European Journal of Cancer 136 (2020) 84–94

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Genetic variants in *TKT* and *DERA* in the nicotinamide adenine dinucleotide phosphate pathway predict melanoma survival

Ning Gu [a,b,c,d,1], Wei Dai [c,d,e,1], Hongliang Liu [c,d], Jie Ge [c,d], Sheng Luo [f], Eunyoung Cho [g,h], Christopher I. Amos [i], Jeffrey E. Lee [j], Xin Li [k], Hongmei Nan [k], Hua Yuan [a,b,l,**], Qingyi Wei [c,d,m,*]

[a] *Jiangsu Key Laboratory of Oral Diseases, Nanjing Medical University, Nanjing, Jiangsu, 210029, China*
[b] *Department of Oral and Maxillofacial Surgery, The Affiliated Hospital of Stomatology, Nanjing Medical University, Nanjing, Jiangsu, 210029, China*
[c] *Duke Cancer Institute, Duke University Medical Center, Durham, NC, 27710, USA*
[d] *Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, 27710, USA*
[e] *Department of Dermatology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, 510515, China*
[f] *Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, 27710, USA*
[g] *Department of Dermatology, Warren Alpert Medical School, Brown University, Providence, RI, 02912, USA*
[h] *Department of Epidemiology, Brown University School of Public Health, Providence, RI, 02912, USA*
[i] *Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX, 77030, USA*
[j] *Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, 77030, USA*
[k] *Department of Epidemiology, Fairbanks School of Public Health, Indiana University, Indianapolis, IN, 46202, USA*
[l] *Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, Jiangsu, 210029, China*
[m] *Department of Medicine, Duke University School of Medicine, Durham, NC, 27710, USA*

Received 14 February 2020; accepted 14 April 2020

Abbreviations: CM, cutaneous melanoma; SNP, single-nucleotide polymorphisms; GWAS, genome-wide association studies; MDACC, The University of Texas MD Anderson Cancer Center; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study; CMSS, cutaneous melanoma-specific survival; HR, hazards ratio; CI, confidence interval; eQTL, expression quantitative trait loci; FPRP, false-positive report probability; LD, linkage disequilibrium; NADPH, nicotinamide adenine dinucleotide phosphate; ROC, receiver operating characteristic; AUC, area under the curve; *TKT*, transketolase; *DERA*, deoxyribose phosphate aldolase; ROS, reactive oxygen species.

* *Corresponding author*: Duke Cancer Institute, Duke University Medical Center and Department of Medicine, Duke School of Medicine, 905 S LaSalle Street, Durham, NC, 27710, USA. Fax: +19 196817386.
** *Corresponding author*: Jiangsu Key Laboratory of Oral Diseases, Nanjing Medical University, Nanjing, Jiangsu, 210029, China. Fax: +(86) 2586516414.
E-mail address: yuanhua@njmu.edu.cn (H. Yuan), qingyi.wei@duke.edu (Q. Wei).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ejca.2020.04.049

KEYWORDS
Cutaneous melanoma;
Nicotinamide adenine dinucleotide phosphate;
Single-nucleotide polymorphism;
Genome-wide association study;
Melanoma-specific survival

Abstract ***Background:*** Cutaneous melanoma (CM) is the most lethal type of skin cancers. Nicotinamide adenine dinucleotide phosphate (NADPH) plays an important role in anabolic reactions and tumorigenesis, but many genes are involved in the NADPH system.
Methods: We used 10,912 single-nucleotide polymorphisms (SNPs) (2018 genotyped and 8894 imputed) in 134 NADPH-related genes from a genome-wide association study (GWAS) of 858 patients from The University of Texas MD Anderson Cancer Center (MDACC) in a single-locus analysis to predict CM survival. We then replicated the results in another GWAS data set of 409 patients from the Nurses' Health Study (NHS) and the Health Professionals Follow-up Study (HPFS).
Results: There were 95 of 858 (11.1%) and 48 of 409 (11.7%) patients who died of CM, respectively. In multivariable Cox regression analyses, we identified two independent SNPs (*TKT* rs9864057 G > A and deoxyribose phosphate aldolase *(DERA)* rs12297652 A > G) to be significantly associated with CM-specific survival [hazards ratio (HR) of 1.52, 95% confidence interval (CI) = 1.18–1.96, $P = 1.06 \times 10^{-3}$ and 1.51 (1.19–1.91, 5.89×10^{-4})] in the meta-analysis, respectively. Furthermore, an increasing number of risk genotypes of these two SNPs was associated with a higher risk of death in the MDACC, the NHS/HPFS, and their combined data sets (P_{trend}<0.001, = 0.004 and <0.001, respectively). In the expression quantitative trait loci analysis, *TKT* rs9864057 G > A and *DERA* rs12297652 A > G were also significantly associated with higher mRNA expression levels in sun-exposed lower-leg skin ($P = 0.043$ and 0.006, respectively).
Conclusions: These results suggest that these two potentially functional SNPs may be valuable prognostic biomarkers for CM survival, but larger studies are needed to validate these findings.

1. Introduction

Cutaneous melanoma (CM) is the most lethal type of skin cancers. Although CM is much less common than basal or squamous cell carcinomas, it accounts for the majority of skin cancer deaths because of its tendency to invade and spread during the course of the disease [1,2]. It has been estimated that 96,480 new CM cases, about 5.5% of all of new cancer cases, will be diagnosed in the United States in 2019, and 7230 patients will die of this disease [3]. Between 2008 and 2014 in the United States, the overall five-years survival rate after diagnosis was 91.8% [4], whereas patients with a late or distant-stage CM generally have a much poorer prognosis with a five-year survival rate of only approximately 18% [2]; survival is particularly poor for those patients with visceral metastasis (a median survival of 6 months) [5]. To better identify high-risk subgroups with a poor survival, it is urgent to identify additional factors, such as genetic variants, that are involved in CM prognosis. Such investigations are likely to help identify clinically relevant mechanisms, allowing for more accurate selection of CM patients for the most effective management and treatment.

Nicotinamide adenine dinucleotide phosphate (NADPH) plays a key role in anabolic reactions in all cellular organisms as a reducing agent [6]. The NADPH system generates free radicals in biological immune cells to destroy pathogens through a process called the respiratory burst [7]. Through the NADPH system, cells produce reactive oxygen species (ROS) that cause cellular and tissue damage by oxidising both DNA and membranes, thus contributing to cellular dysfunction, ageing, neurodegeneration, cell death and cancer [8]. The NADPH system also promotes migration of inflammatory cells and modulates radiation-induced senescent cells by producing ROS [9]. In cancer cells, defects in the mitochondrial oxidative metabolism provide reduced equivalents through NADPH for metabolising hydroperoxides to increase production of superoxide, hydrogen peroxide and hydroperoxide [10]. The NADPH oxidases are an important site of ROS generation that is regulated in cancer cells and T cells in the context of antitumour immunity [11]. Several publications have highlighted the impact of NADPH on CM. For example, melanoma cells are characterised by altered redox signalling, especially in the form of higher levels of ROS, than those required for normal cell signals [12]. Two isoforms of NADPH (NOX1 and NOX4) are overexpressed, producing high ROS levels in melanoma cells, and the abundant ROS then stimulates the proliferation of tumour cells [13–15]. Some studies have shown that ultraviolet-induced chemical excitation of melanin fragments leads to DNA damage by initiating the ROS-dominated NADPH [16–18].

Many genetic variants have been found to be associated with cancer risk or survival of patients with CM

[1,19]. However, the biological relevance of most identified genetic variants, such as single-nucleotide polymorphisms (SNPs), remains unknown. Understanding of genetic factors that contribute to the NADPH pathway offers opportunities for prediction and identification of patients with high-risk CM for personalized management and treatment.

In the present study, therefore, we systematically examined associations between genetic variants in NADPH-related genes and CM survival, by using publicly available genotyping data sets and also evaluated correlations between significant SNPs and their gene expression levels to identify biological mechanisms that may underlie the observed associations.

2. Patients and methods

2.1. Study populations

In the present study, we used two publically available genome-wide association study (GWAS) genotyping data sets: a discovery data set from The University of Texas MD Anderson Cancer Center (MDACC) and a replication data set from the Nurses' Health Study (NHS) and the Health Professionals Follow-up Study (HPFS). There were 95 of 858 (11.1%) and 48 of 409 (11.7%) patients who died of CM, respectively. All patients from the discovery data set gave written informed consent, and the protocol was approved by the MDACC Institutional Review Board. For the replication data set, a written informed consent was also obtained from each subject, and the study protocol was approved by the Institutional Review Boards of Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health, and those of participating registries as required.

2.2. Discovery data set

The discovery data set included 858 non-Hispanic white patients with CM who were recruited between March 1993 and August 2008 to be participants in a hospital-based case-control study [20]. All patients with CM were classified in accordance with the American Joint Committee on Cancer staging system [21] and followed using standardised methods and guidelines [22]. Demographic and clinical variables such as age, sex, Breslow thickness, tumour stage, ulceration and mitotic rate were available in the data sets obtained from the dbGaP database (accession: phs000187.v1.p1) [23] (Supplementary Table 1).The details of genotyping information and data quality control have previously been reported [20]. Using the MACH software program and the 1000 Genomes Project CEU population (March 2010 release) as the reference, data from the MDACC study have been imputed (imputation quality $r^2 \geq 0.8$) [24].

2.3. Replication data set

The replication data set from the NHS/HPFS was produced by merging two subdata sets from two other studies: one having 317 female cases from the NHS and the other having 177 male cases from the HPFS [25]. There were 409 non-Hispanic white patients in the NHS/HPFS GWAS data set. Participants were enrolled in the NHS in 1976 and the HPFS in 1986, and were diagnosed after the enrolment up to the 2008 follow-up cycle for both cohorts. Clinical information on age, sex, survival outcome and genotype data were available. Genotype data were generated using the Illumina HumanHap610 array and imputed using the MACH software program by using the 1000 Genomes Project CEU population (Northern Europeans from Utah) database (phase I v3, March 2012, with imputation quality $r^2 \geq 0.8$) as a reference panel [26,27]. The study protocol was approved by the Institutional Review Boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required.

2.4. Gene and SNP extraction

Because females carry two copies of the X chromosome but males are heterozygous, and there is no standard statistical data based on sex-specific analysis, five genes in the X chromosome (but no pseudogenes and none on Y chromosome) were excluded. Pseudogenes were also excluded because they have no biological function. A total of 134 NADPH pathway genes located on autosomes were extracted from the online Molecular Signatures Database, which includes gene sets extracted from original research publications, as well as and the entire collections from the online resources, such as GO and KEGG (http://software.broadinstitute.org/gsea/msigdb/index.jsp) (Supplementary Table 1). All selected genes were expanded with ±2-kb flanking regions and mapped to all the SNPs identified in the MDACC GWAS data set following standard quality-control criteria, including minor allele frequency ≥ 0.05, genotyping rate $\geq$ 95%, and Hardy-Weinberg equilibrium P value $\geq 1 \times 10^{-5}$. Consequently, 10,912 [with 2018 (18.5%) genotyped and 8894 (81.5%) imputed] common SNPs in the NADPH pathway genes were extracted from the MDACC GWAS data set. A detailed flowchart is shown in Fig. 1.

2.5. Statistical analysis

CM-specific survival (CMSS) was the outcome of interest in the present study, and the survival time was calculated based on the date of CM diagnosis to the last

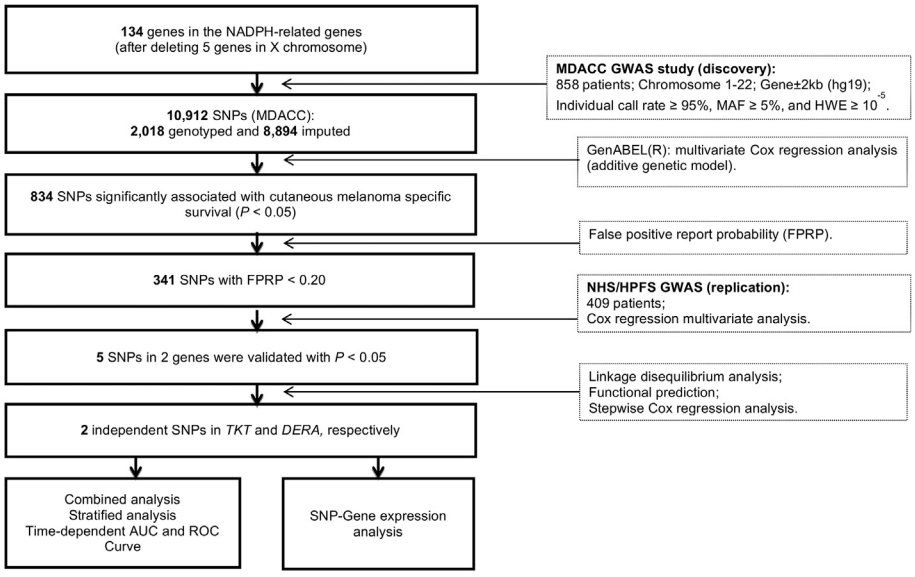

Fig. 1. Research flowchart. NADPH, nicotinamide adenine dinucleotide phosphate; SNP, single-nucleotide polymorphism; MDACC, The University of Texas M.D. Anderson Cancer Center; FPRP, false-positive report probability; *TKT,* transketolase; *DERA,* deoxyribose phosphate aldolase; AUC, area under the curve; ROC, receiver operating characteristic; GWAS, genome-wide association study; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study.

follow-up or CM-related death date. We used multivariate Cox proportional hazards regression models and the GenABEL package of R software to compute the hazards ratio (HR) and 95% confidence interval (CI) for associations between SNPs and CMSS in an additive genetic model. Multivariate analysis of the MDACC data set was adjusted by age, sex, Breslow thickness, tumour stage, ulceration and mitotic rate. Multivariate analysis of the NHS/HPFS data set and the combination of the two data sets were adjusted only for age and sex because clinical factors were not available in the NHS/HPFS data set.

Because most of the SNPs in the present study were imputed based on linkage disequilibrium (LD) among them, false-positive report probability (FPRP) has been suggested as a better probability threshold for these types of data sets for multiple testing correction than the false discovery rate [28]. In the FPRP calculation, a prior probability of 0.10 was selected to detect an HR of 2.0 for an association with variant genotypes or minor alleles of each SNP. SNPs with FPRP <0.2 were selected for replication in the NHS/HPFS data set. We also used a meta-analysis to combine the results from the discovery and replication sets. When there was no heterogeneity between the MDACC and the NHS/HPFS data sets (Cochran's Q test P-value > 0.100 and the heterogeneity statistic (I^2) $< 50\%$), we used a fixed-effects model; otherwise, a random-effects model was used. To identify independent predictive SNPs for CMSS, the validated SNPs together with clinical prognostic variables were included in a multivariate stepwise Cox model only using the MDACC data set that had more covariates available for further adjustment. We used Kaplan-Meier survival curves and log-rank tests to evaluate the effect of selected SNPs on the cumulative probability of CMSS. We also used the chi-square–based Q test with $P < 0.05$ to evaluate effect difference in the stratified analyses. We used the receiver operating characteristic (ROC) curve, which relies on the value of area under the curve (AUC), to illustrate the CMSS prediction with the sensitivity and specificity. Time-dependent AUC and ROC analyses were performed using the two R packages of survival and time ROC [29]. The expression quantitative trait loci (eQTL) analysis for associations with the alleles and genotypes of the significant SNPs was assessed by linear regression analysis using the R software. All analyses were performed with SAS (version 9.3.3; SAS Institute, Cary, NC, USA) unless otherwise specified. All reported P values were two-sided.

3. Results

3.1. Characteristics of study populations

The analyses included 858 patients from the MDACC data set and 409 patients from the NHS/HPFS data set. The age range at diagnosis was between 17 and 94 years with a mean age of 52.4 years (52.4 ± 14.4) in the MDACC data set, compared with 34–87 years and 61.1 years (61.1 ± 10.8 years) in the NHS/HPFS data set. The male/female ratio was 57.8% (496)/42.2% (362) in the MDACC data set, compared with 33.7% (138)/66.3% (271) in the NHS/HPFS data set. Univariate

analysis revealed that age, sex, tumour stage, Breslow thickness, ulceration and mitotic rate were all significantly associated with CMSS in the MDACC data set, while age but not sex was statistically significantly associated with CMSS in the NHS/HPFS data set. The median follow-up time for patients was 81.1 months in the MDACC data set and 179.0 months in the NHS/HPFS data set. Death rates in the MDACC data set (95/858, 11.1%) and the NHS/HPFS data set (48/409, 11.7%) were similar (Supplementary Table 2).

3.2. Gene and SNP extraction

After the exclusion of five genes in the X chromosome, we included 134 NADPH-related genes from the online database (MSigDB) (Supplementary Table 1). Initially, we extracted 10,912 common SNPs from the MDACC GWAS data set (2018 genotyped SNPs and 8894 imputed SNPs). The associations between all these SNPs and CMSS were presented in a Manhattan plot (Supplementary Fig. 1). Using $P < 0.05$ as a threshold, we identified 834 SNPs as significantly associated with CMSS in the discovery phase. After further screening for false-positive findings with a FPRP < 0.2, 341 SNPs were selected for further replication (Fig. 1).

In the replication analyses, we continued to use the Cox regression analysis (including only the two available variables of age and sex) to verify the 341 SNPs in the NHS/HPFS data set. Five SNPs in two genes [rs9864057, rs17234092, rs17306163, rs62255994 in *TKT* (transketolase) and rs12297652 in *DERA* (*deoxyribose phosphate aldolase*)] remained statistically significant ($P < 0.05$) for CMSS (Fig. 1). Subsequently, LD analysis showed a high LD among the four SNPs in *TKT* gene. After meta-analysis of these five SNPs, the same associations remained statistically significant, and there was no statistically significant heterogeneity in the results for these five SNPs between the two data sets (Table 1). However, in the univariate analysis as shown in Supplementary Table 2, sex was a risk factor for the MADCC data set but not for the NHS/HPFS data set, but in the multivariable modelling, sex was no longer a risk factor for the MDACC data set (Table 2).

3.3. Genetic variants in NADPH-related genes as independent survival predictors

The four SNPs in high LD in *TKT* were subsequently annotated and filtered using SNPinfo (http://snpinfo.niehs.nih.gov/snpinfo/snpfunchtml), RegulomeDB (http://www.regulomedb.org/) and F-SNP (http://compbio.cs.queensu.ca/F-SNP) for *in silico* functional prediction. As shown in Supplementary Table 3, of the four LD SNPs on *TKT*, rs9864057 is located at the potential enhancer regions of 16 tissues and the DNase

Table 2
Predictors of CMSS selected by stepwise Cox regression analysis in the MDACC study.

Parameter[a]	Category[b]	Frequency[c]	HR (95% CI)	*P*
Age	Continuous	836	1.02 (1.00–1.03)	0.062
Sex	Female/Male	353/483	1.36 (0.85–2.19)	0.198
Regional/distant metastasis	No/Yes	693/143	3.98 (2.59–6.10)	<0.001
Breslow thickness (mm)	Continuous	836	1.16 (1.10–1.22)	<0.001
Ulceration	No/Yes	681/155	2.51 (1.63–3.88)	<0.001
Mitotic rate (mm)	≤1/>1	271/565	2.83 (1.37–5.87)	0.005
TKT rs9864057 G > A	GG/GA/AA	491/303/42	1.54 (1.12–2.11)	0.008
DERA rs12297652 A > G	AA/AG/GG	316/394/126	1.52 (1.14–2.02)	0.004

CMSS, cutaneous melanoma-specific survival; MDACC, MD Anderson cancer Center; HR, hazards ratio; CI, confidence interval; DERA, deoxyribose phosphate aldolase.
[a] Stepwise analysis included age, sex, regional/distant metastasis, Breslow thickness, ulceration, mitotic rate and two SNPs in two genes (rs9864057 in *TKT*; rs12297652 in *DETA*).
[b] The leftmost was used as the reference.
[c] Twenty-two missing observations of ulceration were excluded; 836 patients remained for the stepwise analysis.

Table 1
Meta-analysis of five validated SNPs using two published melanoma GWAS data sets of MDACC and NHS/HPFS.

SNP	Allele[a]	Gene	Chr	MDACC (n = 858)			NHS/HPFS (n = 409)			Meta-analysis			
				EAF	HR (95% CI)[b]	*P*[b]	EAF	HR (95% CI)[c]	*P*[c]	P_{het}	I^2[b]	HR (95% CI)[d]	*P*[d]
rs12297652	A/G	*DERA*	12p12.3	0.38	1.51 (1.13–2.01)	0.005	0.40	1.51 (1.00–2.27)	0.048	1.000	0.000	1.51 (1.19–1.91)	5.89×10^{-4}
rs9864057	G/A	*TKT*	3p21.1	0.23	1.51 (1.10–2.06)	0.003	0.23	1.55 (1.00–2.40)	0.048	0.924	0.000	1.52 (1.18–1.96)	1.06×10^{-3}
rs17234092	C/T	*TKT*	3p21.1	0.24	1.53 (1.13–2.09)	0.006	0.24	1.65 (1.08–2.52)	0.021	0.778	0.000	1.57 (1.22–2.01)	3.84×10^{-4}
rs17306163	T/A	*TKT*	3p21.1	0.24	1.53 (1.13–2.09)	0.006	0.24	1.65 (1.08–2.52)	0.021	0.778	0.000	1.57 (1.22–2.14)	3.84×10^{-4}
rs62255994	C/T	*TKT*	3p21.1	0.23	1.60 (1.18–2.18)	0.010	0.23	1.58 (1.03–2.43)	0.037	0.963	0.000	1.59 (1.24–2.05)	2.90×10^{-4}

SNP, single-nucleotide polymorphism; EAF, effect allele frequency; HR, hazards ratio; CI, confidence interval; P_{het}, P value for heterogeneity by Cochrane's Q test; GWAS, genome-wide association study; MDACC, The University of Texas M.D. Anderson Cancer Center; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study.
[a] Reference allele/effect allele.
[b] Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in the MDACC data set.
[c] Adjusted for age and sex in the NHS/HPFS data set.
[d] Meta-analysis in a fix-effects model, when the heterogeneity statistic (I^2) < 50% or 0.5.

I sensitive site, whereas *DERA* rs12297652 may modify the biding activity of LXR. Therefore, we selected these two SNPs as the representative SNPs for further analysis. In stepwise Cox regression analysis, including the two tagging SNPs, with adjustment for other clinical covariates in the MDACC data set, these two SNPs (*TKT* rs9864057 G > A and *DERA* rs12297652 A > G) remained significant and independent predictors of CMSS (Table 2). All SNPs, both genotyped and imputed, in *TKT* and *DERA* are shown in regional association plots containing 200-kb up and downstream of rs9864057 and rs12297652, respectively (Supplementary Fig. 2).

3.4. Two independent SNPs as CM survival predictors in the MDACC data set, the NHS/HPFS data set, and the combination of both data sets

We further examined associations between the identified independent SNPs and CM survival in the MDACC data set, the NHS/HPFS data set, and their combination in multivariate analysis. The risk alleles of these two independent SNPs were rs9864057 A in *TKT* and rs12297652 G in *DERA* (Table 3).

Under an additive genetic model, the *TKT* rs9864057 A allele and the *DERA* rs12297652 G allele were significantly associated with CMSS in the MDACC data set (trend test: $P = 0.010$ and 0.005, respectively). Under a dominant genetic model, the *TKT* rs9864057 A and *DERA* rs12297652 G genotypes were both associated with elevated risk of CMSS (HR = 1.82, 95% CI = 1.20–2.77 and $P = 0.005$ for rs9864057 GA + AA vs. GG and HR = 2.09, 95% CI = 1.28–3.41 and $P = 0.003$ for rs12297652 AG + GG vs. AA). Similar results were obtained in the NHS/HPFS data set (trend test: $P = 0.05$ and $P = 0.05$, respectively, and HR = 1.76, 95% CI = 0.99–3.11, $P = 0.053$ and HR = 2.14, 95% CI = 1.07–4.30, $P = 0.03$ in a dominant model, respectively). Furthermore, when these two data sets were combined, the associations persisted (trend test: rs9864057, $P = 0.001$ and rs12297652, $P = 0.001$; rs9864057 GA + AA vs. GG: HR = 1.66, 95% CI = 1.19–2.30, $P = 0.003$; rs12297652 AG + GG vs. AA: HR = 1.93, 95% CI = 1.32–2.84, $P = 0.001$) (Table 3).

3.5. Analysis of combined genotypes of the two independent SNPs

To assess the combined effect of the two SNPs, we combined the risk genotypes of rs9864057 GA + AA and rs12297652 AG + GG into one variable as the number of risk genotypes. The trend test consistently

Table 3
Associations between two independent SNPs and CMSS in the MDACC data set, the NHS/HPFS data set, and the combined data set.

Genotype	MDACC (n = 858)				NHS/HPFS (n = 409)				MDACC + NHS/HPFS (n = 1267)			
	Frequency		Multivariate analysis		Frequency		Multivariate analysis		Frequency		Multivariate analysis	
	All	Death (%)	HR (95% CI)[a]	*P*[a]	All	Death (%)	HR (95% CI)[b]	*P*[b]	All	Death (%)	HR (95% CI)[c]	*P*[c]
***TKT* rs9864057 G > A**												
GG	491	44 (8.96)	1.00		239	22 (9.21)	1.00		746	69 (9.25)	1.00	
GA	303	39 (12.87)	1.84 (1.19–2.86)	0.006	148	22 (14.86)	1.71 (0.94–3.09)	0.077	457	62 (13.57)	1.58 (1.12–2.22)	0.009
AA	42	8 (19.05)	1.74 (0.78–3.88)	0.177	22	4 (18.18)	2.11 (0.72–6.19)	0.176	64	12 (18.75)	2.23 (1.21–4.13)	0.011
Trend test				0.010				0.049				0.001
GA + AA	345	47 (13.62)	1.82 (1.20–2.77)	0.005	170	26 (15.29)	1.76 (0.99–3.11)	0.053	521	74 (14.20)	1.66 (1.19–2.30)	0.003
***DERA* rs12297652 A > G**												
AA	316	21 (6.65)	1.00		145	10 (6.90)	1.00		471	34 (7.22)	1.00	
AG	394	50 (12.69)	2.03 (1.22–3.40)	0.007	200	29 (14.50)	2.11 (1.03–4.34)	0.042	604	80 (13.25)	1.86 (1.24–2.78)	0.002
GG	126	20 (15.87)	2.24 (1.20–4.15)	0.011	64	9 (14.06)	2.23 (0.91–5.50)	0.081	192	29 (15.10)	2.18 (1.33–3.57)	0.002
Trend test				0.005				0.049				0.001
AG + GG	520	70 (13.46)	2.09 (1.28–3.41)	0.003	264	38 (14.39)	2.14 (1.07–4.30)	0.033	796	109 (13.69)	1.93 (1.32–2.84)	0.001
Number of risk genotypes[d]												
0	184	9 (4.89)	1.00		89	6 (6.74)	1.00		281	17 (6.05)	1.00	
1	439	47 (10.71)	2.46 (1.20–5.05)	0.014	206	20 (9.71)	1.42 (0.57–3.55)	0.458	655	69 (10.53)	1.78 (1.05–3.03)	0.033
2	213	35 (16.43)	4.24 (2.02–8.89)	<0.001	114	22 (19.30)	3.08 (1.24–7.62)	0.015	331	57 (17.22)	3.12 (1.81–5.36)	<0.001
Trend test				<0.001				0.004				<0.001
0–1	623	56 (8.99)	1.00		295	26 (8.81)	1.00		936	86 (9.19)	1.00	
2	213	35 (16.43)	2.12 (1.38–3.26)	<0.001	114	22 (19.30)	2.38 (1.35–4.20)	0.003	331	57 (17.22)	2.02 (1.45–2.83)	<0.001

SNP: single-nucleotide polymorphisms; CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas M.D. Anderson Cancer Center; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study; HR, hazards ratio; CI, confidence interval; DERA, deoxyribose phosphate aldolase.

Twenty-two missing observations of ulceration were excluded in the MDACC data set.

[a] Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in Cox models of SNPs and CMSS in the MDACC data set.

[b] Adjusted for age and sex in the NHS/HPFS data set.

[c] Adjusted for age and sex in the combined data set of MDACC and NHS/HPFS.

[d] Risk genotypes were rs9864057 GA + AA, rs12297652 AG + GG.

showed that the risk increased with the number of risk genotypes in the MDACC data set ($P < 0.001$), the NHS/HPFS data set ($P = 0.004$), and the combined data set ($P < 0.001$). Furthermore, when we dichotomised all patients into 0–1 risk genotypes and 2 risk genotypes, patients with 2-risk genotypes had a higher risk of death than those with 0–1 risk genotypes in the MDACC data set (HR = 2.12, 95% CI = 1.38–3.26 and $P < 0.001$), the NHS/HPFS data set (HR = 2.38, 95% CI = 1.35–4.20 and $P = 0.003$), and the combined data set (HR = 2.02, 95% CI = 1.45–2.83, $P < 0.001$). These results were further evaluated by Kaplan-Meier plot to visualise the associations between the number of risk genotypes and CMSS (Fig. 2a–f).

3.6. Stratified analyses for the combined efficacy of risk genotypes on CMSS

We then performed stratified analyses by age and sex for the associations between CMSS and the number of risk genotypes in the MDACC data set and the NHS/HPFS data set using multivariate Cox regression analysis with adjustment where appropriate, looking for possible interaction between these covariates and SNPs. The risk of CM death associated with the 2-risk genotype was statistically higher for males ≤50 years in the MDACC data set and for females >50 years in the NHS/HPFS data set. We found no evidence for any interactions between the subgroups (Supplementary Table 4).

3.7. ROC curve and time-dependent AUC to estimate CMSS survival

We also used inverse probability of censoring weighting estimators for the ROC curve and time-dependent AUC to estimate the improvement in prediction by adding risk genotypes to the model. In the combined data set of both MDACC and NHS/HPFS, we found five-year risk of CMSS with clinical variables (age and sex) had an AUC = 63.58%, whereas the addition of risk genotypes

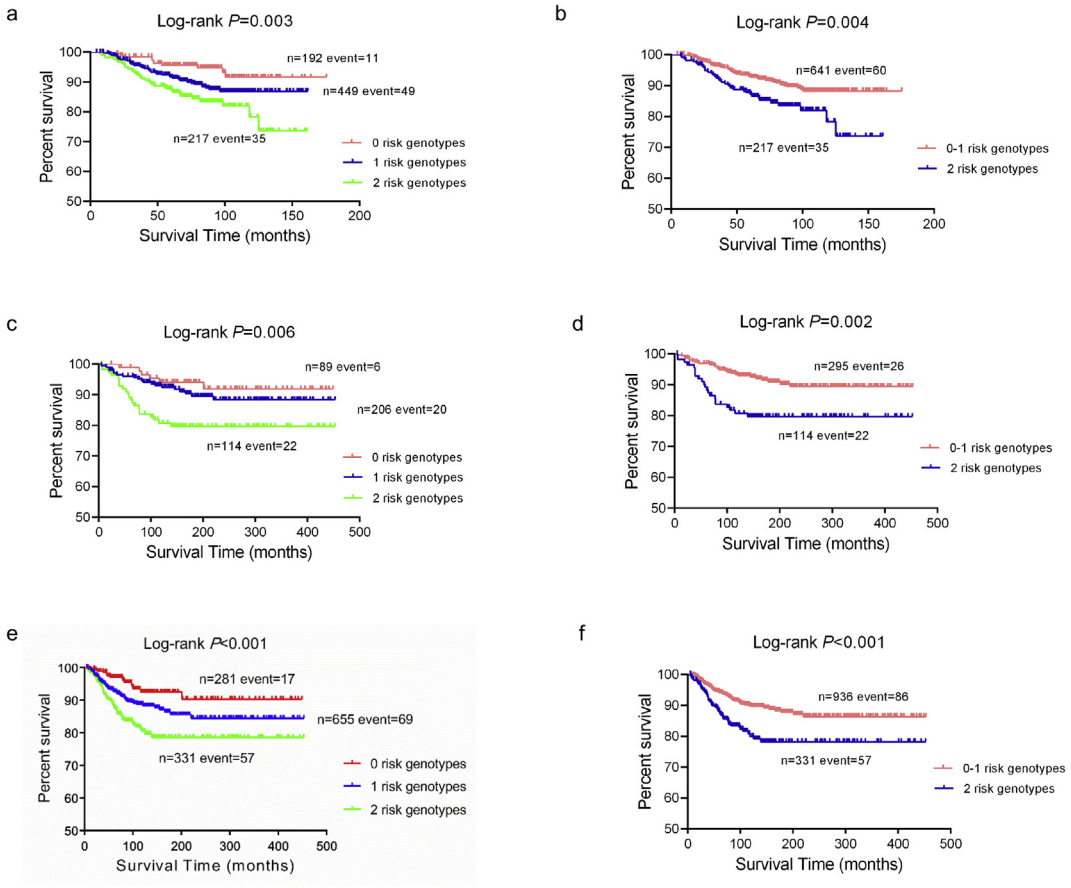

Fig. 2. The independent SNPs and CMSS (*a-f*) Kaplan-Meier survival curves of the exact numbers of risk genotypes (*a*) in the MDACC data set, (*c*) in the NHS/HPFS data set and (*e*) in these two combined data set; dichotomised groups of risk genotypes (*b*) in the MDACC data set, (*d*) in the NHS/HPFS data set and (*f*) in the combined data set. SNP, single-nucleotide polymorphism; CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas M.D. Anderson Cancer Center; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study.

to the model significantly increased the AUC to 69.72% ($P = 0.005$) (Fig. 3a-3b). However, such prediction was not statistically significant for both 5-year and 10-year survival prediction in the MDACC data set with adjustment for age, sex, Breslow thickness, distant/regional metastasis and ulceration (Supplementary Fig. 3a-3b) and in the NHS/HPFS data set with adjustment for age and sex (Supplementary Fig. 3c-3d). These results suggest that the sample sizes of the data sets are not large enough or more SNPs in different pathway genes need to be identified.

3.8. eQTL analysis

To study the correlation between the two independent SNPs and mRNA expression levels of their corresponding genes, we used data on mRNA expression levels of these two SNPs for further statistical analysis using three *in silico* eQTL databases: 373 European descendants from the 1000 Genomes Project, The Cancer Genome Atlas (TCGA) database, and the Genotype-Tissue Expression (GTEx) project.

Through the GTEx project, we found that both *TKT* rs9864057 and *DERA* rs12297652 were significantly correlated with elevated mRNA expression levels in the skin of sun-exposed lower leg ($P = 0.043$ and 0.006, respectively), but not significantly associated with mRNA expression levels in the whole blood or in non−sun-exposed skin (Supplementary Table 5). The lymphoblastoid cell line data from the 373 European descendants and the TCGA database did not demonstrate any correlation between these two SNPs and mRNA expression levels (data not shown).

4. Discussion

NADPH is a key factor in controlling redox reactions in human tissue, including tumour tissue, and NADPH is closely related to cancer cell growth, metabolism, migration, invasion and metastasis [10]. However, few studies have investigated the roles of genetic variants in the NADPH-related genes in predicting outcomes of cancers, including melanoma. In the present study, we found that two independent SNPs (*TKT* rs9864057 G > A and *DERA* rs12297652 A > G) were significantly associated with CMSS. Additional analyses suggested that their variant alleles were correlated with elevated mRNA expression levels in sun-exposed skin, a possible biological mechanism underlying the associations between the variants of NADPH-related genes and CM progression that led to poor survival.

TKT is located on chromosome 3p21.1, and this gene encodes a thiamine-dependent enzyme that plays a role in the pentose phosphate pathway. In mammals, transketolase joins the pentose phosphate pathway and glycolysis, and then transfers sugar phosphates into the main carbohydrate metabolism, an important factor in the production of NADPH for biosynthesis [30]. Because alteration in cellular metabolism is one of the hallmarks of cancers [31], several studies have reported the clinical significance of *TKT* in energy regulatory mechanism, especially in the survival of patients with

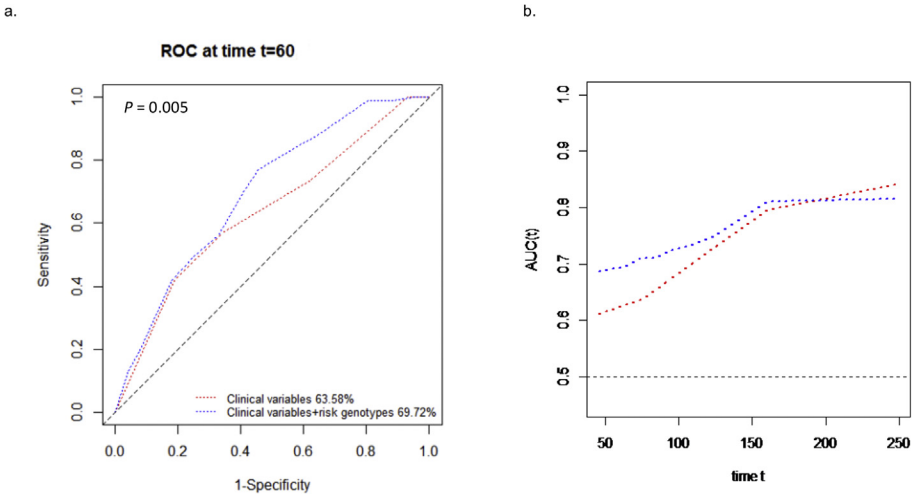

Fig. 3. ROC and time-dependent AUC estimation for prediction of CM survival in the combined data set of MDACC and NHS/HPFS as estimated by inverse probability of censoring weighting approach. (a) Five-year CM survival prediction by ROC curve, (b) Time-dependent AUC estimation: based on age, sex and the combined risk genotypes of the two genes. ROC, receiver operating characteristic; AUC, area under the curve; MDACC, The University of Texas M.D. Anderson Cancer Center; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study; CM, cutaneous melanoma.

cancers. For example, it was reported that *TKT* expression was correlated with tumour size in breast cancer; in addition, *TKT* expression was higher in lymph node metastases than in primary tumour or normal tissues of patients in whom high *TKT* levels were associated with poor survival after breast cancer [32]. *TKT* has also been associated with metastasis of ovarian and oesophageal cancers, predicting poor survival [33,34], but there are no such reports about melanoma survival. Other studies have reported significant associations between SNPs of the *TKT* gene and the morbidity/mortalities of some non-cancer diseases, particularly diabetic nephropathy [35–37]. Because melanoma is a cancer derived from neuroectoderm [38], a link between *TKT* genetic variants and nerve function has been reported [36]. Taken together, the findings in the present study support a significant role for *TKT* genetic variants in predicting CMSS.

DERA is located on chromosome 12p12.3 and encodes the human deoxyribose phosphate aldolase, which is involved in the glycosaminoglycan metabolism and the innate immune system. Deoxyribose phosphate aldolase activity levels have been correlated with *DERA* expression in most cell lines tested, and cells with high *DERA* activity can use deoxy nucleotide as a source of energy [39]. Through *DERA*, thymidine-derived 2-deoxy-D-ribose 5-phosphate enters the glycolytic pathway, affecting cancer cell growth, invasion and metastasis [40]. Thymidine, via the glycolytic pathway, relies on *DERA in vitro* and *in vivo* to convert it into multiple thymidine-derived intermediate metabolites required for the survival of cells under low-glucose conditions; therefore, the number of viable cancer cells is thereby reduced, whereas *DERA* is also reduced in a microenvironment in which glucose was insufficient [41]. Inhibition of thymidine phosphorylase activity suppresses tumor growth by increasing the proportion of apoptotic cells and probably inhibiting angiogenesis in human epidermoid carcinoma cells [42]. These findings suggest that *DERA* plays a central role in the survival and growth of cancer cells, particularly under starvation conditions. Furthermore, enhanced *DERA*-dependent thymidine catabolism has been observed in human gastric cancer [40,43]. *DERA* has not been previously reported to be associated with CM progression and prognosis.

Although we found consistent evidence that two novel SNPs (*TKT* rs9864057 G > A and *DERA* rs12297652 A > G) in two NADPH-related gene predicted CMSS, the present study has some limitations. First, the available clinical variables in the two data sets were different, and the fact there were only two variables (age and sex) available in the NHS/HPFS data set limited adjustment for other potentially relevant clinical covariates. The second limitation is that the predictive model was built on a non-Hispanic white population in the United States, which may limit generalisation of our findings to other populations. Third, SNPs in other relevant genes may have been omitted due to limited knowledge regarding identification of NADPH-related genes. Finally, although differential expression is suggested by the *in silico* analyses, the exact molecular mechanisms of these two SNPs underlying the observed associations remain to be determined. Additional functional studies are needed to explore these newly identified SNPs to confirm their potential use as biomarkers for CM prognosis.

Funding

The MD Anderson Study was supported by NIH / NCI R01 CA100264, 2P50CA093459 and R01CA133996, as well as by The University of Texas MD Anderson Cancer Center Various Donors Melanoma and Skin Cancers Priority Program Fund; the Miriam and Jim Mulva Research Fund; the McCarthy Skin Cancer Research Fund and the Marit Peterson Fund for Melanoma Research. The NHS and the HPFS were in part supported by NIH / NCI R01 CA49449, P01 CA87969, UM1 CA186107 and UM1 CA167552. Q.W. was supported by start-up funds from Duke Cancer Institute, Duke University Medical Center and in part supported by the Duke Cancer Institute as part of the P30 Cancer Center Support Grant (Grant ID: NIH CA014236).

Conflict of interest statement

The authors declare no conflict of interest.

Acknowledgements

The authors wish to thank all of the investigators and funding agencies that enabled the deposition of data in dbGaP that we used in this study (dbGaP #: phs000187.v1.p1). The authors would like to thank the Channing Division of Network Medicine at Brigham and Women's Hospital for their support. The authors also thank the John Hopkins University Center for Inherited Disease Research for conducting high-throughput genotyping for this study. The authors would like to thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data. The authors also thank all the individuals who participated in this project. The results published here are in whole or part based upon data generated by The Cancer Genome Atlas pilot project established by the NCI and NHGRI. Information about TCGA and the investigators and institutions that constitute the TCGA research network can be found at "http://

cancergenome.nih.gov". The authors assume full responsibility for analyses and interpretation of these data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2020.04.049.

References

[1] Law MH, Bishop DT, Lee JE, Brossard M, Martin NG, Moses EK, et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. Nat Genet 2015;47(9):987–95.

[2] Guy Jr GP, Thomas CC, Thompson T, Watson M, Massetti GM, Richardson LC, et al. Vital signs: melanoma incidence and mortality trends and projections - United States, 1982-2030. MMWR Morb Mortal Wkly Rep 2015;64(21):591–6.

[3] Siegel RL, Miller KD, Jemal A. Cancer statistics. CA Cancer J Clin 2019;69(1):7–34. 2019.

[4] Siegel RL, Miller KD, Jemal A. Cancer statistics. CA Cancer J Clin 2018;68(1):7–30. 2018.

[5] Demierre MF. Epidemiology and prevention of cutaneous melanoma. Curr Treat Options Oncol 2006;7(3):181–6.

[6] Hanukoglu I, Rapoport R. Routes and regulation of NADPH production in steroidogenic mitochondria. Endocr Res 1995; 21(1–2):231–41.

[7] Ogawa K, Suzuki K, Okutsu M, Yamazaki K, Shinkai S. The association of elevated reactive oxygen species levels from neutrophils with low-grade inflammation in the elderly. Immun Ageing 2008;5:13.

[8] Hervera A, De Virgiliis F, Palmisano I, Zhou L, Tantardini E, Kong G, et al. Reactive oxygen species regulate axonal regeneration through the release of exosomal NADPH oxidase 2 complexes into injured axons. Nat Cell Biol 2018;20(3):307–19.

[9] Sakai Y, Yamamori T, Yoshikawa Y, Bo T, Suzuki M, Yamamoto K, et al. NADPH oxidase 4 mediates ROS production in radiation-induced senescent cells and promotes migration of inflammatory cells. Free Radic Res 2018;52(1):92–102.

[10] Hegedus C, Kovács K, Polgár Z, Regdon Z, Szabó É, Robaszkiewicz A, et al. Redox control of cancer cell destruction. Redox Biol 2018;16:59–74.

[11] Banerjee R. Introduction to the thematic minireview series: redox metabolism and signaling. J Biol Chem 2018;293(20):7488–9.

[12] Meierjohann S. Oxidative stress in melanocyte senescence and melanoma transformation. Eur J Cell Biol 2014;93(1–2):36–41.

[13] Rouaud F, Boucher JL, Slama-Schwok A, Rocchi S. Mechanism of melanoma cells selective apoptosis induced by a photoactive NADPH analogue. Oncotarget 2016;7(50):82804–19.

[14] Liu F, Gomez Garcia AM, Meyskens Jr FL. NADPH oxidase 1 overexpression enhances invasion via matrix metalloproteinase-2 and epithelial-mesenchymal transition in melanoma cells. J Invest Dermatol 2012;132(8):2033–41.

[15] Yamaura M, Mitsushita J, Furuta S, Kiniwa Y, Ashida A, Goto Y, et al. NADPH oxidase 4 contributes to transformation phenotype of melanoma cells by regulating G2-M cell cycle progression. Cancer Res 2009;69(6):2647–54.

[16] Yuan TA, Yourk V, Farhat A, Ziogas A, Meyskens FL, Anton-Culver H, et al. A case-control study of the genetic variability in reactive oxygen species-metabolizing enzymes in melanoma risk. Int J Mol Sci 2018;19(1).

[17] Raad H, Serrano-Sanchez M, Harfouche G, Mahfouf W, Bortolotto D, Bergeron V, et al. NADPH oxidase-1 plays a key role in keratinocyte responses to UV radiation and UVB-induced skin carcinogenesis. J Invest Dermatol 2017;137(6):1311–21.

[18] Premi S, Wallisch S, Mano CM, Weiner AB, Bacchiocchi A, Wakamatsu K, et al. Photochemistry. Chemiexcitation of melanin derivatives induces DNA photoproducts long after UV exposure. Science 2015;347(6224):842–7.

[19] Manolio TA. Genomewide association studies and assessment of the risk of disease. N Engl J Med 2010;363(2):166–76.

[20] Amos CI, Wang LE, Lee JE, Gershenwald JE, Chen WV, Fang S, et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. Hum Mol Genet 2011;20(24): 5012–23.

[21] Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol 2009;27(36): 6199–206.

[22] Gershenwald JE, Ross MI. Sentinel-lymph-node biopsy for cutaneous melanoma. N Engl J Med 2011;364(18):1738–45.

[23] Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, et al. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 2007;39(10):1181–6.

[24] Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 2010;34(8):816–34.

[25] Song F, Qureshi AA, Zhang J, Amos CI, Lee JE, Wei Q, et al. Exonuclease 1 (EXO1) gene variation and melanoma risk. DNA Repair 2012;11(3):304–9.

[26] Riedel KG, Zwaan J, Kenyon KR, Kolodny EH, Hanninen L, Albert DM. Ocular abnormalities in mucolipidosis IV. Am J Ophthalmol 1985;99(2):125–36.

[27] Biernacka JM, Tang R, Li J, McDonnell SK, Rabe KG, Sinnwell JP, et al. Assessment of genotype imputation methods. BMC Proc 2009;3(Suppl 7):S5.

[28] Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96(6):434–42.

[29] Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. Stat Med 2006; 25(20):3474–86.

[30] Sax CM, Kays WT, Salamon C, Chervenak MM, Xu YS, Piatigorsky J. Transketolase gene expression in the cornea is influenced by environmental factors and developmentally controlled events. Cornea 2000;19(6):833–41.

[31] Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 2009;324(5930):1029–33.

[32] Tseng CW, Kuo WH, Chan SH, Chan HL, Chang KJ, Wang LH. Transketolase regulates the metabolic switch to control breast cancer cell metastasis via the alpha-ketoglutarate signaling pathway. Cancer Res 2018;78(11):2799–812.

[33] Chao YK, Peng TL, Chuang WY, Yeh CJ, Li YL, Lu YC, et al. Transketolase serves a poor prognosticator in esophageal cancer by promoting cell invasion via epithelial-mesenchymal transition. J Canc 2016;7(13):1804–11.

[34] Ricciardelli C, Lokman NA, Cheruvu S, Tan IA, Ween MP, Pyragius CE, et al. Transketolase is upregulated in metastatic peritoneal implants and promotes ovarian cancer cell proliferation. Clin Exp Metastasis 2015;32(5):441–55.

[35] Tanhauserova V, Kuricová K, Pácal L, Bartáková V, Rehořová J, Svojanovský J, et al. Genetic variability in enzymes of metabolic pathways conferring protection against non-enzymatic glycation versus diabetes-related morbidity and mortality. Clin Chem Lab Med 2014;52(1):77–83.

[36] Spallone V. Might genetics play a role in understanding and treating diabetic polyneuropathy? Diabetes Metab Res Rev 2017; 33(4).

[37] Pacal L, Tomandl J, Svojanovsky J, Krusová D, Stepánková S, Rehorová J, et al. Role of thiamine status and genetic variability in transketolase and other pentose phosphate cycle enzymes in the

progression of diabetic nephropathy. Nephrol Dial Transplant 2011;26(4):1229–36.
[38] Little EG, Eide MJ. Update on the current state of melanoma incidence. Dermatol Clin 2012;30(3):355–61.
[39] Salleron L, Magistrelli G, Mary C, Fischer N, Bairoch A, Lane L. DERA is the human deoxyribose phosphate aldolase and is involved in stress response. Biochim Biophys Acta 2014;1843(12):2913–25.
[40] Tabata S, Yamamoto M, Goto H, Hirayama A, Ohishi M, Kuramoto T, et al. Thymidine catabolism as a metabolic strategy for cancer survival. Cell Rep 2017;19(7):1313–21.
[41] Furukawa T, Tabata S, Yamamoto M, Kawahara K, Shinsato Y, Minami K, et al. Thymidine phosphorylase in cancer aggressiveness and chemoresistance. Pharmacol Res 2018;132:15–20.
[42] Matsushita S, Nitanda T, Furukawa T, Sumizawa T, Tani A, Nishimoto K, et al. The effect of a thymidine phosphorylase inhibitor on angiogenesis and apoptosis in tumors. Cancer Res 1999;59(8):1911–6.
[43] Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. Cancer Res 2009;69(11):4918–25.