Additional Studies into the Association of OGG1, XRCC1, XPC, XPD and XPG with Lung Cancer

Authors: A.J.Thornton and R.E.Thornton

Date: 21st October 2008

1. Objectives

Case-control studies[1-46] involving polymorphisms in the DNA repair genes listed above have already been reviewed, on an individual gene basis. The main results of these studies are summarized here in Tables 1-5. Alternative methods for investigating possible associations between the polymorphisms in repair genes and lung cancer are now considered.

2. Methods

The following methods are theoretically possible:

2.1 Entire genome sequencing studies, followed by statistical analysis to seek associations between polymorphisms and specific diseases. So far no such studies have been reported, largely because the technology has only recently become available at affordable cost. However shorter sequences have been sequenced by various methods and associations sought. Those involving case-control studies relating to polymorphisms in the individual repair genes have already been considered in earlier sections. However some are also included in this section if they address wider issues.

2.2 Genome wide association studies, using polymorphisms as markers. Statistical analysis is used to seek associations between genomic regions identified by markers and specific diseases. Sequencing of the marked genomic region is necessary to identify the gene(s) involved.

2.3 Linkage studies based on examining families with a high incidence of a specified disease. Again the result is the identification of a genomic region of interest.

2.4 Measurement of somatic mutations in the cancer genome. Such studies may indirectly reveal DNA repair defects.

2.5 Gene expression studies. Levels of expression of individual genes may indicate the importance of DNA repair genes in response to environmental insults such as tobacco smoke.

2.6 A variety of studies based on models relating to DNA repair processes.

The advantages and disadvantages of the various approaches to discovering polymorphisms conferring multifactorial susceptibility to common diseases have recently been reviewed[47]. This report does not address the effect of any other change in the genomic structure relating to the repair genes (e.g. copy number) and disease susceptibility.

3. Literature Searches

Papers that appeared to fall within the broad outlines described above were sought from:

(i) our in-house database, and

(ii) Medline searches

Twenty-three relevant papers were identified. Terminology varies between the papers and this is clarified in the glossary.

4. Plan

Each paper was allocated to one of the headings listed, judgement being necessary to establish the best fit.

5. Review of Papers

5.1 *Genome Sequencing studies*

One of the most comprehensive studies published so far was carried out in the US[48]. Using currently available technology 127 amino acid substitution variants in 37 DNA repair genes were identified by resequencing. The diverse group of subjects were recruited in four areas of the US and totalled 164. Ethnic groups represented

were: European-American, African-American, Mexican-American, Native-American and Asian-American.

Amino acid substitutions were found in XRCC1 (9 variant alleles), ERCC2 (XPD) (5 variant alleles), XPC (5 variant alleles) and ERCC5 (XPG) (7 variant alleles). Variants of OGG1 were not described. The remaining 101 variants occurred in the other DNA repair genes. Tentative estimates of the frequencies of occurrence of variant alleles (of at least 10%) included XPD, XPC, XRCC1 and two other repair genes MSH3 and POLD1. The most frequently occurring 10 variant alleles accounted for 52% of the genetic variation and 108 variants with frequencies of less than 5% accounted for another 23% of the total variation.

This important study will be discussed further in the conclusions, although it did not include data on mortality, precluding any attempt to look for associations.

In a case-control study from France[40] polymorphisms in 62 DNA repair genes were studied in relation to head and neck and lung cancers. A total of 695 variants were successfully genotyped in 151 lung cancer cases, 251 head and neck cancers and 172 hospital controls. Subjects were restricted to smokers and matched for age, sex and hospital, with 95% of both cases and controls being men. However no polymorphisms in the genes considered in this report were associated with changes in lung cancer risk, although polymorphisms in eight other genes were associated ($p < 0.05$).

In another study by the same group[49] variants in DNA double-strand break repair genes and DNA damage-response genes and associations with lung and head and neck cancers were examined in a case-control study using the same subjects as above. Mean age was 58.4 years for the lung cancer cases and 54.9 years for the controls. A total of 625 variants were detected in 29 out of 30 genes screened. No associations were found for any of the genes being reviewed in this report, although associations were found for other repair genes.

In a German case-control study[50] 463 lung cancer cases (including 204 adenocarcinoma and 212 squamous cell carcinoma) were compared to 460 tumour-

free hospital controls. Six genotypes in five DNA repair genes were determined using two different polymerase chain reaction (PCR) methods. The genes included XPD and XRCC1. One of the objectives of this study was an examination of possible gene-gene interactions. Odds ratios were adjusted for age, gender, smoking and occupational exposure. For a study carried out on a subset of 54 subjects and controls and for squamous cell carcinoma the combination of variant alleles XPA-4AA +XPD-Asn312Asn +XPD-Gln751Gln, compared to XPA-4GG, XPD-Asp312Asp and XPD-Lys751Lys gave an odds ratio (OR) of 5.35 (95% CI 1.49-19.17). The effects for adenocarcinoma were smaller and not statistically significant. In a separate analysis in which individuals with two or three risk alleles were evaluated as one group, for squamous cell carcinoma the OR was 2.83 (95% CI 1.17-6.85) and for adenocarcinoma the OR was 3.05 (95% CI 1.49-6.23).

A further attempt to look at gene-gene interactions has been reported as a case-control study based within the European EPIC study[51]. Subjects included 409 patients with bladder cancer, 116 with lung cancer and 169 with myeloid leukaemia. Controls numbered 757. Matching criteria included gender, age, smoking status, country of recruitment, and follow-up time. Some 36 gene variants in DNA repair and metabolic genes were analysed using a variety of models. The best model was judged to be an association between the XRCC1-Arg399Gln polymorphism compared to XRCC1-Arg399Arg, the double-strand break repair (DSBR) BRCA2-Asn372His polymorphism and levels of air pollution, with an OR of 6.22 (95% CI 3.69-10.47) although an alternative recessive four-loci model involving the same two polymorphisms as above, and the RAD52-2259C>T and NQO1-Pro187Ser polymorphisms actually gave a higher OR of 6.55 (95% CI 3.46-12.40).

In a study from the UK[52], 109 non-synonymous polymorphisms in 50 DNA repair genes were studied in relation to lung cancer survival in 700 lung cancer patients. Although polymorphisms in all of the five DNA repair genes considered in this report were included in the study, only one polymorphism in ERCC5 (XPG)-D1104H was associated with survival.

A large study has been carried out in Japan[24], in which 50 nonsynonymous polymorphisms (associated with amino acid changes) were determined in 36 genes

involved in diverse DNA repair pathways. They were assessed for associations with lung cancer risk in a case-control study consisting of 752 adenocarcinoma cases, 250 squamous cell carcinoma cases and 685 controls. Subjects were recruited from Tokyo and surrounding prefectures. Controls were selected from hospitals (383 subjects) or were volunteers from Keio university, Tokyo. None had any history of cancer during the study period. Associations were found between polymorphisms in six DNA repair genes and lung cancer. Of the genes being considered in this review only the XPG-His1104Asp polymorphism showed any association, using unconditional logistic regression analysis. It was noted that the frequency of the minor allele XPD-Asn312Asn is much lower in the Japanese population than in the US (0.04 v. 0.40), reducing the statistical power of the study. The study contains a reasonably detailed review of allele distributions in different ethnic groups. It showed wide variations in OGG1, XRCC1 and XPG polymorphisms. However the frequency of the XPC-Lys939Gln polymorphisms was similar in all ethnic groups.

In a study from China[53], limited to looking for differences between ethnic groups, polymorphisms in XPC and XPD were determined in blood samples obtained from 150 unrelated and healthy individuals of the Han population in north-eastern China. Allele frequencies XPC-G19007A and XPD-C22541A were examined. The allele frequency for the A allele of the G19007A polymorphism was 0.21 which compares with values of 0.54-0.62 found with Caucasian populations. The other allele was only marginally different.

Based on collaboration with a group in Denmark[54,55], haplotypes of nine polymorphisms in chromosome 19q 13.2-3 have been determined in a Chinese population[56] in Shenyang. This region of chromosome 19 includes the XRCC1, XPC and XPD genes. A total of 247 newly diagnosed lung cancer cases were compared to 253 cancer-free controls. Stratification analysis was defined by gender, age and smoking history. Each haplotype contained a particular set of polymorphisms of these genes and two haplotypes (Hap5 and Hap12) were underrepresented among cases, while three were overrepresented (Hap3, Hap4 and Hap10). For Hap 3, cases = 5.7%, controls = 1.0% (OR: 6.56, 95% CI 1.83-23.54), for Hap 4, cases = 3.5%, controls = 1.0% (OR: 3.77, 95% CI 1.032-13.78), for Hap 10, cases = 13.3%, controls = 5.6% (OR: 2.73, 95% CI 1.51-4.94), for Hap 5, cases = 0.7%, controls = 3.7% (OR: 0.18,

95% CI 0.042-0.76), and for Hap 12, cases = 5.0%, controls = 9.6% (OR: 0.51, 95% CI 0.28-0.96). Because the polymorphisms are in strong linkage disequilibrium it is difficult to determine which are the important ones in relation to lung cancer, but the results were taken to show that this region contains one or more loci of cancer susceptibility.

5.2 *Genome wide association studies*

Three studies into genome wide associations with lung cancer were published together in June 2008[57-59]. All reported that the only susceptibility locus associated with lung cancer occurred at 15q25.1 which contains the nicotinic acetylcholine receptor subunit genes CHRNA3 and CHRNA5. All three studies yielded odds ratios of around 1.3 (typically $p < 1 \times 10^{-17}$) for the two polymorphisms associated with the disease. However, the papers did not agree on the interpretation of these results. Subsequently, one group has reported associations with two other genes[60] but no association was found for any polymorphism in a repair gene and lung cancer. The reasons for these seemingly disappointing results from these large and expensive studies is discussed elsewhere[47].

5.3 *Linkage Studies*

One linkage study has been reported[61], based on 26,108 lung cancer cases recruited from eight centres in the US. A total of 3541 families had at least one first degree relative with lung cancer, but only 52 families had enough biospecimens available to make them informative for linkage analysis. Analysis revealed the highest LOD score at 6q23-25, marker D6S2436. However the identity of the putative susceptibility gene located at this locus has not been established although further work on more families is underway[62], but it is known that the tumour suppressor gene TCF 21 locates at 6q23-24.

Although not a true linkage study, polymorphisms in XPC, XPD, XPG and other repair genes has been examined in familial early-onset lung cancer cases in the UK[63], within the Genetic Lung Cancer Predisposition Study (GELCAPS). Germline DNA was screened in 92 lung cancer patients. Controls consisted of 278 cancer free subjects, matched for age and sex. Some 79 sequence variants were discovered, 38 of which were in introns or other untranslated regions. The remaining 41 variants were

discovered in exons, of which 17 were synonymous and 24 were nonsynonymous variants. Ten were changes that had not been previously discovered, and were all detected in different lung cancer patients. The computer programmes Polyphen and SIFT were then used to predict the impact of these changes on wild type protein function. XPG Q622H was identified as potentially damaging, XPG R819W as probably damaging and XPC R671C and XPG E399V as possibly damaging.

5.4 *Cancer Genome Studies*

Patterns of somatic mutation in human cancer genomes have been reported[64] and indicate substantial variations in the number and patterns of mutations in individual cancers. Early results based on only 20 samples suggest that lung cancer samples have a high incidence of mutations in their DNA, and one of the factors contributing is thought to be DNA repair defects. However, the main conclusion from another study is that many of the genes involved in lung adenocarcinoma remain to be discovered[65].

5.5 *Gene-Expression Studies*

Studies in the US[66] have looked at airway epithelial gene expression in the diagnostic evaluation of smokers with suspected lung cancer. While no indication has been given of possible changes in the expression levels of the genes being reviewed here, the expression of the gene coding for DNA repair protein IC was decreased in the airway epithelium of smokers with lung cancer.

5.6 *Other Studies*

There have been a number of studies that address the effect of polymorphisms in repair genes on biochemical processes thought to be relevant to lung cancer. First there are studies relating to the tumour suppressor gene *TP53*. In the first study reviewed[67], polymorphisms in XPD were considered in relation to the mutation spectra of *TP53*. In a case-only study in the US, the subjects were 206 men and 103 women with lung cancer. Functional polymorphisms in XPD-Asp312Asn, rs1799793 and Lys751Gln, rs1052559 were modestly associated with G:C to T:A mutations in *TP53* in lung tumours. Asp/Asn312 +Asn/Asn312 and/or Lys/Gln751 + Gln/Gln751 versus Asp/Asp312 + Lys/Lys351 gave an OR of 2.73 (95% CI 0.98-7.61). An interaction between the XPD alleles Asn312 and Gln751 and the *TP53* Pro72 allele

was observed for *TP53* mutations (any *TP53* mutation P_{int}= 0.027). For an alternative mechanism linking DNA repair genes and mutations in *TP53* (promoter hypermethylation) see Wu *et al*, 2007[68].

Another approach has been to look at the effect of polymorphisms in repair genes on damage repair capacity (DRC). This can be defined in various ways but in one group from the US[12] it is defined by measuring cellular ability to remove adducts from plasmids transfected into lymphocyte cultures *in vitro* by expression of damaged reporter genes. The mutagen challenge is benzo[a]pyrene, chosen because it occurs in tobacco smoke. DRC (%) is calculated as a ratio of the damaged plasma values to the undamaged plasmid values. For XPD the Lys/Lys751 genotype related to a DRC of 8.21%, for Lys/Gln the DRC was 7.65% and for Gln/Gln 7.20%. *P* for this trend was 0.041. For Asp312Asn genotypes the figures for wild-type homozygotes, heterozygotes and variant homozygotes were 8.37%, 7.50% and 6.84% respectively (p = 0.008). The study was extended to include bleomycin induced chromosome breaks in 612 lung cancer cases and 557 controls. Bleomycin is a source of free oxygen radicals that can produce single- and double-strand breaks in DNA. The mean for cases was 0.78 +/- 0.4 breaks per chromosome and 0.62 +/-0.33 for the controls. This data was used to predict probability of lung cancer by combining DRC, bleomycin sensitivity, and years smoked, with about a 3.5 factor (bleomycin sensitive, DRC poor and 60 years smoking v. bleomycin nonsensitive, DRC good and about 5 years smoking). In a further extension of the study bleomycin was used to study chromosome breaks in 524 lung cancer cases and 524 controls in relation to polymorphisms in XRCC1. Individuals with wild type exon 6 Trp1194Trp exhibited higher numbers of bleomycin-induced chromosome breaks than those with the combined variant genotype. When combined with exon 10 Arg399Arg genotype these differences were enhanced (p for trend = 0.049 for the cases and 0.032 for controls). These patterns remained consistent when the data were stratified by gender, age, and smoking status.

The same group also looked at DRC with regard to polymorphisms in XPC, XPD and XRCC1 in a smaller study[69], using similar techniques, in which 50 male and 52 female healthy non-Hispanic white subjects aged 17-78 years were studied. This study did not show any impact of XRCC1 polymorphisms on DRC, but

polymorphisms in XPC and XPD did relate to changes in DRC. For XPC intron 9, PAT-/- compared to PAT+/+, DRC was 8.79 +/- 2.42% compared to 6.73 +/- 3.18%, p = 0.020, with similar results for XPD exon 23 Lys751Gln. The effects of other polymorphisms were not statistically robust.

In a study from Italy[70] the effect of polymorphisms in the repair genes XPD and XRCC1 on levels of benzo[a]pyrene diolepoxide, the active metabolite of benzo[a]pyrene, adducts were determined. in human lymphocytes. The subjects were 60 caucasian males, all smokers, with incident, histologically confirmed lung cancer. The subjects were also genotyped for the glutathione transferase polymorphism GSTM1, given its role in the detoxification of the adduct. The adduct levels were measured by high-resolution mass spectrometry. It was found that the adduct levels were dependent on the XPD Asp312Asp + Lys751Lys genotype, being lower in carriers of variant alleles (p = 0.02), but only in GTSM1-deficient smokers. No such association was found with XRCC1 variant alleles.

In a three-centre European study[71] micronuclei (MN) were measured in 171 occupationally exposed workers and 132 non-exposed male controls. MN arise from unrepaired DNA damage during nuclear division. Their frequency was determined per 1,000 binucleated cells. Polymorphisms in OGG1, XRCC1 and XRCC3 were determined, and the results analysed by genotype-genotype, genotype-smoking and genotype-exposure interactions, with a mixed regression model. Some complicated relationships were found. There was a significant MN decrease with smoking in subjects carrying the Ser/Ser OGG1 genotype ($p < 0.01$). Significantly lower MN frequencies were observed in carriers of the variant OGG1 326 genotype compared to the Ser326Ser OGG1 genotype in the sub-group of non-smokers with the XRCC3-Thr241Thr phenotype. There was a significant MN increase with smoking in occupationally exposed carriers of the XRCC1-Arg399Gln genotype ($p < 0.001$) compared to wild type XRCC-Arg399Arg.

6. Conclusions

Some 90 years ago, R.A. Fisher suggested that for any trait there are only a few genes with a large effect, but many associated genes with a small effect, although combined these may be substantial[72]. This is known as the "common disease common

variant hypothesis" i.e. common disease liability results from multiple genetic variants of small effect but relatively high frequency. In the "rare variant hypothesis"[47], it is suggested that a significant proportion of the inherited susceptibility to relatively common human chronic disease may be due to the summation of the effects of a series of low frequency dominantly and independently acting variants of a variety of different genes, each conferring a moderate but identifiable increase in relative risk. It is thought that these rare variants will mostly be population specific because of founder effects resulting from genetic drift.

Previous studies have indicated that the development of lung cancer involves both environmental and genetic factors, and since many carcinogenic compounds require metabolic activation to enable them to react with cellular macro-molecules, inter-individual variation in carcinogen metabolism may play an essential role in the development of lung cancer[73]. In the case of lung cancer no genes with a large pre-disposing effect have yet been identified[73]. The genome-wide association studies suggest that there is no dominant polymorphism in a repair gene that leads to lung cancer. Indeed, work carried out in the US[48], and the UK[52] suggests that the role of polymorphisms in DNA repair genes is multifactorial, with a large number of variants contributing. Furthermore, these appear to vary quite widely between individuals. The strength of the association between a specific polymorphism in a repair gene and lung cancer appears to be modest at best, and depends on the particular group of subjects being studied. For example, variations have been noted in the distribution of the polymorphisms on an ethnic basis, and these have been explored in more detail in the reviews for each of the single genes. Some of the studies examined in the previous reviews reported results stratified by smoking status and/or intensity. However, no clear picture of gene-smoking interactions emerged, with the results for any one polymorphism generally being equivocal. However, this was due in part to the small number of studies presenting data for any one given polymorphism. For only one gene, XPC, there was some evidence that as a whole, the ORs for lung cancer associated with the polymorphisms studied were higher in smokers, but no firm conclusions could be drawn for any individual polymorphism.

Although there is generally little evidence of an association with lung cancer when polymorphisms in single genes are examined, ORs reported for combinations of

genes tended to be much higher[50,51,54,55], and were often statistically significant. It is difficult in this type of analysis to determine which might be the important genes in relation to lung cancer, but these results are indicative that cancer susceptibility does have a genetic basis. These findings also suggest that the possible small effects of each of the genes studied may be cumulative.

It has been suggested that for most common variants, the disease-associated variant is unlikely to be functionally relevant[47]. In studies where ORs are in the region of 1.00, and thus the effect of a variant is relatively small, it is likely to be very difficult to establish which of a set of closely linked variants in linkage disequilibrium with each other is most relevant from a functional perspective[47]. It is also possible that common variants may act as significant modifiers of the effects of rare variants. Additionally, the genes for which common variants are found, or nearby genes that may contain the functionally relevant variant, may be considered as candidates for the search for rare variants. They may also help in identifying the functional variant associated with a common disease variant[47].

The reviews of the individual repair genes suggest that some polymorphisms in the OGG1, XPD and XRCC1 genes may be associated with a higher risk of lung cancer, while the Arg399Gln polymorphism of the XRCC1 gene confers a lower risk of the disease. However, these results are far from conclusive, and for most of the polymorphisms studied, no associations were found. Suggested reasons for the failure of case-control studies to clearly identify genetic polymorphisms associated with lung cancer include heterogeneity of the study populations, failure to consider effect modifiers such as environmental exposures, lack of statistical power causing false negatives and multiple testing creating false positive results, and publication bias[73]. Results from haplotype analyses suggest that the cumulative effect of several genes each with a small effect on lung cancer susceptibility may be important in the development of the disease. However, until detailed knowledge of the network of effects crucial to cancer causation exists, it is not possible to determine the importance of any single gene.

Table 1: Overview of association between lung cancer risk and polymorphisms of the OGG1 gene

Author, year	Odds ratio (95% confidence interval)[a]				
	Ser326Cys[b]			Exon 1[c]	
	Ser/Cys	Cys/Cys	All Cys	Allele 2	Allele 3
De Ruyck, 2007[37]	0.51 (0.27-0.95)	0.50 (0.10-2.51)			
Hung, 2005[21]	0.95 (0.82-1.09)	1.34 (0.95-1.88)			
Ishida, 1999[3]				1.33 (0.62-2.85)	2.60 (1.10-6.12)
Ito, 2002[8]	1.06 (0.64-1.76)	0.81 (0.44-1.52)			
Kohno, 1998[1]	0.89 (0.35-2.29)	1.34 (0.41-4.43)	1.01 (0.42-2.42)		
Kohno, 2006[31]	1.20 (0.90-1.60)	1.50 (1.00-2.10)			
Lan, 2004[14]	1.96 (1.10-3.57)	1.85 (0.83-4.11)	1.93 (1.12-3.34)		
Le Marchand, 2002[9]	0.70 (0.50-1.10)	2.10 (1.20-3.70)			
Park, 2004[16]	1.90 (1.20-2.90)	3.80 (1.40-10.60)			
Sorensen, 2006[34]	1.16 (0.83-1.62)	1.05 (0.51-2.16)			
Sugimura, 1999[4]	0.64 (0.39-1.06)	1.31 (0.65-2.62)			
Sunaga, 2002[10]	1.33 (0.80-2.23)	0.90 (0.48-1.70)			
Wikman, 2000[6]		2.20 (0.41-11.79)	0.72 (0.42-1.27)		
Zienolddiny, 2006[35]	1.45 (0.90-2.33)	1.64 (1.06-2.52)			
Overall odds ratio[d]	**1.07 (0.89-1.30)**	**1.40 (1.16-1.69)**	**1.13 (0.59-2.18)**	**1.33 (0.62-2.85)**	**2.60 (1.10-6.12)**

a Using most adjusted odds ratio where available
b Using Ser/Ser individuals as the reference group
c Using allele 1 as the reference group
d Odds ratio estimated by meta-analysis using random effects model

Table 2: Overview of association between lung cancer risk and polymorphisms of the XPC gene

Author, year	Odds ratio (95% confidence interval)[a]							
	-449G>C[b]		-371G>A[b]			-27G>C[b]		
	G/C	C/C	G/A	A/A	All A allele	G/C	C/C	All C allele
Bai, 2007[36]			0.83 (0.69-1.01)	0.76 (0.52-1.11)	0.82 (0.68-0.99)	1.16 (0.82-1.66)	0.78 (0.14-4.34)	1.15 (0.81-1.62)
Lee, 2005[22]	0.90 (0.68-1.20)	0.76 (0.42-1.38)	1.03 (0.77-1.37)	1.47 (0.83-2.62)				1.97 (1.22-3.17)
Overall odds ratio[c]	**0.90 (0.68-1.20)**	**0.76 (0.42-1.38)**	**0.90 (0.73-1.10)**	**1.02 (0.54-1.94)**	**0.82 (0.68-0.99)**	**1.16 (0.82-1.66)**	**0.78 (0.14-4.34)**	**1.47 (0.87-2.48)**

	Ala499Val[d]			Lys939Gln[e]			PAT+/-[f]		
	Ala/Val	Val/Val	All Val allele	Lys/Gln	Gln/Gln	All Gln allele	+/-	+/+	All + allele
Bai, 2007[36]				1.01 (0.83-1.22)	1.17 (0.88-1.56)	1.04 (0.87-1.25)			
Hu, 2005[19]				1.20 (0.85-1.70)	1.28 (0.72-2.28)	1.21 (0.87-1.69)			
Lee, 2005[22]				0.74 (0.55-1.00)	0.97 (0.63-1.48)				
Lopez-Cima, 2007[39]							1.08 (0.79-1.47)	1.28 (0.85-1.92)	
Raasch, 2008[44]				0.91 (0.65-1.28)	1.41 (0.87-2.30)				
Shen, 2005[25]	0.91 (0.52-1.62)	0.98 (0.41-2.38)		0.67 (0.37-1.23)	2.22 (0.86-5.74)				
Wang, 2003[13]							0.85 (0.65-1.10)	0.73 (0.48-1.09)	
Zhang[g], 2008[46]	1.05 (0.84-1.32)	1.16 (0.91-1.47)	1.07 (0.86-1.33)				0.99 (0.69-1.41)	1.24 (0.94-1.65)	1.09 (0.71-1.66)
Overall odds ratio[c]	**1.03 (0.84-1.27)**	**1.15 (0.91-1.45)**	**1.07 (0.86-1.33)**	**0.93 (0.78-1.10)**	**1.20 (0.99-1.46)**	**1.08 (0.92-1.26)**	**0.95 (0.80-1.13)**	**1.07 (0.77-1.49)**	**1.09 (0.71-1.66)**

Table 2 continued

	IVS11-5C>A[h]		12413C>G[h]			+315C>G[h]
	C/A	**A/A**	**C/G**	**G/G**	**All G allele**	**C/G**
Bai, 2007[36]			1.14 (0.88-1.48)	0.90 (0.54-1.50)	1.09 (0.86-1.39)	
Lee, 2005[22]	0.78 (0.58-1.05)	0.90 (0.52-1.58)				
Shen, 2005[25]						0.90 (0.51-1.58)
Overall odds ratio[c]	**0.78 (0.58-1.05)**	**0.90 (0.52-1.58)**	**1.14 (0.88-1.48)**	**0.90 (0.54-1.50)**	**1.09 (0.86-1.39)**	**0.90 (0.51-1.58)**

a Using most adjusted odds ratio where available
b Using G/G individuals as the reference group
c Odds ratio estimated by meta-analysis using random effects model
d Using Ala/Ala individuals as the reference group
e Using Lys/Lys indidivduals as the reference group
f Using -/- individuals as the reference group
g Meta-analysis
h Using C/C individuals as the reference group

Table 3: Overview of the association between lung cancer risk and polymorphisms of the XPD gene

Author, year	Odds ratio (95% confidence interval)[a]								
	Arg156Arg					His201Tyr[b]	Asp312Asn[c]		
	A/A	A/C	C/C	All A allele	All C allele	His/Tyr	Asp/Asn	Asn/Asn	All Asn allele
Buch, 2005[17]							1.50 (0.90-2.20)	1.70 (0.85-3.40)	1.30 (1.00-1.80)
De Ruyck,2007[37]							1.25 (0.67-2.34)	0.96 (0.37-2.48)	
Hu, 2006[29]							1.06 (0.80-1.41)	6.13 (0.67-56.1)	1.10 (0.83-1.46)
Kiyohara[d], 2007[38]							0.95 (0.84-1.07)	1.14 (0.95-1.37)	
Lopez-Cima, 2007[39]							1.01 (0.76-1.35)	1.52 (0.91-2.51)	
Shen, 2005[25]		1.02 (0.52-2.01)[e]	0.47 (0.22-1.01)[e]		0.77 (0.41-1.46)[e]				
Yin, 2005[26]	1.04 (0.75-1.44)[f]	1.12 (0.63-1.97)[f]		1.06 (0.82-1.38)[f]					0.68 (0.20-2.36)[g]
Zienolddiny,2006[35]						1.14 (0.77-1.60)			
Overall odds ratio[h]	**1.04 (0.75-1.44)**	**Not available**	**0.47 (0.22-1.01)**	**1.06 (0.82-1.38)**	**0.77 (0.41-1.46)**	**1.14 (0.77-1.60)**	**1.02 (0.90-1.15)**	**1.25 (1.01-1.54)**	**1.17 (0.96-1.44)**

Table 3 continued

	Lys751Gln[i]			-70C>T[f]		+282A>G[e]		
	Lys/Gln	**Gln/Gln**	**All Gln allele**	**C/T**	**All T allele**	**A/G**	**G/G**	**All G allele**
Buch, 2005[17]			2.05 (1.24-3.28)					
De Ruyck, 2007[37]	0.89 (0.47-1.66)	1.53 (0.56-4.09)						
Escobar, 1999[2]			No association					
Hu, 2006[29]	1.16 (0.89-1.51)	1.64 (0.50-5.36)	1.18 (0.91-1.53)			1.10 (0.86-1.40)	1.15 (0.89-1.51)	1.12 (0.89-1.41)
Kiyohara[d], 2007[38]	1.06 (0.97-1.16)	1.30 (1.13-1.49)						
Lopez-Cima, 2007[39]	1.12 (0.84-1.50)	1.38 (0.85-2.25)						
Shen, 2005[25]				0.44 (0.19-0.99)	0.40 (0.18-0.90)			
Sreeja, 2008[45]	1.80 (1.23-2.81)	1.50 (0.65-3.50)						
Wu, 1999[5]			1.00 (0.30-3.70)[j] 3.00 (0.70-12.50)[k]					
Yin, 2005[26]	2.78 (1.12-6.93)[l]							
Overall odds ratio[h]	**1.20 (0.996-1.45)**	**1.32 (1.16-1.50)**	**1.50 (0.99-2.26)**	**0.44 (0.19-0.99)**	**0.40 (0.18-0.90)**	**1.10 (0.86-1.40)**	**1.15 (0.89-1.51)**	**1.12 (0.89-1.41)**

Table 3 continued

	+58C>T[f]			Rs3916823[m]		
	C/T	**T/T**	**All T allele**	**AAAA/-**	**-/-**	**All - allele**
Hu, 2006[29]	1.12 (0.84-1.49)	6.18 (0.68-56.50)	1.16 (0.87-1.54)	1.10 (0.81-1.25)	1.88 (0.84-4.21)	1.04 (0.84-1.29)
Overall odds ratio[h]	**1.12 (0.84-1.49)**	**6.18 (0.68-56.50)**	**1.16 (0.87-1.54)**	**1.10 (0.81-1.25)**	**1.88 (0.84-4.21)**	**1.04 (0.84-1.29)**

a Using most adjusted odds ratio where available
b Using His/His individuals as the reference group
c Using Asp/Asp individuals as the reference group
d Meta-analysis
e Uaing A/A individuals as the reference group
f Using C/C individuals as the reference group
g Data came from reference [74]
h Odds ratio estimated by meta-analysis using random effects model
i Using Lys/Lys individuals as the reference group
j Mexican Americans
k African Americans
l Data came from reference [75]
m Using AAAA/AAAA individuals as the reference group

Table 4: Overview of association between lung cancer risk and polymorphisms of the XPG gene

Author, year	Odds ratio (95% confidence interval)[a]							
	His46His			Cys529Ser[b]	Leu700Leu[c]	His1104Asp[d]		Rs732321[e]
	C/C	T/C	T/T	Cys/Ser	C/A	His/Asp	Asp/Asp	A/C
Cui, 2006[27]						1.10 (0.80-1.40)	0.65 (0.39-1.10)	
Jeon, 2003[11]						1.20 (0.82-1.75)	0.60 (0.38-0.95)	
Michiels, 2007[40]								2.39 (1.05-5.46)
Sakiyama, 2005[24]						1.00 (0.80-1.30)[g]	1.20 (0.90-1.70)[g]	
						1.10 (0.70-1.60)[h]	1.40 (0.80-2.50)[h]	
Shen, 2005[25]	1.31 (0.55-3.07)[i]	1.54 (0.87-2.74)[i]		1.34 (0.57-3.19)	0.79 (0.30-2.06)	1.08 (0.58-2.01)	1.03 (0.49-2.13)	
Zienolddiny, 2006[35]		0.87 (0.68-1.37)[c]	0.56 (0.38-0.84)[c]					
Overall odds ratio[j]	**1.31 (0.55-3.07)**	**Not available**	**0.56 (0.38-0.84)**	**1.34 (0.57-3.19)**	**0.79 (0.30-2.06)**	**1.08 (0.92-1.26)**	**0.85 (0.57-1.27)**	**2.39 (1.05-5.46)**

	Rs2018836[c]		Rs3759500[c]		Rs3818356[f]		Rs4771436[e]	
	C/T	T/T	C/T	T/T	G/A	A/A	A/C	C/C
Michiels, 2007[40]	0.60 (0.37-0.97)	0.55 (0.22-1.41)	0.62 (0.37-1.03)	0.59 (0.21-1.66)	0.61 (0.37-1.02)	0.58 (0.21-1.64)	0.65 (0.39-1.08)	0.60 (0.21-1.69)
Overall odds ratio[j]	**0.60 (0.37-0.97)**	**0.55 (0.22-1.41)**	**0.62 (0.37-1.03)**	**0.59 (0.21-1.66)**	**0.61 (0.37-1.02)**	**0.58 (0.21-1.64)**	**0.65 (0.39-1.08)**	**0.60 (0.21-1.69)**

a Using most adjusted odds ratio where available
b Using Cys/Cys individuals as the reference group

c Using C/C individuals as the reference group
d Using His/His individuals as the reference group
e Using A/A individuals as the reference group
f Using G/G individuals as the reference group
g Adenocarcinoma cases
h Squamous cell carcinoma cases
i Using T/T individuals as the reference group
j Odds ratio estimated by meta-analysis using random effects model

Table 5: Overview of association between lung cancer risk and polymorphisms in XRCC1 gene

Author, year	Odds ratio (95% confidence interval)[a]								
	-77T>C[b]			Arg194Trp[c]			Pro206Pro[d]		
	C/T	C/C	All C allele	Arg/Trp	Trp/Trp	All Trp allele	A/G	G/G	All G allele
Butkiewicz, 2001[7]						No association			
Chan, 2005[18]				0.52 (0.29-0.94)	0.30 (0.08-1.07)	0.48 (0.27-0.84)			
De Ruyck, 2007[37]	1.12 (0.59-2.12)	1.12 (0.48-2.58)		0.32 (0.12-0.86)					
Hao, 2006[28]	1.44 (1.16-1.80)	1.87 (0.87-4.01)	1.46 (1.18-1.82)	0.98 (0.82-1.18)	1.11 (0.80-1.54)	1.00 (0.84-1.20)			
Hu, 2005[20]	1.51 (1.17-1.94)	2.98 (0.93-9.59)	1.55 (1.21-1.98)	1.01 (0.81-1.26)	1.11 (0.75-1.63)	1.03 (0.83-1.27)			
Kiyohara[e], 2006[30]				0.89 (0.78-1.03)	1.19 (0.76-1.86)				
Li, 2008[43]	1.51 (1.01-2.24)	3.14 (0.96-10.30)	1.61 (1.12-2.39)	0.94 (0.68-1.30)	1.53 (0.84-2.81)				
Matullo, 2006[32]							1.53 (0.90-2.60)	0.81 (0.41-1.60)	
Pachouri, 2007[41]				1.00 (0.75-1.45)	1.30 (0.63-2.92)	1.10 (0.66-2.07)			
Spitz, 2003[12]					1.02 (0.72-1.44)[f]				
Yin, 2007[42]						0.97 (0.67-1.40)			1.96 (1.26-3.06)
Overall odds ratio[g]	**1.45 (1.25-1.69)**	**1.86 (1.17-2.98)**	**1.52 (1.31-1.76)**	**0.91 (0.80-1.04)**	**1.14 (0.93-1.40)[h]**	**0.95 (0.80-1.14)**	**1.53 (0.90-2.60)**	**0.81 (0.41-1.60)**	**1.96 (1.26-3.06)**

Table 5 continued

	Arg280His[c]			Arg399Gln[c]			Gln632Gln[i]
	Arg/His	**His/His**	**All His allele**	**Arg/Gln**	**Gln/Gln**	**All Gln allele**	**All A allele**
Butkiewicz, 2001[7]			No association			No association	
Chan, 2005[18]				1.14 (0.65-2.02)	0.82 (0.25-2.73)	1.09 (0.63-1.90)	
De Ruyck, 2007[37]	0.25 (0.07-0.86)			1.44 (0.76-2.69)	1.62 (0.66-3.98)		
Hao, 2006[28]	0.90 (0.71-1.13)	0.72 (0.27-1.93)	0.89 (0.71-1.12)	0.89 (0.74-1.07)	0.86 (0.62-1.18)	0.88 (0.74-1.05)	
Hu, 2005[20]				0.98 (0.78-1.23)	0.83 (0.54-1.25)	0.95 (0.77-1.18)	
Kiyohara[e], 2006[30]	1.03 (0.88-1.20)		1.06 (0.91-1.23)	0.99 (0.93-1.06)	1.02 (0.88-1.19)		
Li, 2005[23]				0.99 (0.40-2.46)	5.43 (0.99-29.70)	0.73 (0.31-1.72)	
Li, 2008[43]	1.15 (0.80-1.67)	1.78 (0.47-6.75)		1.26 (0.91-1.75)	1.73 (1.01-2.97)		
Liu, 2004[15]				1.04 (0.84-1.29)	1.27 (0.92-1.75)		
Lopez-Cima, 2007[39]				0.86 (0.63-1.16)	0.87 (0.57-1.31)		
Pachouri, 2007[41]				0.30 (0.19-0.67)	0.40 (0.18-1.18)	0.60 (0.46-0.80)	
Ryk, 2006[33]						0.81 (0.52-1.25)	
Spitz, 2003[12]				0.97 (0.74-1.26)	0.80 (0.55-1.16)	0.92 (0.72-1.18)	
Yin, 2007[42]			0.86 (0.69-1.07)			0.97 (0.78-1.20)	0.96 (0.76-1.21)
Overall odds ratio[g]	**0.97 (0.78-1.20)**	**1.01 (0.43-2.38)**	**0.95 (0.83-1.10)**	**0.97 (0.86-1.08)**	**1.00 (0.83-1.20)**	**0.87 (0.78-0.98)**	**0.96 (0.76-1.21)**

a Using most adjusted odds ratio where available

b Using T/T individuals as the reference group
c Using Arg/Arg individuals as the reference group
d Using A/A individuals as the reference group
e Meta-analysis
f Using Arg/Arg and Arg/Trp individuals as the reference group
g Odds ratio estimated by meta-analysis using random effects model
h Excludes one study[12] as the reference group was all subjects with Arg allele rather than Arg homozygotes
i Using G/G individuals as the reference group

GLOSSARY

Unfortunately there are several terminologies in use to describe polymorphisms and related topics: The following is a list of those relating to the reviews.

A **karyotype** is the formula that describes the composition of a chromosome. It is prepared by arresting cells at metaphase and staining with a Giesma stain. Bands are produced. Position of genes along chromosome arms are defined by **region** numbers from the centromere outwards, **band**, **sub-band** and **sub-sub** numbers being allocated, *e.g.* 12q24.32 refers to Chromosome 12, long arm(q), region2, band 4, sub-band 3, sub-sub band 2.

Allele: One of the alternative versions of a gene that may occupy a given locus.

Codon: A triplet of three bases in a DNA or RNA molecule, specifying a single amino acid.

Exon: A transcribed region of a gene that is present in messenger DNA.

Haplotype: A set of alleles of linked genes that tend to be inherited together.

Intron: A segment of a gene that is initially transcribed but is then removed from within the primary RNA transcript by splicing together the sequences (exons) on either side of it.

Linkage Disequilibrium: Co-occurrence of closely-linked alleles in a population more frequently than expected by chance.

Polymorphism: A polymorphism is specified first by the gene, then by the amino acid present in one allele, then by the Codon number, then by the amino acid present in the other allele, for example OGG1-Ser326Cys. Alternatively the polymorphism may be identified by indicating the base change, for example **XPA-4AA** compared to **XPA-AT**. In this case the consequences of the polymorphism are not necessarily known. A final method is to identify the polymorphism by the amino acid derived from the codon of three bases which include the polymorphism . For example one allele can be expressed as **XPGQ622H** using the amino acid code where Q=glutamine and H=Histidine. A polymorphism is **synchronous** if the change does not alter the amino acid derived from the appropriate codon, due to the redundancy of the genetic code. It is **non-synchronous** if it does result in a change. The **rs** number is a **r**eference **s**ingle nucleotide polymorphism used to give an approximate location of a gene of interest, typically in genome wide association studies. For example

rs1799793, an approximate location of XPD-Asp312Asn. By convention a polymorphism is a change in a single nucleotide, and usually refers to inherited nucleotide changes. By convention a mutation is frequently regarded as a change in a nucleotide (or other event) caused by some external factor.

Penetrance is the frequency with which a person carrying a particular genotype will manifest a disease.

References

1. Kohno T, Shinmura K, Tosaka M, Tani M, Kim SR, Sugimura H, *et al*. Genetic polymorphisms and alternative splicing of the *hOGG1* gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA. *Oncogene* 1998;**16**:3219-25.

2. Escobar P, Modugno F, Kanbour-Shakir A, Naus G, Mohrenweiser H, Romkes M. Analysis of the DNA repair enzyme XPD exon 23 polymorphism and cancer risk [Abstract]. *Proc Am Assoc Cancer Res* 1999;**40**:213.

3. Ishida T, Takashima R, Fukayama M, Hamada C, Hippo Y, Fujii T, *et al*. New DNA polymorphisms of human *MMH/OGG1* gene: prevalence of one polymorphism among lung-adenocarcinoma patients in Japanese. *Int J Cancer* 1999;**80**:18-21.

4. Sugimura H, Kohno T, Wakai K, Nagura K, Genka K, Igarashi H, *et al*. *hOGG1* Ser326Cys polymorphism and lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 1999;**8**:669-74.

5. Wu X, Amos CI, Hong WK, Mohrenweiser HW, Spitz MR. Association of a polymorphism in exon 23 of the DNA repair gene, XPD, and lung cancer risk. *Proc Am Assoc Cancer Res* 1999;**40**:89.

6. Wikman H, Risch A, Klimek F, Schmezer P, Spiegelhalder B, Dienemann H, *et al*. *hOGG1* polymorphism and loss of heterozygosity (LOH): significance for lung cancer susceptibility in a caucasian population. *Int J Cancer* 2000;**88**:932-7.

7. 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.

8. Ito H, Hamajima N, Takezaki T, Matsuo K, Tajima K, Hatooka S, *et al*. A limited association of OGG1 Ser326Cys polymorphism for adenocarcinoma of the lung. *J Epidemiol* 2002;**12**:258-65.

9. Le Marchand L, Donlon T, Lum-Jones A, Seifried A, Wilkens LR. Association of the *hOGG1* Ser326Cys polymorphism with lung cancer risk. *Cancer Epidemiol Biomarkers Prev* 2002;**11**:409-12.

10. Sunaga N, Kohno T, Yanagitani N, Sugimura H, Kunitoh H, Tamura T, *et al*. Contribution of the *NQ01* and *GSTT1* polymorphisms to lung adenocarcinoma susceptibility. *Cancer Epidemiol Biomarkers Prev* 2002;**11**:730-8.

11. Jeon HS, Kim KM, Park SH, Lee SY, Choi JE, Lee GY, *et al*. Relationship between XPG codon 1104 polymorphism and risk of primary lung cancer. *Carcinogenesis* 2003;**24**:1677-81.

12. Spitz MR, Wei Q, Dong Q, Amos CI, Wu X. Genetic susceptibility to lung cancer: the role of DNA damage and repair. *Cancer Epidemiol Biomarkers Prev* 2003;**12**:689-98.

13. Wang YG, Xing DY, Tan W, Wang LJ, Tang PZ, Lin DX. (Poly(AT) polymorphism in DNA repair gene XPC and lung cancer risk). *Zhonghua Zhong Liu Za Zhi* 2003;**25**:555-7.

14. Lan Q, Mumford JL, Shen M, Demarini DM, Bonner MR, He X, *et al*. Oxidative damage-related genes *AKR1C3* and *OGG1* modulate risks for lung cancer due to exposure to PAH-rich coal combustion emissions. *Carcinogenesis* 2004;**25**:2177-81.

15. Liu G, Zhou W, Park S, Wang LI, Miller DP, Wain JC, *et al*. The *SOD2 Val/Val* genotype enhances the risk of nonsmall cell lung carcinoma by *p53* and *XRCC1* polymorphisms. *Cancer* 2004;**101**:2802-8.

16. Park J, Chen L, Tockman MS, Elahi A, Lazarus P. The human 8-oxoguanine DNA *N*-glycosylase 1 (hOGG1) DNA repair enzyme and its association with lung cancer risk. *Pharmacogenetics* 2004;**14**:103-9.

17. Buch S, Zhu B, Davis AG, Odom D, Siegfried JM, Grandis JR, *et al*. Association of polymorphisms in the *Cyclin D1* and *XPD* genes and susceptibility to cancers of the upper aero-digestive tract. *Mol Carcinog* 2005;**42**:222-8.

18. Chan EC, Lam SY, Fu KH, Kwong YL. Polymorphisms of the *GSTM1, GSTP1, MPO, XRCC1*, and *NQO1* genes in Chinese patients with non-small cell lung cancers: relationship with aberrant promoter methylation of the *CDKN2A* and *RARB* genes. *Cancer Genet Cytogenet* 2005;**162**:10-20.

19. Hu Z, Wang Y, Wang X, Liang G, Miao X, Xu Y, *et al*. DNA repair gene *XPC* genotypes/haplotypes and risk of lung cancer in a Chinese population. *Int J Cancer* 2005;**115**:478-83.

20. Hu Z, Ma H, Lu D, Zhou J, Chen Y, Xu L, *et al*. A promoter polymorphism (-77T>C) of DNA repair gene *XRCC1* is associated with risk of lung cancer in relation to tobacco smoking. *Pharmacogenet Genomics* 2005;**15**:457-63.

21. Hung RJ, Brennan P, Canzian F, Szeszenia-Dabrowska N, Zaridze D, Lissowska J, *et al*. Large-scale investigation of base excision repair genetic polymorphisms and lung cancer risk in a multicenter study. *J Natl Cancer Inst* 2005;**97**:567-76.

22. Lee GY, Jang J-S, Lee SY, Jeon H-S, Kim KM, Choi JE, *et al*. *XPC* polymorphisms and lung cancer risk. *Int J Cancer* 2005;**115**:807-13.

23. Li MC, Cui ZS, He QC, Zhou BS. (Association of genetic polymorphism in the DNA repair gene XRCC1 with susceptibility to lung cancer in non-smoking women). *Zhonghua Zhong Liu Za Zhi* 2005;**27**:713-6.

24. Sakiyama T, Kohno T, Mimaki S, Ohta T, Yanagitani N, Sobue T, *et al*. Association of amino acid substitution polymorphisms in DNA repair genes *TP53, POLI, REV1* and *LIG4* with lung cancer risk. *Int J Cancer* 2008;**114**:730-7.

25. Shen M, Berndt SI, Rothman N, Demarini DM, Mumford JL, He X, *et al*. Polymorphisms in the DNA nucleotide excision repair genes and lung cancer risk in Xuan Wei, China. *Int J Cancer* 2005;**116**:768-73.

26. Yin J, Li J, Ma Y, Guo L, Wang H, Vogel U. The DNA repair gene *ERCC2/XPD* polymorphism Arg 156Arg (A22541C) and risk of lung cancer in a Chinese population. *Cancer Lett* 2005;**223**:219-26.

27. Cui Y, Morgenstern H, Greenland S, Tashkin DP, Mao J, Cao W, *et al*. Polymorphism of *xeroderma pigmentosum group G* and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. *Int J Cancer* 2006;**118**:714-20.

28. Hao B, Miao X, Li Y, Zhang X, Sun T, Liang G, *et al*. A novel T-77C polymorphism in DNA repair gene *XRCC1* contributes to diminished promoter activity and increased risk of non-small cell lung cancer. *Oncogene* 2006;**25**:3613-20.

29. Hu Z, Xu L, Shao M, Yuan J, Wang Y, Wang F, *et al*. Polymorphisms in the two helicases *ERCC2/XPD* and *ERCC3/XPB* of the transcription factor IIH complex and risk of lung cancer: a case-control analysis in a Chinese population. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:1336-40.

30. Kiyohara C, Yoshimasu K, Takayama K, Nakanishi Y. *EPHX1* polymorphisms and the risk of lung cancer: a HuGE review. *Epidemiology* 2006;**17**:89-99.

31. Kohno T, Kunitoh H, Toyama K, Yamamoto S, Kuchiba A, Saito D, *et al*. Association of the *OGG1*-Ser326Cys polymorphism with lung adenocarcinoma risk. *Cancer Sci* 2006;**97**:724-8.

32. Matullo G, Dunning AM, Guarrera S, Baynes C, Polidoro S, Garte S, *et al*. DNA repair polymorphisms and cancer risk in non-smokers in a cohort study. *Carcinogenesis* 2006;**27**:997-1007.

33. Ryk C, Kumar R, Thirumaran RK, Hou S-M. Polymorphisms in the DNA repair genes *XRCC1*, *APEX1*, *XRCC3* and *NBS1*, and the risk for lung cancer in never- and ever-smokers. *Lung Cancer* 2006;**54**:285-92.

34. Sørensen M, Raaschou-Nielsen O, Hansen RD, Tjønneland A, Overvad K, Vogel U. Interactions between the *OGG1* Ser326Cys polymorphism and intake of fruit and vegetables in relation to lung cancer. *Free Radic Res* 2006;**40**:885-91.

35. Zienolddiny S, Campa D, Lind H, Ryberg D, Skaug V, Stangeland L, *et al*. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. *Carcinogenesis* 2006;**27**:560-7.

36. Bai Y, Xu L, Yang X, Hu Z, Yuan J, Wang F, *et al*. Sequence variations in DNA repair gene *XPC* is associated with lung cancer risk in a Chinese population: a case-control study. *BMC Cancer* 2007;**7**:81.

37. De Ruyck K, Szaumkessel M, De R, I, Dehoorne A, Vral A, Claes K, *et al*. Polymorphisms in base-excision repair and nucleotide-excision repair genes in relation to lung cancer risk. *Mutat Res* 2007;**631**:101-10.

38. Kiyohara C, Yoshimasu K. Genetic polymorphisms in the nucleotide excision repair pathway and lung cancer risk: a meta-analysis. *Int J Med Sci* 2007;**4**:59-71.

39. López-Cima MF, González-Arriaga P, García-Castro L, Pascual T, Marrón MG, Puente XS, *et al*. 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.

40. Michiels S, Danoy P, Dessen P, Bera A, Boulet T, Bouchardy C, *et al*. Polymorphism discovery in 62 DNA repair genes and haplotype associations with risks for lung and head and neck cancers. *Carcinogenesis* 2007;**28**:1731-939.

41. Pachouri SS, Sobti RC, Kaur P, Singh J. Contrasting impact of DNA repair gene *XRCC1* polymorphisms *Arg399Gln* and *Arg194Trp* on the risk of lung cancer in north-Indian population. *DNA Cell Biol* 2007;**26**:186-91.

42. Yin J, Vogel U, Ma Y, Qi R, Sun Z, Wang H. The DNA repair gene *XRCC1* and genetic susceptibility of lung cancer in a northeastern Chinese population. *Lung Cancer* 2007;**56**:153-60.

43. Li M, Yin Z, Guan P, Li X, Cui Z, Zhang J, *et al*. XRCC1 polymorphisms, cooking oil fume and lung cancer in Chinese women nonsmokers [In press]. *Lung Cancer* 2008;
WARNING: doi: 10.1016/j.lungcan.2008.03.002

44. Raaschou-Nielsen O, Sørensen M, Overvad K, Tjønneland A, Vogel U. Polymorphisms in nucleotide excision repair genes, smoking and intake of fruit and vegetables in relation to lung cancer. *Lung Cancer* 2008;**59**:171-9.

45. Sreeja L, Syamala VS, Syamala V, Hariharan S, Raveendran PB, Vijayalekshmi RV, *et al*. Prognostic importance of DNA repair gene polymorphisms of *XRCC1* Arg399Gln and *XPD* Lys751Gln in lung cancer patients from India. *J Cancer Res Clin Oncol* 2008;**134**:645-52.

46. Zhang D, Chen C, Fu X, Gu S, Mao Y, Xie Y, *et al*. A meta-analysis of DNA repair gene XPC polymorphisms and cancer risk. *J Hum Genet* 2008;**53**:18-33.

47. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;**40**:695-701.

48. Mohrenweiser HW, Xi T, Vázquez-Matiás J, Jones IM. Identification of 127 amino acid substitutions variants in screening 37 DNA repair genes in humans. *Cancer Epidemiol Biomarkers Prev* 2002;**11**:1054-64.

49. Danoy P, Michiels S, Dessen P, Pignat C, Boulet T, Monet M, *et al*. Variants in DNA double-strand break repair and DNA damage-response genes and susceptibility to lung and head and neck cancers. *Int J Cancer* 2008;**123**:457-63.

50. Popanda O, Schattenberg T, Phong CT, Butkiewicz D, Risch A, Edler L, *et al*. Specific combinations of DNA repair gene variants and increased risk for non-small cell lung cancer. *Carcinogenesis* 2004;**25**:2433-41.

51. Manuguerra M, Matullo G, Veglia F, Autrup H, Dunning AM, Garte S, *et al*. Multi-factor dimensionality reduction applied to a large prospective investigation on gene-gene and gene-environment interactions. *Carcinogenesis* 2007;**28**:414-22.

52. Matakidou A, el Galta R, Webb EL, Rudd MF, Bridle HB, Eisen T, *et al*. Genetic variation in the DNA repair genes is predictive of outcome in lung cancer. *Hum Mol Genet* 2007;**16**:2333-40.

53. Yin J, Li J, Vogel U, Wang H. Polymorphisms of DNA repair genes: *ERCC1* G19007A and *ERCC2/XPD* C22541A in a northeastern Chinese population. *Biochem Genet* 2005;**43**:543-8.

54. Vogel U, Laros I, Jacobsen NR, Thomsen BL, Bak H, Olsen A, *et al*. Two regions in chromosome 19q13.2-3 are associated with risk of lung cancer. *Mutat Res* 2004;**546**:65-74.

55. Vogel U, Overvad K, Wallin H, Tjønneland A, Nexø BA, Raaschou-Nielsen O. Combinations of polymorphisms in *XPD*, *XPC* and *XPA* in relation to risk of lung cancer. *Cancer Lett* 2005;**222**:67-74.

56. Yin J, Vogel U, Ma Y, Qi R, Wang H. Haplotypes of nine single nucleotide polymorphisms on chromosome 19q13.2-3 associated with susceptibility of lung cancer in a Chinese population. *Mutat Res* 2008;**641**:12-8.

57. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, *et al*. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;**40**:616-22.

58. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, *et al*. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor submit genes on 15q25. *Nature* 2008;**452**:537-8.

59. Thorgeirsson GE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, *et al*. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 2008;**452**:638-42.

60. Brennan P. Lessons from genome-scans-the example of lung cancer. In: *Integrative Molecular Cancer Epidemiology, Lyon, 2008*. 2008; 9-IS.

61. Bailey-Wilson JE, Amos CI, Pinney SM, Petersen GM, de Andrade M, Wiest JS, *et al*. A major lung cancer susceptibility locus maps to chromosome 6q23-25. *Am J Hum Genet* 2004;**75**:460-74.

62. Bailey-Wilson J. Lung cancer susceptibility genes. In: Roth J, Cox J, Hong W, editors. *Lung cancer*, 3rd Ed. Oxford: Blackwell, 2008;20-32.

63. Matakidou A, Eisen T, Fleischmann C, Bridle H, Houlston RS. Evaluation of xeroderma pigmentosum *XPA, XPC, XPD, XPF, XPB, XPG* and *DDB2* genes in familial early-onset lung cancer predisposition. *Int J Cancer* 2006;**119**:964-7.

64. Greenman C, Stephens P, Smith R, Dalgliesh G, Hunter C, Bignell G, *et al*. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;**446**:153-8.

65. Weir B, Woo M, Getz G, Perner S, Ding I, Beroukhim R, *et al*. Characterizing the cancer genome in lung adenocarcinoma. *Nature* 2007;**450**:893-8.

66. Spira A, Beane J, Shah V, Steiling K, Liu G, Schembri F, *et al*. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 2007;**13**:361-6.

67. Mechanic LE, Marrogi AJ, Welsh JA, Bowman ED, Khan MA, Enewold L, *et al*. Polymorphisms in *XPD* and *TP53* and mutation in human lung cancer. *Carcinogenesis* 2005;**26**:597-604.

68. Wu Y-H, Tsai Chang J-H, Cheng Y-W, Wu T-C, Chen C-Y, Lee H. Xeroderma pigmentosum group C gene expression is predominantly regulated by promoter hypermethylation and contributes to p53 mutation in lung cancers. *Oncogene* 2007;**26**:4761-73.

69. Qiao Y, Spitz MR, Shen H, Guo Z, Shete S, Hedayati M, *et al*. Modulation of repair of ultraviolet damage in the host-cell reactivation assay by olymorphic *XPC* and *XPD/ERCC2* genotypes. *Carcinogenesis* 2002;**23**:295-9.

70. Pastorelli R, Cerri A, Mezzetti M, Consonni E, Airoldi L. Effect of DNA repair gene polymorphisms on BPDE-DNA adducts in human lymphocytes. *Int J Cancer* 2002;**100**:9-13.

71. Mateuca RA, roelants M, Iarmarcovai G, Aka PV, Godderis L, tremp A, *et al*. *hOGG1*^{326}, *XRCC1*^{399} and *XRCC3*^{241} polymorphisms influence micronucleus frequencies in human lymphocytes *in vivo*. *Mutagenesis* 2008;**23**:35-41.

72. Petretto E, Edison TL, Aitman TJ. A gene harvest revealing the archeology and complexity of human disease. *Nat Genet* 2007;**39**:1299-301.

73. Risch A, Plass C. Lung cancer epigenetics and genetics. *Int J Cancer* 2008;**123**:1-7.

74. Yin J, Vogel U, Ma Y, Qi R, Sun Z, Wang H. A haplotype encompassing the variant allele of DNA repair gene polymorphism *ERCC2/XPD* Lys751Gln but not the variant allele of Asp312Asn is associated with risk of lung cancer in a northeastern Chinese population. *Cancer Genet Cytogenet* 2007;**175**:47-51.

75. Yin J, Vogel U, Ma Y, Guo L, Wang H, Qi R. Polymorphism of the DNA repair gene *ERCC2* Lys751Gln and risk of lung cancer in a northeastern Chinese population. *Cancer Genet Cytogenet* 2006;**169**:27-32.