Genetic variants in *ELOVL2* and *HSD17B12* predict melanoma-specific survival

Wei Dai¹,²,³, Hongliang Liu²,³,¹, Xinyuan Xu²,³, Jie Ge²,³, Sheng Luo⁴, Dakai Zhu⁵, Christopher I. Amos⁵, Shenying Fang⁶, Jeffrey E. Lee⁶, Xin Li⁷,⁸, Hongmei Nan⁷,⁸, Chunying Li⁵ and Qingyi Wei¹,²,³,⁹

¹Department of Dermatology, Xijing Hospital, Fourth Military Medical University, Xi’an, Shaanxi, China  
²Duke Cancer Institute, Duke University Medical Center, Durham, NC  
³Department of Population Health Sciences, Duke University School of Medicine, Durham, NC  
⁴Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC  
⁵Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX  
⁶Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX  
⁷Department of Epidemiology, Fairbanks School of Public Health, Indiana University, Indianapolis, IN  
⁸Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, MA  
⁹Department of Medicine, Duke University School of Medicine, Durham, NC

Fatty acids play a key role in cellular bioenergetics, membrane biosynthesis and intracellular signaling processes and thus may be involved in cancer development and progression. In the present study, we comprehensively assessed associations of 14,522 common single-nucleotide polymorphisms (SNPs) in 149 genes of the fatty-acid synthesis pathway with cutaneous melanoma disease-specific survival (CMSS). The dataset of 858 cutaneous melanoma (CM) patients from a published genome-wide association study (GWAS) by The University of Texas M.D. Anderson Cancer Center was used as the discovery dataset, and the identified significant SNPs were validated by a dataset of 409 CM patients from another GWAS from the Nurses’ Health and Health Professionals Follow-up Studies. We found 40 noteworthy SNPs to be associated with CMSS in both discovery and validation datasets after multiple comparison correction by the false positive report probability method, because more than 85% of the SNPs were imputed. By performing functional prediction, linkage disequilibrium analysis, and stepwise Cox regression selection, we identified two independent SNPs of *ELOVL2* rs37734398 T>C and *HSD17B12* rs11037684 A>G that predicted CMSS, with an allelic hazards ratio of 0.66 (95% confidence interval = 0.51–0.84 and p = 8.34 × 10⁻⁵) and 2.29 (1.55–3.39 and p = 3.61 × 10⁻⁵), respectively. Finally, the *ELOVL2* rs37734398 variant CC genotype was found to be associated with a significantly increased mRNA expression level. These SNPs may be potential markers for CM prognosis, if validated by additional larger and mechanistic studies.

Key words: cutaneous melanoma, fatty acid synthesis, single-nucleotide polymorphism, genome-wide association study, melanoma-specific survival

Abbreviations: CI: confidence interval; CM: cutaneous melanoma; CMSS: cutaneous melanoma-specific survival; *ELOVL2*: elongation of very long-chain fatty acids 2; FASN: fatty acid synthase; FPRP: false positive report probability; GWAS: genome-wide association study; HPFS: the Health Professionals Follow-up Study; HRadj: adjusted hazards ratio; *HSD17B12*: hydroxysteroid 17-beta dehydrogenase 12; LD: linkage disequilibrium; NHS: the Nurses’ Health Study; SNP: single-nucleotide polymorphism

Additional Supporting Information may be found in the online version of this article.

W.D. and H.L. contributed equally to this work.

Conflict of interest: The authors declare no conflict of interest.

Grant sponsor: NIH Clinical Center; Grant number: R01 CA100264, 2P50CA093459, R01CA133996, R01 CA494; Grant sponsor: China Scholarship Council; Grant sponsor: the State Scholarship Fund; Grant sponsor: the P30 Cancer Center Support Grant; Grant number: CA014236; Grant sponsor: the Marit Peterson Fund for Melanoma Research; Grant numbers: UM1CA167552, UM1CA186107, P01CA87969, R01CA49449, R01CA133996, 2P50CA093459, R01CA100264; Grant sponsor: NIH/NCI

DOI: 10.1002/ijic.32194

History: Received 2 Sep 2018; Accepted 11 Jan 2019; Online 8 Feb 2019

Correspondence to: Chunying Li, Department of Dermatology, Xijing Hospital, Fourth Military Medical University, Xi’an, Shaanxi, China, E-mail: lichying@fmmu.edu.cn; or Qingyi Wei, Duke Cancer Institute, Duke University Medical Center and Department of Population Health Sciences, Duke School of Medicine, 905 S LaSalle Street, Durham, NC 27710, E-mail: qingyi.wei@duke.edu

Int. J. Cancer: 00, 00–00 (2019) © 2019 UICC
What’s new?
Fatty acid synthesis provides metabolic substrates for energy storage, membrane building and signaling transduction and has been linked to cancer prognosis. Here the authors analyzed associations between variants in genes in the fatty acid synthesis pathway and cutaneous melanoma-specific survival using datasets from two published genome-wide association studies. They identified two variants in genes linked to elongation of long-chain fatty acids (ELOVL2 and HSD17B12) that were significantly associated with survival, highlighting their potential as prognostic factors for melanoma patients.

Introduction
Cutaneous melanoma (CM) has the highest mortality rate among all skin cancers, ranking the fifth most common cancer among males and the sixth among females in the United States. In 2018, an estimated 91,270 new CM cases will be diagnosed in the United States (in addition to 87,290 in situ cases), and the CM incidence rate continues to rise. Although many CM patients are considered having an in situ or localized disease, these low-risk cases also comprise a substantial fraction of the overall burden of lethal CM. CM patients can be classified to having a relative low, average or high risk of recurrence and death according to the American Joint Committee on Cancer; however, an estimated 10–20% of the cases will develop an outcome different from the predicted one. Therefore, the identification of alternative prognosis biomarkers is needed.

CM is a disorder of uncontrolled melanocytic cell growth and proliferation, in which cellular metabolism is programmed. For example, high levels of carbon flux through aerobic glycolysis accumulate metabolic intermediates as sources of cellular building blocks, and an increased fatty acid synthesis provides metabolic substrates for energy storage, membrane building and signaling transduction, which have been shown to be strongly associated with cancer prognosis. Furthermore, lipogenic enzymes in the fatty acid synthesis, such as the ATP citrate lyase, fatty acid synthase (FASN) and stearoyl-CoA desaturase, have emerged as potential therapeutic targets in cancer treatment. Chemical inhibition or genetic knock-down of these key enzymes lead to a reduced proliferation and survival of cancer cells in xenograft tumor models. Interestingly, one study found that inhibition of fatty acid desaturation also increased the chemosensitivity of cancer cells that had an induced apoptosis by the mitochondrial pathway, suggesting an important role of the fatty acid metabolism in cancer cell survival and drug resistance. In melanocytes and melanoma cells, fatty acids regulate the degradation of tyrosinase, a critical enzyme associated with melanin biosynthesis. It has also been reported that alterations in the fatty acid synthesis in melanoma cells helped the cells evade apoptosis and sustain survival after ultraviolet A exposure.

Given the importance of fatty acid synthesis in cancer development and progression, we aimed to identify novel genetic variants in the fatty acid synthesis pathway genes in their associations with survival of CM patients by using two published genome-wide association study (GWAS) datasets, which may provide a new clue to novel cancer therapies with interruption of the fatty acid metabolism.

Materials and Methods
Study populations
In the present study, we used 858 CM patients from The University of Texas MD Anderson Cancer Center (MDACC) study as a discovery dataset and 409 CM patients from the Nurses’ Health Study (NHS) and the Health Professionals Follow-up Studies (HPFS) as a validation dataset, and the published GWAS data were available for both discovery and validation studies. Detailed descriptions of subject selection and data collection for both discovery and validation studies were described elsewhere. The approval to perform the present study was granted by Institutional Review Boards at both MD Anderson and Brigham and Women’s Hospital with written informed consent obtained from all participants.

Gene selection and single-nucleotide polymorphism (SNP) genotyping
We selected 149 fatty acid synthesis pathway genes that are located on the autosomes according to the databases of the Molecular Signatures Database v6.2 of Gene Set Enrichment Analysis website (Table S1, Supporting Information). In the MDACC dataset, genomic DNA extracted from the whole blood was genotyped by the Illumina HumanOmniQuad_v1_0_B array using the National Center for Biotechnology Information Database of Genotypes and Phenotypes (accession: phs000187.v1.p1). Genome-wide imputation was performed by using the MACH software based on the 1000 Genomes Project phase 1 v2 CEU. In brief, the typed or imputed common SNPs (with minor allele frequency ≥0.05, genotyping success rate ≥95 %, and Hardy–Weinberg equilibrium p value ≥0.00001, and from imputation for those SNPs with r2 ≥0.8) within genes in the fatty acid synthesis pathway or their ±2 kilobase flanking regions were selected for association analysis. Meanwhile, in the NHS/HPFS study, genotyping was performed using the Illumina HumanHap550 array, HumanHap610 array and Affymetrix 6.0 array. Imputation analysis was based on genotyped SNPs and haplotype information from the 1000 Genomes Phase III data using the program MACH. We selected the SNPs by the same standard used in the discovery dataset.
Statistical methods

The cutaneous melanoma-specific survival (CMSS) time was calculated from the time of diagnosis until death from CM. Specially associations between SNPs and CMSS were assessed by multivariable Cox proportional hazards regression analyses using the GenABEL package of R software with adjustment for age, sex, Breslow thickness, regional/distant metastasis, ulceration and mitotic rate in the MDACC dataset. In the validation analysis from the NHS/HPFS study, only age and sex were available for adjustment.

We used the false positive report probability (FPRP) method to correct for multiple testing, because more than 85% of SNPs included in the present study were imputed and thus in linkage disequilibrium (LD) with other genotyped SNPs. Three factors determine the magnitude of FPRP: the level of p values, the prior probability of a true association of the tested genetic variant with a disease, and the statistical power to detect the odds or hazards ratios of the alternative hypothesis at the given condition. Only the significant results with an FPRP value < 0.2 in both discovery and validation datasets were considered noteworthy. We also used a prior probability of 0.1 to detect a hazards ratio (HR) of 2.0 for an association with variant genotypes or minor alleles of the SNPs with \( p < 0.05 \).

To evaluate the effects of genetic variants on the cumulative probability of CMSS, Kaplan–Meier survival curves and log-rank tests were performed. The establishment of the number of risk genotypes was used to estimate the joint effect of the multi- genetic variants. In the present study, we calculated a genotype score from the number of risk genotypes and performed multivariable Cox regression models to assess the association between the genotype score and CMSS. To assess the SNPs of interest and cumulative incidence of CM-specific death, where death from other causes other than CM was modeled as a competing event. A Fine-Gray competing risk regression model was performed for univariate and multivariable regression analyses, which calculates subdistribution HR adjusted for age, sex, regional/distant metastasis, Breslow thickness, ulceration, and mitotic rate. A Manhattan plot showing the associations between 14,522 SNPs and CMSS is presented in Figure S1, Supporting Information. As a result, 1,042 SNPs were significantly associated with CMSS at \( p < 0.05 \) in an additive genetic model, of which 538 SNPs were still considered noteworthy after the multiple test correction by FPRP, which took into account of the fact that the vast majority of the SNPs under investigation were imputed with a LD approach. Among the 538 SNPs, 40 were validated in the NHS/HPFS dataset and remained significantly associated with CMSS at \( p < 0.05 \) after the correction by an FPRP < 0.2. 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/), 13 of these 40 SNPs were predicted to be putatively functional, including two SNPs in ELOVL2 (the elongation of very long chain fatty acids 2 gene) and eleven SNPs in HSD17B12 (the hydroxysteroid 17-beta dehydrogenase 12 gene) (Table S3, Supporting Information). In the subsequent meta-analysis of the two datasets, the 13 SNPs in ELOVL2 and HSD17B12 remained significant in associations with CMSS (Table 1) without heterogeneity between the two datasets (\( p_{\text{het}} > 0.05 \) for both).

Results

Subject characteristics

In the MDACC dataset, there were slightly more male patients (496, 57.8%) than female patients with an age range between 17 and 94 years at diagnosis (a median age of 53 years); 56.8% of these cases were older than 50 years; and 82.6% (709) had been classified as no regional/distant metastasis. Univariate Cox regression analysis suggested that age, sex, stage, Breslow thickness, ulceration and mitotic rate were significantly associated with CMSS. For the NHS/HPFS study, the dataset only had age, sex, survival outcome and genotype data with an age range between 34 and 87 years at diagnosis (a median age of 60 years), and the majority of the cases were over 50 years old (357, 82.4%) with more female patients (271, 66.3%). The patients from the MDACC dataset had a relatively shorter median follow-up time of 81.1 months with a range between 4.7 to 175.3 months, compared to 179.0 months with a range between 5.0 to 453.0 months for NHS/HPFS patients (Table S2, Supporting Information).

Associations between SNPs in the fatty acid synthesis pathway genes and CMSS

Figure 1 provides a flowchart of study design to illustrate the present study. To assess the associations of 2,161 genotyped and 13,361 imputed SNPs of the fatty acid synthesis pathway genes with CMSS, we performed the single locus analysis by using multivariate Cox proportional hazards regression in the MDACC dataset with adjustments for age, sex, regional/distant metastasis, Breslow thickness, ulceration, and mitotic rate. A Manhattan plot showing the associations between 14,522 SNPs and CMSS is presented in Figure S1, Supporting Information. As a result, 1,042 SNPs were significantly associated with CMSS at \( p < 0.05 \) in an additive genetic model, of which 538 SNPs were still considered noteworthy after the multiple test correction by FPRP, which took into account of the fact that the vast majority of the SNPs under investigation were imputed with a LD approach. Among the 538 SNPs, 40 were validated in the NHS/HPFS dataset and remained significantly associated with CMSS at \( p < 0.05 \) after the correction by an FPRP < 0.2. 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/), 13 of these 40 SNPs were predicted to be putatively functional, including two SNPs in ELOVL2 (the elongation of very long chain fatty acids 2 gene) and eleven SNPs in HSD17B12 (the hydroxysteroid 17-beta dehydrogenase 12 gene) (Table S3, Supporting Information). In the subsequent meta-analysis of the two datasets, the 13 SNPs in ELOVL2 and HSD17B12 remained significant in associations with CMSS (Table 1) without heterogeneity between the two datasets (\( p_{\text{het}} > 0.05 \) for both).
Genetic variants in fatty acid synthesis pathway genes as independent death predictors

We further performed LD analysis of the 13 SNPs in ELOVL2 and HSD17B12 and found that two SNPs of ELOVL2 were in a high LD and that 11 SNPs of HSD17B12 were also in a high LD (Fig. S2, Supporting Information). In consideration of p values, LD and predicted functions, we selected ELOVL2 rs3734398 (genotyped) and HSD17B12 rs11037684 (genotyped) as the independent tagSNPs for further analysis.

An initial stepwise Cox regression analyses of selected clinical variables from the MDACC dataset suggested these two SNPs were independent predictors of CMSS (Table S4, Supporting Information). In multivariate Cox regression analysis using an additive model, we evaluated the effects of these two significant SNPs on death risk with adjustment for clinico-pathological covariates (i.e., age, sex, Breslow thickness, regional/distant metastasis, ulceration of tumor, and tumor cell mitotic rate) in the MDACC dataset but only for age and sex in the NHS/HPFS dataset. In the MDACC study, we observed a statistically significant protective effect of the ELOVL2 rs3734398 C allele (p_{trend} = 0.027) but a risk effect of the HSD17B12 rs11037684 G allele (p_{trend} = 0.007) on CM-specific survival. Similar results were observed for the ELOVL2 rs3734398 C allele in the NHS/HPFS dataset (p_{trend} = 0.005) and the combined dataset of both MDACC and NHS/HPFS (p_{trend} = 0.003). Similarly, the risk effect of the HSD17B12 rs11037684 G allele was observed in the NHS/HPFS dataset (p_{trend} = 0.002) and the combined dataset of both MDACC and NHS/HPFS (p_{trend} = 0.002) (Table 2). To further visualize the HR effects, we used Kaplan–Meier survival curves for the associations between CMSS and risk genotypes of ELOVL2 rs3734398 and HSD17B12 rs11037684 in the combined dataset of both MDACC and NHS/HPFS (Figs. 2a and 2b).

In the Fine-Gray competing-risks regression model, the cumulative incidence of an event of interest was calculated in the presence of competing risks (death not caused by CM). During the follow-up time, 38 and 91 patients died of causes other than CM in the MDACC and NHS/HPFS datasets, respectively. In multivariate competing risks regression models, ELOVL2 rs3734398 was a statistically significant predictor of CMSS, after accounting for the postdiagnosis mortality in both datasets (with subdistribution HR of 0.72 in the MDACC dataset and 0.53 in the NHS/HPFS dataset, respectively); similarly, HSD17B12 rs11037684 was also a significant predictor in the MDACC dataset (subdistribution HR = 1.93 and p = 0.014) and NHS/HPFS dataset (subdistribution HR = 2.56 and p = 0.002). In the subsequent meta-analyses, for both rs3734398 and rs11037684, the direction, magnitude, and significance of subdistribution HR of CMSS were consistent with the cause-specific HR (Table S5, Supporting Information). Furthermore, regional association plots for the MDACC dataset were generated for
<table>
<thead>
<tr>
<th>SNP</th>
<th>Allele</th>
<th>Gene</th>
<th>Position</th>
<th>MDACC (n = 858)</th>
<th>NHS/HPFS (n = 409)</th>
<th>Meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EAF HR (95% CI)</td>
<td>EAF HR (95% CI)</td>
<td>Phet</td>
</tr>
<tr>
<td>rs3734398²</td>
<td>T&gt;C</td>
<td>ELOVL2</td>
<td>6p24.2</td>
<td>0.41 0.72 (0.54–0.96)</td>
<td>0.027 0.195 0.42 (0.33–0.83)</td>
<td>0.005</td>
</tr>
<tr>
<td>rs3798713²</td>
<td>G&gt;C</td>
<td>ELOVL2</td>
<td>6p24.2</td>
<td>0.43 0.68 (0.51–0.91)</td>
<td>0.019 0.076 0.42 (0.33–0.82)</td>
<td>0.005</td>
</tr>
<tr>
<td>rs1103768³</td>
<td>A&gt;C</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.10 (1.22–3.61)</td>
<td>0.007 0.132 0.06 2.52 (1.42–4.46)</td>
<td>0.002</td>
</tr>
<tr>
<td>rs1103768⁴</td>
<td>A&gt;G</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.10 (1.22–3.61)</td>
<td>0.007 0.132 0.06 2.52 (1.42–4.46)</td>
<td>0.002</td>
</tr>
<tr>
<td>rs6188380⁵</td>
<td>A&gt;G</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.23 (1.21–4.10)</td>
<td>0.010</td>
</tr>
<tr>
<td>rs6188380⁶</td>
<td>A&gt;G</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.40 (1.33–4.33)</td>
<td>0.004</td>
</tr>
<tr>
<td>rs6188381⁷</td>
<td>T&gt;C</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.40 (1.33–4.33)</td>
<td>0.004</td>
</tr>
<tr>
<td>rs6188384⁸</td>
<td>T&gt;C</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.34 (1.30–4.23)</td>
<td>0.005</td>
</tr>
<tr>
<td>rs6188384⁹</td>
<td>G&gt;T</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.34 (1.30–4.23)</td>
<td>0.005</td>
</tr>
<tr>
<td>rs1751455⁵</td>
<td>T&gt;A</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.34 (1.30–4.23)</td>
<td>0.005</td>
</tr>
<tr>
<td>rs1759911⁶</td>
<td>C&gt;T</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.34 (1.30–4.23)</td>
<td>0.005</td>
</tr>
<tr>
<td>rs1759868⁷</td>
<td>A&gt;C</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.25 (1.31–3.86)</td>
<td>0.003 0.084 0.06 2.23 (1.21–4.10)</td>
<td>0.010</td>
</tr>
<tr>
<td>rs7773945²</td>
<td>T&gt;A</td>
<td>HSD17B12</td>
<td>11p11.2</td>
<td>0.06 2.22 (1.29–3.82)</td>
<td>0.004 0.091 0.06 2.40 (1.33–4.33)</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Abbreviations: SNP, single-nucleotide polymorphism; GWAS, genome-wide association study; MDACC, The University of Texas MD Anderson Cancer Center; NHS, the Nurses’ Health Study; HPFS, the Health Professionals Follow-up Study; EAF, effect allele frequency; HR, hazards ratio; CI, confidence interval; FPRP, false positive report probability; Phet, P value for heterogeneity by Cochrane’s Q test; ELOVL2, elongation of very long-chain fatty acids 2; HSD17B12, hydroxysteroid dehydrogenase type 12.

1Reference allele/effect allele.
2Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in the additive model.
3FPRP was used for multiple test correction because 85.1% of the analyzed SNPs in MDACC dataset were imputed with a high level of linkage disequilibrium.
4Adjusted for age and sex in an additive genetic model.
5Meta-analysis in the fix-effect model.
6Genotyped SNPs in the MDACC dataset.
7Imputed SNPs in the MDACC dataset.
ELOVL2 and HSD17B12, including the 200-kb regions flanking the neighborhoods of these two genes (Fig. S3, Supporting Information).

Survival of CM patients with combined risk genotypes

To better estimate the joint effect of the two tagSNPs on risk of death, we combined the risk genotypes (those associated with an increased death risk) of ELOVL2 rs3734398 TT and HSD17B12 rs11037684 AG + GG into one variable as a genetic score. We then categorized all the patients into three groups with 0, 1 and 2 risk genotype. As illustrated in Table 2, we observed a risk-genotype dose–response effect; that is, the effect on CMSS increased as the number of risk genotypes increased in the MDACC dataset ($p_{\text{trend}} = 0.007$), the NHS/HPFS dataset ($p_{\text{trend}} < 0.0001$) and the combined dataset of both MDACC and NHS/HPFS ($p_{\text{trend}} < 0.0001$) after adjustments for covariates where appropriate. We next dichotomized all patients into the 0 risk genotype group and the 1–2 risk genotype groups and found that, compared to the 0 risk genotype group, the 1–2 risk genotypes group had a higher CM-death risk in the MDACC dataset (adjusted hazards ratio [HR$_{\text{adj}}$] = 1.66, 95% CI = 1.09–2.53 and $p = 0.019$), the NHS/HPFS dataset (2.82, 1.56–5.10 and 0.0006) and the combined dataset of both MDACC and NHS/HPFS (1.79, 1.29–2.50 and 0.0005). Figure 2c shows the Kaplan–Meier curves for the associations between risk genotypes and CMSS.

Stratified analyses for the effect of combined risk genotypes on CMSS

We further conducted stratified analyses to investigate whether the joint effect of risk genotypes on CMSS was modified by clinicalpathologic variables including age, sex, distant/regional metastasis, Breslow thickness, ulceration and mitotic rate in the MDACC dataset and age and sex in the NHS/HPFS dataset. As a result, patients with the 1–2 risk genotypes group, compared to the 0 risk genotype group, showed a substantially increased risk of CM-associated death in the presence of clinical variables, which were more evident in the subgroups of age ≤50, male subjects and those with tumor cell mitotic rate of ≤1/mm$^2$ in the MDACC dataset and the subgroups age >50 and female subjects in the NHS/HPFS dataset. However, no significant interaction was found among all the subgroups (Table S6, Supporting Information).

ROC and AUC estimation for CMSS prediction

To assess the ability of risk genotypes to predict CMSS, we compared the model with ROC for clinical variables where appropriate to that of ROC for both clinical variables and risk genotypes. Consistently, the AUC of the five-year CMSS improved prediction performance in the MDACC dataset, the NHS/HPFS dataset and the combined dataset of both MDACC and NHS/HPFS with the addition of risk genotypes to the model (Supporting Information Figs. S4a, S4c and S4e). Only the AUC of the five-year CMSS in the NHS/HPFS dataset significantly increased from 54.05% to 73.51% ($p = 0.022$) with the addition of risk genotypes to the model. In addition, the time-dependent AUC curves were also provided to assess the ability of risk genotypes to predict CMSS through the entire follow-up period in the above-mentioned three datasets (Supporting Information Figs. S4b, S4d and S4f).

Genotype–phenotype correlation analyses

We further evaluated the correlations between SNPs and their corresponding mRNA expression levels using publically available RNA-seq data of 373 lymphoblastoid cell lines from the 1000 Genomes Project. Notably, the rs3734398 C allele was significantly correlated with mRNA expression levels of ELOVL2 in an additive model ($p = 0.024$, Fig. 2d). We also performed expression quantitative trait loci (eQTL) analysis using genomic data from the Genotype-Tissue Expression (GTEx) Project (http://www.gtexportal.org/home), which includes ELOVL2 rs3734398 in transformed fibroblasts from 300 donors. We found that rs3734398 C allele was associated with a significantly increased ELOVL2 mRNA expression level ($p = 7.3 \times 10^{-7}$) in an additive genetic model (Fig. 2e), which is consistent with our initial findings. However, there was no significant correlation between rs11037684 genotypes and HSD17B12 mRNA expression levels ($p = 0.911$, 0.988 and 0.547 for additive, dominant and recessive models, respectively) (Fig. S5, Supporting Information) in the 1000 Genomes Project nor in the GTEx. No significant associations between selected SNPs and their corresponding mRNA expression levels were observed in the normal skin tissues from the sun exposed lower leg and the unexposed suprapubic (Table S7, Supporting Information) from the 1000 Genomes Project (http://www.internationalgenome.org/). We also evaluated the correlations between SNPs and their corresponding mRNA expression levels using experimental data from the ENCODE Project (Fig. S6, Supporting Information), we found the two SNPs (i.e., rs3734398 and rs11037684) to be located in a DNase I hypersensitive site, where the DNase hypersensitivity and histone modification H3K27 acetylation indicated some signals for active enhancer and promoter functions. The evidence from the DNase cluster and transcription factor CHIP-seq data suggests that rs3734398 is located on the neighborhood of these two genes (Fig. S3, Supporting Information).

Discussion

In the present study, we found that genetic variants ELOVL2 rs3734398 and HSD17B12 rs11037684 were likely to independently or jointly modulate the survival of CM patients. We also observed a dose–response effect of their combined risk genotypes on CMSS. Moreover, the rs3734398 C allele was correlated with an increase in ELOVL2 mRNA expression level in lymphoblastoid cell lines derived from 373 European descendants from the 1000 Genomes Project. These findings are biologically plausible, because the fatty acid synthesis pathway contributes to membrane biosynthesis, energy storage and the regulation of oncogenic signaling.
Table 2. Associations between two independent SNPs in the fatty acid synthesis pathway genes and CMSS of patients in the MDACC dataset, the NHS/HPFS dataset and the combined dataset of both MDACC and NHS/HPFS

<table>
<thead>
<tr>
<th>Genotype</th>
<th>MDACC (n = 858)</th>
<th>NHS/HPFS (n = 409)</th>
<th>MDACC + NHS/HPFS (n = 1,267)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Multivariate analysis</td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td>All Death (%)</td>
<td>HR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td><strong>ELOVL2 rs3734398 T&gt;C</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>310</td>
<td>40 (12.90)</td>
<td>1.00</td>
</tr>
<tr>
<td>TC</td>
<td>385</td>
<td>42 (10.91)</td>
<td>0.86 (0.55–1.35)</td>
</tr>
<tr>
<td>CC</td>
<td>163</td>
<td>13 (7.98)</td>
<td>0.45 (0.23–0.89)</td>
</tr>
<tr>
<td>Trend test</td>
<td></td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>TC + CC</td>
<td>548</td>
<td>55 (10.04)</td>
<td>0.72 (0.47–1.10)</td>
</tr>
<tr>
<td><strong>HSD17B12 rs11037684 A&gt;G</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>757</td>
<td>79 (10.44)</td>
<td>1.00</td>
</tr>
<tr>
<td>AG</td>
<td>99</td>
<td>16 (16.16)</td>
<td>2.23 (1.28–3.91)</td>
</tr>
<tr>
<td>GG</td>
<td>2</td>
<td>0 (0.00)</td>
<td>-</td>
</tr>
<tr>
<td>Trend test</td>
<td></td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>AG + GG</td>
<td>101</td>
<td>16 (15.84)</td>
<td>2.21 (1.26–3.86)</td>
</tr>
<tr>
<td>Combined number of risk genotypes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>489</td>
<td>46 (9.41)</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>327</td>
<td>42 (12.84)</td>
<td>1.56 (1.00–2.41)</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>7 (16.67)</td>
<td>2.70 (1.18–6.18)</td>
</tr>
<tr>
<td>Trend test</td>
<td></td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>489</td>
<td>46 (9.41)</td>
<td>1.00</td>
</tr>
<tr>
<td>1–2</td>
<td>369</td>
<td>49 (13.28)</td>
<td>1.66 (1.09–2.53)</td>
</tr>
</tbody>
</table>

Abbreviations: SNP, single-nucleotide polymorphism; CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas MD Anderson Cancer Center; NHS, the Nurses’ Health Study; HPFS, the Health Professionals Follow-up Study; HR, hazards ratio; CI, confidence interval; ELOVL2, elongation of very long-chain fatty acids 2; HSD17B12, hydroxysteroid dehydrogenase type 12.

1Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in the MDACC dataset.
2Adjusted for age and sex in the NHS/HPFS dataset.
3Adjusted for age and sex in the combined dataset of both MDACC and NHS/HPFS.
4Risk genotypes include **ELOVL2** rs3734398 TT and **HSD17B12** rs11037684 AG + GG.
A deregulated fatty acid synthesis can affect drug resistance and cancer risk, prognosis and recurrence. For example, several studies have shown that overexpression of FASN is associated with a poor prognosis and drug resistance in breast cancer and gastrointestinal stromal tumors as well as associated with a higher risk of recurrence of human cancers, including cancers of the breast, prostate and bladder. Furthermore, blocking the fatty acid synthesis overcomes tumor regrowth and metastasis after withdrawal of the antiangiogenic therapy in breast and colon cancer cells. When restricted to hepatocellular carcinoma patients receiving surgery treatment, genetic variants of FASN could predict recurrence risk. Importantly, evidence also exists that fatty acid synthesis inhibitors may induce apoptosis and also reduce metastases and angiogenesis in melanoma cells. Consistently, CM patients with high expression levels of fatty-acid metabolic signature genes resulted in a significant decrease in survival rates of CM patients, supporting a role of the fatty acid metabolism in CM progression.

We report here some striking significant associations of CMSS with genetic variants in ELOVL2 and HSD17B12. CM patients with an increasing number of risk variant genotypes had a worse survival. Importantly, the risk effect was consistent across different analyses and the majority of subgroup comparisons, suggesting a strong association of a genetic effect on CM survival. We believe that these results are likely biologically plausible, since the genotype–phenotype correlation demonstrates that ELOVL2 expression levels may be modulated by rs3734398T>C change, although additional investigation is needed to unravel molecular mechanisms underlying the observed correlation.

ELOVL2 is located on chromosome 6p24.2, encoding for a transmembrane protein that controls the elongation of polyunsaturated fatty acids (PUFA), which modulates energy production, and influences inflammation and cell membrane integrity. For patients with breast cancer, ELOVL2 can hormonally regulate the PUFA synthesis and thus may have a potential implication on the endocrine therapy. Deletion of
ELOVL2 in a mouse model leads to a decrease in Foxp3+ regulatory T cells, suggesting its potential role in the adaptive immunity.\textsuperscript{26} GWAS have identified ELOVL2 variants to be associated with serum metabolic profile,\textsuperscript{29} aging process and DNA methylation.\textsuperscript{30} Recently, ELOVL2 rs3734398 has been reported to be significantly associated with plasma eicosapentaenoic and docosahexaenoic acid concentrations after fish oil supplement, which provides evidence on personalized dietary recommendations for reducing cardiovascular disease risk based on the genotype of this SNP.\textsuperscript{31} To date, ELOVL2 has not been reported to be associated with CM progression and prognosis. In light of our results and in the consideration that PUFAs are involved in crucial biological functions and that rs3734398 may regulate ELOVL2 expression, it is possible that genetic variants in ELOVL2 may be utilized in managing CM progression and prognosis in the future precision medicine, once validated by additional studies.

HSD17B12, located on chromosome 11p11.2, is a multifunctional isozyme, catalyzing the elongation of long chain fatty acids, particularly the conversion of palmitic to arachidonic acid.\textsuperscript{32} The latter is the precursor of prostaglandin E2, an important mediator of inflammation, linking HSD17B12 expression levels to inflammation and cancer.\textsuperscript{33} HSD17B12 expression levels were also shown to be associated with adipocyte differentiation\textsuperscript{34} as well as embryogenesis and differentiation.\textsuperscript{35} HSD17B12 also is believed to act as an oncogene involved in multiple cancers. For example, immunohistochemical analyses indicated that cytoplasmic staining of HSD17B12 was enhanced along with the severity of ovarian cancer, whereas HSD17B12 weak expression was correlated to a better overall survival and a longer time to first tumor recurrence.\textsuperscript{36} For breast cancer cases, HSD17B12 expression was significantly higher in tumor tissues than in normal tissues,\textsuperscript{37} leading to an increased risk of recurrence and adverse clinical outcome.\textsuperscript{38} Furthermore, HSD17B12 variants were found to be significantly associated with biochemical recurrence in patients with localized prostate cancer in one study\textsuperscript{39} and with less aggressive form of neuroblastoma in another study.\textsuperscript{40}

The present study has some strengths and limitations. A major strength of the present study is the comprehensive analysis of associations between SNPs in all genes involved in the fatty acid synthesis pathway and survival of CM as well as the use of two published GWAS datasets with a relative long median follow-up time and strict quality control procedures. The effects of risk genotypes of the two novel SNPs on CMSS were consistent in two different GWAS datasets. However, a potential weakness was the lack of information about different treatment, which should have been adjusted for the possible effect on CM patients’ outcomes. The samples of the two GWAS studies were not large enough to allow for the false discovery rate test, a more desired multiple test correction method, although the FPRP was more appropriate for highly correlated SNPs under investigations as a result of imputation in the present study. Finally, further functional investigation should be conducted to provide mechanistic insights into the mechanisms underlying the CM-death association with these two novel SNPs.

Acknowledgements

The authors would like to thank Ning Gu, Danwen Qian and Xiaomeng Wang for their technical support and insightful discussions and all participants and staffs of the Nurses’ Health Study and Health Professionals Follow-Up Study for their time and generosity, as well as all of the following state cancer registries for their support: 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. We also thank the John Hopkins University Center for Inherited Disease Research for conducting high-throughput genotyping for our study. The results published herein are fully or partly based upon data from 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.” This work was supported in whole or part by NIH/NCI R01CA100264, 2P50CA93459, R01CA133996, R01CA49449, P01CA87969, UM1CA186107, UM1CA167552, 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. Qingyi Wei was partly supported by start-up funds from Duke Cancer Institute, Duke University Medical Center and also partly by Duke Cancer Institute as part of the P30 Cancer Center Support Grant (Grant ID: NIH CA014236). Wei Dai was sponsored by the State Scholarship Fund, the China Scholarship Council for studying at Duke University.

References


