 98 men) from South Tunisia were enrolled after approval from the Ethics Committee of the Institute. Their age ranged from 20 to 60 years, with a mean ± SD of 37.5 ± 11. Information about social habits and health problems of each individual were gathered through standardized questionnaire. We excluded those who had a history of cancer.

DNA extraction and sample genotyping

Blood samples (5 ml) were collected after getting proper written informed consent and further processed for DNA extraction using phenol–chloroform standard procedure [32]. The primers for *XRCC3* were 5′-GGTCGAGTGA CAGTCCAAAC-3′and 5′-TGCAACGGCTGAGGGTC T-3′, which generated a 456 bp fragment [33]. The primers for *XPD* were 5′-GCCCGCTCTGGATTATACG-3′ and 5′-CTATCATCTCCTGGCCCCC-3′, which generated a 436 bp fragment [34].

These fragments were amplified separately, but under the following conditions: a 50 µl reaction mixture containing approximately 50 ng of genomic DNA, 10 µM of each primer, and 2 mM MgCl2, 10 mM mix dNTP, 1× PCR buffer, and 1 U Taq DNA polymerase. The mixtures were amplified with a GeneAmp PCR System 9700 Thermal cycler (Applied Biosystems). The PCR profile consisted of an initial melting step of 95 °C for 5 min, followed by 35 cycles with melting at 94 °C for 50 s, annealing at 63 °C (*XRCC3*) for 60 s, 64 °C (*XPD*) for 50 s, and elongation at 72 °C for 50 s with a final elongation step of 72 °C for 10 min. The PCR products were checked on a 2 % agarose gel and photographed using Gel doc and were then subjected to RFLP analysis.

The restriction enzyme NlaIII (Fermentas, EU) was used to distinguish the *XRCC3* C → T polymorphism of exon 7, in which a create NlaIII restriction site occurs in the polymorphic allele. PCR–RFLP patterns resulted in two bands of 140 and 316 bp in the homozygous wild-type (Thr/Thr), whereas in the homozygous mutant type (Met/Met), three bands at 140, 211, and105 bp were produced. The restriction enzyme PstI (Fermentas, EU) was used to distinguish the *XPD* T → G polymorphism of exon 23, in which a create RsaI restriction site occurs in the polymorphic allele. PCR–RFLP patterns resulted in two bands of 146 and 290 bp in the homozygous wild-type (Lys/Lys), whereas in the homozygous mutant Gln/Gln), three bands at 146, 227 and 63 bp were produced.

Digestion of the PCR product was carried out using 10U of NlaIII or PstI and the 10× buffer supplied with each restriction enzyme at 37 C overnight. The digestion products were separated on a 3.5 % agarose gel and photographed using Gel doc film (BIO-RAD, USA). A few samples from each of the three genotypes were sequenced to confirm the PCR–RFLP results for the *XRCC3* and *XPD* loci.

Bioinformatics analysis

The HapMap (see footnote 1) database was used to identify the allele and genotype frequencies of *XRCC3* and *XPD* polymorphisms in various ethnic groups such as Utah residents with Northern and Western European ancestry from the CEPH collection (CEU), Tuscan in Italy (TSI), Han Chinese in Beijing, China (HCB), Chinese in Metropolitan Denver, Colorado (CHD), Gujarati Indians in Houston, Texas (GIH), Japanese in Tokyo, Japan (JPT), Luhya in Webuye, Kenya (LWK), Maasai in Kinyawa, Kenya (MKK), Yoruban in Ibadan, Nigeria (YRI), Mexican ancestry in Los Angeles, California (MEX), African ancestry in Southwest USA (ASW).

Knowing that each amino acid substitution potentially affects protein function, we used the sorting intolerant from tolerant (SIFT)[3] software (version 1.03) developed by Ng and Henikoff group [35, 36] to predict whether the p.Thr241>Met and p.Lys751>Gln substitutions may have an impact on their protein function's. To assess their effects, SIFT assumes that important positions in a protein sequence have been conserved throughout evolution and therefore substitutions at these positions may affect protein function. The median sequence conservation ranges from 0 to 4.32, ideally the number would be between 2.75 and 3.25. A warning will occur if this is greater than 3.25 because this indicates that the prediction was based on closely related sequences or there were not enough sequences. The SIFT score ranges from 0 to 1. The amino acid substitution is predicted damaging if the score is $\leq$0.05, possibly damaging if the score ranges between 0.051 and 0.1, and tolerated if the score is >0.1.

The SIFT predictions were verified by PolyPhen-2 (Polymorphism Phenotyping, version 2.2) software[4] which predicts possible impact of amino acid substitutions on the structure and function of human proteins using straightforward physical and evolutionary comparative considerations.

[3] http://sift.jcvi.org/www/SIFT_dbSNP.html.

[4] http://genetics.bwh.harvard.edu/pph2/index.shtml.

Statistical analysis

The allele frequency was calculated from the genotype frequency. Hardy–Weinberg equilibrium was examined using a Chi-square ($\chi 2$) test with one degree of freedom. The comparison of genotype frequency distributions as a gender function was performed by $\chi 2$ test. Pair-wise Chi square ($\chi 2$) tests were also performed between Tunisian population (TUN) and HapMap populations using the allele frequencies in a 2 × 2 contingency table to find whether the TUN population was significantly different from others. These calculations were conducted by computer SPSS program (Version 17.0). p value <0.05 was considered statistically significant.

Results and discussion

The discordance in genetic variation observed across ethnicities highlights the necessity for researchers to establish a specific genotype profile for each population. In the current study, two non-synonyms polymorphisms in *XRCC3* and *XPD* genes have been studied in the same individuals of a Tunisian population.

The genotype and allele frequencies of *XRCC3* p.Thr241>Met and *XPD* p.Lys751>Gln substitutions are summarized in Table 1. The genotype frequencies were found to be consistent with Hardy–Weinberg expectations ($p > 0.05$, Table 1). There was no difference regarding the genotypes distribution as a gender function.

The sequencing results confirmed the SNPs genotypes obtained by PCR–RFLP.

Genotype and allele frequencies of *XRCC3* C > T (p.Thr241>Met)

The SNP rs861539: C > T at exon 7 causes Threonine to Methionine substitution at codon 241 of *XRCC3* protein (p.Thr241>Met). The observed (CC) Thr/Thr, (CT) Thr/Met and (TT) Met/Met genotype frequencies in Tunisian population were 0.312, 0.428, and 0.260, respectively (Table 2). The wild-type (C) and the variant (T) alleles frequencies were 0.526 and 0.474, respectively (Table 2).

Regarding HapMap populations, the variant allele frequency ranged between 0.041 among Chinese population in Colorado (CHD), and 0.461 among Tuscan population in Italy (TSI) (Table 2). Tunisian population (TUN) was found to be significantly different from all HapMap populations included in this study, except the Caucasian ones such as the Utah residents with Northern and Western European ancestry from the CEPH collection (CEU) and the Tuscans in Italy (TSI) ($p > 0.05$, Table 2).

Table 1 Genotype distributions and allele frequencies of *XRCC*3 and *XPD* SNPs among Tunisian population

Genotype	Total		Women		Men		Allele frequency		HWE[b] (*p* value)
	N	Freq.[a]	N	Freq.	N	Freq.	Wild-type allele	variant allele	
rs861539: C > T (p.Thr241>Met)									
Total	154		56		98				
C/C	48	0.312	16	0. 286	33	0.337			
C/T	66	0.428	26	0.464	40	0.408	(C) 0.526	(T) 0.474	0.08
T/T	40	0.260	14	0.250	26	0.265[c]			
rs13181: T > G (p.Lys751>Gln)									
Total	154		56		98				
T/T	56	0.364	22	0.393	34	0.347			
T/G	73	0.474	25	0.446	48	0.490	(T) 0.600	(G) 0.400	0.88
G/G	25	0.162	9	0.161	16	0.163[c]			

N number of subjects

[a] Frequency

[b] Hardy–Weinberg equilibrium (there was no deviation from the expected frequencies, $p > 0.05$)

[c] $p > 0.05$, no statistically difference as a function of gender

Genotype and allele frequencies of *XPD* T > G (p.Lys751>Gln)

The SNP rs13181: T > G at exon 23 causes Lysine to Glutamine substitution at codon 751 of the *XPD* protein (p.Lys751>Gln). The (TT) Lys/Lys, (TG) Lys/Gln and (GG) Gln/Gln genotype distribution in Tunisian population were 0.364, 0.474, and 0.162, respectively (Table 3). The wild-type (T) and the variant (G) allele frequencies were 0.600 and 0.400, respectively (Table 3). The variant allele frequency exhibited a range between 0.062 among Japanese population (JPT), and 0.426 among Tuscan population in Italy (TSI) from HapMap project.

Overall, similarly to *XRCC3* polymorphism, the allele frequencies of p.Lys751>Gln polymorphism in Tunisian population (TUN) were in line with those reported in Caucasians from Northern and Western European ancestry according to the CEPH collection (CEU) and Tuscans in Italy (TSI) ($p < 0.05$, Table 3). These findings may be explained by the position of Tunisia in the heart of the Mediterranean which allowed exchanges with the European populations.

Interestingly, in our study, an unexpected relationship was observed for *XPD* p.Lys751>Gln polymorphism between Tunisians (TUN) and Gujarati Indians living in Houston, Texas (GIH) which whose ancestors came from the state of Gujarat in India. This result may be explained by the genetic similarity between Indians and European (CEU) which has been observed in the Indian Genome Variation Consortium [37].

However, a high discordance was observed between Tunisians (TUN) and Americans with Mexican (MEX) and African (ASW) ancestries as well as with Sub-Saharans Africans (LWK, MKK, and YRI). This genetic differentiation could be explained by the geographic isolation, especially, from the Sub-Saharan populations, since the Sahara desert is a geographic barrier.

Combined genotypes analysis

The combined genotype distributions are shown in Table 4. Six genotype combinations showed frequencies greater than 10 %. Approximately 22 percent of the studied subjects carried double heterozygous genotype (TG/CT) which is the most frequently genotype combination in our population. However, among the studied population (154 subjects), only two subjects have double homozygous mutant genotype GG/TT.

The association between cancer and the combined genotypes of p.Thr241>Met and p.Lys751>Gln substitutions is not common practice and its effect is still unknown. Indeed, these data would be useful in the epidemiologic studies and/or in the occupational and environmental genotoxic exposure. However, largest population studies might be more suitable to assess the effect of each combined genotype.

Table 2 Genotype and allele frequencies of *XRCC3* p.Thr241>Met in HapMap and Tunisian (TUN) populations

Population	Genotype frequency						Allele frequency		Pair-wise *p* value between TUN and other populations
	Geno.	Freq.	Geno.	Freq.	Geno.	Freq.	Wild-type allele	Variant allele	
XRCC3 C > T (p.Thr241>Met)							(C)	(T)	
CEU (n = 226)	C/C	0.310	C/T	0.522	T/T	0.168	0.571	0.429	0.52
TSI (n = 204)	C/C	0.324	C/T	0.431	T/T	0.245	0.539	0.461	0.85
HCB (n = 274)	C/C	0.876	C/T	0.124	T/T	0.000	0.938	0.062	9×10^{-6}
CHD (n = 218)	C/C	0.917	C/T	0.083	T/T	0.000	0.959	0.041	9×10^{-6}
GIH (n = 202)	C/C	0.574	C/T	0.356	T/T	0.069	0.752	0.248	8×10^{-4}
JPT (n = 224)	C/C	0.777	C/T	0.214	T/T	0.009	0.884	0.116	9×10^{-6}
LWK (n = 220)	C/C	0.582	C/T	0.364	T/T	0.055	0.764	0.236	4×10^{-4}
MKK (n = 310)	C/C	0.645	C/T	0.310	T/T	0.045	0.800	0.200	4×10^{-5}
YRI (n = 288)	C/C	0.660	C/T	0.319	T/T	0.021	0.819	0.181	9×10^{-6}
MEX (n = 114)	C/C	0.719	C/T	0.246	T/T	0.035	0.842	0.158	2×10^{-6}
ASW (n = 114)	C/C	0.684	C/T	0.281	T/T	0.035	0.825	0.175	6×10^{-6}
TUN (n = 154)	C/C	0.312	C/T	0.428	T/T	0.260	0.526	0.474	

Population data source: (www.hapmap.org)

n number of individuals

CEU Utah residents with Northern and Western European ancestry from the CEPH collection, *TSI* Tuscan in Italy, *HCB* Han Chinese in Beijing, China, *CHD* Chinese in Metropolitan Denver, Colorado, *GIH* Gujarati Indians in Houston, Texas, *JPT* Japanese in Tokyo, Japan, *LWK* Luhya in Webuye, Kenya, *MKK* Maasai in Kinyawa, Kenya, *YRI* Yoruban in Ibadan, Nigeria, *MEX* Mexican ancestry in Los Angeles, California, *ASW* African ancestry in Southwest USA

$p < 0.05$, a significant difference between TUN and other populations

$p > 0.05$, no significant difference between TUN and other populations

Effect of p.Thr241>Met substitution on protein function

Threonine (T) to Methionine (M) substitution at 241 codon was predicted as possibly damaging to the protein function (SIFT score at 0.08, Table 5). In our analysis, the SIFT prediction was based on 13 homologous proteins in the alignments. SIFT prediction was verified by PolyPhen-2 software, and, therefore a possible damaging effect (with a score at 0.541) was also observed for p.Thr241>Met substitution (data not shown). To date, few systematic functional and structural studies of *XRCC3* were investigated, and therefore, the specific functions of the different regions of the *XRCC3* protein have not yet fully identified. Only two potential ATP-binding domains, the Walker boxes A and B, have been identified in highly conserved domains in *XRCC3* [38] (See Fig. 1). The p.Thr241>Met substitution occurs near to the Walker box B [38] (Fig. 1), and, therefore the transition from hydrophilic Threonine to a hydrophobic Methionine could affect ATP-binding and DNA repair efficiency.

It has been reported that cells lacking functional *XRCC3*, such as the hamster cell line irs1SF, exhibit decreased homologous recombination HRR resulting in a high rate of spontaneous and induced chromosomal aberrations [39]. They are also sensitive to ionizing radiation and highly sensitive to DNA cross-linking agents such as cisplatin and mitomycin-C [39]. In addition, *XRCC3*-deficient cells were found to be unable to form Rad51 foci after radiation damage and demonstrated genetic instability and increased sensitivity to DNA damaging agents [6, 7]. These phenotypes are a result of failure to initiate HRR and aberrant processing of HRR intermediates [3, 40, 41]. These findings highlight the damaged effect p.Thr241>Met substitution and support its association with cancers risk [10–17].

Effect of p.Lys751>Gln substitution on protein function

Lysine (K) to Glutamine (Q) substitution at 751 codon was predicted as tolerated to the protein function (SIFT score at 0.78, Table 5). This data was verified by PolyPhen-2 prediction and a benign effect was also observed for p.Lys751>Gln substitution (data not shown). According to Holstege report, this substitution is located within C-terminal domain, about 50 bases upstream from

Table 3 Genotype and allele frequencies of *XPD* p.Lys751>Gln in HapMap and Tunisian (TUN) populations

Population	Genotype frequency						Allele frequency		Pair-wise *p* value between TUN and other populations
	Geno.	Freq.	Geno.	Freq.	Geno.	Freq.	Wild-type allele	Variant allele	
XPD SNP T > G (p.Lys751>Gln)							(T)	(G)	
CEU (n = 226)	T/T	0.407	G/T	0.522	G/G	0.071	0.668	0.332	0.31
TSI (n = 202)	T/T	0.347	G/T	0.455	G/G	0.198	0.574	0.426	0.70
HCB (n = 274)	T/T	0.810	G/T	0.190	G/G	0.000	0.905	0.095	10^{-6}
CHD (n = 218)	T/T	0.862	G/T	0.128	G/G	0.009	0.927	0.073	10^{-7}
GIH (n = 220)	T/T	0.376	G/T	0.505	G/G	0.119	0.629	0.371	0.67
JPT (n = 226)	T/T	0.885	G/T	0.106	G/G	0.009	0.938	0.062	10^{-8}
LWK (n = 220)	T/T	0.645	G/T	0.327	G/G	0.027	0.809	0.191	10^{-3}
MKK (n = 312)	T/T	0.667	G/T	0.295	G/G	0.038	0.814	0.186	8×10^{-4}
YRI (n = 294)	T/T	0.667	G/T	0.313	G/G	0.002	0.823	0.177	5×10^{-4}
MEX (n = 116)	T/T	0.655	G/T	0.310	G/G	0.034	0.810	0.190	10^{-3}
ASW (n = 114)	T/T	0.544	G/T	0.386	G/G	0.007	0.737	0.263	0.039
TUN (n = 154)	T/T	0.364	G/T	0.474	G/G	0.162	0.600	0.400	

Population data source: (www.hapmap.org)

n number of individuals

CEU Utah residents with Northern and Western European ancestry from the CEPH collection, *TSI* Tuscan in Italy, *HCB* Han Chinese in Beijing, China, *CHD* Chinese in Metropolitan Denver, Colorado, *GIH* Gujarati Indians in Houston, Texas, *JPT* Japanese in Tokyo, Japan, *LWK* Luhya in Webuye, Kenya, *MKK* Maasai in Kinyawa, Kenya, *YRI* Yoruban in Ibadan, Nigeria, *MEX* Mexican ancestry in Los Angeles, California, *ASW* African ancestry in Southwest USA

$p < 0.05$, a significant difference between TUN and other populations

$p > 0.05$, no significant difference between TUN and other populations

Table 4 The frequency of combined genotype distributions for *XRCC*3 and *XPD* polymorphisms

rs13181: T > G (p.Lys751>Gln)	rs861539: C > T (p.Thr241>Met)	Frequency	
		N[a]	Freq[b].
TT	CC	16	0.104
	CT	21	0.137
	TT	19	0.123
TG	CC	20	0.130
	CT	34	0.220
	TT	19	0.123
GG	CC	12	0.780
	CT	11	0.714
	TT	2	0.129
Total			1.000

[a] Number of individuals

[b] Frequency

the poly (A); therefore, it affects an ATP-binding site of *XPD* and destroys its helicase activity, which is important for NER, but did not affect its transcriptional activity [42]. In addition, it has been reported that the T → G polymorphism of *XPD* gene leads to a change in the configuration of the coded protein and may alter the *XPD* protein's interaction with helicase activator p44 protein inside the TFIIH complex [43]. Also, Benhamou and Sarasin showed that the codon 751 involves in interactions with the substrate of *XPD*, thus any substitution at this residue might produce changes in its function which could impair the DNA repair capacity [44]. These findings highlight an eventual deleterious effect of p.Lys751>Gln polymorphism and may explain its association with cancer risk [24–29].

In the overall, our present study establishes the prevalence and the possible effects of *XRCC*3 and *XPD* polymorphisms. Tunisian population showed eventual genetic background homogeneity with European population. These data will provide a basic database for our future clinical and genetic studies pertaining to variability and the defect in DNA repair capacity. Such studies are in progress in our laboratory regarding the effect of *XRCC3* and *XPD* polymorphisms on occupational genotoxic exposure.

Table 5 SIFT evolutionary conservation status of p.Thr241>Met and p.Lys751>Gln substitutions

dbSNPs	Substitution	Protein ID	Aa[a]	SIFT Prediction	SIFT score	Median info	Number of sequence at position
rs861539	T241M	NP_001093588	T	TOLERATED	0.69	3.21	16
			M	POSSIBLY DAMAGING	0.08	3.21	16
rs13181	K751Q	NP_000391	K	TOLERATED	0.66	3.96	8
			Q	TOLERATED	0.78	3.96	8

SIFT score: the amino acid substitution is predicted DAMAGING if SIFT score $\leq$0.05, POSSIBLY DAMAGING if 0.05 < SIFT score $\leq$0.1 and TOLERATED if SIFT score >0.1

Median info: this is used to measure the diversity of the sequences used for prediction; it ranges from 0 to 4.32; ideally the number would be between 2.75 and 3.25

If the median score >3.25, it indicate that there is not enough sequences to make our prediction

The number of sequences at position: The number of homologue sequences that have an amino acid at the position prediction (SIFT automatically chooses the sequences)

[a] Amino acid

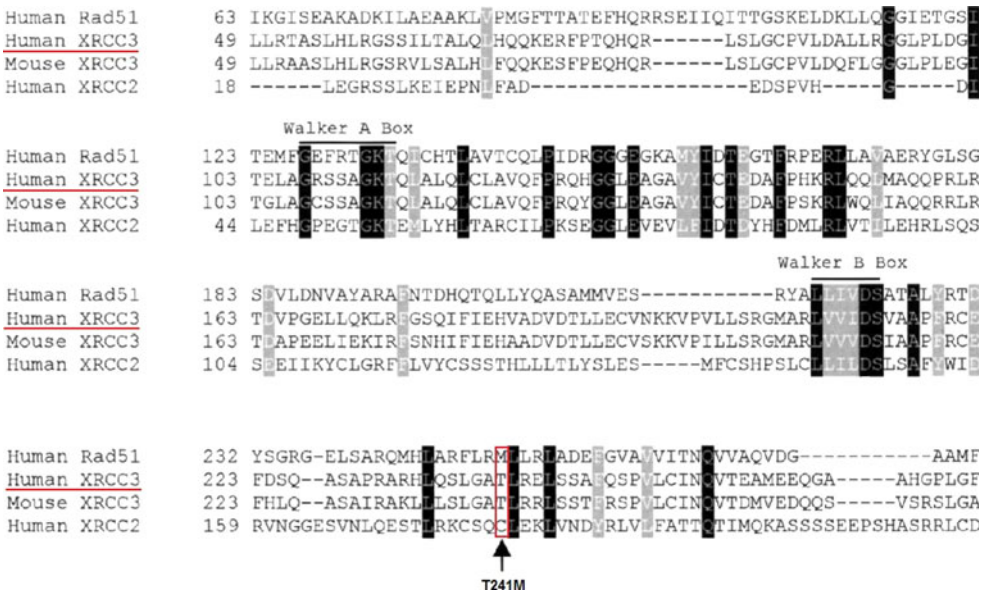

Fig. 1 Amino acid sequence alignment. Amino acid sequences of the relevant regions of human *RAD51*, human *XRCC3*, *mouse XRCC3* and human *XRCC2* (*XPD*) are shown. *Black boxes* indicate amino acid identity and *grey boxes* indicate conservation of charge. The positions of the Walker boxes A and B are shown by a *horizontal line* and the position of p.Thr241>Met substitution is indicated by *vertical arrow*

Acknowledgments We would like to thank the volunteers, for their cooperation. We were additionally grateful to Pr. Jamil JAOUA, Founder and Former Head on the English Department at the Sfax Faculty of Science, for proofreading this paper. This work was supported by the Ministry of Higher Education and Scientific Research in Tunisia and Chemical Tunisian Group.

Conflict of interest We declare that we have not any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations concerning this work which be published in your journal.

References

1. Wood RD, Mitchell M, Sgouros J, Lindahl T (2001) Human DNA repair genes. Science 5507:1284–1289
2. Branzei D, Foiani M (2008) Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol 9:297–308
3. Liu N, Lamerdin JE, Tebbs RS, Schild D, Tucker JD, Shen MR, Brookman KW, Siciliano MJ, Walter CA, Fan W, Narayana LS, Zhou ZQ, Adamson AW, Sorensen KJ, Chen DJ, Jones NJ, Thompson LH (1998) XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. Mol Cell 1:783–793
4. Pierce AJ, Johnson RD, Thompson LH, Jasin M (1999) XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev 13:2633–2638
5. Brenneman MA, Weiss AE, Nickoloff JA, Chen DJ (2000) XRCC3 is required for efficient repair of chromosome breaks by homologous recombination. Mutat Res 459:89–97
6. Suwaki N, Klare K, Tarsounas M (2011) RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis. Semin Cell Dev Biol 22:898–905
7. Griffin CS (2002) Aneuploidy, centrosome activity and chromosome instability in cells deficient in homologous recombination repair. Mutat Res 504:149–155
8. Shen MR, Jones IM, Mohrenweiser H (1998) Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. Cancer Res 58:604–608

9. Matullo G, Palli D, Peluso M, Guarrera S, Carturan S, Celentano E, Krogh V, Munnia A, Tumino R, Polidoro S, Piazza A, Vineis P (2001) XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. Carcinogenesis 22:1437–1445
10. Matullo G, Guarrera S, Carturan S, Peluso M, Malaveille C, Davico L, Piazza A, Vineis P (2001) DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. Int J Cancer 92: 562–567
11. Kuschel B, Auranen A, McBride S, Novik KL, Antoniou A, Lipscombe JM, Day NE, Easton DF, Ponder BA, Pharoah PD, Dunning A (2002) Variants in DNA double-strand break repair genes and breast cancer susceptibility. Hum Mol Genet 11: 1399–1407
12. Improta G, Sgambato A, Bianchino G, Zupa A, Grieco V, La Torre G, Traficante A, Cittadini A (2008) Polymorphisms of the DNA repair genes XRCC1 and XRCC3 and risk of lung and colorectal cancer: a case-control study in a Southern Italian population. Anticancer Res 28:2941–2946
13. He XF, Wei W, Su J, Yang ZX, Liu Y, Zhang Y, Ding DP, Wang W (2012) Association between the XRCC3 polymorphisms and breast cancer risk: meta-analysis based on case-control studies. Mol Biol Rep 39:5125–5134
14. Vral A, Willems P, Claes K, Poppe B, Perletti G, Thierens H (2011) Combined effect of polymorphisms in Rad51 and XRCC3 on breast cancer risk and chromosomal radiosensitivity. Mol Med Rep 5:901–912
15. Hamdy MS, El-Haddad AM, Bahaa El-Din NM, Makhlouf MM, Abdel-Hamid SM (2011) RAD51 and XRCC3 gene polymorphisms and the risk of developing acute myeloid leukemia. J Investig Med 7:1124–1130
16. Butkiewicz D, Rusin M, Sikora B, Lach A, Chorąży M (2011) An association between DNA repair gene polymorphisms and survival in patients with resected non-small cell lung cancer. Mol Biol Rep 8:5231–5241
17. Krupa R, Sliwinski T, Wisniewska-Jarosinska M, Chojnacki J, Wasylecka M, Dziki L, Morawiec J, Blasiak J (2011) Polymorphisms in RAD51, XRCC2 and XRCC3 genes of the homologous recombination repair in colorectal cancer—a case control study. Mol Biol Rep 4:2849–2854
18. Sung P, Bailly V, Weber C, Thompson LH, Prakash L, Prakash S (1993) Human xeroderma pigmentosum group D gene encodes a DNA helicase. Nature 365:852–855
19. Lehmann AR (2001) The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. Genes Dev 15:15–23
20. Weber CA, Salazar EP, Stewart SA, Thompson LH (1988) Molecular cloning and biological characterization of a human gene, ERCC2, that corrects the nucleotide excision repair defect in CHO UV5 cells. Mol Cell Biol 8:1137–1146
21. Wang HY, Xiong GF, Zhang JX, Xu H, Guo WH, Xu JJ, Xiong XY (2012) The role of XPD in cell apoptosis and viability and its relationship with p53 and cdk2 in hepatoma cells. Med Oncol 1:161–167
22. Lunn RM, Helzlsouer KJ, Parshad R, Umbach DM, Harris EL, Sanford KK, Bell DA (2000) XPD polymorphisms: effects on DNA repair proficiency. Carcinogenesis 21:551–555
23. Matullo G, Peluso M, Polidoro S, Guarrera S, Munnia A, Krogh V, Masala G, Berrino F, Panico S, Tumino R, Vineis P, Palli D (2003) Combination of DNA repair gene single nucleotide polymorphisms and increased levels of DNA adducts in a population-based study. Cancer Epidemiol Biomarkers Prev 12: 674–677
24. Spitz MR, Wu X, Wang Y, Wang LE, Shete S, Amos CI, Guo Z, Lei L, Mohrenweiser H, Wei Q (2001) Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. Cancer Res 61:1354–1357
25. Hou SM, Falt S, Angelini S, Yang K, Nyberg F, Lambert B, Hemminki K (2002) The XPD variant alleles are associated with increased aromatic DNA adduct level and lung cancer risk. Carcinogenesis 23:599–603
26. Tang D, Cho S, Rundle A, Chen S, Phillips D, Zhou J, Hsu Y, Schnabel F, Estabrook A, Perera FP (2002) Polymorphisms in the DNA repair enzyme XPD are associated with increased levels of PAH-DNA adducts in a case-control study of breast cancer. Breast Cancer Res Treat 75:159–166
27. Liang G, Xing D, Miao X, Tan W, Yu C, Lu W, Lin D (2003) Sequence variations in the DNA repair gene XPD and risk of lung cancer in a Chinese population. Int J Cancer 105:669–673
28. Kertat K, Rosdahl I, Sun XF, Synnerstad I, Zhang H (2008) The Gln/Gln genotype of XPD codon 751 as a genetic marker for melanoma risk and Lys/Gln as an important predictor for melanoma progression: a case control study in the Swedish population. Oncol Rep 20:179–183
29. Zhang J, Gu SY, Zhang P, Jia Z, Chang JH (2010) ERCC2 Lys751Gln polymorphism is associated with lung cancer among Caucasians. Eur J Cancer 13:2479–2484
30. Skjelbred CF, Saebo M, Wallin H, Nexo BA, Hagen PC, Lothe IM, Aase S, Johnson E, Hansteen IL, Vogel U, Kure EH (2006) Polymorphisms of the XRCC1, XRCC3 and XPD genes and risk of colorectal adenoma and carcinoma, in a Norwegian cohort: a case control study. BMC Cancer 6:67
31. Zhang Y, Ding D, Wang X, Zhu Z, Huang M, He X (2011) Lack of association between XPD Lys751Gln and Asp312Asn polymorphisms and colorectal cancer risk: a meta-analysis of case–control studies. Int J Colorectal Dis 26:1257–1264
32. Lewin HA, Stewart-Haynes JA (1992) A simple method for DNA extraction from leukocytes for use in PCR. Biotechniques 4:522–524
33. Andreassi MG, Foffa I, Manfredi S, Botto N, Cioppa A, Picano E (2009) Genetic polymorphisms in XRCC1, OGG1, APE1 and XRCC3 DNA repair genes, ionizing radiation exposure and chromosomal DNA damage in interventional cardiologists. Mutat Res 666:57–63
34. Lopez-Cima MF, Gonzalez-Arriaga P, Garcia-Castro L, Pascual T, Marron MG, Puente XS, Tardon A (2007) Polymorphisms in XPC, XPD, XRCC1, and XRCC3 DNA repair genes and lung cancer risk in a population of Northern Spain. BMC Cancer 7:162
35. Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. Genome Res 11:863–874
36. Ng PC, Henikoff S (2002) Accounting for human polymorphisms predicted to affect protein function. Genome Res 12:436–446
37. Consortium IGV (2008) Genetic landscape of the people of India: a canvas for disease gene exploration. J Genet 87:3–20
38. Walker JE, Saraste M, Runswick MJ, Gay N (1982) Distantly related sequences in the α and β subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and in common nucleotide binding fold. EMBO J 1:945–951
39. Tebbs RS, Zhao Y, Tucker JD, Scheerer JB, Siciliano MJ, Hwang M, Liu N, Legerski RJ, Thompson LH (1995) Correction of chromosomal instability and sensitivity to diverse mutagens by a cloned c DNA of the XRCC3 DNA repair gene. Proc Natl Acad Sci USA 92:6354–6358
40. Brenneman MA, Wagener BM, Miller CA, Allen C, Nickoloff JA (2002) XRCC3 controls the fidelity of homologous recombination: roles for XRCC3 in late stages of recombination. Mol Cell 10:387–395
41. Masson JY, Stasiak AZ, Stasiak A, Benson FE, West SC (2001) Complex formation by the human RAD51C and XRCC3 recombination repair proteins. Proc Natl Acad Sci USA 98:8440–8446

42. Holstege FC, Van der Vliet PC, Timmers HT (1996) Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. EMBO J 15:1666–1677
43. Coin F, Bergmann E, Tremeau-Bravard A, Egly JM (1999) Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. EMBO J 18:1357–1366
44. Benhamou S, Sarasin A (2002) ERCC2/XPD gene polymorphisms and cancer risk. Mutagenesis 17:463–469