The Journal of Investigative Dermatology

Prediction of melanoma risk in a Southern European population based on a weighted genetic risk score

Journal:	*Journal of Investigative Dermatology*
Manuscript ID	JID-2015-0859
Manuscript Type:	Original Article
Date Submitted by the Author:	27-Sep-2015
Complete List of Authors:	Kypreou, Katerina; University of Athens, A. Sygros Hospital, Dermatology Stefanaki, Irene; University of Athens, A. Sygros Hospital, Dermatology Antonopoulou, Kyriaki; University of Athens, A. Sygros Hospital, Dermatology Karagianni, Fani; University of Athens, A. Sygros Hospital, Dermatology Ntritsos, George; University of Ioannina, Department of Hygiene and Epidemiology, School of Medicine Zaras, Alexios; A.Sygros Hospital , Dermatology Vasiliki, Nikolaou; A.Sygros Hospital , Dermatology Kalfa, Iro; Laikon Hospital, Blood Donation Unit Hasapi, Vasiliki; A.Sygros Hospital , Dermatology Polydorou, Dorothea; A.Sygros Hospital , Dermatology Gogas, Helen; University of Athens, Medical Oncology, Medical School Spyrou, George; Academy of Athens, Biomedical Informatics Unit, Biomedical Research Foundation Bertram, Lars; Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics; Imperial College London, Faculty of Medicine Lill, Christina; Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics; University Medical Center of the Johannes Gutenberg University Mainz, Department of Neurology, Focus Program Translational Neuroscience Ioannidis, John; Stanford University, Department of Medicine and Department of Health Research and Policy Antoniou, Christina; University of Athens, Dermatology and Venereology, Medical School Evangelou, Evangelos; Imperial College London, St Mary's Campus, Department of Epidemiology and Biostatistics; University of Ioannina, Department of Hygiene and Epidemiology, School of Medicine Stratigos, Alexander; University of Athens, A. Sygros Hospital, Dermatology
Key Words:	cutaneous melanoma, genetic association, GWAS, genetic risk score, multivariable prediction model

Prediction of melanoma risk in a Southern European population based on a weighted genetic risk score

Katerina P. Kypreou[1*], Irene Stefanaki[1*], Kyriaki Antonopoulou[1], Fani Karagianni[1], Georgios Ntritsos[2], Alexios Zaras[1], Vasiliki Nikolaou[1], Iro Kalfa[3], Vasiliki Hasapi[1], Dorothea Polydorou[1], Helen Gogas[4], George M. Spyrou[5], Lars Bertram[6,7], Christina M. Lill[8], John P.A. Ioannidis[9], Christina Antoniou[1], Evangelos Evangelou[2,10**] and Alexander I. Stratigos[1**]

[1]Department of Dermatology, University of Athens, School of Medicine, Andreas Sygros Hospital, Athens, Greece; [2]Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, School of Medicine, University of Ioannina, Ioannina, Greece, GR; [3]Blood Donation Unit, Laikon Hospital, Athens, Greece; [4]Department of Internal Medicine, University of Athens, Laikon Hospital, Athens, Greece; [5]Biomedical Informatics Unit, Biomedical Research Foundation, Academy of Athens, Athens, Greece; [6]Platform for Genome Analytics, Institutes of Neurogenetics & Integrative and Experimental Genomics, University of Lübeck, Lübeck, Germany; [7]School of Public Health, Faculty of Medicine, Imperial College London, London, UK; [8]Institute of Neurogenetics, University of Lübeck, Lübeck, Germany; [9]Stanford Prevention Research Center, Department of Medicine and Department of Health Research and Policy, Stanford University School of Medicine, and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, USA; [10]Department of Epidemiology and Biostatistics, Imperial College London, St Mary's Campus, London, UK.

Word count for the abstract: 200

Word count for the text (excluding references, figures, and tables):3,485

Figures: 1

Tables: 2

Supplementary Figures: 0

Supplementary Tables: 6

The study was approved by the Scientific and Ethics Committee of Andreas Sygros Hospital.

Short title: SNPs and risk score in Greek melanoma cohort

Abbreviations: AUC, area under the receiver-operating characteristics curve; GWAS, genome wide association study; GRS, genetic risk score; CM, cutaneous melanoma; GWS, genome wide significant; OR, odds ratio; SNP, single nucleotide polymorphism.

*these authors contributed equally to this work

**these authors contributed equally to this work

Corresponding author: Alexander J. Stratigos, MD

Department of Dermatology,

University of Athens, A. Sygros Hospital

5 I. Dragoumi Street, Athens 16121, Greece

alstrat@hol.gr

Abstract

Background: Many single-nucleotide polymorphisms (SNPs) have been described as putative risk factors for melanoma.

Objective: To validate the most prominent genetic risk loci in an independent Greek melanoma case-control dataset and to assess their cumulative effect solely or combined with established phenotypic risk factors on individualized risk prediction.

Methods: We genotyped 59 SNPs in 800 patients and 800 controls and tested their association with melanoma using logistic regression analyses. We constructed a weighted genetic risk score (GRS_{GWS}) based on SNPs that showed genome-wide significant (GWS) association with melanoma in previous studies and assessed their impact on risk prediction

Results: Fifteen independent SNPs from 12 loci were significantly associated with melanoma ($p < 0.05$). Risk score analyses yielded an odds ratio: OR=1.36 per standard deviation increase of the GRS_{GWS} ($p=1.1x10^{-7}$). Individuals in the highest 20% of the GRS_{GWS} had a ~1.88-fold increase in melanoma risk compared with those in the middle quintile. By adding the GRS_{GWS} to a phenotypic risk model, including eye, hair and skin color, phototype, tanning ability, sex and age, the C-stastistic increased from 0.764 to 0.775 ($p=0.007$).

Conclusion: The GRS_{GWS} is associated with melanoma risk and achieves a modest improvement in risk prediction when added in the phenotypic risk model.

Keywords: cutaneous melanoma, genetic association, GWAS, genetic risk score, multivariable prediction model, non-genetic risk factors, risk assessment.

Introduction

Cutaneous melanoma (CM) is a potentially lethal skin malignancy, showing a continuously increasing incidence rate in Caucasians worldwide. The development of melanoma is a complex process involving the interplay of environmental, phenotypic and genetic risk factors. The role of genetic factors in melanomagenesis has been recognized since the identification of CDKN2A (Hussussian *et al.*, 1994; Kamb *et al.*, 1994) and CDK4 (Puntervoll *et al.*, 2013; Soufir *et al.*, 1998; Zuo *et al.*, 1996) as high penetrance susceptibility genes. Recent efforts have contributed to the discovery of an additional number of high risk genes, such as BAP1, MITF, TERT, POT1 and other shelterin complex GENES (ACD and TERF2IP) (Aoude *et al.*, 2015; Bertolotto *et al.*, 2011; Harbour *et al.*, 2010; Horn *et al.*, 2013; Robles-Espinoza *et al.*, 2014; Shi *et al.*, 2014; Yokoyama *et al.*, 2011). Genetic association studies, i.e., genome-wide association studies (GWAS) and candidate-gene studies have also revealed numerous common SNPs exerting more modest risk effects with more than 20 loci, including 5 new, that surpassed the genome wide significance threshold (i.e. $p < 5x10^{-8}$) for association with CM in recent GWAS (Law *et al.*, 2015). These studies have established the association of CM with pigmentation (MC1R, TYR and SLC45A2) and nevi-associated genes (MTAP, PLA2G6), as well as with loci potentially implicated in apoptosis (CASP8), DNA repair (PARP-1, ATM), metabolism (FTO) and more recently, telomerase length (TERT/CLPTM1L) (Barrett *et al.*, 2011; Iles *et al.*, 2013; Ward *et al.*, 2012). Most reported genetic variants have been summarized in an updated field synopsis of published genetic association studies (Antonopoulou *et al.*, 2015; Chatzinasiou *et al.*, 2011;).

This growing list of melanoma risk loci needs to be validated in large and independent datasets from other populations. In this context, the Greek population is of particular interest since it has a reportedly low incidence of melanoma compared to other European countries despite a high

degree of ambient ultraviolet exposure year-round (Ferlay *et al.*, 2013). The aim of this study was to validate the extensive set of SNPs that have been previously associated with CM risk in an independent sample of melanoma patients and healthy controls from Greece. In addition, we assessed the cumulative impact of the genetic variants on melanoma risk prediction by calculating a weighted GRS and combined this GRS with non-genetic, phenotypic risk factors.

RESULTS

Demographics and phenotypic traits of the 800 patients with CM and 800 control subjects included in this study are shown in Table S1. Fifty-five of 59 SNPs were genotyped with call rates ≥97%. One SNP deviated from Hardy-Weinberg equilibrium in the control population (rs1129038, $p=1.6x10^{-4}$, in HERC2) and one SNP (rs149617956 in MITF) was monomorphic. These SNPs were excluded from subsequent analyses. Fifty-three SNPs were considered in the final analysis, of which 26 were genome-wide significantly associated with CM based on the MelGene meta-analysis or from independent GWAS if they had not been included in the MelGene meta-analysis. Calculation of the linkage disequilibrium r^2 metric and conditional analyses revealed that all SNPs represent independent loci (data not shown).

The median power to detect the original effects as reported previously for the 53 SNPs based on the observed risk allele frequency in the control group was 0.455 (interquartile range 0.226 to 0.725) and the mean power was 0.495. For the 26 SNPs that were found to be GWS, the median and mean of the power estimates were 0.668 (0.380 to 0.906) and 0.634, respectively. Based on power calculation, it is expected that our study yielded 26 statistically significant associations

among the 53 tested SNPs. Fifteen SNPs thereof were statistically significantly associated with CM in our study. Sixteen of the 26 robustly GWS variants were expected to be associated in our study, and 11 SNPs were indeed nominally significant (p=0.07 for probability test).

Association between putative risk SNPs and melanoma

Logistic regression analyses assuming an additive model revealed 15 SNPs with nominally significant ($p<0.05$) effect size estimates showing the same direction of effect as previously described (Table 1, Table S2). This included 10 SNPs that had been reported to be associated with CM with GWS, specifically rs16891982, rs1805007, rs401681, rs1885120, rs4636294, rs10931936, (Antonopoulou *et al.*, 2015) as well as rs12918773, rs10739221, rs4778138, rs17119490 (Barrett *et al.*, 2011; Bishop *et al.*, 2009; Law *et al.*, 2015). Among the five new loci identified in the most recent GWAS meta-analysis (Law *et al.*, 2015) the intergenic SNP with rs10739221 near TMEM38B, ZNF462 and RAD23B as well as the SNP with rs4778138 in OCA2 at 15q13.1 were significantly associated with CM in our dataset showing effect estimates into the same direction as in Law et al. (Law *et al.*, 2015) (rs10739221: OR=1.209, p=0.015, rs4778138: OR=0.833, p=0.014, Table 1).

Figure S1 and Table S2 summarize the additive ORs of the eligible SNPs with melanoma risk in our study as well as the ORs reported in the original reference source. Overall, we observed a modest correlation of our effect size estimates and those reported previously (r^2=0.41, p=0.038 for the previously GWS SNPs; r^2=0.34, p=0.0130 for all 53 SNPs). The median risk allele frequency for the 53 risk alleles was 40.95% (IQR, 14.14-64.72%) in the Greek population.

Compared to a set of European populations derived from the 1KG project panel the correlation of risk allele frequencies was very high (r^2=0.97, P<0.0001) (Figure S2, Table S3).

Association between the GRS and melanoma

Risk score analyses yielded an OR=1.36 (95% CI: 1.21-1.52) per standard deviation increase of the GRS_{GWS} ($p=1.1x10^{-7}$). The magnitude and the strength of the association were comparable for GRS_{ALL} (OR = 1.39 (95% CI: 1.23-1.55, $p=3.2x10^{-8}$); Table S4). The adjusted ORs for melanoma showed a linear relationship with increasing percentiles of the GRS (trend test result for quintiles of GRS_{GWS}: $p=1.4x10^{-7}$, GRS_{ALL}: $p=3.2x10^{-9}$) (Figure 1, Table S5). The OR for individuals in the lowest quintile was 0.73 (95% CI: 0.50-1.05) and for participants in the highest quintile 1.88 (95% CI: 1.29-2.74) compared with study participants in the middle quintile (Table S5).

The discriminative ability of the GRS_{GWS} as measured by the C-statistic was 0.575 (95% CI: 0.549-0.604). When we considered traditional non-genetic risk factors only (i.e. sex, age, eye, hair and skin colour, phototype and tanning ability) the C-statistic was 0.764 (95% CI: 0.741-0.787). Upon combination of all genetic and non-genetic risk factors the C-statistic including GRS_{GWS} increased to 0.775 (95% CI: 0.752-0.797, p for area under the receiver-operating characteristics curve (AUC) comparison=0.007). The results were similar when GRS_{All} was considered (Table 2). The root mean square error (RMSE) in the 5-fold cross-validation approach ranged from 0.453 to 0.465 for the non-genetic model. When the GRS_{GWS} was added root mean square error ranged from 0.442 to 0.486. In both models cross-validation indicates a very good

model fit. Moreover, calibration assessment revealed that the predicted probabilities agree with the observed probabilities (Hosmer-Lemeshow test p-value=0.77).

A sensitivity analysis excluding all participants with missing values yielded comparable results. A total of 1,285 participants were considered and the C-statistic for the non-genetic model was 0.728 (95% CI: 0.701-0.755). The model including GRS_{GWS} yielded a C-statistic of 0.741 (95% CI: 0.741-0.767).

The age-stratified association results of the GRS and CM are summarized in Table S6. As shown, the interaction between GRS and age was not significant ($p \geq 0.05$).

DISCUSSION

We comprehensively assessed over 50 putative melanoma risk SNPs in a large and independent Greek dataset. Furthermore, we showed that the inclusion of common genetic variants in a CM prediction model leads to a modest improvement of its predictive abilities compared to a risk prediction model based on non-genetic factors only.

The selection of variants was mainly based on the MelGene field synopsis of genetic associations of melanoma, which systematically curates and meta-analyzes all published melanoma-associated variants (Antonopoulou *et al.*, 2015). Most of the variants that showed significant effects in our dataset pertained to genes controlling pigmentary traits. This can be explained by the fact that the

majority of these variants seem to exert moderate or even large effects on disease risk, thus our study had sufficient power to detect them. Eleven of the 26 variants reported as genome wide significant ($p < 5 \times 1\ 10^{-8}$) in the original GWAS (Amos *et al.*, 2011; Barrett *et al.*, 2011; Bishop *et al.*, 2009; Brown *et al.*, 2008; Iles *et al.*, 2013; Macgregor *et al.*, 2011; Teerlink *et al.*, 2012) or subsequent meta-analysis in MelGene were replicated in our cohort at a nominal significant level. Among the 5 newly identified genetic loci in a recent two-stage GWAS meta-analysis (Law *et al.*, 2015) involving 15,990 cases and 26,409 controls, 1 intergenic locus at 9q31.2 (rs10739221), at the proximity of TMEM38B, ZNF462 and the nucleotide excision repair gene RAD23B, was replicated (Masutani *et al.*, 1994). In addition, the SNP in OCA2 at 15q13.1, a potential determinant of eye color, that was found GWS for melanoma in Law et al (Law *et al.*, 2015), was also replicated in our study.

Several risk prediction models for melanoma using conventional phenotypic or clinical factors have been published, in an effort to better assess individual risk and develop more targeted prevention plans (Olsen *et al.*, 2015; Vuong *et al.*, 2014). Most of these prediction tools achieve modest discriminatory efficacy, yet their performance is variable upon independent validation due to poor calibration, lack of reproducible standardized assessment items, or heterogeneity in the definition of predictor variables (Olsen *et al.*, 2015). The discovery of multiple genetic variants that are associated with melanoma risk along with the constantly decreasing genotyping costs, have led to the development of genetic risk models with the potential advantage of identifying individuals at risk who may not be considered as so based on phenotypic characterization or exposure data. In the present study, we attempted to summarize the available genetic information by constructing a GRS using evidence from SNPs that have been associated with melanoma. We found that the risk for melanoma was associated with GRS even when adjusting for traditional risk factors, such as skin, hair and eye color. The results were similar for our primary model

analysis (GRS consisting of only GWS variants) and a secondary model consisting of all 53 analyzed variants. Our multivariable prediction model combining the most robust genetic factors by means of GWS association and phenotypic or non-genetic factors yielded a summary C-statistic of 0.775. The statistically significant, but marginal increase of 0.011 for the C statistic achieved by the addition of genetic susceptibility variants to a non-genetic model, does not strongly support the clinical utility of genetic variant profiling in individualized risk prediction.

Previous risk models using various clinical risk factors yielded AUCs ranging from 0.62 to 0.86 (Davies *et al.*, 2015; Vuong *et al.*, 2014). However, there are limited published prediction models incorporating genetic factors in CM (Cust *et al.*, 2013; Fang *et al.*, 2013; Penn *et al.*, 2014; Stefanaki *et al.*, 2013; Whiteman and Green, 2005). Three studies focused on the effect of MC1R in melanoma prediction (Cust *et al.*, 2013; Penn *et al.*, 2014; Whiteman and Green, 2005). Whiteman and Greene found that MC1R variants substantially increased melanoma risk when present in persons of olive skin color (Whiteman and Green, 2005). Cust et al concluded that MC1R is a better predictor than pigmentation characteristics in early-onset melanoma (Cust *et al.*, 2013), while Penn et al reported that the addition of MC1R genotype information to the baseline model resulted in a slight but statistically significant improvement in risk prediction, especially in nevus-prone patients (Penn *et al.*, 2014). In our previous study (Stefanaki *et al.*, 2013) the addition of 8 SNPs with nominal significance to a clinical model did not substantially improve melanoma risk prediction. In the present study, as well as in a recently published study of a GRS based on 11 SNPs tested in 1,804 melanoma patients and 1,026 controls (Fang *et al.*, 2013), the discrimination ability of the conventional phenotypic risk model increased when the GRS was incorporated to this model (C-statistic reaching 0.775 in our study and 0.69 in the study by Fang et al). Although differences between the two studies with regards to study design and population do not allow for direct comparison, the association of GRS with CM risk was significant in both

cases (OR=1.36, 95% CI: 1.21-1.59 for our GRS_{GWS} model compared to 1.12, 95% CI: 1.06-1.18, adjusting for similar risk factors) (Fang *et al.*, 2013).

Our sample represents the largest series of melanoma patients studied in Greece. We constructed our main GRS model based on established independent signals that showed genome-wide significance in previous studies, regardless of how they performed in our Greek case-control study. To this end, we avoided data overfitting by models that take in account only those variants with statistical significance in our population. In addition, we applied the checklist of the TRIPOD (Collins *et al.*, 2015) and GRIPS (Janssens *et al.*, 2011) statement recommendations, which aim to strengthen the transparency and homogeneity of reporting of multivariable and genetic risk prediction models among studies. Certain limitations apply to our study, with foremost the small size of our cohort. Several values concerning phenotypic characteristics, including the number of nevi, are missing due to variations in the information and questionnaires used by the participating centers. In addition, we did not include family history as a risk factor since this information was not available for the vast majority of our control samples. Risk prediction algorithms in other cancers, i.e., breast cancer suggest that the inclusion of family history in a polygenic risk score leads to further substantial improvement of the risk prediction model (Mavaddat *et al.*, 2015; So *et al.*, 2011). In addition, we did not take into account possible gene-environment interactions or gene-gene interactions. Incorporating SNPs with a stronger evidence of association after fine mapping of relevant genomic regions (Barrett *et al.*, 2015), in combination with intermediate or high risk genes might further improve the risk stratification of the GRS. Although we tested the internal validity of our prediction models, genetic predictive models for melanoma would benefit from additional external validation testing in similar (southern European) or other populations.

In conclusion, we replicated several genetic variants that confer susceptibility for melanoma in our population, confirming the polygenic nature of melanoma. We also explored the predictive capability of a GRS, which incorporated several GWS variants reported in the literature. The GRS was not superior from a phenotypic risk model, and its combination with phenotypic risk variables only slightly enhanced the discriminatory ability of our model. Based on our results, we cannot support the implementation of genetic variant profile in risk prediction models of melanoma. Independent studies in other populations will be required to adequately validate these findings.

MATERIALS AND METHODS

Study Population

The Greek sample consisted of patients with a histologically confirmed diagnosis of invasive melanoma at A. Sygros Hospital, a large referral center of melanoma and skin cancer in Athens and a collaborating oncological center (Laiko Hospital, Oncology Clinic), from 2000 to 2014. Both centers receive the majority of melanoma patients from Athens, thus consisting a representative sample size of the Greek population.

The control subjects were blood donors from a blood donation center in Athens and individuals with minor skin diseases and no history of skin malignancy, attending the outpatient service of our hospital. All subjects were above the age of 18.

Demographic variables, pigmentation traits (eye, hair, and skin color), skin phototype and tanning ability were obtained through a questionnaire that was given to the participants and clinical examination by a certified dermatologist. The study protocol was approved by the Scientific and Ethics Committee of A. Sygros Hospital; all participating individuals gave written informed consent prior to participating in the study.

SNP selection

Fifty nine variants were genotyped. Most of the SNPs (n=52) to be genotyped were selected from MelGene (www.melgene.org), a continuously updated database that collects all SNPs associated with melanoma risk (Antonopoulou *et al.*, 2015; Athanasiadis *et al.*, 2014; Chatzinasiou *et al.*, 2011). We further included in our study 7 GWAS SNPs from arecent GWAS meta-analysis, which were tested in our cohort as part of the replication phase (Law *et al.*, 2015). A MelGene SNP should have a p-value <0.05 and strong evidence of credibility using Venice criteria (grade A) (Ioannidis *et al.*, 2008) or should be GWS ($p<5x10^{-8}$) if it had emerged from a GWA study to be included in our study. Thirteen out of the 52 variants selected from MelGene database were included in the analysis because of their strong biological correlation with important melanoma pathways, even if they did not meet the above criteria.

DNA isolation, Genotyping and Quality control

Genomic DNA was isolated from peripheral blood using the QIAamp DNA blood mini kit (Qiagen). A total of 100ng from each DNA sample were used to genotype the selected SNPs, using the Sequenom iPLEX assay (Sequenom, Hamburg, Germany) (Gabriel *et al.*, 2009). Our quality control criteria included the inclusion of SNPs with a genotype call rate of 97% or higher and no deviation from Hardy-Weinberg equilibrium ($p< 8{,}5x10^{-4}$). We also excluded participants that had available <90% of SNPs genotyped.

Statistical Analysis

The association of each SNP with CM was computed using logistic regression and assuming an additive model. Adjustment for multiple testing was conducted using Bonferroni correction for the effective number of SNPs included in the analysis (cut off: $p=8.5x10^{-4}$). Additionally, we estimated: a) the correlation of risk allele frequencies between the Europeans from a panel derived from the 1000 Genomes (1KG) project ("EUR" population, Phase 3 v5) and the Greek population and b) the correlation of the effect size estimates found in the Greek population with those reported previously. Minor allele frequencies from the 1KG panel were extracted from SNiPA (Arnold *et al.*, 2015), a genetic variant-centered annotation browser.

Finally, we calculated Linkage Disequilibrium metrics (r^2) using PLINK 1.07, for SNPs located in the same locus. We considered SNPs with $r^2<0.6$ as independent. For SNP pairs with r^2 ranging

from 0.3 to 1 we performed additional conditional logistic regression analyses to ensure independence.

Power Calculation

The QUANTO software was used for power calculations (http://biostats.usc.edu/Quanto.html). For every SNP, the power Gi to detect each of the described effects at a = 0.05 level given the observed risk allele frequency in the Greek sample, was calculated assuming an additive (per-allele) genetic model. The sum of the power estimates corresponds to the number of variants that would be expected to replicate.

GRS calculation and melanoma risk prediction analyses

We constructed two different weighted GRS. Primary GRS was based on SNPs that have been found GWS from MelGene meta-analysis (n=11), 7 SNPs from Law et al. (Law et al, 2015) and 8 SNPs from independent GWAS that did not have sufficient datasets to be meta-analyzed in MelGene (GRS_{GWS}). A secondary GRS consisted of all analysed SNPs (GRS_{ALL}) (n=53 successfully genotyped of the 59) (Table S2). The GRS represents a sum of the number of effect alleles weighted by their effect size estimates, specifically by their beta coefficients. The effect estimates were derived from the MelGene meta-analysis or independent published GWAS (Table S2). Each weighted GRS was standardized per unit increase in the control population.

For each GRS we calculated the association with CM adjusted for sex, age and a list of traditional risk factors including eye color, hair color, skin color, phototype (according to the Fitzpatrick scale) and tanning ability. In case of missing values of the predictors we created an indicator variable for missingness and that was incorporated into the model as a separate covariate. We also performed a sensitivity analysis including only sex, age and the relevant GRS and an analysis limited to variables with non-missing values.

We also assessed the performance of the predictive capability of the GRS by calculating the AUC. The AUC was calculated based on the covariates described above with and without the GRS. Bootstrapping (n=1,000) was used to calculate the p-values for the comparisons of the AUCs. In order to assess the internal validity of our predictive models we calculated the root mean square error, which error represents the differences between predicted and observed values, in 5-fold validation splits with 1,000 replications. Small values with a narrow range indicates good validation. The calibration of the model was also assessed by calculating the distribution of expected values and compared with the observed ones using Hosmer-Lemeshow test.

Finally, quantiles of the GRS were created and ORs were calculated and compared in 5 different categories using the 3rd category as a reference. Moreover, we stratified the dataset into quartiles of age (i.e. age at onset for cases and age at examination for controls) and we calculated the OR within each age group.

Reporting of study results

The report followed the recommendations by two consensus publications aiming to enhance the quality of articles focusing on multivariable and genetic risk prediction models, i.e. the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement (Collins *et al.*, 2015) and the Genetic Risk Prediction Studies (GRIPS) statement (Janssens *et al.*, 2011) respectively.

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGEMENTS:

Funding sources for the work: This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program ‘‘Education and Lifelong Learning’’ of the National Strategic Reference Framework (NSRF) - Research Funding Program: Aristeia I - 1094.

REFERENCES

Amos CI, Wang LE, Lee JE, *et al.* (2011) Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet* 20:5012-23.

Antonopoulou K, Stefanaki I, Lill CM, *et al.* (2015) Updated Field Synopsis and Systematic Meta-Analyses of Genetic Association Studies in Cutaneous Melanoma: The MelGene Database. *The Journal of investigative dermatology* 135:1074-9.

Aoude LG, Pritchard AL, Robles-Espinoza CD, *et al.* (2015) Nonsense mutations in the shelterin complex genes ACD and TERF2IP in familial melanoma. *Journal of the National Cancer Institute* 107.

Arnold M, Raffler J, Pfeufer A, *et al.* (2015) SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* 31:1334-6.

Athanasiadis EI, Antonopoulou K, Chatzinasiou F, *et al.* (2014) A Web-based database of genetic association studies in cutaneous melanoma enhanced with network-driven data exploration tools. *Database : the journal of biological databases and curation* 2014.

Barrett JH, Iles MM, Harland M, *et al.* (2011) Genome-wide association study identifies three new melanoma susceptibility loci. *Nature genetics* 43:1108-13.

Barrett JH, Taylor JC, Bright C, *et al.* (2015) Fine mapping of genetic susceptibility loci for melanoma reveals a mixture of single variant and multiple variant regions. *International journal of cancer Journal international du cancer* 136:1351-60.

Bertolotto C, Lesueur F, Giuliano S, *et al.* (2011) A SUMOylation-defective MITF germline mutation predisposes to melanoma and renal carcinoma. *Nature* 480:94-8.

Bishop DT, Demenais F, Iles MM, *et al.* (2009) Genome-wide association study identifies three loci associated with melanoma risk. *Nature genetics* 41:920-5.

Brown KM, Macgregor S, Montgomery GW, *et al.* (2008) Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nature genetics* 40:838-40.

Chatzinasiou F, Lill CM, Kypreou K, *et al.* (2011) Comprehensive field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma. *Journal of the National Cancer Institute* 103:1227-35.

Collins GS, Reitsma JB, Altman DG, *et al.* (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC medicine* 13:1.

Cust AE, Goumas C, Vuong K, *et al.* (2013) MC1R genotype as a predictor of early-onset melanoma, compared with self-reported and physician-measured traditional risk factors: an Australian case-control-family study. *BMC cancer* 13:406.

Davies JR, Chang YM, Bishop DT, *et al.* (2015) Development and validation of a melanoma risk score based on pooled data from 16 case-control studies. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 24:817-24.

Fang S, Han J, Zhang M, *et al.* (2013) Joint effect of multiple common SNPs predicts melanoma susceptibility. *PloS one* 8:e85642.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, *et al.* (2013) Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *European journal of cancer* 49:1374-403.

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current protocols in human genetics / editorial board, Jonathan L Haines [et al]* Chapter 2:Unit 2 12.

Harbour JW, Onken MD, Roberson ED, *et al.* (2010) Frequent mutation of BAP1 in metastasizing uveal melanomas. *Science* 330:1410-3.

Horn S, Figl A, Rachakonda PS, *et al.* (2013) TERT promoter mutations in familial and sporadic melanoma. *Science* 339:959-61.

Hussussian CJ, Struewing JP, Goldstein AM, *et al.* (1994) Germline p16 mutations in familial melanoma. *Nature genetics* 8:15-21.

Iles MM, Law MH, Stacey SN, *et al.* (2013) A variant in FTO shows association with melanoma risk not due to BMI. *Nature genetics* 45:428-32, 32e1.

Ioannidis JP, Boffetta P, Little J, *et al.* (2008) Assessment of cumulative evidence on genetic associations: interim guidelines. *International journal of epidemiology* 37:120-32.

Janssens AC, Ioannidis JP, van Duijn CM, *et al.* (2011) Strengthening the reporting of genetic risk prediction studies: The GRIPS Statement. *Annals of internal medicine* 154:421-5.

Kamb A, Shattuck-Eidens D, Eeles R, *et al.* (1994) Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nature genetics* 8:23-6.

Law MH, Bishop TD, Lee EJ, *et al.* (2015) Genome-wide meta-analysis of 5,990 melanoma cases and 26,409 controls reveals novel susceptibility loci *Nature genetics*.

Macgregor S, Montgomery GW, Liu JZ, *et al.* (2011) Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nature genetics* 43:1114-8.

Masutani C, Sugasawa K, Yanagisawa J, *et al.* (1994) Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *The EMBO journal* 13:1831-43.

Mavaddat N, Pharoah PD, Michailidou K, *et al.* (2015) Prediction of breast cancer risk based on profiling with common genetic variants. *Journal of the National Cancer Institute* 107.

Olsen CM, Neale RE, Green AC, *et al.* (2015) Independent validation of six melanoma risk prediction models. *The Journal of investigative dermatology* 135:1377-84.

Penn LA, Qian M, Zhang E, *et al.* (2014) Development of a melanoma risk prediction model incorporating MC1R genotype and indoor tanning exposure: impact of mole phenotype on model performance. *PloS one* 9:e101507.

Puntervoll HE, Yang XR, Vetti HH, *et al.* (2013) Melanoma prone families with CDK4 germline mutation: phenotypic profile and associations with MC1R variants. *J Med Genet* 50:264-70.

Robles-Espinoza CD, Harland M, Ramsay AJ, *et al.* (2014) POT1 loss-of-function variants predispose to familial melanoma. *Nature genetics* 46:478-81.

Shi J, Yang XR, Ballew B, *et al.* (2014) Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. *Nature genetics* 46:482-6.

So HC, Kwan JS, Cherny SS, *et al.* (2011) Risk prediction of complex diseases from family history and known susceptibility loci, with applications for cancer screening. *American journal of human genetics* 88:548-65.

Soufir N, Avril MF, Chompret A, *et al.* (1998) Prevalence of p16 and CDK4 germline mutations in 48 melanoma-prone families in France. The French Familial Melanoma Study Group. *Hum Mol Genet* 7:209-16.

Stefanaki I, Panagiotou OA, Kodela E, *et al.* (2013) Replication and predictive value of SNPs associated with melanoma and pigmentation traits in a Southern European case-control study. *PloS one* 8:e55712.

Teerlink C, Farnham J, Allen-Brady K, *et al.* (2012) A unique genome-wide association analysis in extended Utah high-risk pedigrees identifies a novel melanoma risk variant on chromosome arm 10q. *Human genetics* 131:77-85.

Vuong K, McGeechan K, Armstrong BK, *et al.* (2014) Risk prediction models for incident primary cutaneous melanoma: a systematic review. *JAMA dermatology* 150:434-44.

Ward KA, Lazovich D, Hordinsky MK (2012) Germline melanoma susceptibility and prognostic genes: a review of the literature. *Journal of the American Academy of Dermatology* 67:1055-67.

Whiteman DC, Green AC (2005) A risk prediction tool for melanoma? *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 14:761-3.

Yokoyama S, Woods SL, Boyle GM, *et al.* (2011) A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. *Nature* 480:99-103.

Zuo L, Weger J, Yang Q, *et al.* (1996) Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nature genetics* 12:97-9.

Table 1. Statistical significant results from the univariable analysis of the 53 eligible SNPs.

			Univariable Analysis		
SNP	**Nearest Gene[1]**	**MAF**	**P**	**OR (95% CI)**	**Function**
rs12918773	(CDK10)	0.031	$1.63x10^{-6}$	2.28 (1.61, 3.22)	Pigmentation
rs16891982	SLC45A2	0.135	$3.82x10^{-6}$	0.59 (0.47, 0.74)	Pigmentation
rs1805007	MC1R	0.024	$8.22x10^{-6}$	2.34 (1.59, 3.43)	Pigmentation
rs11547464	MC1R	0.009	$1.04x10^{-4}$	3.13 (1.71, 5.75)	Pigmentation
rs401681	CLPTM1L	0.416	$2.23x10^{-4}$	1.30 (1.13, 1.50)	Nevi
rs12913832	HERC2	0.368	$7.78x10^{-4}$	1.28 (1.11, 1.47)	Pigmentation
rs1805005	MC1R	0.141	$2.56x10^{-3}$	1.34 (1.11, 1.62)	Pigmentation
rs1885120	MYH7B	0.019	$3.09x10^{-3}$	1.94 (1.24, 3.04)	Pigmentation
rs35390	SLC45A2	0.089	$3.46x10^{-3}$	0.67 (0.51, 0.88)	Pigmentation
rs10739221[2]	(TMEM38B, ZNF462, RAD23B)	0.271	0.015	1.21 (1.04, 1.41)	Intergenic locus
rs4778138[2]	OCA2	0.370	0.014	0.83 (0.72, 0.96)	Pigmentation
rs3768080	NID1	0.4095	0.026	1.17 (1.02, 1.35)	Basement membrane
rs10931936	CASP8	0.307	0.030	1.18 (1.02, 1.37)	Apoptosis
rs17119490	LOC101927549	0.01757	0.033	1.67 (1.04, 2.68)	Intergenic locus
rs4636294	MTAP	0.4044	0.030	0.85 (0.74, 0.98)	Nevi

Abbreviations: MAF=minor allelic frequency, OR=Odds Ratio, CI=Confidence Intervals

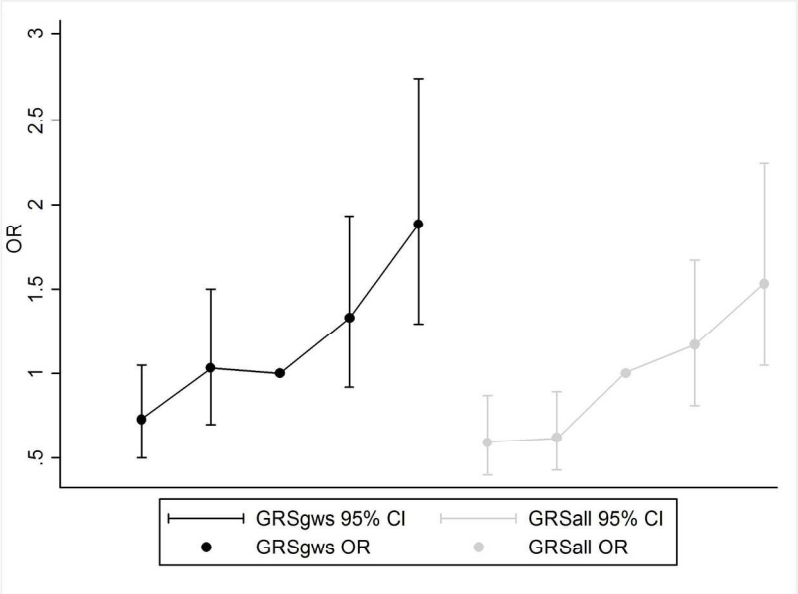
[1]"Nearest Gene" denotes the gene in the respective locus or one proximal gene in the respective locus (denoted with parenthesis) if the SNP itself does not map into a gene region. It should be noted that these genes are not necessarily the genes that are functionally affected by the genetic association finding in this locus.

[2]SNPs derived from GWAS meta-analysis (Law et. al, 2015) and replicated to our cohort.

Table2. Risk prediction performance for the four different models of predictors in the Greek data set.		
	AUC	**95% CI**
Phenotypic Risk factors only[1]	0.764	0.741-0.787
Phenotypic Risk factors + GRS_{GWS}	0.775	0.752-0.797
Phenotypic Risk factors + GRS_{ALL}	0.775	0.752-0.798
Abbreviations: AUC, area under the receiver operating characteristic curve, CI=Confidence Intervals, GRS=genetic risk score, 1Risk factors= sex, age, eye color, hair color, skin color, phototype and tanning ability.		

Figure Legends:

1. Associations between GRS and melanoma in different quintile groups for GRS_{GWS} and GRS_{ALL}.



Associations between GRS and melanoma in different quintile groups for GRSGWS and GRSALL.
190x142mm (300 x 300 DPI)

Supplementary Table 1. Demographic characteristics and pigmentary phenotype of melanoma cases and control subjects.

	Patients (n=800)	**Controls (n=800)**	**P[1]**
Median age (years) (IQR; range)	53 (41-66; 17-97)	41 (31-53; 19-80)	**0.005**
Missing (N)	40	33	
Sex, N (%)			0.201
Men	394 (49.25%)	408 (51.00%)	
Women	406 (50.75%)	365 (45.63%)	
Missing	0	27 (3.38%)	
Hair color			0.082
Blonde	79 (9.88%)	47 (5.88%)	
Red	21 (2.63%)	25 (3.13%)	
Light Brown	216 (27.00%)	245 (30.63%)	
Dark Brown	278 (34.75%)	333 (41.63%)	
Black	74 (9.25%)	97 (12.13%)	
Missing	132 (16.50%)	53 (6.63%)	
Eye color			**0.007**
Grey/Blue	87 (10.88%)	73 (9.13%)	
Green	144 (18.00%)	119 (14.88%)	
Light Brown	183 (22.88%)	226 (28.25%)	
Dark brown	232 (19.00%)	316 (39.50%)	
Black	3 (0.38%)	11 (1.38%)	
Missing	151 (18.88%)	55 (6.88%)	
Skin color			0.200
White	372 (46.50%)	294 (36.75%)	
Light Brown	277 (34.63%)	325 (40.63%)	

Dark	20 (2.50%)	124 (15.50%)	
Missing	131 (16.38%)	57 (7.13%)	
Phototype			0.075
Phototype I	33 (4.13%)	43 (5.38%)	
Phototype II	303 (37.88%)	243 (30.38%)	
Phototype III	234 (29.25%)	316 (39.50%)	
Phototype IV	98 (12.25%)	127 (15.88%)	
Missing	132 (16.50%)	71 (8.88%)	
Tanning ability[2]			**0.048**
Burn	96 (12%)	122 (15.25%)	
Minimal tan	287 (35.88%)	258 (32.25%)	
Burn than tan	207 (25.88%)	254 (31.75%)	
Deep tan	73 (9.13%)	81 (10.13%)	
Missing	137 (17.3%)	85 (10.63%)	

[1]Results from a Wilcoxon-Mann-Whitney test for the comparison of age between cases and controls; results from a chi-square test for the comparison of all other variables between cases and controls.
[2] Represents the answers to the question "How your skin reacts when you sunbathe during the first weeks of your vacation".

Supplementary Table 2: Location, Original Source and Genotype Results of the 59 selected SNPs.

SNP	Chr	BP	Nearest Gene[1]	Minor Allele	MAF	P	OR (95%CI)	OR source/Selection source of SNP	
rs7412746[2]	1	150860471	LOC100996521	T	0.4768	0.1618	0.905 (0.787, 1.041)	1.14	MacGregor et al., 2011
rs3219090	1	226564691	PARP1	A	0.3528	0.09918	0.883 (0.763, 1.024)	0.86	MelGene meta-analysis
rs3768080	1	236179869	NID1	G	0.4095	0.02578	1.174 (1.019, 1.351)	1.07	Nan et al., 2011
rs6750047[3]	2	38276549	RMDN2	A	-	-	-	-	Law et al., 2015
rs10931936[2]	2	202143928	CASP8	T	0.307	0.02993	1.180 (1.016, 1.370)	1.15	MelGene meta-analysis
rs1035142[2]	2	202153078	(ALS2CR12 and CASP8)	T	0.4546	0.1466	1.109 (0.9644, 1.275)	1.14	MelGene meta-analysis
rs149617956[4]	3	70014091	MITF						MelGene meta-analysis
rs13097028[5]	3	169464942	(ACTRT3)	T	0.2895	0.4274	0.939 (0.805, 1.096)	0.89	Song et al., 2014
rs12696304	3	169481271	(TERC)	G	0.2707	0.3378	0.926 (0.790, 1.084)	0.91	Law et al., 2015
rs4698934[5]	4	106139387	TET2	C	0.1335	0.4521	0.924 (0.751, 1.136)	0.85	Song et al., 2014
rs401681[2]	5	1322087	CLPTM1L	T	0.4159	0.000223	1.302 (1.132, 1.498)	1.19	MelGene meta-analysis
rs16891982[2]	5	33951693	SLC45A2	C	0.1355	$3.8x10^{-6}$	0.587 (0.467, 0.737)	0.42	MelGene meta-analysis
rs35390[5]	5	33955326	SLC45A2	C	0.08908	0.003462	0.672 (0.515, 0.879)	0.36	Barrett et al., 2011
rs12203592[5]	6	396321	IRF4	T	0.05451	0.4569	0.887 (0.648, 1.216)	1.16	MelGene meta-analysis
rs872071[5]	6	411064	IRF4	G	0.4385	0.8636	1.012 (0.879, 1.165)	0.93	Barrett et al., 2011
rs6914598[2]	6	21163919	CDKAL1	C	0.3331	0.6419	1.036 (0.894, 1.200)	1.10	Law et al., 2015
rs1636744[2]	7	16984280	(AGR3)	A	0.3695	0.8981	1.009 (0.874, 1.166)	1.09	Law et al., 2015
rs1408799	9	12672097	(TYRP1)	T	0.3319	0.4665	1.056 (0.912, 1.224)	0.91	MelGene meta-analysis
rs4636294[2]	9	21747804	MTAP	G	0.4044	0.03023	0.854 (0.739, 0.985)	0.83	MelGene meta-analysis
rs10757257[2]	9	21806562	MTAP	A	0.2976	0.06139	0.863 (0.739, 1.007)	0.81	MelGene meta-analysis
rs7023329[2]	9	21816528	MTAP	G	0.394	0.8437	0.986 (0.855, 1.137)	0.83	MelGene meta-analysis
rs3088440	9	21968159	CDKN2A	A	0.0801	0.515	1.087 (0.845, 1.397)	1.27	MelGene meta-analysis
rs11515[5]	9	21968199	CDKN2A	G	0.1809	0.518	0.942 (0.784, 1.130)	1.05	MelGene meta-analysis
rs1011970[5]	9	22062134	(CDKN2A)	T	0.1769	0.1826	1.129 (0.944, 1.351)	1.18	Maccioni et al., 2013
rs10739221[2]	9	109060830	(TMEM38B,	T	0.271	0.01536	1.209 (1.037, 1.409)	1.13	Law et al., 2015

			ZNF462, RAD23B)						
rs2995264[2]	10	105668843	OBFC1	G	0.1179	0.2326	1.137 (0.921, 1.403)	1.17	Law et al., 2015
rs17119490[2]	10	107522927	LOC101927549	A	0.01757	0.03287	1.668 (1.038, 2.683)	8.4	Teerlink et al., 2011
rs1485993	11	69362414	(CCND1)	T	0.4211	0.07393	1.137 (0.988, 1.308)	1.09	MelGene meta-analysis
rs1042602	11	88911696	TYR	A	0.4855	0.5016	1.049 (0.912, 1.206)	0.94	MelGene meta-analysis
rs1847142[2]	11	89021574	TYR	A	0.2199	0.2166	1.110 (0.941, 1.310)	1.31	Bishop et al., 2009
rs1801516[2]	11	108175462	ATM	A	0.1395	0.1383	0.856 (0.696, 1.052)	0.84	MelGene meta-analysis
rs1544410	12	48239835	VDR	A	0.4266	0.6725	0.970 (0.843, 1.117)	0.9	MelGene meta-analysis
rs17655	13	103528002	XPG	G	0.2748	0.2088	0.904 (0.772, 1.058)	0.91	MelGene meta-analysis
rs1800407	15	28230318	OCA2	A	0.06078	0.1572	1.223 (0.925, 1.616)	1.38	MelGene meta-analysis
rs4778138[2]	15	28355820	OCA2	G	0.3698	0.01417	0.833 (0.719, 0.964)	0.84	Law et al., 2015
rs1129038[6]	15	28356859	HERC2	A	-	-	-	-	Amos et al., 2011
rs12913832	15	28365618	HERC2	G	0.3676	0.000778	1.276 (1.107, 1.471)	1.11	Amos et al., 2011
rs16953002[3]	16	54114824	FTO	A	-	-	-	-	Iles et al., 2013
rs7188458[2]	16	89726484	C16orf55	A	0.3199	0.2367	1.094 (0.943, 1.268)	1.30	Bishop et al., 2009
rs12918773[2]	16	89741403	(CDK10)	A	0.03082	$1.6x10^{-6}$	2.281 (1.615, 3.223)	1.87	Bishop et al., 2009
rs258322[3]	16	89755903	CDK10	T	-	-	-	-	MelGene meta-analysis
rs1805005	16	89985844	MC1R	T	0.1414	0.002556	1.339 (1.107, 1.619)	1.14	MelGene meta-analysis
rs1805006	16	89985918	MC1R	A	0.003145	0.2556	0.399 (0.077, 2.058)	1.53	MelGene meta-analysis
rs2228479	16	89985940	MC1R	A	0.04255	0.1566	1.266 (0.913, 1.755)	1.08	MelGene meta-analysis
rs11547464	16	89986091	MC1R	A	0.008794	0.000104	3.133 (1.707, 5.750)	1.47	MelGene meta-analysis
rs1805007[2]	16	89986117	MC1R	T	0.02453	$8.2x10^{-6}$	2.339 (1.594, 3.433)	1.8	MelGene meta-analysis
rs1805009[2]	16	89986546	MC1R	C	0.001252	0.4138	2.003 (0.366, 10.95)	1.89	MelGene meta-analysis
rs4238833[2]	16	90050689	AFG3L1	G	0.3218	0.2432	1.092 (0.942, 1.266)	1.32	Bishop et al., 2009
rs4785763[2]	16	90066936	AFG3L1	A	0.2972	0.1812	1.108 (0.953, 1.289)	1.35	MelGene meta-analysis
rs8059973[2]	16	90079534	DBNDD1	A	0.1814	0.9167	1.010 (0.843, 1.209)	0.74	Bishop et al., 2009
rs17305657[2]	20	31806588	C20orf71	C	0.02324	0.2198	1.312 (0.849, 2.025)	1.58	Brown et al., 2008

rs4911414	20	32729444	(ASIP)	T	0.2535	0.2643	0.912 (0.775, 1.072)	1.16	MelGene meta-analysis
rs6058017[5]	20	32856998	ASIP	G	0.1409	0.000412	1.406 (1.163, 1.699)	0.91	MelGene meta-analysis
rs17305573[3]	20	33180152	PIGU	C	-	-	-	-	
rs4911442	20	33355046	NCOA6	G	0.04887	0.05595	1.343 (0.992, 1.819)	1.28	MelGene meta-analysis
rs1885120[2]	20	33576989	MYH7B	C	0.01884	0.003086	1.944 (1.242, 3.041)	1.55	MelGene meta-analysis
rs1015362[5]	20	37738612	(ASIP)	A	0.2895	0.1906	0.902 (0.772, 1.053)	0.95	MelGene meta-analysis
rs45430[2]	21	42746081	MX2	G	0.4143	0.1772	0.907 (0.787, 1.045)	0.88	Barrett et al., 2011
rs6001027	22	38545619	PLA2G6	G	0.3785	0.8839	0.989 (0.857, 1.142)	0.86	MelGene meta-analysis

Abbreviations: MAF=minor allelic frequency, OR=Odds Ratio, CI=Confidence Intervals, BP=base pairs, Chr=chromosome

[1]"Nearest Gene" denotes the gene in the respective locus or one proximal gene in the respective locus (denoted with parenthesis) if the SNP itself does not map into a gene region. It should be noted that these genes are not necessarily the genes that are functionally affected by the genetic association finding in this locus.

[2]SNPs included in the GRS_{GWS}.

[3]SNP not included in the analysis due to call rate<0.97.

[4]SNP with rs149617956 was excluded from the analysis since it was monomorphic.

[5]SNPs selected from MelGene due to their biological significance.

[6]Deviation from Hardy-Weinberg equilibrium.

Supplementary Table 3. Risk allele frequency in the Greek sample and European sample from the 1000 genomes (1KG) panel for the 53 eligible SNPs.

SNP	Risk allele in the Greek sample	Risk allele frequency in the Greek sample (95% CI)	Risk allele frequency in the EU sample from 1KG (95% CI)
rs1011970	T	0.177 (0.15-0.203)	0.155 (0.122-0.188)
rs1015362	G	0.711 (0.679-0.742)	0.723 (0.683-0.763)
rs1035142	T	0.455 (0.42-0.489)	0.386 (0.342-0.429)
rs1042602	A	0.486 (0.451-0.52)	0.372 (0.329-0.415)
rs10739221	T	0.271 (0.24-0.302)	0.244 (0.205-0.282)
rs10757257	G	0.702 (0.671-0.734)	0.612 (0.568-0.655)
rs10931936	T	0.307 (0.275-0.339)	0.283 (0.243-0.323)
rs11515	C	0.819 (0.792-0.846)	0.875 (0.845-0.905)
rs11547464	A	0.009 (0.002-0.015)	0.009 (0.0002-0.018)
rs12203592	C	0.945 (0.93-0.961)	0.884 (0.855-0.913)
rs12696304	C	0.729 (0.699-0.76)	0.735 (0.695-0.774)
rs12913832	G	0.368 (0.334-0.401)	0.636(0.593-0.679)
rs12918773	A	0.031 (0.019-0.043)	0.084 (0.059-0.109)
rs13097028	C	0.711 (0.679-0.742)	0.664 (0.623-0.706)
rs1408799	T	0.332 (0.299-0.365)	0.346 (0.303-0.388)
rs1485993	T	0.421 (0.387-0.455)	0.365 (0.322-0.408)
rs1544410	G	0.573 (0.539-0.608)	0.596 (0.552-0.639)
rs1636744	A	0.370 (0.336-0.403)	0.407 (0.363-0.451)
rs16891982	G	0.865 (0.841-0.888)	0.938(0.916-0.960)
rs17119490	A	0.018 (0.008-0.027)	0.011 (0.0008-0.021)
rs17305657	C	0.023 (0.013-0.034)	0.065 (0.0425-0.087)
rs17655	C	0.725 (0.694-0.756)	0.750 (0.711-0.789)
rs1800407	A	0.061 (0.044-0.077)	0.076 (0.052-0.100)
rs1801516	G	0.861 (0.836-0.885)	0.838 (0.805-0.871)
rs1805005	T	0.141 (0.117-0.166)	0.112 (0.083-0.140)
rs1805006	C	0.997 (0.993-1.00)	0.990 (0.980-0.999)
rs1805007	T	0.025 (0.014-0.035)	0.072 (0.048-0.095)
rs1805009	C	0.001 (-0.001-0.004)	0.008 (-0.0007-0.017)
rs1847142	A	0.22 (0.191-0.249)	0.299 (0.258-0.340)
rs1885120	C	0.019 (0.009-0.028)	0.042 (0.023-0.060)
rs2228479	A	0.043 (0.029-0.057)	0.069 (0.04-0.092)
rs2995264	G	0.118 (0.096-0.14)	0.089 (0.063-0.115)
rs3088440	A	0.08 (0.061-0.099)	0.079 (0.054-0.103)
rs3219090	G	0.647 (0.614-0.68)	0.676 (0.634-0.718)
rs35390	A	0.911 (0.891-0.931)	0.965 (0.948-0.982)
rs3768080	G	0.41 (0.375-0.444)	0.494 (0.449-0.538)
rs401681	T	0.416 (0.382-0.45)	0.441 (0.397-0.485)
rs4238833	G	0.322 (0.289-0.354)	0.322 (0.280-0.363)
rs45430	A	0.586 (0.552-0.62)	0.622 (0.579-0.665)
rs4636294	A	0.596 (0.562-0.63)	0.504 (0.459-0.549)

rs4698934	T	0.867 (0.843-0.89)	0.815 (0.780-0.849)
rs4778138	A	0.63 (0.597-0.664)	0.831 (0.797-0.865)
rs4785763	A	0.297 (0.266-0.329)	0.299 (0.258-0.340)
rs4911414	G	0.747 (0.716-0.777)	0.700 (0.659-0.741)
rs4911442	G	0.049 (0.034-0.064)	0.087 (0.06-0.113)
rs6001027	A	0.622 (0.588-0.655)	0.636 (0.592-0.679)
rs6058017	G	0.141 (0.117-0.165)	0.103 (0.075-0.130)
rs6914598	C	0.333 (0.3-0.366)	0.313 (0.271-0.354)
rs7023329	A	0.606 (0.572-0.64)	0.520 (0.475-0.564)
rs7188458	A	0.32 (0.288-0.352)	0.393 (0.349-0.437)
rs7412746	C	0.523 (0.489-0.558)	0.477 (0.432-0.521)
rs8059973	A	0.181 (0.155-0.208)	0.183 (0.148-0.218)
rs872071	G	0.439 (0.404-0.473)	0.474 (0.429-0.518)

Abbreviations: CI=confidence intervals, N/A=not aplicable

Supplementary Table 4. Association between GRS and melanoma risk.

	OR[1]	95% CI	P
GRS_{GWS}	1.36	1.21-1.52	$1.1x10^{-7}$
GRS_{ALL}	1.39	1.23-1.55	$3.2x10^{-8}$

Abbreviations: OR=Odds Ratio, CI=Confidence Intervals, GRS=genetic risk score.
[1]OR for association between the GRS, coded as a continuous variable, and melanoma risk adjusted for sex, age, eye color, hair color, skin color, phototype and tanning ability.

Supplementary Table 5. Quintiles of GRS_{GWS} , GRS_{ALL}.

Analysis	**GRS_{GWS}**			**GRS_{GWS} adjusted for risk factors**		
Quintiles	**OR[1]**	**95% CI**	**P**	**OR[2]**	**95% CI**	**P**
1	0.77	0.56-1.05	0.095	0.73	0.50-1.05	0.095
2	1.01	0.74-1.38	0.937	1.03	0.70-1.50	0.881
3 (ref)	1	.	.	1	.	.
4	1.32	0.97-1.80	0.082	1.33	0.92-1.93	0.129
5	1.72	1.26-2.36	0.001	1.88	1.29-2.74	0.001
Analysis	**GRS_{ALL}**			**GRS_{ALL} adjusted for risk factors**		
1	0.65	0.49-0.91	0.007	0.59	0.48-0.87	0.007
2	0.68	0.50-0.92	0.014	0.62	0.43-0.89	0.01
3 (ref)	1	.	.	1	.	.
4	1.09	0.80-1.49	0.579	1.17	0.81-1.67	0.404
5	1.52	1.11-2.09	0.009	1.53	1.05-2.24	0.029

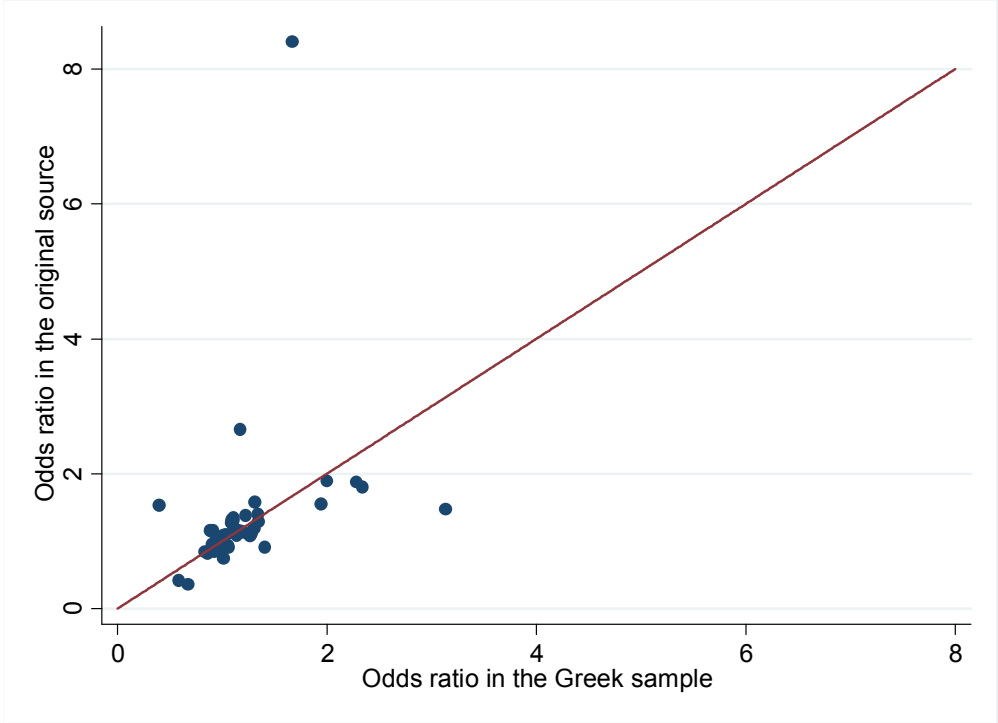
Abbreviations:OR=Odds Ratio, CI=Confidence Intervals, GRS=genetic risk score
[1]Odds ratios are for different quintiles of the genetic GRS relative to the middle quintile (40% to 60%) of the GRS
[2]Odds ratios are for different quintiles of the genetic GRS relative to the middle quintile (40% to 60%) of the GRS, adjusted for sex, age, eye color, hair color, skin color, phototype and tanning ability

Supplementary Table 6. Association between GRS and melanoma risk in different age groups (quartiles of age).

	GRS_{GWS}		GRS_{ALL}	
Age group[1]	**OR[2] (95% CI)**	**P**	**OR[2] (95% CI)**	**P**
<36	1.54 (1.21-1.98)	0.001	1.59 (1.20-2.08)	0.001
36-47	1.43 (1.13-1.80)	0.003	1.37 (1.07-1.73)	0.012
48-61	1.31 (1.04-1.64)	0.020	1.31 (1.05-1.63)	0.015
>61	1.20 (0.93-1.54)	0.171	1.32 (1.01-1.72)	0.041
	Interaction OR[3] (95% CI)		**Interaction OR[3] (95% CI)**	
Interaction between GRS and age	0.97 (0.91, 1.04)		0.98 (0.92, 1.05)	
$P_{interaction}$	0.392		0.649	

Abbreviations: OR=Odds Ratio, CI=Confidence Intervals, GRS=genetic risk score
[1]Age at diagnosis for melanoma patients, age at interview for controls.
[2]OR for association between the GRS and melanoma risk adjusted for sex, age, eye color, hair color, skin color, phototype and tanning ability.
[3]OR per 10 years for interaction between GRS and age.
Each weighted GRS was standardized per unit increase in the control population.

Supplementary Figure 1. Correlation of the effect sizes found in the Greek sample and those derived from MelGene, original publication or the Law et al., 2015.



Supplementary Figure 2. Correlation of the risk allele frequencies found in the Greek sample and the frequencies of the same alleles from the European sample from the 1000 genomes (1KG) panel.