

Association of interferon regulatory factor-4 (*IRF4*) polymorphism rs12203592 with divergent melanoma pathways

David C. Gibbs, Irene Orlow, Jennifer I. Bramson, Peter A. Kanetsky, Li Luo, Anne Kricker, Bruce K. Armstrong, Hoda Anton-Culver, Stephen B. Gruber, Loraine D. Marrett, Richard P. Gallagher, Roberto Zanetti, Stefano Rosso, Terence Dwyer, Ajay Sharma, Emily La Pilla, Lynn From, Klaus J. Busam, Anne E. Cust, David W. Ollila, Colin B. Begg, Marianne Berwick, Nancy E. Thomas, on behalf of the GEM Study Group.

Affiliations of authors:

Department of Dermatology, University of North Carolina, Chapel Hill, NC (DCG, NET); Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC (NET, DWO); Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, NY (IO, AS, ELP, KJB, CBB); Department of Surgery, University of North Carolina, Chapel Hill, NC (JIB, DWO); Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL (PAK); Department of Internal Medicine, University of New Mexico Cancer Center, University of New Mexico, Albuquerque, NM (LL, MB); Sydney School of Public Health, University of Sydney, Sydney, New South Wales, Australia (AEC, AK, BKA); Department of Epidemiology, University of California, Irvine, CA (HAC); USC Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA (SBG); Department of Population Studies and Surveillance, Cancer Care Ontario, Toronto, Ontario, Canada (LDM); Cancer Control Research, British Columbia Cancer Agency, Vancouver, British Columbia, Canada (RPG); Piedmont Cancer Registry, Centre for Epidemiology and Prevention in Oncology in Piedmont, Turin, Italy (RZ, SR); The George Institute for Global Health, Oxford Martin School of Public Health, University of Oxford, Oxford, United Kingdom (TD); Department of Pathology, Women's College Hospital, Toronto, Ontario, Canada (LF).

Correspondence to: Nancy E. Thomas, MD PhD, Department of Dermatology, University of North Carolina, 405 Mary Ellen Jones Building, CB#7287, Chapel Hill, NC 27599. Phone: (919) 966-0785; Fax: (919) 966-6460; Email: nthomas@med.unc.edu.

Running Title: *IRF4* rs12203592 and divergent melanoma pathways

Keywords: Melanoma; risk factors; population-based; genotype; single nucleotide polymorphism; sun exposure; age distribution; dermatology; melanin; melanocytes; interferon regulatory factors; nevus; hair color; eye color

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Abbreviations: CI, confidence interval; GEM, Genes, Environment, and Melanoma Study; IQR, interquartile range; NR, neval remnants; OR, odds ratio; SE, solar elastosis; SNP, single nucleotide polymorphism.

Background Solar elastosis and neval remnants are histologic markers characteristic of divergent melanoma pathways linked to differences in age at onset, host phenotype, and sun exposure. However, the association between these pathway markers and newly identified low-penetrance melanoma susceptibility loci remains unknown.

Methods In the Genes, Environment and Melanoma (GEM) Study, 2,103 Caucasian participants had first primary melanomas that underwent centralized pathology review. For 47 single nucleotide polymorphisms (SNPs) previously identified as low-penetrant melanoma risk variants, we used multinomial logistic regression to compare melanoma with solar elastosis and melanoma with neval remnants simultaneously to melanoma with neither of these markers, excluding melanomas with both markers. All statistical tests were two-sided.

Results *IRF4* rs12203592 was the only SNP to pass the false discovery threshold in baseline models adjusted for age, sex, and study center. rs12203592*T was associated positively with melanoma with solar elastosis (OR = 1.47; 95% CI = 1.18 to 1.82) and inversely with melanoma with neval remnants (OR = 0.65; 95% CI = 0.48 to 0.87) compared to melanoma with neither marker ($P_{\text{global}} = 3.78 \times 10^{-08}$). Adjusting for phenotypic characteristics and total sun exposure hours did not materially affect rs12203592's associations. Distinct early- and late-onset age distributions were observed in patients with *IRF4* rs12203592 [CC] and [TT] genotypes, respectively.

Conclusions Our findings suggest a role of *IRF4* rs12203592 in pathway-specific risk for melanoma development. We hypothesize that *IRF4* rs12203592 could underlie in part the bimodal age distribution reported for melanoma and linked to the divergent pathways.

Introduction

Melanoma incidence has increased in fair-skinned populations worldwide in recent decades [1, 2]. Understanding melanoma etiology is critical for prevention; however, risk factors may differ by the pathway along which melanoma develops. The divergent pathway hypothesis postulates at least two separate pathways, one related to chronic sun exposure and the other related to the host's tendency to develop nevi, with melanomas marked, respectively, by the presence of histologic solar elastosis (SE) and neval remnants (NR) in adjacent tissue [3-5]. Predictors of SE and NR adjacent to melanomas include host phenotypic characteristics; however, the association of these pathway markers with recently identified germline variants in melanoma-associated risk loci has yet to be explored [5-7].

The divergent pathway model postulates that melanoma risk associated with UV exposure differs by host characteristics and results in at least two causal pathways linked to differences in age at onset, body site, and histopathologic features [4, 5, 8]. In individuals with low propensity to develop nevi, melanomas seem to require high levels of cumulative sun exposure [4, 9]. These chronic sun exposure-related melanomas, marked by SE, tend to occur at older ages, in individuals with fewer nevi, and on sun exposed body sites like the head/neck [3, 5, 9, 10]. Conversely, melanomas in individuals with a high propensity to develop nevi seem to require lower levels of sun exposure and arise from an existing nevus [4]. These nevogenic melanomas, marked by NR, tend to occur at younger ages, in individuals with higher nevus counts, and on less sun exposed sites like the trunk [11-15].

Recent genome wide association studies (GWAS) and candidate pathway studies have identified several low-penetrant genetic variants associated with cutaneous melanoma [16]. The majority are in loci associated with pigmentation (e.g., *TYRP1*, *TYR*, *HERC2/OCA2*, *SLC45A2*, and *ASIP*), nevi (e.g., *PLA2G6*, *MTAP*, and *NID1*), or both (e.g., *IRF4*) [17-24]; however, variants in other identified loci (e.g., *ATM*, *MX2*, *PARP1*, *ARNT*, and *CASP8*) may not be associated with phenotypic risk traits for melanoma [21, 25].

Within the large international population-based Genes, Environment, and Melanoma (GEM) Study, we previously investigated 47 single nucleotide polymorphisms (SNPs) in 21 of these putative low-penetrance melanoma susceptibility loci and found variants in the *TERT*, *TYRP1*, *MTAP*, *TYR*, *NCOA6*, and *MX2* gene regions and a *PARP1* haplotype to be associated with multiple primary melanoma compared to single primary melanoma [26]. The purpose of this study was to determine whether these 47 SNPs influence susceptibility to melanomas arising from the chronic sun exposure or nevogenic pathways by examining their associations with the exclusive presence of SE or NR adjacent to primary melanoma.

Materials and Methods

Study Population

The population-based case-only GEM Study enrolled 3,579 patients diagnosed between 1998 and 2003 with incident primary melanoma in Australia, Canada, Italy, and the United States [27, 28]. Of the 2,953 (83% of 3,567) Caucasian patients with centralized pathology review, 2,103 (86% of 2,458) were diagnosed with a first primary melanoma and are included in the analyses reported here. The Institutional Review Boards of participating institutions approved the protocol and informed consent was obtained from participants.

Hair and eye color, ability to tan, number of back nevi, residential histories, and sun exposure information were collected from phone interviews and self-administered questionnaires as described previously [9, 29, 30]. Number of back nevi is significantly correlated with whole-body nevus density diagrams [31] in GEM, and the suitability of using this variable as a proxy for total body nevus counts has been previously reported [32-34]. Total hours of sun exposure from the age of 5 until the age at diagnosis were calculated as the sum of reported outdoor sun exposure hours between 9 a.m. and 5 p.m. on working and non-working days in warmer and cooler months (based on residential calendars) estimated for each decade of life [29]. Average

annual sun exposure was calculated as the total exposure hours divided by the participant's age at diagnosis minus 5 years [29].

Three dermatopathologists scored adjacent SE as previously described [9] and adjacent NR as present when benign nevus cells were in the epidermis or dermis immediately adjacent to or below the melanoma cells. Kappa statistics, using the three categories, were 0.60 and 0.64 for scoring NR and SE, respectively, in a test set of 19 H&Es reviewed by the three dermatopathologists.

SNP Selection, Genotyping, and Principal Component Analysis

We selected 47 SNPs from 21 loci based on evidence they were low-penetrant risk variants for melanoma in other studies [26]. DNA was collected from buccal swabs. SNPs were genotyped using the MassArray iPLEX chemistry and platform (Agena Bioscience) with quality control measures as described previously [35]. Proxy SNPs ($r^2 > 0.95$; 1000 Genomes, CEU population; Proxy SNP; Broad Institute) rs6735656 and rs12278954 were substituted, respectively, for *CASP8* rs10931936 and *ATM* rs1801516 [21] - two SNPs of interest rejected during the assay design steps.

We performed principal component analysis (PCA) of the 47 SNPs with and without *MC1R* (available to this dataset [36, 37]) to detect potential population structure within our data. Using an eigen-value decomposition approach, we transformed and projected our data into lower dimensional spaces by extracting the top principal components (PC) that explain most of the total variance.

Statistical Analyses

A joint outcome variable was created with three levels: melanomas with SE+/NR- (adjacent SE but not NR), SE-/NR+ (adjacent NR but not SE), or SE-/NR- (neither SE nor NR). Melanomas with SE+/NR+ (adjacent SE and NR) were excluded because our goal was to determine the

associations of genetic variants separately for melanomas with one marker or the other. Multinomial logistic regression was used to estimate the per allele odds ratios (ORs) and 95% confidence intervals (95% CIs) of each SNP for SE+/NR- and SE-/NR+ melanomas compared simultaneously to SE-/NR- melanomas. Initial models used an additive model of inheritance of the minor allele of each SNP and were adjusted for baseline features: age (tertiles), sex, and study center. The false discovery threshold adjusted for multiple comparisons was computed using a resampling method that takes into account the linkage disequilibrium information among SNPs evaluated [38, 39].

For the one SNP that passed false discovery (*IRF4* rs12203592), different models of inheritance were compared using the Akaike information criterion statistic corrected for the number of model parameters (AICc). We selected the model with the lowest value indicating superior fit of the data [40]. Next, the associations of *IRF4* rs12203592 with phenotypic characteristics were estimated using multinomial logistic regression models adjusted for baseline features and limited to patients with no missing phenotypic data. We then estimated the associations of rs12203592, each phenotypic characteristic, and total sun exposure hours with SE+/NR- and SE-/NR+ melanomas compared simultaneously to SE-/NR- melanomas in separate baseline adjusted models limited to patients with no missing data for rs12203592, phenotype, or sun exposure hours. The associations of each variable with SE+/NR- and SE-/NR+ melanomas were then estimated in a multivariable model.

Age density distributions were plotted by the presence/absence of SE/NR and by the *IRF4* rs12205932 genotype using R, version 3.2.0 (R package 'sm': nonparametric smoothing methods (version 2.2-5.4)). All other statistical analyses were performed in SAS (SAS Institute) version 9.4. The Wilcoxon rank-sum test was used to compare the median age at onset of melanomas with SE+/NR- and SE-/NR+ melanomas each relative to SE-/NR- melanomas. All significance tests were two-sided with a significance threshold of 0.05.

Results

Of the 2,103 Caucasian patients in GEM with centrally reviewed incident first primary melanomas, the median age was 55 years and 51.7% were male (Table 1). Data for SE were available for 2,048 (97.4% of 2,103) and NR for 2,090 (99.4% of 2,103) of the melanomas, with missing data mainly a result of insufficient surrounding dermal/stromal tissue to score SE and NR. Data for both SE and NR were available for 2,039 (97.0% of 2,103) patients, of whom 285 (13.6% of 2,103) had both SE and NR present adjacent to their melanoma and were excluded in the analyses, as our primary goal was to investigate the SE and NR pathways separately.

SNP locations, GEM minor allele frequencies (MAF) overall and by study center, 1000 genomes MAFs [41], and RegulomeDB [42] predicted functional impact are in Supplementary Table 1 (available online). The SNP associations were adjusted for center to account for possible MAF differences among the study centers. To further evaluate potential confounding by genetic ancestry, we performed PCA of the 47 SNPs with and without *MC1R* to detect population structure. Scatterplots for PC1 and PC2 with study center color-coded are shown in Supplementary Figure 1 (available online). We observed similar PCA loadings for participants in different study centers. The first three principal components were not significantly associated with either solar elastosis or neval remnants and thus did not show evidence of potential confounding by population structure.

The associations using an additive model of inheritance of the 47 SNPs genotyped with SE+/NR- and SE-/NR+ relative to SE-/NR- melanoma are presented in Figure 1 and Supplementary Table 2 (available online). *IRF4* rs12203592 was the only SNP to pass the false discovery threshold ($P = .0015$). rs12203592 was positively associated with SE+/NR- (OR = 1.47; 95% CI = 1.18 to 1.82) and inversely associated with SE-/NR+ (OR = 0.65; 95% CI = 0.48 to 0.87) melanomas ($P_{\text{global}} = 3.78 \times 10^{-08}$). Comparing other models of inheritance for rs12203592, we found that the additive model had the lowest AICc value indicating the best fit of the data (data not shown). *IRF4* rs12203592*T was significantly associated ($P < .05$) with

having ≤ 10 back nevi, dark hair color, light eye color, and decreased ability to tan in multinomial models adjusted for baseline features (Supplementary Table 3, available online).

The associations of *IRF4* rs12203592, phenotypic characteristics, and total sun exposure hours with SE+/NR- and SE-/NR+ relative to SE-/NR- melanoma in baseline and multivariable models limited to participants with complete data are shown in Table 2. rs12203592*T, older age, increased total sun exposure hours, ≤ 10 back nevi, and light eye color were each significantly associated ($P < .05$) with SE+/NR- compared to SE-/NR- melanomas in baseline and multivariable models. Only rs12203592*C was significantly associated ($P < .05$) with SE-/NR+ melanomas compared to SE-/NR- melanomas in baseline and multivariable models. Adjusting for phenotypic characteristics and total sun exposure hours did not materially affect rs12203592's association with SE+/NR- or SE-/NR+ melanomas relative to SE-/NR- melanomas. Re-analyses were performed of the fully adjusted model, adding either self-reported ancestry, log Breslow, or AJCC tumor stage [43], or substituting average annual total sun exposure hours for total exposure hours, but none of these changes materially affected rs12203592's associations (results not shown).

Using all cases, including those with SE+/NR+ tumors, probability density functions revealed a bimodal distribution for age at diagnosis with an early-onset peak around 50 years of age and a late-onset peak around 75 years of age (Figure 2, A). Age density plots stratified by the presence/absence of SE/NR revealed SE+/NR- melanomas to have a predominantly late-onset distribution, SE-/NR- or SE-/NR+ melanomas to have an early-onset distribution, and SE+/NR+ melanomas to have a distribution falling between the early-and late-onset density peaks (Figure 2, B). The median age at onset for SE+/NR- melanomas was significantly higher than SE-/NR- melanomas, while the median age at onset of SE-/NR+ and SE-/NR- melanomas did not significantly differ (Table 2).

Patients with the *IRF4* rs12203592 [CC] genotype had a predominantly early-onset distribution peaking around 45 years of age while patients with the [TT] genotype had a

predominantly late-onset distribution peaking around 75 years of age. Patients with the rs12203592 [CT] genotype had a bimodal age distribution with approximately equal density peaks around 50 and 70 years of age (Figure 2, C). The age distributions differed by the rs12203592 genotype, even when stratified by Australia and North America (data not shown).

Discussion

IRF4 encodes interferon regulatory factor-4, a transcription factor involved in B- and T-cell development, pigmentation, and potentially melanocyte differentiation and proliferation [44-47]. *IRF4* (also known as *MUM1*) is predominantly expressed in immune cells, but is also present in melanocytes, nevi, and primary melanomas [44-47]. The variant T allele of rs12203592, in intron 4 of *IRF4*, is a functional SNP shown to increase *IRF4* promoter activity in Burkitt Lymphoma B-cells and repress promoter activity in human epidermal melanocytes, and thus may increase or decrease *IRF4* expression in different contexts [45, 48, 49]. In mouse melanocytes, decreased expression of *IRF4* as a result of rs12203592*T also led to decreased levels of certain downstream effectors of IRF4, such as the melanin-synthesizing enzyme TYR [45]. Also, RegulomeDB [42] predicts that *IRF4* rs12203592 is likely to affect binding.

We found that *IRF4* rs12203592*T was strongly associated with light eye color, poor tanning ability, dark hair color, and lower nevus counts, consistent with previous epidemiological studies [18, 50-52]. Decreased *TYR* expression in melanocytic cells as a result of rs12203592*T might explain, at least in part, its association with certain fair pigmentary traits. Adjusting for pigmentary traits and nevi, however, did not materially change the OR for the association of rs12203592 with the exclusive presence of SE or NR adjacent to melanomas. This suggests that rs12203592 could influence the development of divergent melanoma pathways independent of host phenotypic traits associated with melanoma risk. Experimental studies are needed to elucidate this influence, which may be linked to differences in expression of *IRF4* and

its downstream targets, regulated in part by the rs12203592 genotype in melanocytes and/or immune system cells.

IRF4 rs12203592's association with histologic SE and NR adjacent to melanomas, representing divergent causal pathways, may also explain why this variant has been inconsistently associated with melanoma risk in recent epidemiological studies. In a combined analysis of Australian, UK, and Swedish subjects, rs12203592*C was positively associated with melanoma, most significantly with melanoma on the trunk [50]. Similarly, in a UK study rs12203592*T was inversely associated with melanoma, most significantly with melanoma on the trunk [53]. In contrast, rs12203592*T was positively associated with melanoma risk in two US studies [52, 54]. In another Australian study, rs12203592*T was inversely associated with melanoma in children/adolescents, but was not associated with melanoma in adults [55]. Further, rs12203592 was not significantly associated with melanoma in a combined analysis of GenoMEL data with patients from Europe and Israel [21], nor was it associated with multiple compared to single primary melanoma in a recent GEM study [26]. The differing associations in these studies may depend upon the proportion of melanomas that arose via the chronic sun exposure versus nevogenic pathway, given rs12203592's strong and opposing association with the exclusive presence of SE and NR found in the present study.

We are not aware of another study that has reported associations of *IRF4* rs12203592 with SE or NR adjacent to melanomas. Previous GEM investigations and independent studies have found older age, fewer nevi, and higher self-reported sun exposure hours to be associated with SE, with ORs in directions consistent with those reported here [3, 5, 9].

We also found a bimodal age distribution in GEM with distinct early- and late-onset peaks consistent with those reported previously for melanoma and which have been linked to the divergent causal pathways [56, 57]. This bimodal distribution has been described as displaying an early-onset peak frequency for trunk melanomas and a late-onset peak frequency for face/ear melanomas using Surveillance, Epidemiology and End Results (SEER) data [56]. In GEM,

melanomas without histologic SE displayed an early-onset distribution similar to that found in SEER for trunk melanomas while melanomas without histologic NR but with SE had a late-onset distribution similar to that found in SEER for face/ear melanomas [56]. A similar early-onset distribution was also observed in patients with the *IRF4* rs12203592 [CC] genotype, while a similar late-onset distribution was observed in patients with the rs12203592 [TT] genotype. This suggests that the rs12203592 genotype could be an inherited characteristic that, in part, underlies the bimodal age distribution of melanomas related to divergent causal pathways linked to differences in age at onset, body site, and histopathologic features. Further work could be done to investigate this possibility.

Major advantages of our study are its large sample size, population-based ascertainment, and centralized pathology review. The majority of GEM participants (70%) reported Northern European ancestry and the remaining persons divided by center into Southern European and other groups. While population stratification could have been an issue, the PCA results did not show evidence of a population substructure confounding effect. We adjusted for study center in our analysis and further adjustment for self-reported ancestry had little effect on the associations in the fully adjusted model. Melanomas with NR have also been reported as having lower Breslow thickness [13, 15], possibly because thicker melanomas overgrow and obscure any NR; however, adding log Breslow or AJCC tumor stage to the fully adjusted multivariable model did not change the ORs for rs12203592. The effect of age on cumulative sun exposure was of potential concern but substituting average annual total sun exposure hours for total exposure hours did not materially affect rs12203592's associations. Another limitation could be bias in the selection of cases for pathology review; however, none of the predictor variables used in the fully adjusted multivariable model were associated with missing versus non-missing pathology review. Also we recognize that our SE and NR measures must contain some misclassification as evidenced by the modest agreement between our pathologists, but misclassification typically has the effect of attenuating observed associations. Power may also have been insufficient to detect

associations of SNPs with lower MAFs (e.g. SNPs in *SLC45A2*) in GEM. Lastly, the age distributions presented could potentially be affected by age structure and secular trends for sun exposure among the different populations; however, the age distributions among the GEM populations are quite similar, and the age distributions shifted by the rs12203592 genotype, even when stratified by Australia and North America.

IRF4 rs12203592's independent yet contrasting associations with SE+/NR- and SE-/NR+ melanoma suggest that this inherited genetic variant may play a pivotal role underlying susceptibility to divergent pathways in melanoma development. We are not aware of a similar strong germline genetic crossover effect for another cancer, with an inverse association with one subtype and positive association with another subtype of cancer. Knowledge of the genetic etiology of divergent melanoma pathways with resultant differences in host characteristics and tumor features among melanoma patients could inform future risk models and prevention efforts for this complex and heterogeneous disease.

Funding

This work was supported by National Cancer Institute (NCI) grants R01CA112243, R01CA112524, R01CA112243-05S1, R01CA112524-05S2, CA098438, U01CA83180, R33CA10704339, P30CA016086, P30CA008748, and P30CA014089; National Institute of Environmental Health Sciences (P30ES010126); University of Sydney Medical Foundation Program grant (Bruce Armstrong); Michael Smith Foundation for Health Research Infrastructure Award (Richard Gallagher).

Acknowledgements

The study was conducted by The GEM Study Group: Coordinating Center, Memorial Sloan-Kettering Cancer Center, New York, NY: Marianne Berwick, M.P.H., Ph.D. (Principal Investigator (PI), currently at the University of New Mexico), Colin B. Begg, Ph.D. (co-PI),

Irene Orlow, Ph.D. (co-Investigator), Klaus J. Busam, M.D. (Dermatopathologist), Anne S. Reiner, M.P.H. (Biostatistician), Pampa Roy, Ph.D. (Laboratory Technician), Ajay Sharma, M.S. (Laboratory Technician), Emily La Pilla (Laboratory Technician). University of New Mexico, Albuquerque: Marianne Berwick, M.P.H., Ph.D. (PI), Li Luo, Ph.D. (Biostatistician), Kirsten White, MSc (Laboratory Manager), Susan Paine, M.P.H. (Data Manager). Study centers included the following: The University of Sydney and The Cancer Council New South Wales, Sydney, Australia: Bruce K. Armstrong M.B.B.S.; D.Phil., (PI), Anne Krickler, Ph.D. (co-PI), Anne E. Cust, Ph.D. (co-Investigator); Menzies Research Institute Tasmania, University of Tasmania, Hobart, Australia: Alison Venn, Ph.D. (current PI), Terence Dwyer, M.D. (PI, currently at University of Oxford, United Kingdom), Paul Tucker, M.D. (Dermatopathologist); British Columbia Cancer Research Centre, Vancouver, Canada: Richard P. Gallagher, M.A. (PI), Donna Kan (Coordinator); Cancer Care Ontario, Toronto, Canada: Loraine D. Marrett, Ph.D. (PI), Elizabeth Theis, M.Sc. (co-Investigator), Lynn From, M.D. (Dermatopathologist); CPO, Center for Cancer Prevention, Torino, Italy: Roberto Zanetti, M.D (PI), Stefano Rosso, [M.D.](#), [M.Sc.](#) (co-PI); University of California, Irvine, CA: Hoda Anton-Culver, Ph.D. (PI), Argyrios Ziogas, Ph.D. (Statistician); University of Michigan, Ann Arbor, MI: University of Michigan, Ann Arbor: Stephen B. Gruber, M.D., M.P.H., Ph.D. (PI, currently at University of Southern California, Los Angeles, CA), Timothy Johnson, M.D. (Director of Melanoma Program), Duveen Sturgeon, M.S.N. (co-Investigator, joint at USC-University of Michigan); University of North Carolina, Chapel Hill, NC: Nancy E. Thomas, M.D., Ph.D. (PI), Robert C. Millikan, Ph.D. (previous PI, deceased), David W. Ollila, M.D. (co-Investigator), Kathleen Conway, Ph.D. (co-Investigator), Pamela A. Groben, M.D. (Dermatopathologist), Sharon N. Edmiston, B.A. (Research Analyst), Honglin Hao (Laboratory Specialist), Eloise Parrish, MSPH (Laboratory Specialist), David C. Gibbs, B.S. (Research Assistant), Jill S. Frank, M.S. (Research Assistant), Jennifer I. Bramson (Research Assistant); University of Pennsylvania, Philadelphia, PA: Timothy R. Rebbeck, Ph.D. (PI), Peter A. Kanetsky, M.P.H., Ph.D. (co-Investigator, currently at

H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida); UV data consultants: Julia Lee Taylor, Ph.D. and Sasha Madronich, Ph.D., National Centre for Atmospheric Research, Boulder, CO.

References

1. Diepgen T, Mahler V. The epidemiology of skin cancer. *British Journal of Dermatology* 2002;146(s61):1-6.
2. Garbe C, Leiter U. Melanoma epidemiology and trends. *Clinics in dermatology* 2009;27(1):3-9.
3. Kvaskoff M, Pandeya N, Green AC, *et al.* Solar elastosis and cutaneous melanoma: A site-specific analysis. *Int J Cancer* 2014; 10.1002/ijc.29335.
4. Whiteman DC, Watt P, Purdie DM, *et al.* Melanocytic nevi, solar keratoses, and divergent pathways to cutaneous melanoma. *J Natl Cancer Inst* 2003;95(11):806-12.
5. Lee EY, Williamson R, Watt P, *et al.* Sun exposure and host phenotype as predictors of cutaneous melanoma associated with neval remnants or dermal elastosis. *International journal of cancer* 2006;119(3):636-642.
6. Law MH, MacGregor S, Hayward NK. Melanoma genetics: recent findings take us beyond well-traveled pathways. *Journal of Investigative Dermatology* 2012;132(7):1763-1774.
7. Ward KA, Lazovich D, Hordinsky MK. Germline melanoma susceptibility and prognostic genes: a review of the literature. *Journal of the American Academy of Dermatology* 2012;67(5):1055-1067.
8. Mishima Y. Melanocytic and nevocytic malignant melanomas. Cellular and subcellular differentiation. *Cancer* 1967;20(5):632-649.
9. Thomas NE, Kricker A, From L, *et al.* Associations of cumulative sun exposure and phenotypic characteristics with histologic solar elastosis. *Cancer Epidemiol Biomarkers Prev* 2010;19(11):2932-41.

10. English DR, Heenan PJ, Holman CDAJ, *et al.* Melanoma in Western Australia in 1980-81: incidence and characteristics of histological types. *Pathology* 1987;19(4):383-392.
11. Lee EY, Williamson R, Watt P, *et al.* Sun exposure and host phenotype as predictors of cutaneous melanoma associated with neval remnants or dermal elastosis. *Int J Cancer* 2006;119(3):636-42.
12. Cho E, Rosner BA, Colditz GA. Risk factors for melanoma by body site. *Cancer Epidemiol Biomarkers Prev* 2005;14(5):1241-4.
13. Purdue MP, From L, Armstrong BK, *et al.* Etiologic and other factors predicting nevus-associated cutaneous malignant melanoma. *Cancer Epidemiol Biomarkers Prev* 2005;14(8):2015-22.
14. Olsen CM, Zens MS, Stukel TA, *et al.* Nevus density and melanoma risk in women: a pooled analysis to test the divergent pathway hypothesis. *Int J Cancer* 2009;124(4):937-44.
15. Bevona C, Goggins W, Quinn T, *et al.* Cutaneous melanomas associated with nevi. *Archives of dermatology* 2003;139(12):1620-1624.
16. Law MH, Montgomery GW, Brown KM, *et al.* Meta-analysis combining new and existing data sets confirms that the TERT-CLPTM1L locus influences melanoma risk. *J Invest Dermatol* 2012;132(2):485-7.
17. Gudbjartsson DF, Sulem P, Stacey SN, *et al.* ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 2008;40(7):886-91.
18. Han J, Kraft P, Nan H, *et al.* A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008;4(5):e1000074.
19. Bishop DT, Demenais F, Iles MM, *et al.* Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 2009;41(8):920-5.
20. Nan H, Xu M, Zhang J, *et al.* Genome-wide association study identifies nidogen 1 (NID1) as a susceptibility locus to cutaneous nevi and melanoma risk. *Hum Mol Genet* 2011;20(13):2673-9.

21. Barrett JH, Iles MM, Harland M, *et al.* Genome-wide association study identifies three new melanoma susceptibility loci. *Nat Genet* 2011;43(11):1108-13.
22. Amos CI, Wang LE, Lee JE, *et al.* Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet* 2011;20(24):5012-23.
23. Jannot AS, Meziani R, Bertrand G, *et al.* Allele variations in the OCA2 gene (pink-eyed-dilution locus) are associated with genetic susceptibility to melanoma. *Eur J Hum Genet* 2005;13(8):913-20.
24. Fernandez LP, Milne RL, Pita G, *et al.* SLC45A2: a novel malignant melanoma-associated gene. *Hum Mutat* 2008;29(9):1161-7.
25. Macgregor S, Montgomery GW, Liu JZ, *et al.* Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat Genet* 2011;43(11):1114-8.
26. Gibbs DC, Orlow I, Kanetsky PA, *et al.* Inherited genetic variants associated with occurrence of multiple primary melanoma. *Cancer Epidemiology Biomarkers & Prevention* 2015;24(6):992-7.
27. Millikan RC, Hummer A, Begg C, *et al.* Polymorphisms in nucleotide excision repair genes and risk of multiple primary melanoma: the Genes Environment and Melanoma Study. *Carcinogenesis* 2006;27(3):610-8.
28. Begg CB, Hummer AJ, Mujumdar U, *et al.* A design for cancer case-control studies using only incident cases: experience with the GEM study of melanoma. *Int J Epidemiol* 2006;35(3):756-64.
29. Kricker A, Armstrong BK, Goumas C, *et al.* Ambient UV, personal sun exposure and risk of multiple primary melanomas. *Cancer Causes Control* 2007;18(3):295-304.
30. Thomas NE, Edmiston SN, Alexander A, *et al.* Number of nevi and early-life ambient UV exposure are associated with BRAF-mutant melanoma. *Cancer Epidemiol Biomarkers Prev* 2007;16(5):991-7.

31. Marrett L, King W, Walter S, *et al.* Use of host factors to identify people at high risk for cutaneous malignant melanoma. CMAJ: Canadian Medical Association Journal 1992;147(4):445.
32. English J, Swerdlow A, MacKie R, *et al.* Site - specific melanocytic naevus counts as predictors of whole body naevi. British Journal of Dermatology 1988;118(5):641-644.
33. English DR, Armstrong BK. Melanocytic nevi in children I. Anatomic sites and demographic and host factors. American journal of epidemiology 1994;139(4):390-401.
34. Autier P, Boniol M, Severi G, *et al.* The body site distribution of melanocytic naevi in 6–7 year old European children. Melanoma research 2001;11(2):123-131.
35. Orlow I, Roy P, Reiner AS, *et al.* Vitamin D receptor polymorphisms in patients with cutaneous melanoma. Int J Cancer 2012;130(2):405-18.
36. Kanetsky PA, Rebbeck TR, Hummer AJ, *et al.* Population-based study of natural variation in the melanocortin-1 receptor gene and melanoma. Cancer Res 2006;66(18):9330-7.
37. Taylor NJ, Reiner AS, Begg CB, *et al.* Inherited variation at MC1R and ASIP and association with melanoma-specific survival. Int J Cancer 2015;136(11):2659-67.
38. Lin DY. An efficient Monte Carlo approach to assessing statistical significance in genomic studies. Bioinformatics 2005;21(6):781-7.
39. He Q, Avery CL, Lin DY. A general framework for association tests with multivariate traits in large-scale genomics studies. Genet Epidemiol 2013;37(8):759-67.
40. Burnham KP, Anderson DR. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media; 2002.
41. Genomes Project C, Abecasis GR, Auton A, *et al.* An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491(7422):56-65.
42. Boyle AP, Hong EL, Hariharan M, *et al.* Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22(9):1790-7.

43. Balch CM, Gershenwald JE, Soong SJ, *et al.* Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009;27(36):6199-206.
44. Shaffer AL, Emre NC, Romesser PB, *et al.* IRF4: Immunity. Malignancy! Therapy? *Clin Cancer Res* 2009;15(9):2954-61.
45. Praetorius C, Grill C, Stacey SN, *et al.* A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway. *Cell* 2013;155(5):1022-1033.
46. Natkunam Y, Warnke RA, Montgomery K, *et al.* Analysis of MUM1/IRF4 protein expression using tissue microarrays and immunohistochemistry. *Modern Pathology* 2001;14(7):686-694.
47. Sundram U, Harvell JD, Rouse RV, *et al.* Expression of the B-cell proliferation marker MUM1 by melanocytic lesions and comparison with S100, gp100 (HMB45), and MelanA. *Mod Pathol* 2003;16(8):802-10.
48. Do TN, Ucisik-Akkaya E, Davis CF, *et al.* An intronic polymorphism of IRF4 gene influences gene transcription in vitro and shows a risk association with childhood acute lymphoblastic leukemia in males. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 2010;1802(2):292-300.
49. Visser M, Palstra R-J, Kayser M. Allele-specific transcriptional regulation of IRF4 in melanocytes is mediated by chromatin-looping of the intronic rs12203592-enhancer to the IRF4-promoter. *Human molecular genetics* 2015:ddv029.
50. Duffy DL, Iles MM, Glass D, *et al.* IRF4 variants have age-specific effects on nevus count and predispose to melanoma. *The American Journal of Human Genetics* 2010;87(1):6-16.
51. Duffy DL, Zhao ZZ, Sturm RA, *et al.* Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma. *Journal of Investigative Dermatology* 2010;130(2):520-528.

52. Zhang M, Song F, Liang L, *et al.* Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Human molecular genetics* 2013;ddt142.
53. Newton-Bishop JA, Chang Y-M, Iles MM, *et al.* Melanocytic nevi, nevus genes, and melanoma risk in a large case-control study in the United Kingdom. *Cancer Epidemiology Biomarkers & Prevention* 2010;19(8):2043-2054.
54. Han J, Qureshi AA, Nan H, *et al.* A germline variant in the interferon regulatory factor 4 gene as a novel skin cancer risk locus. *Cancer research* 2011;71(5):1533-1539.
55. Kvaskoff M, Whiteman DC, Zhao ZZ, *et al.* Polymorphisms in nevus-associated genes MTAP, PLA2G6, and IRF4 and the risk of invasive cutaneous melanoma. *Twin Research and Human Genetics* 2011;14(05):422-432.
56. Lachiewicz AM, Berwick M, Wiggins CL, *et al.* Epidemiologic support for melanoma heterogeneity using the surveillance, epidemiology, and end results program. *Journal of Investigative Dermatology* 2008;128(5):1340.
57. Anderson WF, Pfeiffer RM, Tucker MA, *et al.* Divergent cancer pathways for early - onset and late - onset cutaneous malignant melanoma. *Cancer* 2009;115(18):4176-4185.

Figure Legends

Figure 1. A) Association of SNPs with melanomas with adjacent solar elastosis but without neval remnants (SE+/NR-). **B)** Association of SNPs with melanomas with adjacent neval remnants but without solar elastosis (SE-/NR+). Per allele odds ratios (ORs), color-coded by chromosome, and 95% confidence intervals (95% CIs) were estimated using multinomial logistic regression models comparing melanomas with SE-/NR+ and SE+/NR- simultaneously to melanomas with neither marker (SE-/NR-). Excluded were melanomas with both markers (SE+/NR+). All models were adjusted for age (tertiles), sex, and study center (Supplementary

Table 2, available online). All statistical tests were two-sided. Square markers and * denotes *IRF4* rs12205392 - the only SNP to pass false discover for its association with SE+/NR- and SE-/NR+ melanomas compared to melanomas with neither marker ($P_{\text{global}} = 3.78 \times 10^{-08}$) (ORs and 95% CIs provided in the figure).

Figure 2. Age distributions of patients with first primary melanomas scored for both solar elastosis (SE) and neval remnants (NR) in the GEM Study ($n = 2,103$) shown for (A) all cases, (B) histologic markers of divergent melanoma pathways, and (C) *IRF4* rs12203592 genotype. The y-axis shows the smoothed density estimate of the proportion of patients who were diagnosed with melanoma at a given age plotted in years on the x-axis. GEM = Genes, Environment and Melanoma; SE+/NR- = melanomas with solar elastosis but without neval remnants; SE-/NR+ = melanomas with neval remnants but without solar elastosis; SE-/NR- = melanomas without solar elastosis or neval remnants; SE+/NR+ = melanoma with solar elastosis and neval remnants.

Table 1. Clinical characteristics among patients with first primary melanomas that underwent centralized pathology review in the GEM Study ($n = 2,103$)*

Characteristic	No. (%)
Age at diagnosis, years	
Median (IQR)	55 (25)
Sex	
Male	1087 (51.7)
Female	1016 (48.3)
Solar elastosis	
Absent	720 (34.2)
Present	1328 (63.2)
Missing	55 (2.6)
Neval remnants	
Absent	1523 (72.4)
Present	567 (27.0)
Missing	13 (0.6)
Solar elastosis / neval remnants	
SE-/NR-	444 (21.1)
SE+/NR-	1037 (49.3)
SE-/NR+	273 (13.0)
SE+/NR+	285 (13.6)
Missing	64 (3.0)
Number of back nevi	
0-10	1186 (56.4)
>10	883 (42.0)
Missing	34 (1.6)
Hair color	
Dark hair (dark brown, black)	643 (30.6)
Light hair (light brown, blonde)	1251 (59.5)
Red	186 (8.8)
Missing	23 (1.1)
Eye color	
Dark eyes (brown, black)	405 (19.3)
Light eyes (blue, grey, green, hazel)	1678 (79.8)
Missing	20 (1.0)
Ability to tan	
Deep / moderate tan	1213 (57.7)
Mild / no tan	840 (39.9)
Missing	50 (2.4)

* Non-Caucasians with single primary melanoma that underwent pathology review ($n = 7$) were excluded. GEM = Genes, Environment and Melanoma; IQR = Interquartile range; NR = neval remnants; SE = solar elastosis; SE+/NR- = melanomas with solar elastosis but without neval remnants; SE-/NR+ = melanomas with neval remnants but without solar elastosis; SE-/NR- = melanomas without solar elastosis or neval remnants; SE+/NR+ = melanomas with solar elastosis and neval remnants.

Table 2. Associations of *IRF4* rs12203592, patient characteristics and sun exposure with histologic solar elastosis (SE) and neval remnants (NR) adjacent to first primary melanomas in the GEM Study ($n = 1,590$)*

Variable	Histologic markers of divergent melanoma pathways			Baseline adjusted †					Fully adjusted ‡				
	SE-/NR- n = 400	SE+/NR- n = 943	SE-/NR+ n = 247	Compared to SE-/NR-					Compared to SE-/NR-				
				SE+/NR-		SE-/NR+		P _{global}	SE+/NR-		SE-/NR+		P _{global}
				OR (95% CI)	P	OR (95% CI)	P		OR (95% CI)	P	OR (95% CI)	P	
<i>IRF4</i> rs12203592													
Per T allele	400	943	247	1.43 (1.15 to 1.79)	.001	0.60 (0.44 to 0.82)	.001	<.001	1.30 (1.02 to 1.65)	.03	0.61 (0.44 to 0.85)	.003	<.001
Age at diagnosis, years													
Median (IQR)	48 (22)	60 (24)	47 (20)		<.001		.68		-		-		
<49	214	264	135	1.00 (referent)	<.001§	1.00 (referent)	.15§	<.001	1.00 (referent)	<.001§	1.00 (referent)	0.11§	<.001
50-70	141	385	89	2.80 (2.09 to 3.76)		0.85 (0.60 to 1.22)			2.09 (1.52 to 2.86)		0.82 (0.56 to 1.21)		
>70	45	294	23	6.73 (4.55 to 9.97)		0.67 (0.38 to 1.18)			4.10 (2.65 to 6.36)		0.62 (0.34 to 1.14)		
Sex													
Male	180	504	121	1.00 (referent)	.69	1.00 (referent)	.26	.53	1.00 (referent)	.67	1.00 (referent)	.43	.51
Female	220	439	126	0.95 (0.73 to 1.23)		0.83 (0.59 to 1.15)			1.06 (0.80 to 1.41)		0.87 (0.61 to 1.24)		
Total sun exposure, hours													
<28,418	213	297	124	1.00 (referent)	<.001§	1.00 (referent)	0.44§	<.001	1.00 (referent)	<.001§	1.00 (referent)	0.47§	<.001
28,418-50,050	125	309	80	1.43 (1.05 to 1.97)		1.01 (0.75 to 1.61)			1.47 (1.07 to 2.03)		1.08 (0.73 to 1.59)		
>50,050	61	337	43	2.36 (1.57 to 3.55)		1.23 (0.72 to 2.08)			2.49 (1.64 to 3.77)		1.22 (0.72 to 2.09)		
Number of back moles													
0-10	198	591	118	1.00 (referent)	.006	1.00 (referent)	.62	.004	1.00 (referent)	.03	1.00 (referent)	.94	.04
>10	202	352	129	0.69 (0.53 to 0.90)		1.09 (0.78 to 1.51)			0.73 (0.56 to 0.96)		1.01 (0.73 to 1.41)		
Hair color													
Dark hair	121	311	61	1.00 (referent)	.74	1.00 (referent)	.20	.19	1.00 (referent)	.58	1.00 (referent)	.63	.52
Light hair	245	551	168	0.91 (0.69 to 1.21)		1.38 (0.95 to 2.00)			0.84 (0.61 to 1.16)		1.18 (0.78 to 1.78)		
Red	34	81	18	1.05 (0.64 to 1.72)		1.06 (0.55 to 2.05)			0.85 (0.50 to 1.45)		0.95 (0.47 to 1.91)		
Eye color													
Dark eyes	94	182	51	1.00 (referent)	.04	1.00 (referent)	.53	.13	1.00 (referent)	.04	1.00 (referent)	.50	.12
Light eyes	306	761	196	1.38 (1.01 to 1.88)		1.14 (0.77 to 1.68)			1.43 (1.02 to 2.00)		1.16 (0.76 to 1.77)		
Ability to tan													
Deep / moderate tan	245	528	155	1.00 (referent)	.09	1.00 (referent)	.64	.09	1.00 (referent)	.17	1.00 (referent)	.87	.27
Mild / no tan	155	415	92	1.26 (0.97 to 1.64)		0.92 (0.66 to 1.29)			1.22 (0.92 to 1.62)		0.97 (0.68 to 1.38)		

* Excluded were participants with both solar elastosis and neval remnants adjacent to their melanoma ($n = 285$), non-Caucasians ($n = 7$), and participants with 1 or more missing data point for any of the variables presented in the table. Multinomial logistic regression was used to estimate the odds ratios and 95% confidence intervals comparing melanomas with solar elastosis but without neval remnants (SE+/NR-) and melanomas with neval remnants but without solar elastosis (SE-/NR+) simultaneously to melanomas without either marker (SE-/NR-). CI = confidence interval; GEM = Genes, Environment and Melanoma; OR = odds ratio; NR = neval remnants; SE = solar elastosis. All statistical tests were two-sided.

† Sex is adjusted for age (tertiles) and study center. Age is adjusted for sex and study center. All other variables are adjusted for age (tertiles), sex, and study center.

‡ Adjusted for age (tertiles), sex, study center, total sun exposure hours (tertiles) and all other variables in the table.

§ Where noted, linear trend was tested using the Wald statistic when the tertiles of the variable were treated as a single ordinal variable.

|| P values calculated using Wilcoxon rank-sum test comparing the median age of patients with SE+/NR- vs. SE-/NR- melanomas and SE-/NR+ vs. SE-/NR- melanomas.

