Published in final edited form as:

Br J Dermatol. 2020 October ; 183(4): 719–728. doi:10.1111/bjd.18878.

Genetic variants in the folate metabolic pathway genes predict melanoma-specific survival

W. Dai^{1,2,3,4,†}, H. Liu^{2,3,†}, Y. Liu^{2,3,†}, X. Xu^{2,3}, D. Qian^{2,3}, S. Luo⁵, E. Cho^{6,7,8}, D. Zhu⁹, C. I. Amos⁹, S. Fang¹⁰, J. E. Lee¹⁰, X. Li^{8,11}, H. Nan^{8,11}, C. Li⁴, Q. Wei^{2,3,12}

¹Department of Dermatology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, 510515, China

²Duke Cancer Institute, Duke University Medical Center, Durham, NC, 27710, USA

³Department of Medicine, Duke University School of Medicine, Durham, NC, 27710, ssUSA

⁴Department of Dermatology, Xijing Hospital, Fourth Military Medical University, Xi'an, Shaanxi, 710032, China

⁵Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, 27710, USA

⁶Department of Dermatology, Warren Alpert Medical School, Brown University, Providence, RI, 02912, USA

⁷Department of Epidemiology, Brown University School of Public Health, Providence, RI, 02912, USA

⁸Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, 02115, USA

⁹Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX, 77030, USA

¹⁰Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, 77030, USA

¹¹Department of Epidemiology, Fairbanks School of Public Health, Indiana University, Indianapolis, IN, 46202, USA

¹²Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, 27710, USA

Summary

Background: Folate metabolism plays an important role in DNA methylation and nucleic acid synthesis and thus may function as a regulatory factor in cancer development. Genome-wide

Corresponding author: Chunying Li, qingyi.wei@duke.edu.

[†]These authors contributed equally to this work.

Supporting Information

Supplementary Figures and Tables and Acknowledgement can be found in the online version of this article at the publisher's website.

Conflicts of Interest None declared.

association studies (GWAS) have identified some single-nucleotide polymorphisms (SNPs) associated with cutaneous melanoma-specific survival (CMSS), but no SNPs were found in genes involved in the folate metabolic pathway.

Objective: To examine associations between SNPs in folate metabolic pathway genes and CMSS.

Methods: We comprehensively evaluated 2,645 (422 genotyped and 2,223 imputed) common SNPs in folate metabolic pathway genes from a published GWAS of 858 patients from The University of Texas M.D. Anderson Cancer Center and performed the validation in another GWAS of 409 patients from the Nurses' Health Study and Health Professionals Follow-up Study, in which 95/858 (11.1%) and 48/409 (11.5%) patients died of cutaneous melanoma, respectively.

Results: We identified two independent SNPs (*MTHFD1* rs1950902 G>A and *ALPL* rs10917006 C>T) to be associated with CMSS in both datasets, and their meta-analysis yielded an allelic hazards ratio of 1.75 (95% confidence interval=1.32–2.32, $P=9.96\times10^{-5}$) and 2.05 (1.39–3.01, $P=2.84\times10^{-4}$), respectively. The genotype-phenotype correlation analyses provided additional support for biologic plausibility of these two variants' roles in tumour progression, suggesting that variation in SNP-related mRNA expression levels is likely to be the mechanism underlying the observed associations with CMSS.

Conclusion: Two possibly functional genetic variants, *MTHFD1* rs1950902 and *ALPL* rs10917006, were likely to be independently or jointly associated with CMSS, which may add to personalized treatment in the future, once further validated.

Keywords

genome-wide association study; cutaneous melanoma-specific survival; folate metabolism; single-nucleotide polymorphism

Introduction

Cutaneous melanoma (CM) is the most lethal form of skin cancers because of its high metastatic potential. In 2019, there will be an estimated 96,480 new cases and 7,230 deaths from CM in the United States ¹. The CM incidence rate has risen rapidly over the past 30 years, although the death rate has declined by 1% per year in adults 50 years of age and older between 2006 and 2015 ¹. While the conventional staging system can subdivide patients with clinically localized CM into those with low risk of recurrence and death, about 10–20% of the patients will defy the prediction ². Therefore, it remains a critical need to identify prognostic biomarkers for CM.

Folate metabolism comprises an interlinked network of reactions that transfer one-carbon units for numerous cellular functions, including the biosynthesis of purines and pyrimidines for DNA and RNA as well as methyl groups for DNA methylation ^{3,4}. Numerous studies have investigated alterations in DNA methylation and disruption of DNA integrity and repair, all of which have been observed with folate depletion ^{4,5}. These aberrations may lead to tumour initiation and progression, linking folate metabolism to tumorigenesis, which associated with patients' survival ⁶. However, results from epidemiological studies on folate

intake and CM risk are controversial^{7,8}. These studies varied between in method of folate status determination, and did not account for the association between effect of genetic factors and survival of CM. Since folate can either suppress the development of early lesions in normal tissue or facilitate the proliferation of neoplastic cells⁹. This duality, coupled with the observation that folic acid (a synthetic form of folate) can affect cellular phototoxicity and photogenotoxicity, provides a strong rationale for evaluating the relationship between folate intake and skin cancer risk¹⁰. Furthermore, the folate receptor-targeting ligand molecule in the liposomal formulation may alter the mode of anticancer action of the encapsulated drug, suggesting a new possibility for folate-mediated cancer treatment¹¹. It has also been shown that immunoglobulin conjugated folate complexes can enhance the immune response by natural killer cells against folate receptor-positive CM tumour cells *in vitro*, suggesting their potential role in the host-rejection of tumour cells¹².

Most of the published genome-wide association studies (GWASs) have mainly focused on single-nucleotide polymorphisms (SNPs) that reached genome-wide significance, most of which was associated with risk susceptibility and did not have very long follow-up period¹³. We take a targeted pathway-based, multigene approach to identify genetic variation in a metabolism-related signaling pathway and its association with cancer survival¹⁴. Since the number of SNPs to be tested is greatly reduced, the typically highly stringent GWAS significance threshold could be much relaxed for such a targeted pathway-based approach. Given the role of folate metabolism in melanoma development, we hypothesized that genetic variants, i.e., SNPs, in genes involved in the folate metabolic pathway are associated with the survival of CM patients. We tested this hypothesis by using two independently published GWAS datasets. We focused our analysis on SNPs that may have potential biological functions and thus are most likely provide additional insights about the complex mechanisms of melanoma progression.

Materials and methods

Study populations

The analysis used the GWAS dataset from The University of Texas M.D. Anderson Cancer Center (MDACC) study as a discovery dataset and the Nurses' Health Study (NHS) and Health Professionals Follow-up Study (HPFS) as a validation dataset. All participants were from a hospital-based case-control study of CM, for which cases were recruited from among non-Hispanic white patients. A detailed description of subject selection and data collection for these two GWAS datasets has been published elsewhere^{15,16}, and all patients provided a written informed consent under a protocol approved by Institutional Review Boards at both MD Anderson and Brigham and Women's Hospital.

Gene selection and SNP genotyping

According to the Molecular Signatures Database of GSEA website (<http://software.broadinstitute.org/gsea/msigdb/search.jsp>), we selected 27 folate metabolic pathway genes located on autosomes (Table S1). In the MDACC dataset, genomic DNA extracted from whole blood was genotyped by the Illumina HumanOmni-Quad_v1_0_B array, in which the National Center for Biotechnology Information Database of Genotypes

and Phenotypes (dbGaP Study Accession: phs000187.v1.p1) was used. Based on the 1000 Genomes Project, we performed phase I v2 CEU Genome-wide imputation using the MACH software (March 2010 release). Briefly, the typed or imputed common SNPs (minor allele frequency > 0.05 , genotyping success rate $> 95\%$, and Hardy-Weinberg equilibrium P value $> 1 \times 10^{-5}$, and from imputation for those SNPs with $r^2 > 0.8$) within genes in the folate metabolic pathway or their ± 2 kilobase flanking regions were included in association analysis. In the NHS/HPFS dataset, genotyping was performed with the Illumina HumanHap550 array, HumanHap610 array and Affymetrix 6.0 array, and imputation analysis was based on genotyped SNPs and haplotype information from phase II HapMap CEU data using the program (MACH March 2012 release). Only SNPs with imputation quality $r^2 > 0.8$ were included.

Results

Patient characteristics

The present study included a discovery dataset of 858 CM patients from MDACC and a validation dataset of 409 CM patients from NHS/HPFS. The baseline characteristics of both patient cohorts have been published elsewhere^{15,16}. In the MDACC study, CM patients were aged between 17 and 94 years at diagnosis (mean age $\pm$ standard deviation, 52.4 ± 14.4 years), with a median follow-up time of 81.1 months. There were more males (496, 57.8%) than females (362, 42.2%) and many more diagnosed with stages I/II (709, 82.6%) than with stages III/IV (149, 17.4%). In the NHS/HPFS study, patient age ranged between 34 and 87 years at diagnosis (61.1 ± 10.8 years), and 66.3% (271) were women. These patients had a relatively longer median follow-up time (179.0 months). However, death rates were similar between the MDACC (95/858, 11.1%) and NHS/HPFS (48/409, 11.5%) studies (Table S2). We did not adjust for principal components in both discovery and validation dataset, because none of principal components was significantly associated with melanoma survival, indicating that no detectable population stratification in both MDACC and NHS/HPFS datasets was observed.

Multivariate analyses of associations between SNPs and CMSS

Figure 1 provides a flowchart of the present study design. We first performed a single-locus analysis in the MDACC dataset with multivariate Cox proportional hazards regression to assess the associations between 422 genotyped (Table S3) and 2,223 imputed SNPs in the 27 folate metabolic pathway genes with CMSS, adjusting for age, sex, tumour stage, Breslow thickness, ulceration and mitotic rate. Manhattan plots of associations between these variants and CMSS and quantile-quantile plot of the observed P values were shown (Figure S1 and S2). Among these SNPs, 202 were significantly associated with CMSS at $P < 0.05$ in an additive genetic model. After multiple test correction (Table S4) by a Bayesian false-discovery probability (BFDP) method, 190 SNPs were still considered noteworthy. As a result, 181 SNPs were identified in the NHS/HPFS validation dataset, of which 21 SNPs in two genes were validated and remained significantly associated with CMSS after multiple-testing correction by BFDP < 0.8 . On the basis of the *in silico* functional prediction by using SNPinfo (<http://SNPinfo.niehs.nih.gov/SNPinfo/SNPfunc.html>) and RegulomeDB (<http://www.regulomedb.org/>), five of these 21 SNPs were predicted to be functional, including one

SNP in methylenetetrahydrofolate dehydrogenase 1 (*MTHFD1*) and four SNPs in alkaline phosphatase (*ALPL*) (Table S5). The subsequent meta-analysis using both datasets for these five SNPs showed that the same associations remained statistically significant, without any evidence for heterogeneity between the two datasets (Table 1). By using the Versatile Gene-based Association Study method, we performed the gene-based test and found three genes (i.e., *MTHFD1L*, *QDPR* and *ALPL*) with an empirical P value < 0.05 . Although no significant gene-based test statistic was found for *ALPL* and *MTHFD1*, the top SNPs in both of these two genes had a significant P value < 0.01 (Table S6).

Further analysis found all four SNPs in *ALPL* to be in high linkage disequilibrium (LD) (Figure S3); therefore, we selected the genotyped rs10917006 in *ALPL* as an independent tagSNP, having considered P values, LD and predicted functions, along with rs1950902 in *MTHFD1* (also a genotyped SNP), for further analyses. These two SNPs together with other variables were included in a multivariable stepwise Cox model, in which both (rs1950902 and rs10917006) remained significantly associated with CMSS at $P < 0.05$ (Table S7).

In multivariable Cox regression analyses using an additive model, we evaluated the effects of these two significant SNPs on risk of death, with adjustment for covariates where appropriate. In the MDACC dataset, we observed a significant risk effect of the *MTHFD1* rs1950902 A allele ($P_{\text{trend}} = 0.005$) and the *ALPL* rs10917006 T allele ($P_{\text{trend}} = 0.023$) on CM survival; these were also observed in the NHS/HPFS dataset ($P_{\text{trend}} = 0.007$ for the *MTHFD1* rs1950902 A allele and $P_{\text{trend}} = 0.004$ for the *ALPL* rs10917006 T allele) and in the MDACC and NHS/HPFS combined dataset ($P_{\text{trend}} = 0.0003$ for the *MTHFD1* rs1950902 A allele and $P_{\text{trend}} = 0.0007$ for the *ALPL* rs10917006 T allele) (Table 2). We presented Kaplan-Meier curves to illustrate the association of CMSS and risk genotypes of *MTHFD1* rs1950902 and *ALPL* rs10917006 (Figure 2a–2f). For illustrative purposes, regional association plots 50 kb up- and down-stream of the gene regions for *MTHFD1* and *ALPL* in the MDACC dataset were also generated (Figure S4).

Survival of CM patients with combined risk genotypes

To better estimate the joint association between the two independent SNPs and CMSS, we combined the risk genotypes (those associated with increased risk of death) of *MTHFD1* rs1950902 GA+AA (1 vs. 0 for GG) and *ALPL* rs10917006 CT+TT (1 vs 0 for CC) into one variable as a genetic score. First, we categorized all the patients into three groups according to the number of risk genotype (i.e., 0, 1 and 2). We found that the trend test indicated a risk-genotype dose-response effect. Specifically, the effect on CMSS increased as the number of risk genotypes increased in the MDACC dataset ($P_{\text{trend}} = 0.0001$), the NHS/HPFS dataset ($P_{\text{trend}} < 0.0001$) and the MDACC and NHS/HPFS combined dataset ($P_{\text{trend}} < 0.0001$) after adjustment for covariates where appropriate. We then dichotomized all patients into the 0 risk genotype group and the 1–2 risk genotypes group. As shown in Table 2, compared with the 0 risk genotype group, the 1–2 risk genotypes group had a significant higher CM-death risk in the MDACC dataset (adjusted Hazards Ratio [HR_{adj}] = 1.95; 95% confidence interval [CI] = 1.28–2.96, $P = 0.002$), the NHS/HPFS dataset ($HR_{\text{adj}} = 2.73$; 95% CI = 1.51–4.96, $P = 0.0009$) and the MDACC and NHS/HPFS combined dataset ($HR_{\text{adj}} = 2.09$; 95% CI =

1.49–2.92, $P < 0.0001$). We also used Kaplan-Meier curves to illustrate the association between the number of risk genotypes and CMSS (Figure 2g–2i).

Stratified analyses for the effect of combined risk genotypes on CMSS

Next, we performed stratified analyses to investigate whether the joint effect of risk genotypes on CMSS was modified by other variables including age, sex, tumour stage, Breslow thickness, ulceration and mitotic rate in the MDACC dataset but only age and sex in the NHS/HPFS dataset. Compared with those with the 0 risk genotype, patients in the 1–2 risk genotypes group exhibited significantly poorer survival; this was evident in all the subgroups except for the no ulceration subgroup, those with Breslow thickness ≤ 1 mm, and those with mitotic rate ≤ 1 in the MDACC dataset and the subgroups age ≤ 50 and male subjects in the NHS/HPFS dataset. However, no heterogeneity was found among these subgroups. (Table S8).

Receiver operating characteristic (ROC) curve and internal validation

To assess the predictive accuracy of the two independent SNPs, we generated the time-dependent area under receiver curve (AUC) of the ROC curve for CM patients by stage in the MDACC dataset in the presence of other variables. Consistently, the predictive performance of the ROC of 5-year CMSS was significantly improved by the addition of risk genotypes to the model (Figure 3a) with an increase in area from 86.70% to 89.15% ($P = 0.036$). In the time-dependent AUC, we found that predictive performance was improved by the addition of risk genotypes from the beginning of follow-up and persisting over times in CM patients with stages I/II (Figure 3b) but not in those with stages III/IV (data not shown), indicating the observed effect in early stage CM patients. We did not perform further ROC analysis in the NHS/HPFS dataset, because only age and sex were available.

Genotype-phenotype correlation analyses

Two independent SNPs showed some evidence of functional relevance based on the online prediction tools, including SNPinfo and RegulomeDB. For example, *MTHFD1* rs1950902 is predicted to be a missense variant, which might play a role in exonic splicing enhancer regulation. *ALPL* rs10917006 is located in an intron and predicted to affect translation of protein USF2. To investigate possible correlations between the SNP genotypes and their corresponding mRNA expression levels, we conducted expression quantitative trait loci analysis with data from The Cancer Genome Atlas database. Although there were no significant associations between genetic variants of *MTHFD1* rs1950902 and corresponding mRNA levels, another SNP rs4902284 in strong LD with rs1950902 ($r^2 = 0.97$ in HaploReg v4.1) was significantly associated with mRNA expression levels of *MTHFD1* in primary CM tissue (P value in additive and dominant models were 0.041 and 0.035, respectively) (Figure S5a–b). We also found that the minor *ALPL* rs10917006 T allele was associated with an increased mRNA expression of *ALPL* in normal skin from the sun-exposed lower leg and unexposed suprapubic area, but a borderline in the whole blood from the Genotype-Tissue Expression (GTEx) Portal (<http://www.gtexportal.org/home/>) ($P = 0.0007$, 0.014 and 0.062, respectively, Table S9). However, no significant associations were observed between genotypes of these two SNPs and their mRNA expression levels in blood samples from 373 Europeans from the 1000 Genomes Project (data not shown). Moreover, according to

experimental data from the ENCODE project, two SNPs (rs1950902 and rs10917006) are located in a DNase I hypersensitive site, where histone modification H3K27 acetylation predicted some active enhancer and promoter functions (Figure S6). These results support the possibility that these two independent SNPs have an effect on their gene expression at the transcription level.

Discussion

Folate, as a cofactor in *de novo* nucleotide synthesis, is critically required for cell division, growth and DNA repair. Mutations in folate metabolism genes, such as methionine synthase reductase, cause epigenetic instability on mouse development¹⁷. In cancer cells, where DNA replication and cell division occur at a much faster rate, the blockade of folate metabolism causes inhibition of tumour proliferation and growth, forming the rational for using antifolate drugs in chemotherapy⁹. Recently, genetic variants in the folate metabolism genes have been linked to risk of cancer, such as pancreatic cancer, meningioma and adult glioma^{18,19}. For example, lung cancer risk was also reported to be associated with genetic variants in the folate metabolism genes²⁰, and these associations may be modulated by dietary folate intake^{21,22}. Another study found that folate intake affected ovarian cancer survival²³. Although epidemiologic studies have generated contradictory results on the protection of folate intakes or concentrations against risk of cancers²⁴, higher intake of folate from food was found to be associated with a modestly increased risk of CM⁷ and skin cancers (including melanoma and keratinocyte cancer)²⁵. These lines of evidence suggested that folate intake and its related metabolism might affect progression of melanoma, and this association is important for incorporating folate intervention into clinical practice, but no such studies have been reported. The present study, however, does provide the support for a link of genetic variants in *MTHFD1* and *ALPL* to CM patient survival.

In the present study, we provided additional evidence for associations of CMSS with genetic variants in two folate metabolic genes, *MTHFD1* and *ALPL*. Specifically, CM patients with a higher number of risk genotypes of these two genes had worse survival. It is worth mentioning that the findings were consistent across different analyses and the majority of subgroup comparisons, suggesting a robust association between the number of risk genotypes and CMSS. In particular, the genotype-phenotype correlation demonstrates that *MTHFD1* expression levels may be modulated by rs1950902 through another SNP in high LD and that *ALPL* expression levels may be regulated by rs10917006, suggesting that these findings are biologically plausible.

MTHFD1, located on chromosome 14q23.3, is a gene encoding a trifunctional enzyme involved in nuclear *de novo* thymidylate biosynthesis in the folate-dependent one-carbon metabolism²⁶. Recently, *MTHFD1* has been reported to interacted with the histone acetyl reader bromodomain-containing protein 4, and thus to maintain a pool of nuclear folate metabolites and to control gene expression in leukemia²⁷. Since the folate pool is maintained by *MTHFD1*, alterations in its activity due to genetic variants could modify disease susceptibility²⁸. More importantly, knockdown of *MTHFD1* inhibited distant metastasis of melanoma cells *in vivo*, and metastasizing melanoma cells might reversibly increase expression of folate pathway enzymes, including *MTHFD1*, inhibition of which

might selectively impair metastasis via oxidative stress²⁹. Furthermore, though GWASs have associated *MTHFD1* variants with risk of colon cancer³⁰ and postmenopausal breast cancer³¹, *MTHFD1* variants were also associated with event-free survival of patients with childhood acute lymphoblastic leukemia³². However, no published reports have linked *MTHFD1* variants to CM patient survival. Given that *MTHFD1* is a key enzyme providing one-carbon units for thymidylate biosynthesis³³, the present study suggests that genetic variants in *MTHFD1* may play a role in CM progression and prognosis.

ALPL, a gene that encodes a tissue non-specific isozyme, is located on chromosome 1p36.12. This enzyme is a defining marker of osteoblast activity and an important regulator of epithelial plasticity that influences biological processes including regulation of extracellular ATP levels and cell apoptosis^{34–36}. However, *ALPL* also acts as an oncogene in certain cancers. For example, a recent study showed that tumour-derived alkaline phosphatase, often elevated in metastatic cancer tissues, was associated with reduced disease-free survival in prostate cancer³⁶. Together with other stem-cell markers, a high expression of *ALPL* was significantly associated with shorter survival in refractory glioblastoma³⁷. To date, there is no report of altered *ALPL* expression in CM tumour tissues, and the present study suggests that genetic variants in *ALPL* may influence CM progression by regulating its mRNA expression levels, although the exact mechanisms underlying the observed association remain unknown.

The present study has some limitations. One of the limitations is that the Q-Q plots deviated considerably between the observed and expected ones across the continuum of $-\log_{10}(p)$, which may indicate some selection bias in the MDACC dataset. One potential weakness was the incomplete adjustment in the NHS/HPFS dataset used in the validation, because unlike the MDACC dataset used in discovery, the NHS/HPFS GWAS dataset contained only age and sex. However, the results of significant association with risk of death across different datasets and strata were consistent, suggesting that our findings might not be affected by the incomplete adjustment. Another limitation is the lack of information about different treatments, which may have had an effect on outcomes. Nevertheless, it is highly unlikely that treatment plans varied systematically by germline genotype, nor did we observe differences in death risk by tumour stage, Breslow thickness, ulceration, or mitosis, which suggests minimal bias from different treatments, if any. Although participants in the NHS/HPFS validation dataset have folate intake information, the discovery dataset lack such detailed information. The validation dataset, which has a relative smaller sample size, was used as validation for genetic data, not for folate intake, which was also a potential study limitation. Another caveat is that the sample sizes in the two GWAS studies were not large enough to allow for the false discovery rate test, a more desirable multiple test correction method; however, the BFDP might be more appropriate for the highly correlated SNPs included in the analyses, as more than 84% of these SNPs were imputed based on the LD method. These imputed SNPs were also a limitation. Finally, because the GTEx portal and other biological function-prediction databases are insufficient in evaluating mechanistic function of SNPs, further functional investigation is needed to validate our findings.

In summary, variants of genes involved in folate metabolism may play an important role in determining cancer susceptibility, also affecting the response to therapy with antifolate drugs

and eventually become useful in optimizing the treatment of CM patients. Given the importance of the folate metabolism pathway in biological processes and the number of risk genotypes of these two validated SNPs in folate metabolic pathway genes, these SNPs may be useful in personalized prognosis and treatment of CM patients, but it requires a large validation cohort to substantiate these findings in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding sources

The MD Anderson Study was supported by National Institutes of Health/National Cancer Institute (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 Harvard NHS/HPFS Study was in part supported by National Institutes of Health/National Cancer Institute (R01 CA49449, P01 CA87969, UM1 CA186107 and UM1 CA167552). This work was also partly supported by National Natural Science Foundation of China (No.81625020, 81402736). Qingyi Wei was supported by start-up funds from Duke Cancer Institute, Duke University Medical Center and also in part supported by the Duke Cancer Institute as part of the P30 Cancer Center Support Grant (Grant ID: NIH CA014236). Wei Dai was sponsored by the China Postdoctoral Science Foundation funded project (2019M662982).

Abbreviations

ALPL	alkaline phosphatase
AUC	area under curve
BFDP	Bayesian false-discovery probability
CI	confidence interval
CM	cutaneous melanoma
CMSS	cutaneous melanoma-specific survival
GTE_x	Genotype-Tissue Expression
GWAS	genome-wide association study
HR	hazards ratio
HR_{adj}	adjusted hazards ratio
LD	linkage disequilibrium
MDACC	The University of Texas MD Anderson Cancer Center
MTHFD1	methylenetetrahydrofolate dehydrogenase 1
NHS	the Nurses' Health Study
HPFS	the Health Professionals Follow-up Study

ROC	receiver operating characteristic
SNP	single-nucleotide polymorphism

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018; 68: 7–30. [PubMed: 29313949]
2. Ankeny JS, Labadie B, Luke J et al. Review of diagnostic, prognostic, and predictive biomarkers in melanoma. *Clin Exp Metastasis* 2018.
3. Choi SW, Mason JB. Folate status: effects on pathways of colorectal carcinogenesis. *J Nutr* 2002; 132: 2413S–8S.
4. Choi SW, Mason JB. Folate and carcinogenesis: an integrated scheme. *J Nutr* 2000; 130: 129–32. [PubMed: 10720158]
5. Duthie SJ, Hawdon A. DNA instability (strand breakage, uracil misincorporation, and defective repair) is increased by folic acid depletion in human lymphocytes in vitro. *FASEB J* 1998; 12: 1491–7. [PubMed: 9806758]
6. Widschwendter M, Jones PA. DNA methylation and breast carcinogenesis. *Oncogene* 2002; 21: 5462–82. [PubMed: 12154408]
7. Yang K, Fung T, Nan H. An epidemiological review of diet and cutaneous malignant melanoma. *Cancer Epidemiol Biomarkers Prev* 2018.
8. Vollset SE, Clarke R, Lewington S et al. Effects of folic acid supplementation on overall and site-specific cancer incidence during the randomised trials: meta-analyses of data on 50,000 individuals. *Lancet* 2013; 381: 1029–36. [PubMed: 23352552]
9. Smith AD, Kim YI, Refsum H. Is folic acid good for everyone? *Am J Clin Nutr* 2008; 87: 517–33. [PubMed: 18326588]
10. Butzbach K, Epe B. Photogenotoxicity of folic acid. *Free Radic Biol Med* 2013; 65: 821–7. [PubMed: 23973753]
11. Elechalawar CK, Sridharan K, Pal A et al. Cationic folate-mediated liposomal delivery of bis-arylidene oxindole induces efficient melanoma tumor regression. *Biomater Sci* 2017; 5: 1898–909. [PubMed: 28715002]
12. Skinner CC, McMichael EL, Jaime-Ramirez AC et al. Folate-conjugated immunoglobulin targets melanoma tumor cells for NK cell effector functions. *Melanoma Res* 2016; 26: 329–37. [PubMed: 27035691]
13. Sud A, Kinnarsley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer* 2017; 17: 692–704. [PubMed: 29026206]
14. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet* 2007; 81: 1278–83. [PubMed: 17966091]
15. Amos CI, Wang LE, Lee JE et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet* 2011; 20: 5012–23. [PubMed: 21926416]
16. Song F, Qureshi AA, Zhang J et al. Exonuclease 1 (EXO1) gene variation and melanoma risk. *DNA repair* 2012; 11: 304–9. [PubMed: 22230721]
17. Padmanabhan N, Jia D, Geary-Joo C et al. Mutation in folate metabolism causes epigenetic instability and transgenerational effects on development. *Cell* 2013; 155: 81–93. [PubMed: 24074862]
18. Chittiboyina S, Chen Z, Chiorean EG et al. The role of the folate pathway in pancreatic cancer risk. *PLoS One* 2018; 13: e0193298.
19. Chen D, Dong J, Huang Y et al. Folate metabolism genetic polymorphisms and meningioma and glioma susceptibility in adults. *Oncotarget* 2017; 8: 57265–77.
20. Durda K, Kaklewski K, Gupta S et al. Serum folate concentration and the incidence of lung cancer. *PLoS One* 2017; 12: e0177441.

21. Shi Q, Zhang Z, Li G et al. Polymorphisms of methionine synthase and methionine synthase reductase and risk of lung cancer: a case-control analysis. *Pharmacogenet Genomics* 2005; 15: 547–55. [PubMed: 16006998]
22. Shi Q, Zhang Z, Li G et al. Sex differences in risk of lung cancer associated with methylene-tetrahydrofolate reductase polymorphisms. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 1477–84. [PubMed: 15941959]
23. Dixon SC, Ibiebele TI, Protani MM et al. Dietary folate and related micronutrients, folate-metabolising genes, and ovarian cancer survival. *Gynecol Oncol* 2014; 132: 566–72. [PubMed: 24368279]
24. Mason JB. Unraveling the complex relationship between folate and cancer risk. *Biofactors* 2011; 37: 253–60. [PubMed: 21915934]
25. Donnenfeld M, Deschasaux M, Latino-Martel P et al. Prospective association between dietary folate intake and skin cancer risk: results from the Supplementation en Vitamines et Mineraux Antioxydants cohort. *Am J Clin Nutr* 2015; 102: 471–8. [PubMed: 26156743]
26. West MG, Horne DW, Appling DR. Metabolic role of cytoplasmic isozymes of 5,10-methylenetetrahydrofolate dehydrogenase in *Saccharomyces cerevisiae*. *Biochemistry* 1996; 35: 3122–32. [PubMed: 8608153]
27. Sdelci S, Rendeiro AF, Rathert P et al. MTHFD1 interaction with BRD4 links folate metabolism to transcriptional regulation. *Nature genetics* 2019; 51: 990–8. [PubMed: 31133746]
28. Field MS, Kamynina E, Stover PJ. MTHFD1 regulates nuclear de novo thymidylate biosynthesis and genome stability. *Biochimie* 2016; 126: 27–30. [PubMed: 26853819]
29. Piskounova E, Agathocleous M, Murphy MM et al. Oxidative stress inhibits distant metastasis by human melanoma cells. *Nature* 2015; 527: 186–91. [PubMed: 26466563]
30. Moruzzi S, Guarini P, Udali S et al. One-carbon genetic variants and the role of MTHFD1 1958G>A in liver and colon cancer risk according to global DNA methylation. *PLoS One* 2017; 12: e0185792.
31. Stevens VL, McCullough ML, Pavluck AL et al. Association of polymorphisms in one-carbon metabolism genes and postmenopausal breast cancer incidence. *Cancer Epidemiol Biomarkers Prev* 2007; 16: 1140–7. [PubMed: 17548676]
32. Krajcinovic M, Lemieux-Blanchard E, Chiasson S et al. Role of polymorphisms in MTHFR and MTHFD1 genes in the outcome of childhood acute lymphoblastic leukemia. *Pharmacogenomics J* 2004; 4: 66–72. [PubMed: 14647408]
33. Kamynina E, Lachenauer ER, DiRisio AC et al. Arsenic trioxide targets MTHFD1 and SUMO-dependent nuclear de novo thymidylate biosynthesis. *Proc Natl Acad Sci U S A* 2017; 114: E2319–E26.
34. Diez-Zaera M, Diaz-Hernandez JI, Hernandez-Alvarez E et al. Tissue-nonspecific alkaline phosphatase promotes axonal growth of hippocampal neurons. *Mol Biol Cell* 2011; 22: 1014–24. [PubMed: 21289095]
35. Millan JL. The role of phosphatases in the initiation of skeletal mineralization. *Calcif Tissue Int* 2013; 93: 299–306. [PubMed: 23183786]
36. Rao SR, Snaith AE, Marino D et al. Tumour-derived alkaline phosphatase regulates tumour growth, epithelial plasticity and disease-free survival in metastatic prostate cancer. *Br J Cancer* 2017; 116: 227–36. [PubMed: 28006818]
37. Iwadata Y, Suganami A, Tamura Y et al. The Pluripotent Stem-Cell Marker Alkaline Phosphatase is Highly Expressed in Refractory Glioblastoma with DNA Hypomethylation. *Neurosurgery* 2017; 80: 248–56. [PubMed: 28173571]

What's already known about this topic?

- Wide range of survival rates exists among melanoma patients with similar clinical characteristics, therefore development of complementary biomarkers with specific prognostic potential is needed.
- Hypothesis-driven approach by pooling the effects of single nucleotide polymorphisms in a specific biological pathway as genetic risk scores can increase their prognostic utility and genetic polymorphisms in folate metabolism have been associated with cancer risk.

What does this study add?

- Two genetic variants in the folate metabolic pathway genes, *MTHFD1* rs1950902 and *ALPL* rs10917006, are significantly associated with cutaneous melanoma-specific survival.
- The risk prediction model enables estimation of cutaneous melanoma-specific survival with the addition of risk genotypes in genes of folate metabolic pathway, and can provide personalized patient education.

Figure 1. Research workflow for SNPs in the folate metabolic pathway genes.

Abbreviations: *ALPL*, alkaline phosphatase; AUC, area under curve; BFDP, Bayesian false-discovery probability; CMSS, cutaneous melanoma-specific survival; GO, Gene Ontology; GWAS, genome wide association study; HWE, Hardy Weinberg equilibrium; KEGG, Kyoto Encyclopedia of Genes and Genomes; MAF, minor allele frequency; MDACC, The University of Texas M.D. Anderson Cancer Center; *MTHFD1*, methylenetetrahydrofolate dehydrogenase 1; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study; PID, Pathway Interaction Database; ROC, receiver operating characteristic; SNP, single nucleotide polymorphism.

Figure 2. Selected SNPs and survival prediction.

Kaplan-Meier curves of cutaneous melanoma-specific survival (CMSS) stratified by *MTHFD1* rs1950902, assuming a dominant model in (a) the MDACC, (b) the NHS/HPFS and (c) the MDACC and NHS/HPFS combined dataset. Kaplan-Meier curves of CMSS stratified by *ALPL* rs10917006 in (d) the MDACC, (e) the NHS/HPFS and (f) the MDACC and NHS/HPFS combined dataset. Kaplan-Meier survival curves of the combined risk genotypes on CMSS: dichotomized 0 risk genotype group and 1–2 risk genotypes group in (g) the MDACC, (h) the NHS/HPFS and (i) the MDACC and NHS/HPFS combined dataset. Abbreviations: SNP, single nucleotide polymorphism; CMSS, cutaneous melanoma-specific survival; *MTHFD1*, methylenetetrahydrofolate dehydrogenase 1; MDACC, The University of Texas M.D. Anderson Cancer Center; NHS, the Nurses' Health Study; HPFS, the Health Professionals Follow-up Study; *ALPL*, alkaline phosphatase.

Figure 3. ROC and AUC curves for CMSS prediction.

ROC curve (a) and time-dependent AUC (b) estimation for five-year CMSS prediction in cutaneous melanoma patients with stages I/II from MDACC dataset. Clinical variables include age, sex, tumour stage, Breslow thickness, ulceration and mitotic rate. Abbreviations: ROC, receiver operating characteristic; AUC, area under receiver curve; CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas M.D Anderson Cancer Center.

Table 1.

Meta-analysis of five validated SNPs in the folate metabolic pathway genes using two independently published melanoma GWAS datasets

SNP	Allele ¹	Gene	Discovery-MDACC (n=858)				Validation-NHS/HPFS (n=409)				Combined-Meta-analysis (n=1267)			
			EAF	HR (95% CI) ²	P ²	BFD ³	EAF	HR (95% CI) ⁴	P ⁴	BFD ³	P _{het}	I ²	HR (95% CI) ⁵	P ⁵
rs1950902 ⁶	G>A	<i>MTHFD1</i>	0.19	1.68 (1.17–2.41)	0.005	0.462	0.18	1.87 (1.18–2.94)	0.007	0.555	0.718	0	1.75 (1.32–2.32)	9.96×10 ⁻⁵
rs12086386 ⁷	C>T	<i>ALPL</i>	0.05	2.07 (1.19–3.59)	0.010	0.649	0.06	2.17 (1.19–3.95)	0.012	0.686	0.910	0	2.12 (1.41–3.18)	3.04×10 ⁻⁴
rs56222534 ⁷	T>C	<i>ALPL</i>	0.06	1.94 (1.16–3.25)	0.011	0.662	0.07	2.34 (1.32–4.16)	0.004	0.541	0.634	0	2.12 (1.44–3.09)	1.32×10 ⁻⁴
rs10917006 ⁶	C>T	<i>ALPL</i>	0.06	1.83 (1.09–3.09)	0.023	0.746	0.07	2.34 (1.32–4.16)	0.004	0.541	0.534	0	2.05 (1.39–3.01)	2.84×10 ⁻⁴
rs11586977 ⁷	C>A	<i>ALPL</i>	0.06	1.83 (1.09–3.09)	0.023	0.746	0.07	2.34 (1.32–4.16)	0.004	0.541	0.534	0	2.05 (1.39–3.01)	2.84×10 ⁻⁴

¹Reference allele/effect allele

²Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in an additive genetic model

³BFD³ was used for multiple test correction with detected a highest HR of 2.0 and a prior probability of 0.1

⁴Adjusted for age and sex in an additive genetic model

⁵Meta-analysis in a fix-effects model

⁶Genotyped SNPs in the MDACC dataset

⁷Imputed SNPs in the MDACC dataset.

Abbreviations: SNP, single-nucleotide polymorphism; GWAS, genome-wide association study; MDACC, The University of Texas MD Anderson Cancer Center; NHS/HPFS, the Nurses' Health Study/Health Professionals Follow-up Study; EAF, effect allele frequency; HR, hazards ratio; CI, confidence interval; BFD³, Bayesian false-discovery probability; P_{het} , P value for heterogeneity by Cochrane's Q test; I^2 , heterogeneity statistic; *MTHFD1*, methylenetetrahydrofolate dehydrogenase 1; *ALPL*, alkaline phosphatase

Table 2.

Associations between two independent SNPs in the folate metabolic pathway genes and CMSS of patients in the MDACC dataset, the NHS/HPFS dataset and the MDACC and NHS/HPFS combined dataset

Genotype	MDACC (n=858)				NHS/HPFS (n=409)				MDACC + NHS/HPFS (n=1267)			
	Frequency		Multivariate analysis ¹		Frequency		Multivariate analysis ²		Frequency		Multivariate analysis ³	
	All	Death (%)	HR (95% CI)	P	All	Death (%)	HR (95% CI)	P	All	Death (%)	HR (95% CI)	P
<i>MTHFD1</i> rs1950902 G>A												
GG	554	50 (9.03)	1.00		275	24 (8.73)	1.00		829	74 (8.93)	1.00	
GA	281	43 (15.30)	1.92 (1.26–2.92)	0.002	123	22 (17.89)	2.27 (1.27–4.06)	0.006	404	65 (16.09)	1.97 (1.41–2.74)	0.049
AA	23	2 (8.70)	1.51 (0.36–6.25)	0.573	11	2 (18.18)	2.13 (0.50–9.10)	0.308	34	4 (11.76)	1.51 (0.55–4.15)	0.007
Trend test				0.005				0.007				0.0003
GA+AA	304	45 (14.80)	1.90 (1.25–2.87)	0.003	134	24 (17.91)	2.26 (1.28–3.99)	0.005	438	69 (15.75)	1.93 (1.39–2.68)	<0.001
<i>ALPL</i> rs10917006 C>T												
CC	751	77 (10.25)	1.00		353	35 (9.92)	1.00		1104	112 (10.14)	1.00	
CT	107	18 (16.82)	1.83 (1.09–3.09)	0.023	54	12 (22.22)	2.34 (1.21–4.53)	0.012	161	30 (18.63)	1.92 (1.28–2.87)	<0.001
TT	0	0 (0.00)	-	-	2	1 (50.00)	5.44 (0.74–40.32)	0.097	2	1 (50.00)	4.94 (0.68–35.79)	0.420
Trend test				0.023				0.004				0.0007
CT+TT	107	18 (16.82)	1.83 (1.09–3.09)	0.023	56	13 (23.21)	2.45 (1.29–4.65)	0.006	163	31 (19.02)	1.95 (1.31–2.91)	0.001
Combined number of risk genotypes⁴												
0	482	41 (8.51)	1.00		236	17 (7.20)	1.00		718	58 (8.08)	1.00	
1	341	45 (13.20)	1.75 (1.13–2.71)	0.012	156	25 (16.03)	2.44 (1.31–4.52)	0.005	497	70 (14.08)	1.89 (1.33–2.68)	0.0004
2	35	9 (25.71)	4.32 (2.02–9.23)	0.0002	17	6 (35.29)	5.78 (2.25–14.84)	0.0003	52	15 (28.85)	4.18 (2.35–7.42)	<0.001
Trend test				0.0001				<0.001				<0.001
0	482	41 (8.51)	1.00		236	17 (7.20)	1.00		718	58 (8.08)	1.00	
1–2	376	54 (14.36)	1.95 (1.28–2.96)	0.002	173	31 (17.92)	2.73 (1.51–4.96)	0.0009	549	85 (15.48)	2.09 (1.49–2.92)	<0.001

¹ Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in the MDACC dataset

² Adjusted for age and sex in the NHS/HPFS dataset

³ Adjusted for age and sex in the MDACC and NHS/HPFS combined dataset

⁴ Risk genotypes include *MTHFD1* rs1950902 GA+AA and *ALPL* rs10917006 CT+TT.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Abbreviations: SNP, single-nucleotide polymorphism; CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas MD Anderson Cancer Center; NHS/HPFS, the Nurses' Health Study/Health Professionals Follow-up Study; HR, hazards ratio; CI, confidence interval; *MTHFD1*, methylenetetrahydrofolate dehydrogenase 1; *ALPL*, alkaline phosphatase.