DNA Repair 10 (2011) 111–116

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

Brief report

Genetic variation in DNA repair pathway genes and melanoma risk☆

Mingfeng Zhang[a,b], Abrar A. Qureshi[a,c], Qun Guo[c], Jiali Han[a,c,d,*]

[a] Clinical Research Program, Department of Dermatology, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA
[b] Department of Epidemiology and Biostatistics, Cancer Center, Nanjing Medical University, Nanjing, China
[c] Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA
[d] Department of Epidemiology, and the Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, MA, USA

ARTICLE INFO

Article history:
Received 14 April 2010
Received in revised form 22 July 2010
Accepted 16 August 2010
Available online 15 September 2010

Keywords:
DNA repair
Melanoma
Etiology
Polymorphism

ABSTRACT

Reduced DNA repair capacity has been proposed as a predisposing factor for melanoma. We comprehensively evaluated 1463 genetic variants across 60 DNA repair–related pathway genes in relation to melanoma risk in a nested case–control study of 218 melanoma cases (20% on head and neck) and 218 matched controls within the Nurses' Health Study (NHS). We then genotyped the two variants with the smallest *P* value in two replication sets: 184 melanoma cases (28% on head and neck) and 184 matched controls in the Health Professionals Follow-Up Study (HPFS); and 183 melanoma cases (10% on head and neck) and 183 matched controls in the NHS. The SNP rs3219125 in the *PARP1* gene was significantly associated with melanoma risk in the discovery set (odds ratio (OR) 3.14; 95% confidence interval (CI) 1.70–5.80) and in the HPFS replication set (OR, 1.92; 95% CI, 1.05–3.54) but not in the NHS replication set (OR, 1.07; 95% CI, 0.58–1.97). In the joint analysis, the OR was 1.89 (95% CI, 1.34–2.68) for this polymorphism, and this increased risk was more pronounced among patients with lesions in head/neck (OR, 3.19; 95% CI, 1.77–5.73 for head/neck, and OR, 1.54; 95% CI, 1.03–2.30 for other sites, *P* value for heterogeneity test = 0.036). Our findings suggest the possible involvement of the *PARP1* variant in melanoma development, especially for sites with high sun exposure. Further work on fine-mapping and on the functional characterization of this and linked SNPs in this region is required.

1. Introduction

Skin cancer is the most common neoplasm in Caucasians in the United States, and melanoma, the most serious type of skin cancer, causes about 79% of skin cancer deaths [1] Ultraviolet (UV) radiation is an established environmental carcinogen for skin cancer, including melanoma [2–4]. Because UV radiation is capable of causing a wide range of lesions in DNA, one important defense mechanism against melanoma is the ability to repair DNA damage induced by UV light. Reduced DNA repair capacity is a risk factor for melanoma and may contribute to susceptibility to UV-induced melanoma among the general population [5]. Melanoma was reported in 5% of patients with xeroderma pigmentosum [6]. Even though a variety of factors modulate the path from genotype to phenotype, there are substantial correlations between DNA repair gene variants and DNA repair capacity [7–10]. Thus, genetic variants in relevant DNA repair genes may confer genetic susceptibility to melanoma. However, very few of them have been evaluated previously for associations with melanoma risk.

We conducted a comprehensive and systematical evaluation of genetic variation across 60 DNA repair-related pathway genes in relation to melanoma risk in a nested case–control study of 218 prospectively ascertained cutaneous malignant melanoma cases and 218 matched controls within the Nurses' Health Study (NHS), and replicated in both NHS and the Health Professionals Follow-Up Study (HPFS), with a total of 367 melanoma cases and 367 matched controls. These pathways/genes included direct reversion repair (*MGMT*); base excision repair (*APE1*, *LIG3*, *NEIL1*, *NEIL2*, *OGG1, PARP1*, *XRCC1*); nucleotide excision repair (*XPA*, *ERCC3*, *XPC*, *ERCC2*, *ERCC4*, *ERCC5*, *ERCC1*, *LIG1*, *ERCC6*, *ERCC8*, *RPA1*, *RPA2*, *RPA3*); double-strand break repair via (1) homologous recombination (*RAD50*, *RAD51*, *RAD52*, *XRCC2*, *XRCC3*, *NBN*, *MRE11A*) or (2) non-homologous end-joining (*XRCC4*, *XRCC5*, *XRCC6*, *ARTEMIS*, *PRKDC*, *LIG4*); DNA polymerases, nucleases, and helicases (*POLB*, *POLD1*, *POLE*, *POLI*, *POLK*, *PCNA*, *FEN1*, *BLM*); DNA-cross-link repair (*FANCA*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*); mismatch repair (*MSH2*, *MSH3*, *MSH6*, *MLH1*, *MLH3*, *PMS1*, *PMS2*); and genes involved in DNA damage recognition and response (*ATM*, *ATR*, *CHEK1*, *CHEK2*, *TP53*) [11–13].

Abbreviations: OR, odds ratio; CI, confidence interval; BER, base excision repair; NER, nucleotide excision repair; DSB, double-strand break repair; HR, homologous recombination; NHEJ, non-homologous end-joining; MMR, mismatch repair.

☆ The work is supported by NIH grants CA122838. The Broad Institute Center for Genotyping and Analysis is supported by grant U54 RR020278-01 from the National Center for Research Resources.

* Corresponding author at: 181 Longwood Ave, 3rd Floor, Boston, MA 02115, USA. Tel.: +1 617 525 2098; fax: +1 617 525 2088.
E-mail address: jiali.han@channing.harvard.edu (J. Han).

1568-7864/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.dnarep.2010.08.005

2. Materials and methods

2.1. Study population

Our study population consisted of participants in the Nurses' Health Study (NHS) and the Health Professionals Follow-Up Study (HPFS). The NHS was established in 1976, when 121,700 female registered nurses between the ages of 30 and 55, residing in 11 larger US states, completed and returned the initial self-administered questionnaire on their medical histories and baseline health-related exposures. Updated information was obtained by questionnaires every 2 years, and blood samples were collected from 32,826 participants in the NHS cohort between May 1989 and September 1990. For the HPFS, in 1986, 51,529 men from all 50 US states in health professions (dentists, pharmacists, optometrists, osteopath physicians, podiatrists, and veterinarians) aged 40–75 years answered a detailed mailed questionnaire, forming the basis of the HPFS, and during 1993–1994, 18,159 study participants provided blood samples. All the cases and controls in our study were from sub-cohorts of the NHS and HPFS who had given a blood specimen. Details of the two cohorts have been described previously [14,15].

For the discovery set, eligible cases consisted of women in the NHS with pathologically confirmed incident melanoma, diagnosed any time after blood collection up to June 1, 2000 and with no previously diagnosed cancer. Cases for replication consisted of pathologically confirmed melanoma cases diagnosed after the baseline up to 2006 follow-up cycle (for both cohorts) who had no previously diagnosed cancer. One control per case was randomly selected from participants who were free of diagnosed melanoma up to and including the questionnaire cycle in which the cases were diagnosed. Controls were matched to cases by age (±1 year). The cases and their matched controls were selected in the same cohort. All subjects were US non-Hispanic Caucasians. Finally, we recruited 585 melanoma cases and 585 matched controls. The study protocol was approved by the Committee on Use of Human Subjects of the Brigham and Women's Hospital, Boston, MA.

2.2. Exposure data

We obtained information regarding melanoma risk factors from prospective biennial questionnaires. Information on natural hair color at age 20, mole count on the left arm, childhood and adolescent tanning tendency, number of lifetime severe sunburns, and family history of melanoma in first-degree relatives were collected in both the NHS and HPFS prospective questionnaires.

2.3. Single nucleotide polymorphism (SNP) selection

We selected a high density of 1463 common SNPs across the promoter regions, untranslated regions (UTRs), and coding and non-coding regions of 60 DNA repair genes. For each gene, we first included all coding SNPs, followed by the selection of additional tagging-SNPs by a well-accepted tagger program [16], which combines pair-wise r^2 methods [17] with the potential efficiency of multi-marker approaches [18]. The detailed description on the tagging-SNP selection for these 60 DNA repair genes was presented previously elsewhere [11]. Briefly, genotype data were collected from seven population samples, including 20 CEPH trios (60 individuals in total), which are a subset of the 30 trios used in the HapMap and 70 White subjects from the Multiethnic Cohort (MEC) study [19]. In total, 3072 SNPs have been genotyped across these 60 genes, including a high density of common SNPs ($n > 2700$, minor allele frequency $\geq 5\%$) selected from the public dbSNP database and all known missense SNPs (>300, minor allele frequency $\geq 1\%$) identified through gene resequencing from the Environmental Genome Project (http://egp.gs.washington.edu/); the average spacing of common SNPs across each locus is 1.7 kb. In the selection of tag-SNPs for Caucasians ($r^2 > 0.8$), the 3072 SNPs genotyped in-house in the 20 CEPH trios and the HapMap phase I data of the same 60 Caucasians were combined to achieve a much higher density of SNP markers. The patterns of linkage disequilibrium (LD) in these individuals should provide an accurate estimate of the patterns in our study population [20]. In brief, 91% of HapMap phase II SNPs were predicted by this panel with 80% or greater multi-allelic r^2.

2.4. Genotyping assay

In the discovery stage, high-throughput genotyping was performed using the Illumina high-multiplex BeadArray genotyping system at the MIT Broad Institute, Center for Genotyping and Analysis. The assay employs allele-specific extension methods and universal PCR amplification reactions conducted at 1536 loci. DNA samples were processed through the highly multiplexed GoldenGate protocol using bar-coded microwell plates and robust automation systems. Eight pairs of blinded duplicate samples were included. Among the 1536 SNPs, there were 1463 SNPs in 60 DNA repair genes, as described above.

For the replication study, we genotyped the top two SNPs in discovery stage with the smallest *P* value for the association with melanoma risk, using the OpenArrayTM SNP Genotyping System (BioTrove, Woburn, MA). Laboratory personnel were blinded to case–control status, and blinded quality control samples were inserted to validate genotyping procedures; concordance for the blinded samples was 100%. Primers, probes, and conditions for genotyping assays are available upon request.

2.5. Statistical analysis

We used a χ^2 test to assess whether the genotypes were in Hardy–Weinberg equilibrium. Unconditional logistic regression was employed to calculate odds ratios (ORs) and 95% confidence intervals (CIs). The test for screening the main effects of 1067 SNPs was based on the additive model, treating genotype as an ordinal variable (wildtype coded as 0, heterozygote as 1, and homozygotes variant as 2). For the top two SNPs with the smallest *P* values, odds ratios (ORs) and 95% confidence intervals (95% CIs) were calculated in dominant models due to the relatively low minor allele frequency. The *P* value for heterogeneity according to body sites was tested by a χ^2 test in the cases only. All *P* values were two-sided.

For the correction of multiple testing in the discovery stage, the Bonferroni correction, which is the most commonly used method to adjust type I error, α, treats each single-SNP test as an independent test and is overly conservative for SNPs that are in LD, because it ignores the correlation among SNPs. To address this limitation, we calculated the effective number of independent SNPs, $M_{\text{eff},i}$, for each candidate gene *i*, on the basis of the spectral decomposition (SpD) of matrices of pair-wise LD between SNPs [21,22]. M_{eff} provides a simple correction for multiple testing of non-independent SNPs in LD with each other. For each SNP for candidate gene *i*, the multiplicity-adjusted point-wise α (α_{p}) was then calculated as $\alpha/M_{\text{eff},i}$.

3. Results

3.1. Characteristics of cases and controls

Basic characteristics of cases and controls in our study are presented in Table 1. The mean age at diagnosis of melanoma cases was 63.5 years old in the discovery set, and 62.9 and 57.6 years old in the HPFS and NHS replication sets, respectively. Melanoma cases were more likely to have a family history of melanoma and to have