Polymorphisms within base and nucleotide excision repair pathways and risk of differentiated thyroid carcinoma.

Monica Cipollini[a*], Gisella Figlioli[a*], Giuseppe Maccari[b], Sonia Garritano[c], Chiara De Santi[a], Ombretta Melaiu[a], Elisa Barone[a], Franco Bambi[d], Stefano Ermini[d], Giovanni Pellegrini[e], Alfonso Cristaudo[f], Rudy Foddis[f], Alessandra Bonotti[f], Cristina Romei[f], Agnese Vivaldi[e], Laura Agate[f], Eleonora Molinari[f], Roberto Barale[a], Asta Forsti[g,h], Kari Hemminki[g,h], Rossella Elisei[f*], Federica Gemignani[a*]*, Stefano Landi[a*].

Affiliations:

[a] Department of Biology, University of Pisa, Pisa, Italy.

[b] Center for Nanotechnology and Innovation @NEST, Istituto Italiano di Tecnologia, Piazza San Silvestro Pisa, Italy.

[c] Center for Integrated Biology, University of Trento, Trento, Italy.

[d] Blood Centre of University Hospital of Meyer, Florence, Italy.

[e] Operative Unit of laboratory of Clinical Chemistry Analyses, University Hospital of Cisanello, Pisa, Italy.

[f] Department of Endocrinology and Metabolism, Orthopaedics and Traumatology, Occupational Medicine, University of Pisa, Pisa, Italy.

[g] Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.

[h] Center for Primary Health Care Research, Clinical Research Center, Lund University, Malmö, Sweden.

* equal contribution

Corresponding Author: Stefano Landi, University Of Pisa, via Derna, 1 Pisa, 56126, Italy.

Phone: 39-050-2211528; Fax: 39-050-2211527;

E-mail: stefano.landi@unipi.it

ABSTRACT

The thyrocytes are exposed to high levels of oxidative stress which could induce DNA damages. Base excision repair (BER) is one of the principal mechanisms of defense against oxidative DNA damage, however recent evidences suggest that also nucleotide excision repair (NER) could be involved. The aim of present work was to identify novel differentiated thyroid cancer (DTC) risk variants in BER and NER genes. For this purpose, the most strongly associated SNPs within NER and BER genes found in our previous GWAS on DTC were selected and replicated in an independent series of samples for a new case-control study. Although a positive signal was detected at the nominal level of 0.05 for rs7689099 (encoding for an aminoacid change proline to arginine at codon 117 within *NEIL3*), none of the considered SNPs (i.e. rs7990340 and rs690860 within *RFC3*, rs3744767 and rs1131636 within *RPA1*, rs16962916 and rs3136166 in *ERCC4*, and rs17739370 and rs7689099 in *NEIL3*) was associated with the risk of DTC when the correction of multiple testing was applied. In conclusion, a role of NER and BER pathways was evoked in the susceptibility to DTC. However, this seemed to be limited to few polymorphic genes and the overall effect size appeared weak.

Keywords: Polymorphism, Differentiated thyroid cancer risk, BER, NER

1. INTRODUCTION

Thyroid carcinoma (TC) is a rare tumor but it is the most frequent malignancy of the endocrine system. DTC is the most prevalent type [1] and its incidence is increasing worldwide [2]. Within DTC, the papillary type (PTC) is the most common, whereas the follicular type (FTC) is observed only in 5-10% of cases. During the biosynthesis of thyroid hormones a high quantity of hydrogen peroxide (H_2O_2) is generated on the apical membrane of thyrocytes by dual oxidase [3] and reactive oxygen species (ROS) are generated by NOX4 in the endoplasmic reticulum and nuclear membrane [4]. Thus, it is conceivable that thyrocytes are among the most exposed types of cells to oxidative stress. For example, Maier et al. (2006) showed that the rat thyroid gland has a high level of DNA oxidative damage in comparison with other tissues, such as liver, spleen, and lung [5]. This study also suggested that the tumor-initiating somatic mutations could be due to the oxidative microenvironment to which thyroid cells are subjected during their life. The carcinogenic effects of H_2O_2 on thyrocytes have been clearly demonstrated in different *in vitro* studies. When a rat thyroid cell line (PCCL3) was incubated with nonlethal H_2O_2 concentrations, the number of single- and double-strand breaks in DNA increased as did the phosphorylation of histone H2AX, a marker of double-strand breaks [6]. Moreover, H_2O_2 exposure induced *RET/PTC* rearrangements in human thyroid cell lines, which was suppressed when catalase was added to the incubation medium [7]. Thus, it could be hypothesized that the metabolic stress linked to thyroid hormones biosynthesis causes somatic mutations and, then, constitutes a risk factor for developing DTC in humans [8].

Base Excision Repair (BER) is one of the most important lines of defense against oxidative DNA damages in nucleus and mitochondria [9]. It removes oxidized bases producing a-basic sites and it has a role also in repairing single-strand breaks in DNA. However, cumulating evidences show that also Nucleotide Excision Repair (NER) could play a role [10]. Defects in Cockayne syndrome proteins A and B (CSA and CSB, respectively) are thought to cause an accumulation of ROS, thereby contributing to the premature aging phenotype of the disease [11]. Moreover, CSB-

defective cell lines from rodents or patients appeared hypersensitive to ionizing radiations as well as various types of oxidative stress [12-14]. In addition, rodent Xp-C defective cells were more sensitive, in terms of survival and mutations' accumulation, to ROS than wild-type cells [15]. Furthermore, studies using human primary keratinocytes and fibroblasts from XP-C patients showed that XPC protects human skin cells from the killing effects of oxidants and X-rays [16]. Finally, in normal human keratinocytes the down-regulation of XPC resulted in increased intracellular ROS levels and genomic oxidation [17].

Previous studies suggested an association between DTC and SNPs within the NER genes *ERCC2* [18], *CCNH*, *XPC*, and *ERCC5* [19]. Moreover, a meta-analysis has shown an association between DTC risk and rs25489 (G>A) within the BER gene *XRCC1* [20]. However, the role played by polymorphisms within DNA repair genes in relation to the risk of DTC has not been fully evaluated. Positive signals within NER/BER genes in a Genome-Wide Association Study (GWAS) (carried out by our research group, [21]) did not exceed the stringent genome-wide statistical significance threshold, thus the results remained inconclusive (see suppl. table 1). In the present work we performed a hypothesis-driven study aimed to replicate the strongest associations of the GWAS with the goal of better defining whether polymorphic genes within NER and BER pathways could play a role in modulating the risk of DTC in humans.

2. MATERIAL AND METHODS

2.1 Study population

Cases were DTC patients followed at the Department of Endocrinology of the University Hospital of Pisa, Italy. Controls were blood donors collected at Meyer Hospital of Florence (Italy) and healthy subjects working at the University Hospital of Pisa recruited during their routine visits in the context of a program of surveillance performed by the Occupational Medicine unit. In the present study, consecutively recruited 1,500 cases and 1,500 controls were collected with participation rates of about 95% and 85%, respectively. The genotyping for replicating the results of

the GWAS were carried out on the last consecutive 450 cases and 450 controls. The eligibility criteria were the same for cases and controls, and it included a minimal age of 18 years. Volunteers affected by any malignancy, chronic inflammatory disease, or related diseases in the past were excluded. A further exclusion criterion for controls was the presence of thyroid nodules or other benign thyroid diseases, when known. Both cases and controls should not have familial relationship with each other. All participants were of Caucasian origins. According to the Helsinki declaration, both healthy and affected volunteers gave their written informed consent to participate in the study and the study protocol was cleared by the local Ethical Committee.

2.2 Candidate genes and SNPs selection

In order to retrieve a list of genes coding for BER and NER involved proteins, three publicly available online databases were investigated: KEGG (Kyoto Encyclopedia of Genes and Genome, http://www.genome.jp/kegg), CGAP (Cancer Genome Anatomy Project, (http://cgap.nci.nih.gov) and Gene Ontology (http://www.geneontology.org). Among a broad list of 144 genes, 104 were described as belonging to these pathways in at least two databases and 979 SNPs, belonging to these latter genes, were analyzed in association with DTC in a GWAS performed by our research group [21]. For each gene a regional plot was generated using LocusZoom (http://csg.sph.umich.edu/locuszoom/). Genes with at least five haplotype-tagging SNPs (not in pair-wise linkage disequilibrium, $r^2<0.8$, with each other) associated with the risk of DTC at allelic P-value level of 0.05 and minor allele frequency (MAF)>0.15 were selected and, for each of them, the two SNPs with the lowest allelic P-value were replicated (i.e. *RFC3*: rs7990340, rs690860; *RPA1*: rs3744767, rs1131636; *ERCC4*: rs16962916, rs3136166; *NEIL3*: rs17739370, rs7689099).

2.3 DNA extraction and SNP genotyping

DNA was extracted from peripheral blood samples otherwise disposed following the analyses of clinical routine. Puregene Blood Kit (Gentra Systems, Inc., Minneapolis, MN) was used for the extractions. Samples from cases and controls were randomized and mixed on PCR plates, so that an

equal number of cases and controls could be analyzed simultaneously by personnel blinded for the case/control status. Genotyping for all the SNPs was carried out using pre-designed TaqMan SNP Genotyping Assays, according to protocol specified by the manufacturer (Life Technologies Inc., Grand Island, New York, USA). Two percent of DNA samples were re-genotyped as quality control.

2.4 Statistical analyses

Hardy–Weinberg (H-W) equilibrium in controls was tested for each polymorphism by the Chi-square (χ^2) test. A logistic regression analysis was used to examine the associations between the genotypes and the considered risk factors. The association analyses were based on the estimation of the Odd Ratios (ORs) and of their 95% confidence intervals (CIs). Genotypes were analyzed with a multivariate logistic regression (MLR) model, allowing the ORs to be adjusted for covariates (age and sex) as linear variables (the OR_{adj}). Each genotype category was compared using the common homozygotes as reference category and the P-value of the association (P_{ass}) was calculated separately for heterozygotes and homozygotes. The False Discovery Rate (FDR) was employed to correct for multiple testing. The software Statgraphics Centurion (StatPoint Inc., USA) was used.

2.5 Computational analyses

To assess the possible functional role of each SNP in relation to putative cis-eQTL, we examined the eQTL dataset available for thyroid tissues within the GTEx Portal (http://www.gtexportal.org/home/). For *in silico* analysis of rs7689099, the functional effect on the encoded protein was evaluated with: Polyphen2 (http://genetics.bwh.harvard.edu/pph2/) [22], SIFT (PROVEAN) (http://sift.jcvi.org/), [23] and SNPs3D (http://www.snps3d.org/) [24].

3. RESULTS

3.1 Case-control association study

The characteristics of the population employed in the case-control association study are reported in table 1. The group of patients showed a higher proportion of females (OR=5.11; 95% CI=4.23-6.18, $P_{ass}<10^{-6}$) and a lower average age when compared to controls (45.8 and 50.5 and, respectively, $P<10^{-5}$). These covariates were taken into account in the MLR analysis, thus providing adequately adjusted ORs. Genotypes were in agreement with H-W equilibrium and passed the quality control (>99.5% of the re-analyzed samples confirmed their genotype). The results of the previous GWAS are summarized in figure 1, where top 13 genes were ranked using the SNPs with the lowest P-association for each of them (the remaining genes are reported as supplementary table 1). The OR_{adj}, the 95% CIs, and their P_{ass} following MLR analysis of the additive, dominant, and recessive model are reported in table 2 for the replication study. The heterozygotes CT for rs1131636 within the 3'UTR of *RPA1* were associated with a reduced risk of DTC (OR=0.73, 95% CI 0.54-0.98; P_{ass}=0.04), but the homozygotes CC showed a risk very similar to the reference category (TT homozygotes) in the additive model, suggesting that the variant allele rs1131636-C is very unlikely associated with the risk of DTC.

However, two SNPs appeared to be statistically significant when using the classical threshold of 0.05 as type-I error. TT homozygotes of rs17739370 within *NEIL3* showed an increased risk of DTC: the OR was 1.76 (95% CI=1.13-2.72; P_{ass}=0.01) when the common homozygotes CC were used as reference category in the additive model, or it was 1.92 (95% CI=1.07-2.43; P_{ass}=0.02) when the CC+CT grouped genotypes were used as reference category in the recessive model. Moreover, GG homozygotes of rs7689099 within the coding region of *NEIL3* showed a reduced risk of DTC in both the additive (OR=0.55, 95%CI=0.31-0.98; P_{ass}=0.042) and recessive (OR=0.56, 95%CI=0.31-1.00; P_{ass}=0.05) models. However, after applying the FDR correction for multiple testing, none of the SNPs reached a statistical significance.

3.2 Computational analysis

The genotyped polymorphisms were evaluated for their association with the mRNA expression levels (intended as quantitative trait loci, cis-eQTL) in 112 thyroid tissues available within GTex portal. None of the SNPs was associated with a differential mRNA expression when the correction for multiple testing was applied. Moreover, rs7689099 within *NEIL3* is a coding SNP causing the aminoacid change Pro>Arg at the codon 117 and this change was predicted by Polyphen2, SIFT (PROVEAN), and SNPs3D to affect the protein function.

4. DISCUSSION

GWAS promised to reveal common variants associated with disease risk. However, the initial enthusiasm was tempered by the fact that truly positive associations could be neglected once stringent thresholds of statistical significance at genome-wide level were applied, greatly reducing the power of these studies [25, 26]. Often, the effect of SNPs is too small to be individually detected [27] and an approach based on pathways could be more helpful to reveal positive signals hidden in the statistical noise. Following this reasoning, we screened all SNPs located in the NER and BER pathway genes for their association in our previous GWAS, under the hypothesis that DNA repair is very important in the defense of thyrocytes towards their own high level of oxidative stress. We selected the most strongly associated SNPs for the present replication study even though they did not reach the genome-wide statistical significance in the GWAS. The replication study failed to find statistically significant associations for the considered SNPs when a correction of multiple testing is applied. Thus, the results of the present study, together with the results obtained in the previous GWAS, suggest that NER and BER pathways are unlikely involved as important determinant in the predisposition to DTC. Overall, we are keen to conclude that if any effect of polymorphisms within NER and BER genes is present this should be very weak and difficult to be ascertained.

However, it should be also noticed that the variant allele rs7689099, encoding for an aminoacid change Pro>Arg at codon 117 of *NEIL3* showed an association with a reduced risk of DTC at the

nominal threshold level of 0.05. Moreover, the ORs in the replication study had the same direction of that found in the GWAS. Therefore, it cannot be ruled out that GWAS and replication study, separately, did not have enough statistical power to pinpoint weak associations.

NEIL3 (endonuclease VIII-like 3) gene (within 4q34.3) encodes for a 605-amino acid Fpg/Nei-like DNA glycosylase [28-29], sharing homologies with other members of the same superfamily (i.e. *NEIL1* and *NEIL2*) [30-33]. The protein has a role in the first step of BER by cleaving bases damaged by ROS and introducing an a-basic site [34]. While SNPs within *NEIL1* and *NEIL2* have already been evaluated in relation to different types of cancer [35-38], SNPs within *NEIL3* have been poorly investigated. In one study, rs12645561 was associated with glioblastoma among Chinese [39] whereas rs7689099 was found associated with a reduced risk of prostate cancer among Caucasians [40], in agreement with the results obtained here for DTC. Theoretically, the aminoacid change Pro117Arg should affect the conformation of the protein and the activity, since proline, due to its cyclic structure, typically confers a reduced flexibility to the carbamidic bond and the *in silico* predictions seem to confirm this hypothesis. Thus, rs7689099 could be a functional polymorphism playing a role in several types of cancer, including TC. Future studies on large sample sets could be warranted for better ascertain the association of rs7689099 with the risk of DTC.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

This work has received financial support from the Istituto Toscano Tumori and AIRC (Associazione Italiana Ricerca Cancro) that supported the study with an investigator grant (year 2008).

References

1. A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA Cancer J. Clin. 61 (2011) 69-90.
2. E. Grande, J.J. Díez, C. Zafon, J.Capdevila, Thyroid cancer: molecular aspects and new therapeutic strategies, J. Thyroid Res. (2012) doi: 10.1155/2012/847108.
3. U. Weyemi, B. Caillou, M. Talbot, R. Ameziane-El-Hassani, L. Lacroix, O. Lagent-Chevallier, A. Al Ghuzlan, D. Roos, J.M. Bidart, A. Virion, M. Schlumberger, C. Dupuy, Intracellular expression of reactive oxygen species-generating NADPH oxidase NOX4 in normal and cancer thyroid tissues, Endocr. Relat.Cancer. 29 (2010) 27-37.
4. D. Carvalho, C. Dupuy, Role of the NADPH Oxidases DUOX and NOX4 in Thyroid Oxidative Stress, Eur Thyroid J 2 (2013) 160–167.
5. J. Maier, H. van Steeg, C. van Oostrom, S. Karger, R. Paschke, K. Krohn: Deoxyribonucleic acid damage and spontaneous mutagenesis in the thyroid gland of rats and mice, Endocrinology 147 (2006) 3391–3397.
6. N. Driessens, S. Versteyhe, C. Ghaddhab, A. Burniat, X. De Deken, J. Van Sande, J.E. Dumont, F. Miot, B. Corvilain, Hydrogen peroxide induces DNA single- and double-strand breaks in thyroid cells and is therefore a potential mutagen for this organ, Endocr Relat Cancer (16) 2009 845–856.
7. R. Ameziane-El-Hassani, M. Boufraqech, O. La-gente-Chevallier, U. Weyemi, M. Talbot, D. Métivier, F. Courtin, J.M. Bidart, M. El Mzibri, M. Schlumberger, C. Dupuy: Role of H_2O_2 in RET/PTC1 chromosomal rearrangement produced by ionizing radiation in human thyroid cells, Cancer Res 70 (2010) 4123–4132.
8. K. Krohn, J. Maier, R. Paschke, Mechanisms of disease: hydrogen peroxide, DNA damage and mutagenesis in the development of thyroid tumors, Nat Clin Pract Endocrinol Metab. 3 (2007) 713-20.

9. G. Slupphaug, B. Kavli, H.E. Krokan, The interacting pathways for prevention and repair of oxidative DNA damage, Mutat. Res. 531 (2003) 231-51.
10. B. Pascucci, M. D'Errico, E. Parlanti, S. Giovannini, E. Dogliotti, Role of nucleotide excision repair proteins in oxidative DNA damage repair: an updating, Biochemistry (Mosc). 76 (2011) 4-15.
11. J. P. Melis, H. van Steeg, M. Luijten, Oxidative DNA Damage and Nucleotide Excision Repair Antioxid Redox Signal. 20 (2013) 2409-19.
12. H. de Waard, J. de Wit, J.O. Andressoo, C.T. van Oostrom, B. Riis, A. Weimann, H.E. Poulsen, H. van Steeg, J.H. Hoeijmakers, G.T. van der Horst, Different effects of CSA and CSB deficiency on sensitivity to oxidative DNA damage, Mol Cell Biol 24 (2004) 7941–7948.
13. H. de Waard, J. de Wit, T.G. Gorgels, G. van den Aardweg, J.O. Andressoo, M. Vermeij, H. van Steeg, J.H. Hoeijmakers, G.T. van der Horst. Cell type-specific hypersensitivity to oxidative damage in CSB and XPA mice. DNA Repair (Amst) 2 (2003) 13–25.
14. G. Spivak, P.C. Hanawalt, Host cell reactivation of plasmids containing oxidative DNA lesions is defective in Cockayne syndrome but normal in UV-sensitive syndrome fibroblasts. DNA Repair (Amst) 5 (2006) 13–22.
15. J.P. Melis, S.W. Wijnhoven, R.B. Beems, M. Roodbergen, J. van den Berg, H. Moon, E. Friedberg, G.T. van der Horst, J.H. Hoeijmakers, J. Vijg, H. van Steeg, Mouse models for xeroderma pigmentosum group A and group C show divergent cancer phenotypes. Cancer Res 68 (2008) 1347–1353.
16. M. D’Errico, E. Parlanti, M. Teson, B.M. de Jesus, P. Degan, A. Cal-cagnile, P. Jaruga, M. Bjoras, M. Crescenzi, A.M. Pedrini, J.M. Egly, G. Zambruno, M. Stefanini, M. Di zdaroglu, E. Dogliotti. New functions of XPC in the protection of human skin cells from oxidative damage. EMBO J 25 (2006) 4305–4315.

17. H.R. Rezvani, R. Rossignol, N. Ali, G. Benard, X. Tang, H.S. Yang, T. Jouary, H. de Verneuil, A. Taieb, A.L. Kim, F. Mazurier. XPC silencing in normal human keratinocytes triggers metabolic alterations through NOX-1 activation-mediated reactive oxygen species.

Biochim Biophys Acta 1807 (2011) 609–619.

18. S.N. Silva, O.M. Gil, V.C. Oliveira, M.N. Cabral, A.P. Azevedo, A. Faber, I. Manita, T.C. Ferreira, E. Limbert, J.E. Pina, J. Rueff, J. Gaspar, Association of polymorphisms in ERCC2 gene with non-familial thyroid cancer risk, Cancer Epidemiol. Biomarkers Prev. 14 (2005) 2407-12.

19. L.S. Santos, B.C. Gomes, R. Gouveia, S.N. Silva, A.P. Azevedo, V. Camacho, I. Manita, O.M. Gil, T.C. Ferreira, E. Limbert, J. Rueff, J.F. Gaspar, The role of *CCNH* Val270Ala (rs2230641) and other nucleotide excision repair polymorphisms in individual susceptibility to well-differentiated thyroid cancer, Oncol. Rep. 30 (2013) 2458-66.

20. Y. Bao, L. Jiang, J.Y. Zhou, J.J. Zou, J.Y. Zheng, X.F. Chen, Z.M. Liu, Y.Q. Shi. *XRCC1* gene polymorphisms and the risk of differentiated thyroid carcinoma (DTC): a meta-analysis of case-control studies, PLoS One. 22 (2013) e64851.

21. A. Köhler, B. Chen, F. Gemignani, R. Elisei, C. Romei, G. Figlioli, M. Cipollini, A. Cristaudo, F. Bambi, P. Hoffmann, S. Herms, M. Kalemba, D. Kula, S. Harris, P. Broderick, R. Houlston, S. Pastor, R. Marcos, A. Velázquez, B. Jarzab, K. Hemminki, S. Landi, A. Försti. Genome-wide association study on differentiated thyroid cancer, J. Clin. Endocrinol. Metab. 98 (2013). doi: 10.1210/jc.2013-1941

22. I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, Nat. Methods 7 (2010) 248-249.

23. P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4 (2009) 1073-81.

24. P. Yue, E. Melamud, J. Moult, SNPs3D: candidate gene and SNP selection for association studies BMC Bioinformatics. 22 (2006) 166.
25. F. Dudbridge, A. Gusnanto, Estimation of significance thresholds for genomewide association scans. Genet. Epidemiol. 32 (2008) 227–234.
26. N. Risch, K. Merikangas, The future of genetic studies of complex human diseases. Science 273 (1996) 1516–1517
27. C. Correia, Y. Diekmann, A.M. Vicente, JB Pereira-Leal, Hope for GWAS: relevant risk genes uncovered from GWAS statistical noise, Int J Mol Sci. 15 (2014) 17601-21.
28. M. Takao, Y. Oohata, K. Kitadokoro, K. Kobayashi, S. Iwai, A. Yasui, S. Yonei, Q.M. Zhang, Human Nei-like protein NEIL3 has AP lyase activity specific for single-stranded DNA and confers oxidative stress resistance in Escherichia coli mutant. Genes Cells. 14 (2009) 261-70
29. M. Liu, V. Bandaru, A. Holmes, A.M. Averill, W. Cannan, S.S. Wallace. Expression and purification of active mouse and human NEIL3 proteins, Protein Expr. Purif. 84 (2012) 130-9.
30. M. Takao, S. Kanno, K. Kobayashi, Q.M. Zhang, S. Yonei, G.T. van der Horst, A. Yasui, A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue, J Biol Chem. 277 (2002) 42205-13.
31. V. Bandaru, S. Sunkara, S.S. Wallace, J.P. Bond, A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII, DNA Repair (Amst) 1 (2002) 517-29.
32. I. Morland, V. Rolseth, L. Luna, T. Rognes, M. Bjørås, E. Seeberg, Human DNA glycosylases of the bacterial Fpg/MutM superfamily: an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA, Nucleic Acids Res. 15 (2002) 4926-36.

33. P. Jaruga, M. Birincioglu, T.A. Rosenquist, M. Dizdaroglu, Mouse NEIL1 protein is specific for excision of 2,6-diamino-4-hydroxy-5-formamidopyrimidine and 4,6-diamino-5-formamidopyrimidine from oxidatively damaged DNA. Biochemistry. 21 (2004) 15909-14.
34. T.L. Scott, S. Rangaswamy, C.A. Wicker, T. Izumi, Repair of oxidative DNA damage and cancer: recent progress in DNA base excision repair. Antioxid Redox Signal. 20 (2014) 708-26.
35. H. Wei, A. Kamat, M. Chen, H.L. Ke, D.W. Chang, J. Yin, H.B. Grossman, C.P. Dinney, X. Wu, Association of polymorphisms in oxidative stress genes with clinical outcomes for bladder cancer treated with Bacillus Calmette-Guérin, PLoS One. 7 (2012) e38533.
36. S. Dey, A.K. Maiti, M.L. Hegde, P.M. Hegde, I. Boldogh, P.S. Sarkar, S.Z. Abdel-Rahman, A.H. Sarker, B. Hang, J. Xie, A.E. Tomkinson, M. Zhou, B. Shen, G. Wang, C. Wu, D. Yu, D. Lin, V. Cardenas, T.K. Hazra, Increased risk of lung cancer associated with a functionally impaired polymorphic variant of the human DNA glycosylase NEIL2. DNA Repair (Amst). 11 (2012) 570-8.
37. M. Goto, K. Shinmura, H. Tao, S. Tsugane, H. Sugimura, Three novel NEIL1 promoter polymorphisms in gastric cancer patients, World J. Gastrointest. Oncol. 15 (2010) 117-20.
38. X. Zhai, H. Zhao, Z. Liu, L.E. Wang, A.K. El-Naggar, E.M. Sturgis, Q. Wei, Functional variants of the NEIL1 and NEIL2 genes and risk and progression of squamous cell carcinoma of the oral cavity and oropharynx, Clin. Cancer Res. 14 (2008) 4345-52.
39. T.B. Jin, X.L. Li, H. Yang, M. Jiri, X.G. Shi, D.Y. Yuan, L.L. Kang, S.Q. Li, Association of polymorphisms in FLT3, EGFR, ALOX5, and NEIL3 with glioblastoma in the Han Chinese population, Med. Oncol. 30 (2013) 718.
40. K.H. Barry, S. Koutros, S.I. Berndt, G. Andreotti, J.A. Hoppin, D.P. Sandler, L.A. Burdette, M. Yeager, L.E. Freeman, J.H. Lubin, X. Ma, T. Zheng, M.C. Alavanja, Genetic

variation in base excision repair pathway genes, pesticide exposure, and prostate cancer risk, Environ. Health Perspect. 119 (2011) 1726-32.

Legend to supplementary table 1.

An extract of the original PLINK file from previous genome-wide case-control association study on the risk of DTC (see reference [21]). Only the DNA repair genes considered in this study are showed, with the exception of the 13 genes displayed in figure 1.

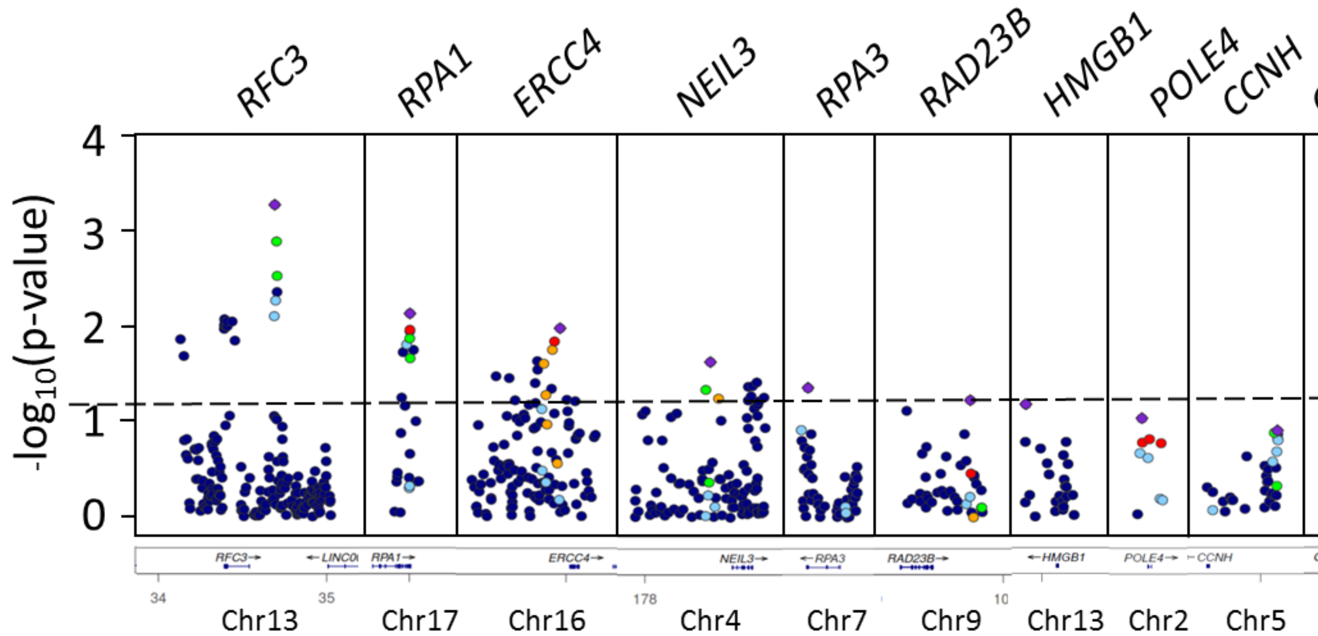

Figure 1. The figure summarizes the most strongly associated 13 loci among 104 loci evaluated for NER and BER genes. Genes were ranked according to their most strongly associated SNP and the two most strongly associated SNPs for each of the first four genes were selected for replication. Dotted line shows the statistical threshold of 0.05.

Table 1. Characteristics of DTC patients and controls. The OR and its 95% confidence interval (CI) were calculated with a logistic regression analysis. The difference between cases and controls for continuous variables was calculated with the student's *t*-test.

	Controls N(%)	**Cases N(%)**	**OR[a] (95% CI)**	***P*-value**
Total	450	450		
Gender:				
Male	292 (64.8 %)	119 (26.5 %)	Ref	
Female	158 (35.2 %)	331(73.5 %)	5.11 (4.23-6.18)	$<10^{-6}$
Age (years)				
Average ± st.err.	50.5±0.52	45.8±0.49		$<10^{-5}$[b]
Median	49.0	45.0		

[a]Crude Odd Ratio from logistic regression analysis.; [b] P-value from student's *t*-test.

Table 2: Statistical analyses of 8 polymorphisms in BER and NER genes in the replication set.

SNP	**Location**	**Gene**	**Risk Allele**	**Risk allele frequency (cases)**	**Risk allele frequency (controls)**	**Additive (Het) OR (95% CI)[b]**	**P-value**	**Additive (Hom) OR (95% CI)[b]**	**P-value**	**Dominant OR (95% CI)[b]**	**P-value**	**Recessive OR (95% CI)[b]**	**P-value**
rs7990340	flanking_3'-UTR	*RFC3*	G	0.36	0.33	1.07 (0.80-1.43)	0.65	1.36 (0.87-2.13)	0.18	1.12 (0.86-1.48)	0.41	1.32 (0.86-2.1)	0.22
rs690860	flanking_3'-UTR	*RFC3*	C	0.43	0.44	0.84 (0.62-1.15)	0.27	1.01 (0.69-1.49)	0.96	0.89 (0.67-1.19)	0.43	1.12 (0.79-1.58)	0.52
rs3744767	3'UTR	*RPA1*	C	0.22	0.24	0.89 (0.67-1.19)	0.43	0.90 (0.48-1.70)	0.74	0.89 (0.68-1.17)	0.4	0.94 (0.50-1.76)	0.85
rs1131636	3'-UTR	*RPA1*	C	0.36	0.39	0.73 (0.54-0.98)	0.04	0.93 (0.61-1.41)	0.73	0.77 (0.58-1.01)	0.06	1.10 (0.74-1.62)	0.63
rs16962916	flanking_5'-UTR	*ERCC4*	C	0.27	0.27	1.09 (0.82-1.44)	0.55	0.90 (0.53-1.54)	0.12	1.06 (0.81-1.38)	0.67	0.87 (0.52-1.46)	0.6
rs3136166	intron	*ERCC4*	G	0.36	0.37	0.94 (0.73-1.20)	0.63	1.18 (0.84-1.65)	0.34	0.99 (0.79-1.25)	0.93	1.22 (089-1.67)	0.22
rs17739370	flanking_3'-UTR	*NEIL3*	T	0.38	0.33	1.16 (0.87-1.55)	0.31	1.76 (1.13-2.72)	0.01	1.27 (0.96-1.66)	0.09	1.62 (1.07-2.43)	0.02
rs7689099	coding	*NEIL3*	G	0.17	0.18	0.88 (0.71-1.09)	0.24	0.55 (0.31-0.98)	0.042	0.85 (0.69-1.05)	0.13	0.56 (0.31-1.00)	0.05

[b] For each SNP, ORs adjusted for age and sex with corresponding 95% CIs and P values are shown.