Washington University School of Medicine
Digital Commons@Becker

Open Access Publications

2012

Melanocortin-1 receptor, skin cancer and phenotypic characteristics (M-SKIP) project: Study design and methods for pooling results of genetic epidemiological studies

Lynn Cornelius
Washington University School of Medicine in St. Louis

et al

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
Cornelius, Lynn and et al, ,"Melanocortin-1 receptor, skin cancer and phenotypic characteristics (M-SKIP) project: Study design and methods for pooling results of genetic epidemiological studies." BMC Medical Research Methodology.12,. 116. (2012).
http://digitalcommons.wustl.edu/open_access_pubs/1225

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.

Raimondi et al. BMC Medical Research Methodology 2012, 12:116
http://www.biomedcentral.com/1471-2288/12/116

BMC
Medical Research Methodology

STUDY PROTOCOL Open Access

Melanocortin-*1* receptor, skin cancer and phenotypic characteristics (M-SKIP) project: study design and methods for pooling results of genetic epidemiological studies

Sara Raimondi[1,2*], Sara Gandini[1], Maria Concetta Fargnoli[3], Vincenzo Bagnardi[1,4], Patrick Maisonneuve[1], Claudia Specchia[5], Rajiv Kumar[6], Eduardo Nagore[7], Jiali Han[8,9,10], Johan Hansson[11], Peter A Kanetsky[12], Paola Ghiorzo[13], Nelleke A Gruis[14], Terry Dwyer[15], Leigh Blizzard[16], Ricardo Fernandez-de-Misa[17], Wojciech Branicki[18], Tadeusz Debniak[19], Niels Morling[20], Maria Teresa Landi[21], Giuseppe Palmieri[22], Gloria Ribas[23], Alexander Stratigos[24], Lynn Cornelius[25], Tomonori Motokawa[26], Sumiko Anno[27], Per Helsing[28], Terence H Wong[29], Philippe Autier[30], José C García-Borrón[31], Julian Little[32], Julia Newton-Bishop[33], Francesco Sera[34], Fan Liu[35], Manfred Kayser[35], Tamar Nijsten[36] and GEM Study Group on behalf of the M-SKIP Study Group

Abstract

Background: For complex diseases like cancer, pooled-analysis of individual data represents a powerful tool to investigate the joint contribution of genetic, phenotypic and environmental factors to the development of a disease. Pooled-analysis of epidemiological studies has many advantages over meta-analysis, and preliminary results may be obtained faster and with lower costs than with prospective consortia.

Design and methods: Based on our experience with the study design of the Melanocortin-1 receptor (*MC1R*) gene, SKin cancer and Phenotypic characteristics (M-SKIP) project, we describe the most important steps in planning and conducting a pooled-analysis of genetic epidemiological studies. We then present the statistical analysis plan that we are going to apply, giving particular attention to methods of analysis recently proposed to account for between-study heterogeneity and to explore the joint contribution of genetic, phenotypic and environmental factors in the development of a disease. Within the M-SKIP project, data on 10,959 skin cancer cases and 14,785 controls from 31 international investigators were checked for quality and recoded for standardization. We first proposed to fit the aggregated data with random-effects logistic regression models. However, for the M-SKIP project, a two-stage analysis will be preferred to overcome the problem regarding the availability of different study covariates. The joint contribution of *MC1R* variants and phenotypic characteristics to skin cancer development will be studied via logic regression modeling.

Discussion: Methodological guidelines to correctly design and conduct pooled-analyses are needed to facilitate application of such methods, thus providing a better summary of the actual findings on specific fields.

Keywords: Genetic epidemiology, Melanoma, Meta-analysis, Pooled-analysis, Skin cancer, Study design

* Correspondence: sara.raimondi@ieo.it
[1]Division of Epidemiology and Biostatistics, European Institute of Oncology, Via Ramusio 1, Milan 20141, Italy
[2]Department of Occupational Health, University of Milan, Milan, Italy
Full list of author information is available at the end of the article

BioMed Central

© 2012 Raimondi et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background

Since millions of Single Nucleotide Polymorphisms (SNPs) were identified by the SNP Consortium [1], a growing number of studies have reported the association of SNPs in candidate genes with several diseases. However individual studies of typical size usually have low statistical power to find true associations given the polygenic nature of most common diseases, leaving alone the various forms of potential interactions between genetic, phenotypic and environmental factors. The advent of genome-wide association studies allowed genotyping of hundreds of thousands of SNPs across the genome on a usually large number of subjects, but information on a wide spectrum of epidemiological and lifestyle factors were seldom collected, although the role of these factors in complex diseases is undoubtedly crucial.

Meta-analysis of genetic epidemiological studies has been adopted to increase the power of smaller candidate gene studies by summarizing results from multiple studies. However the lack of access to individual data precludes in-depth investigations, including analyses of gene-gene and gene-environment interaction, and appropriate stratified analyses. This may potentially lead to false-positive or false-negative results, or biased magnitudes of associations, as previously pointed out [2].

Pooled-analysis of the primary data has been shown to have critical methodological advantages over meta-analysis [3,4] and has been applied successfully in the genetic epidemiology field [4-11]. Pooled-analysis uses standardized definitions of cases, outcomes and covariates, as well as the same analytical methods, thus limiting potential sources of heterogeneity across different studies. It also allows investigators to better control for confounding factors, evaluate alternative genetic models and estimate the joint effect of multiple genes. Finally, population-specific effect and gene-gene and gene-environment interactions may be better assessed using pooled-analysis [12]. The pooling of data from observational studies has become more common recently, and different approaches of data analysis have been applied [13]. Methodological guidelines to correctly design and conduct pooled-analyses are needed to facilitate application of such methods, thus providing a better summary of the actual findings on specific fields. Moreover, the awareness of the potential problems connected with the establishment of international collaborations and data pooling might help investigators to avoid or overcome them.

We describe here our experience with the study design of an international pooled-analysis on Melanocortin-1 receptor gene, SKin cancer and Phenotypic characteristics (M-SKIP project). In the first part of the paper, we explain the procedures that were used to identify studies and to collect and standardize data. In the second part we describe the statistical analysis plan that we are going to apply, giving particular attention to methods of analysis recently proposed to account for between-study heterogeneity and to explore the joint contribution of genetic, phenotypic and environmental factors in the development of a disease.

The M-SKIP project: rationale and aims

Melanocortin-1-receptor (*MC1R*, MIM#155555) is one of the major genes that determine skin pigmentation and it has been reported to be associated with risk of melanoma [14], possibly through the determination of the tanning response of skin to UV radiation [15-17]. However the relationship between some *MC1R* variants and melanoma also in darkly-pigmented European populations suggests that *MC1R* signaling may have an additional role in skin carcinogenesis beyond the UV-filtering differences between dark and fair skin [18]. In previous meta-analyses [14,19,20] authors found evidence of a significant association between melanoma, red hair and fair skin and the five *MC1R* variants R151C, R160W, D294H, D84E and R142H, and suggested a possible role in melanoma development, via non-pigmentary pathways, for I155T and R163Q variants. However, the specific contribution of each *MC1R* variant to melanoma development via pigmentary and non-pigmentary pathways could not be evaluated in meta-analyses due to the lack of individually joint information on *MC1R* variants and phenotypic characteristics.

The aim of the M-SKIP project is therefore to perform a pooled-analysis of individual data on sporadic skin cancer cases and controls with information on *MC1R* variants, in order to: 1) assess the association of *MC1R* variants with melanoma, basal cell carcinoma (BCC) and squamous cell carcinoma (SCC); 2) assess the association between *MC1R* variants and phenotypic characteristics, including hair and eye color, skin color, skin type, common and atypical nevi, freckles, and solar lentigines; and 3) perform stratified analyses on *MC1R* variants and skin cancer by phenotypic characteristics, and evaluate *MC1R*-phenotype interaction in skin cancer risk.

Data collection and creation of the standardized dataset

The identification of data sets and data collection

Published epidemiological studies on *MC1R* variants, melanoma, non-melanoma skin cancer (NMSC) and phenotypic characteristics associated with melanoma [21,22] were searched until April 2010 in the following databases: PubMed, ISI Web of Science (Science Citation Index Expanded) and Embase, using the keywords

"MC1R" and "melanocortin 1 receptor" alone and in combination with the terms "melanoma", "basalioma", "basal cell carcinoma", "squamous cell carcinoma", "skin cancer", "hair color", "skin color", "skin type", "eye color", "nevi", "freckles", and "solar lentigines", with no search restriction. The computer search was supplemented by consulting the bibliographies of the articles and reviews. We also tried to identify unpublished datasets by personal communication with participant investigators, members of the Advisory Committee, and with attendees of scientific meetings. Unpublished datasets were evaluated by an internal peer-review process before inclusion.

We selected papers according to the following inclusion criteria: 1) observational studies on single-primary sporadic skin cancer cases with information on any *MC1R* variant or 2) control series with information on any *MC1R* variant and at least one phenotypic characteristic under study. Permanent exclusion criteria were: 1) populations selected for *MC1R* status or for other genetic factors, 2) studies including only familial and/or multiple-primary melanoma cases, because we wanted to study *MC1R*-melanoma association at a population level, therefore excluding cases for whom the role of genetics is probably stronger. In the first step of the project, we also excluded genome-wide association studies (GWAS), because their different study design and genotyping methodology would significantly increase the heterogeneity of our data; however GWAS with epidemiological data would be included in a next step of the project and their results would be compared with those of classical genetic epidemiological studies.

The original search provided 748 papers, among them 111 were considered potentially interesting and full-text articles were retrieved and evaluated. We excluded 49 articles for the following reasons: duplicate populations (N = 20), no data on outcome (case/control status or any of the studied phenotypic characteristics) or on *MC1R* variants (N = 12), case reports, commentaries or reviews (N = 6), GWAS (N = 6), populations selected for genetic factors (N = 4) and multiple primary melanoma cases only (N = 1). The remaining 62 independent studies were considered eligible for inclusion in the pooled-analysis.

For each independent study, we identified the corresponding investigator and retrieved his/her contact information. Each investigator was invited to join the M-SKIP project: this required them to sign a participation form and a document attesting to approval of the study guidelines, and then to provide their data in electronic form without restrictions on format. A detailed list of variables relevant for skin cancer was provided and, for each available variable in the list, the authors were required to compile a form with a clear and complete description on how it was collected and coded. Investigators did not send any personal identifier with data, but only identification codes. Finally, investigators were asked to send a signed statement declaring that the original study was approved by an Ethic Committee and/or that study subjects provided a written consent to participate in the original study.

Data collection started in May 2009 and was closed in December 2010. During this period, 43 investigators were contacted and invited to share data. Thirty-one (72%) agreed to participate and provided data on 28,998 subjects, including 13,511 skin cancer cases (10,182 melanomas) and 15,477 controls from 37 independent published [19,23-62] and 2 unpublished studies. Both the unpublished datasets came from investigators who were originally contacted for their published data and who had further data of (still) unpublished studies. Among the 12 non-participant investigators, seven did not reply to our invitation letter, three were not able to retrieve the original dataset and two were not interested in the project. The total number of skin cancer cases and controls from the 25 independent studies [63-95] of non-participant investigators was 5,135 and 8,262, respectively. The study design was case–control for 13 studies, control-only for 11 studies, and case-only for one study.

Quality control, data coding and creation of the standardized dataset

We inspected the data for completeness and resolved inconsistencies with the investigator of each study. A number of subjects were excluded due to the following reasons: multiple-primary melanoma cases (N = 1596), missing data on *MC1R* variants (N = 1081), non-skin melanoma cases (N = 150), subjects with atypical mole syndrome and no skin cancer (N = 58), non first-primary melanoma cases (N = 24), familial melanoma cases, defined as subjects with two first-degree relatives or three or more any-degree relatives with melanoma (N = 25), other reasons including: unknown case/control status, duplicate subjects, or inappropriate controls (N = 232).

The following study-related variables were recoded uniformly: study country, study design, source of controls, application of case–control matching, methods to define phenotypic characteristics, genotyping methodology, whether genotyping was done in the same center for cases and controls and was blinded for case/control status, and DNA source. These variables were not used to assign a quality score to each study, but will be taken into account in meta-regression and sensitivity analyses. In addition, the variables listed in Table 1 were retrieved from each study if available,

Table 1 List of the main variables, number of original studies and related subjects per variable

Variable	Studies (%) N = 39	Melanoma cases (%) n = 7806	NMSC cases (%) n = 3151	Controls (%) n = 14875
Age	37 (95%)	7761 (99%)	3150 (100%)	14550 (98%)
Gender	39 (100%)	7801 (100%)	3151 (100%)	14853 (100%)
Ethnicity	38 (97%)	6770 (87%)	3142 (100%)	13833 (93%)
Body mass index	8 (21%)	557 (7%)	1380 (44%)	2226 (15%)
Smoking status	6 (15%)	2266 (29%)	419 (13%)	2286 (15%)
Intermittent sun exposure	21 (54%)	4493 (58%)	1266 (40%)	2286 (15%)
Continuous sun exposure	21 (54%)	4909 (62%)	741 (24%)	1938 (13%)
Sunburns	25 (64%)	4210 (54%)	1288 (41%)	2968 (20%)
Artificial UV exposure	16 (41%)	3842 (49%)	298 (9%)	1058 (7%)
Family history of skin cancer	27 (69%)	6660 (85%)	1289 (41%)	3318 (22%)
Family history of cancer other than skin	19 (49%)	4445 (57%)	371 (12%)	1630 (11%)
Melanoma body site	24 (62%)	6271 (80%)	NA	NA
Melanoma histology	19 (49%)	4868 (62%)	NA	NA
Breslow thickness	24 (62%)	5907 (76%)	NA	NA
Hair color	34 (87%)	6841 (88%)	2590 (82%)	11889 (80%)
Eye color	31 (79%)	5990 (77%)	2456 (78%)	10720 (72%)
Skin color	23 (59%)	3517 (45%)	826 (26%)	2963 (20%)
Skin type	31 (79%)	6590 (84%)	1992 (63%)	4540 (31%)
Common nevi	19 (49%)	3817 (49%)	442 (14%)	1181 (8%)
Atypical nevi	11 (28%)	2681 (34%)	642 (20%)	1447 (10%)
Freckles	21 (54%)	4028 (52%)	737 (23%)	2333 (16%)
Solar lentigines	6 (15%)	1419 (18%)	442 (14%)	1088 (7%)

NA = not applicable; *NMSC* = non melanoma skin cancer.

checked for quality, recoded in a standardized manner and entered in the main database. Finally, data on *MC1R* variants were entered for each subject. Quality controls and data coding were performed independently by two investigators, and inconsistencies were solved via consensus.

Some variables were collected in different ways in different studies. We report here as an example the rules we used to standardize sun exposure variables, in order to provide suggestions on how to recode variables with highly heterogeneous assessment among studies.

Intermittent and continuous sun exposure was coded as hours of exposure per day if the information was structured in this way. If not, we converted it to hours/day on a scale of 0 as no exposure and 6 as the maximum hours of exposure per day. For example, for datasets with four classes of exposure (never, seldom, often, always), we recoded the classes as 0, 2, 4, 6 hours/day. If individual sun exposure was collected over different time periods, we calculated the average exposure weighting for years of exposure in each time period. Other continuous variables (i.e. days of exposure per year, average hours of exposure per year) were converted to hours/day using the following algorithm:

1) calculate the variable mean on all the study subjects as:

$$\mu = \sum_{i=1}^{n} x_i / n \tag{1}$$

where x_i is the measure of the continuous variable on subject i, and n is the study sample size;

2) calculate the average hours of exposure/day (ν) over all the datasets with the variable coded (or recoded) in this way as in 1);
3) recode each observation basing on the proportion $x_i : \mu = \hat{x}_i : \nu$ as:

$$\hat{x}_i = \nu x_i / \mu \tag{2}$$

4) set as 6 (maximum hours of exposure per day) the value of all calculated values greater than 6.

The assumption underlying this coding was that the average sun exposure pattern for study subjects was

similar for different studies (and countries). Since we will use this variable only for confounding adjustment and/or effect modifier analyses, the purpose was to regroup subjects with a similar pattern of sun exposure, although the precise individual amount of sun exposure could not be estimated.

As a general rule, when a variable (i.e. common nevi count) was collected into classes, we recoded each class by using its median. The maximum numbers for open categories were chosen according to the available M-SKIP data.

Brief description of the collected data and statistical power

The final dataset was created in June 2011 and included data on 7,806 melanoma cases, 3,151 NMSC cases (2,211 BCC, 788 SCC and 152 with both), and 14,875 controls.

Distribution of data according to study country in which the study was performed is presented in Table 2. The majority of data came from Europe, especially from southern European populations. There was no significant difference in participation rate according to study area (Fisher exact test p-value: 0.25).

The main characteristics of the studies included in the M-SKIP database are described in Table 3. The majority are case–control studies (54%) with population or healthy controls and case–control matching. Phenotypic characteristics were frequently assessed by self-administered questionnaire (41%) or examination by a dermatologist or research nurse (36%). The majority of studies sequenced the entire coding region of the *MC1R* (67%) and used blood as DNA source (62%).

We calculated that the minimum required sample size to find a statistically significant association between a *MC1R* variant and melanoma assuming a similar association to that observed in our previous meta-analysis [14] (Odds Ratio (OR) = 1.5) is around 7,500 cases and 7,500 controls for rare variants (1-2% allele frequency in controls), and 1,400 cases and 1,400 controls for common variants (8-10% allele frequency in controls), with 90% statistical power. Sample size for gene-environment interaction analysis was also calculated with the program POWER, version 3.0 [96]. Considering the study of a simple two-way interaction between an environmental factor and a rare *MC1R* variant, around 5,000 cases and 5,000 controls would be needed to observe a multiplicative interaction effect of 2.0, arising to 16,000 cases and 16,000 controls to observe a smaller multiplicative effect of 1.5, both with 90% statistical power. For common *MC1R* variants, the same gene-environment interaction effects of 2.0 and 1.5 could be observed with around 1,200 cases and 1,200 controls, and with around 3,500 cases and 3,500 controls, respectively. Our sample size therefore is appropriate for the purpose of the analysis, and large enough to allow stratified and interaction analyses, especially to find even small interaction effects with the most frequent variants, and larger interaction effects for less common variants.

Statistical analysis plan

Appropriateness and representativeness of the collected data

Comparability of the main study population characteristics and results between studies included and excluded from the pooled-analysis will be assessed. Funnel plots to evaluate participation bias will be drawn and Egger's test [97] will be performed.

Departure of genotype frequencies of each *MC1R* variant from expectation under Hardy-Weinberg equilibrium will be assessed by Chi Square test among controls for each study, in order to detect any possible genotyping error or stratification problem in the datasets.

Table 2 Summary of data included in the M-SKIP project by geographical location

Study area	Invited investigators	Participant investigators (studies)	Melanoma cases	NMSC cases	Controls
Africa	1	0 (0)	0	0	0
Asia	3	2 (2)	0	0	345
Australia	4	2 (3)	744	298	290
Northern Europe[a]	8	6 (6)	858	1629	8095
Central Europe[b]	6	3 (4)	977	639	2398
Southern Europe[c]	9	8 (12)	2,747	0	2263
North America	13	11 (12)	2,480	585	1484
TOTAL	**43 [d]**	**31[d] (39)**	**7808**	**3151**	**14875**

NMSC = non melanoma skin cancer.
[a] includes Denmark, Norway, Sweden, The Netherlands, UK.
[b] includes France, Germany, Poland.
[c] includes Greece, Italy, Spain.
[d] one investigator collected data for two different areas (North America and Asia).

Table 3 Main characteristics of the included studies

	Studies (%) N = 39	Melanoma cases (%) n = 7806	NMSC cases (%) n = 3151	Controls (%) n = 14875
Study design				
Case–control	21 (54%)	5092 (65%)	2052 (65%)	6852 (46%)
Case only	11 (28%)	2646 (34%)	0	0
Control only	6 (15%)	0	0	1464 (10%)
Cohort	1 (3%)	68 (1%)	1099 (35%)	6559 (44%)
Source of controls				
Hospital	6 (21%)	509 (10%)	1169 (37%)	1847 (12%)
Population or healthy[a]	21 (75%)	4651 (90%)	1982 (63%)	12872 (87%)
Mixed	1 (4%)	0	0	156 (1%)
Case–control matching [b]				
No	10 (45%)	3151 (61%)	1739 (55%)	9578 (71%)
Yes	12 (55%)	2009 (39%)	1412 (45%)	3833 (29%)
Phenotype assessment				
Self-administered questionnaire	16 (41%)	2768 (35%)	672 (21%)	1875 (13%)
Examination by an expert	14 (36%)	3970 (51%)	1380 (44%)	4392 (30%)
Instrumental measure	2 (5%)	0	0	222 (1%)
Mixed	5 (13%)	297 (4%)	1099 (35%)	7247 (49%)
No measure	2 (5%)	771 (10%)	0	1139 (8%)
Genotype assessment				
Sequencing analysis	26 (67%)	5942 (76%)	1059 (34%)	4813 (32%)
Others[c]	13 (33%)	1864 (24%)	2092 (66%)	10062 (68%)
DNA source				
Blood	24 (62%)	4645 (60%)	2743 (87%)	13304 (89%)
Buccal cells	14 (36%)	3161 (40%)	408 (13%)	1326 (9%)
Tissue	1 (3%)	0	0	245 (2%)

NMSC = non melanoma skin cancer.
[a] healthy subjects are blood donors, friends or relatives of cases.
[b] individual or frequency.
[c] includes RFLP, SNaPshot, allele discrimination assay.

Combining data into a single dataset with random effects models

A first analysis using data combined into one dataset, is to fit them with logistic regression models with random slope. Considering a dominant model, let π_{ik}/X be the probability of skin cancer for subject i ($i = 1,\ldots,n_k$) of study k ($k = 1,\ldots,K$) conditional on the presence of a certain *MC1R* variant (X). We will account for the fixed *MC1R* effect and the random selection of studies, assuming a model that relates *MC1R* and study effects linearly to the logit of the probabilities:

$$\mathrm{logit}(\pi_{ik}/X) = \alpha + \beta X_{ik} + b_k X_{ik} \qquad (3)$$

In this model the transformed regression coefficient $\exp(\beta)$ is the odds of skin cancer for a subject with the *MC1R* variant compared with a subject without the *MC1R* variant, and the b_k are the study-specific coefficients accounting for the random selection of studies, with $b_k \sim N(0, \sigma_b^2)$, where σ_b^2 represents the between study variance of the *MC1R* effect.

The logistic regression model above described could be applied to different inheritance models and could include covariates, in order to adjust the studied associations by possible confounding factors. In order to include the available information from all the studies, missing values could be estimated in the model with multiple imputation and/or the creation of a missing-data indicator variable. However, when the majority of missing data are the results of non-availability of certain variables in some studies, as for the M-SKIP project, the use of both multiple imputation and the missing-data indicator would be

likely to introduce a bias in comparison with the complete case method [98,99] and a two-stage approach would be preferred.

Two-stage analysis with random effects models

The two-stage analysis method [100] will allow us to overcome the problem of the availability of different study covariates. The pooled-estimates of the association of *MC1R* variants with each skin cancer type and each phenotypic characteristic will be calculated as follows.

First, study-specific ORs will be calculated by applying logistic regression to the data from each study to estimate the odds of skin cancer conditional on the presence of a *MC1R* variant (X), controlling for confounders Z_k. For study k ($k = 1,\ldots,K$), assuming just one confounder, the model is written as:

$$\text{logit}(\pi_{ik}/X) = \alpha_k + \beta_k X_{ik} + \gamma_k Z_{ik} \tag{4}$$

where π_{ik} is the conditional probability of skin cancer for subject i ($i = 1,\ldots,n_k$) of study k. Although *MC1R* variants were uniformly defined across studies, the confounders Z_k may be specific to a particular study. Analyses with original covariates and with recoded data will be performed and compared. The exposure log-odds ratio for study k is denoted β_k, the confounding log-odds ratio is denoted γ_k, and the α_k are the study-specific intercepts. The β_k are assumed to vary across studies according to the second-stage model:

$$\beta_k = \beta + b_k + e_k \tag{5}$$

where β is the pooled-exposure log-odds ratio, b_k are random effects with $b_k \sim N(0, \sigma_b^2)$, where σ_b^2 represents the variability of the study-specific exposure effects β_k about the population mean β, and e_k are independent errors with $e_k \sim N(0, \sigma_k^2)$, where σ_k^2 describes the within-study variation of the β_k. In the first stage $\hat{\beta}_k$ and its variance $\hat{\sigma}_k 2$ are estimated from equation 4, separately for each study.

The two-stage estimator of the pooled exposure effect β is a weighted average of the $\hat{\beta}_k$, weighted by the inverse marginal variances of the $\hat{\beta}_k$, denoted $w_k = \left(\hat{\sigma}_k^2 + \sigma_b{}^2\right)^{-1}$. Thus:

$$\hat{\beta} = \left(\sum_k w_k \hat{\beta}_k\right) / \sum_k w_k \tag{6}$$

$$\text{var}\left(\hat{\beta}\right) = \left(\sum_k w_k\right)^{-1} \tag{7}$$

Two methods [100] are frequently used to estimate the random effects variance σ_b^2 in equations 6 and 7. These methods are pseudo-maximum likelihood and moment estimation.

Investigation of heterogeneity among studies

Homogeneity among the study estimates will be measured by Q statistic and I-Square [101], the latter representing the percentage of total variation across studies that is attributable to heterogeneity rather than to chance. Meta-regression analysis will be performed to investigate heterogeneity among study estimates, by evaluating the role of methodological characteristics of the studies and the characteristics of study populations.

Joint association of MC1R and phenotypic characteristics with skin cancer risk

Stratified analysis for the association of *MC1R* variants with each skin cancer type will be performed for different phenotypic characteristics. The hypothesis of homogeneity of ORs among strata will be verified using the Breslow-Day test [102].

In order to identify combinations of *MC1R* variants and phenotypic characteristics associated with each skin cancer type, we will perform logic regression, a recently proposed tree-based statistical method intended for binary predictors [103]. This approach is particularly useful for detecting subpopulations at high or low risk of disease, characterized by high-order interactions among covariates, and thus the methodology could be well applied to the study of complex diseases like cancer. First, we will dichotomize continuous and categorical variables by choosing appropriate thresholds and by creation of dummy variables. For phenotypic characteristics we will define dummy variables in order to 1) have as much differentiation as possible in hair and eye color, and 2) separate the extreme classes of skin type and common nevi count from intermediate classes, because it has been suggested [21] that in the meta- and pooled-analysis setting misclassification affects the intermediate classes of exposure more than the extreme ones. Let $X_1, X_2, \ldots, X_p$ be the binary predictors obtained by dichotomization, and let Y be the binary response variable (case/control status). The appropriate logic regression model can be written as:

$$\text{logit}(Y = 1 | X_1, X_2, \ldots, X_n) = \alpha + \sum_{j=1}^{k} \beta_j L_j \tag{8}$$

where L_j is a Boolean expression of the predictors X_i, such as $L_j = X_4^C \wedge \left(X_5 \vee X_1 \vee X_3^C\right)$ with $\wedge$ = AND, $\vee$ = OR and ^{C} = NOT. Logic regression could be generally applied to any type of regression outcome as long as the proper scoring function is specified. For the logic regression model in equation 8, the goal is to find the Boolean expressions L_j that minimize the binomial deviance, with the parameters β_j and the Boolean expressions L_j estimated simultaneously. The output from logic regression

is represented as a series of trees, one for each Boolean predictor L_j, and the associated regression coefficient. An example of a logic tree that may be applied to our pooled analysis is shown in Figure 1. Using this representation it is possible to start from a logic tree and obtain any other logic tree by a finite number of operations such as growing of branches, pruning of branches, changing of predictors and/or operators.

The searching for the optimum combinations of *MC1R* variants and phenotypic characteristics mostly associated with skin cancer will be undertaken by a (stochastic) simulated annealing algorithm [104-106]. This algorithm has a good chance to find a model that has the best or close to best possible score but, in the presence of noise in the data, typically overfits data. In order to select the best model, application of a combination of cross-validation and randomization tests has been suggested [104,105].

In an explanatory setting, at risk gene-phenotype combinations will be identified among a very large number of possible combinations by logic regression-based methods recently proposed [106,107]. The skin cancer risk of the identified subpopulations will be estimated within the pooled-analysis context using the two-stage analysis previously described.

Structural equation models will be also applied to eventually clarify the independent and dependent role of *MC1R* variants on skin cancer by phenotypic characteristics.

Finally, the role of environmental exposure will be investigated by entering new covariates in the models, by subgroup analyses and by studying gene-environment and phenotype-environment interactions using traditional and new proposed methodologies [108].

Use of MC1R data

In all the proposed analyses, each of the nine most frequently investigated *MC1R* variants (V60L, D84E, V92M, R142H, R151C, I155T, R160W, R163Q, D294H), as well as known rare mutations affecting *MC1R* function [109] will be evaluated assuming different inheritance models and choosing the one that fits the data best. Haplotype frequencies will be estimated using the iterative Expectation-Maximization algorithm [110,111], and their association with each skin cancer type and phenotypic characteristics will be evaluated. Moreover, for the studies that sequenced the entire gene, we will evaluate the impact on skin cancer and phenotypic characteristics of the total number of *MC1R* variant alleles and of the scores obtained from appropriate classification of *MC1R* variants [112].

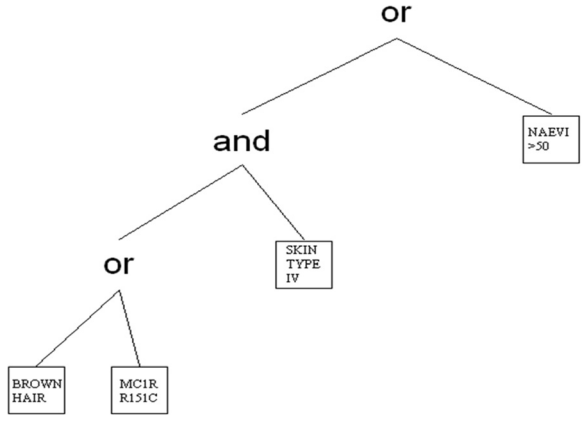

Figure 1 Example of a logic tree representing the Boolean expression "*more than 50 common naevi* V [*skin type IV* Λ (*MC1R R151C* V *brown hair*)]".

Discussion

Based on our experience with the study design of the M-SKIP project, we have described here the most important steps in planning, conducting and analyzing pooled individual data from genetic epidemiological studies. A previously published commentary highlighted the advantages and limitations of this kind of analyses, but did not describe the statistical methods that could be used to pool datasets [4]. Some methods for pooling results of epidemiological studies were suggested [10,100,112-114], but specific problems related to genetic epidemiology – such as the evaluation of different genotyping methodology, the Hardy-Weinberg equilibrium testing, the hereditary model assumption, and the assessment of gene-phenotype and gene-environment interaction – were not discussed.

Within the M-SKIP project, we collected a large amount of data in which multiple hypotheses can be examined with greater statistical power than is possible in individual studies. The response rate of invited investigators was high (72%), probably due to the well defined criteria of data collection and use, the clear publication policy, and the presence of an Advisory Committee tasked with monitoring adherence to project guidelines and scientific quality. Another strength of the pooled-analysis here described is the carefully-planned approach to standardizing the demographic, epidemiological and phenotypic information obtained from individual studies, giving the opportunity to perform appropriate and detailed subgroup and interaction analyses. Because the inclusion of an individual study in a particular analysis is not dependent on whether those investigators have published findings on that association, and because of inclusion of unpublished datasets, our pooled-analysis should not be affected by publication bias, as it might a meta-analysis of the published literature. Finally, we plan to analyze data by conventional and recently proposed statistical methods, and will compare and integrate the results obtained with these different approaches.

The main limitation of a pooled analysis, especially with respect to prospective consortia, is that it was planned retrospectively, and hence there was no *a priori* standardization of data collection. On the other hand, pooled-analysis may be feasible with fewer funds than those required for a prospective consortium, and it takes shorter time to obtain results because the original data have already been collected. The quality of genotype methodology may be heterogeneous among different participant laboratories. We will take into account this possible problem both by calculation of Hardy-Weinberg equilibrium and by meta-regression analysis. Finally, while we will try to assess the existence of participation bias, we cannot completely rule out that the results could be affected by the exclusions of the studies from the investigators who refused to participate in this pooled analysis.

In conclusion, the data collected within the M-SKIP project are a valuable resource for investigating associations between *MC1R* variants and skin cancer, particularly for population subgroups, and may be an appropriate setting to better investigate the genetics of sporadic skin cancer. A pooled-analysis of epidemiological studies is feasible, has many advantages over meta-analysis in making it possible to adjust for confounders and assess interactions, and in addition preliminary results may be obtained with lower costs and shorter time than with prospective consortia. We are convinced that its success depends upon the initial definition and approval of clear guidelines necessary for conducting such studies. The diffusion of pooled-analysis in genetic epidemiology field will assist epidemiologists and other health professionals in synthesizing the vast amount of available data on specific gene-disease associations and a common data-base would be the source of possible future investigations.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SR designed the study and directed its implementation, including quality assurance and control, wrote the paper and was the project leader. SG contributed to the design and supervision of the study and advised on recoding epidemiological data. MCF advised on recoding genetic data, is a member of the Advisory Committee of the project and contributed data to the study. VB, CS helped to carry out the statistical analysis plan of the study. PM contributed to the design and supervision of the study. RK, EN, JH, PG, NAG, TD, LB, RFdM, WB, TD, NM, GP, GR, AS, LC, TM, SA, PH, THW, FL, MK, TN and the GEM Study Group contributed data to the study. JH, PAK, MTL are members of the Advisory Committee of the project and contributed data to the study. PA, GCGB, JL, JNB, FS are members of the Advisory Committee of the project. All the members of the Advisory Committee were involved in the design and supervision of the study. All authors have revised and approved the manuscript.

Acknowledgements

The M-SKIP study was supported by the Italian Association for Cancer Research [MFAG 11831]. The GEM study was supported by National Cancer Institute [R01 CA112243, R01 CA112243-05 S1].

The M-SKIP study group consists of the following members: Principal Investigator: Sara Raimondi (European Institute of Oncology, Milan, Italy); Advisory Committee members: Philippe Autier (International Prevention Research Institute, Lyon, France), Maria Concetta Fargnoli (University of L'Aquila, Italy), José C. García-Borrón (University of Murcia, Spain), Jiali Han (Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA), Peter A. Kanetsky (Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA), Maria Teresa Landi (National Cancer Institute, NIH, Bethesda, MD, USA), Julian Little (University of Ottawa, Canada), Julia Newton-Bishop (University of Leeds, UK), Francesco Sera (UCL Institute of Child Health, London, UK); Consultants: Saverio Caini (ISPO, Florence, Italy), Sara Gandini and Patrick Maisonneuve (European Institute of Oncology, Milan, Italy); Participant Investigators: Albert Hofman, Manfred Kayser, Fan Liu, Tamar Nijsten and Andre G. Uitterlinden (Erasmus MC University Medical Center, Rotterdam, The Netherlands), Rajiv Kumar and Dominique Scherer (German Cancer Research Center, Heidelberg, Germany), Eduardo Nagore (Instituto Valenciano de Oncologia, Valencia, Spain), Johan Hansson and Veronica Hoiom (Karolinska Institutet, Stockholm, Sweden), Paola Ghiorzo and Lorenza Pastorino (University of Genoa, Italy), Nelleke A. Gruis (Leiden University Medical Center, The Netherlands), Terry Dwyer (Murdoch Childrens Research Institute, Victoria, Australia), Leight Blizzard and Jennifer Cochrane (Menzies Research Institute, Hobart, Australia), Ricardo Fernandez-de-Misa (Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain), Wojciech Branicki (Institute of Forensic Research, Krakow, Poland), Tadeusz Debniak (Pomeranian Medical University, Polabska, Poland), Niels Morling and Peter Johansen (University of Copenhagen, Denmark), Ruth Pfeiffer (National Cancer Institute, NIH, Bethesda, MD, USA), Giuseppe Palmieri (Istituto di Chimica Biomolecolare, CNR, Sassari, Italy), Gloria Ribas (Fundacion Investigation Hospital Clinico Universitario de Valencia- INCLIVA, Spain), Alexander Stratigos and Katerina Kypreou (University of Athens, Andreas Sygros Hospital, Athens, Greece), Anne Bowcock and Lynn Cornelius (Washington University, St. Louis, MO, USA), M. Laurin Council (St. Louis University, St. Louis, MO, USA), Tomonori Motokawa (POLA Chemical Industries, Yokohama, Japan), Sumiko Anno (Shibaura Institute of Technology, Tokyo, Japan), Per Helsing and Per Arne Andresen (Oslo University Hospital, Norway), Terence H. Wong (University of Edinburgh, UK), and the GEM Study Group.

Participants in the GEM Study Group are as follows: Coordinating Center, Memorial Sloan-Kettering Cancer Center, New York, NY, USA: Marianne Berwick (PI, currently at the University of New Mexico), Colin Begg (Co-PI), Irene Orlow (Co-Investigator), Urvi Mujumdar (Project Coordinator), Amanda Hummer (Biostatistician), Klaus Busam (Dermatopathologist), Pampa Roy (Laboratory Technician), Rebecca Canchola (Laboratory Technician), Brian Clas (Laboratory Technician), Javiar Cotignola (Laboratory Technician), Yvette Monroe (Interviewer). Study Centers: The University of Sydney and The Cancer Council New South Wales, Sydney (Australia): Bruce Armstrong (PI), Anne Kricker (co-PI), Melisa Litchfield (Study Coordinator). Menzies Centre for Population Health Research, University of Tasmania, Hobart (Australia): Terence Dwyer (PI), Paul Tucker (Dermatopathologist), Nicola Stephens (Study Coordinator). British Columbia Cancer Agency, Vancouver (Canada): Richard Gallagher (PI), Teresa Switzer (Coordinator). Cancer Care Ontario, Toronto (Canada): Loraine Marrett (PI), Beth Theis (Co-Investigator), Lynn From (Dermatopathologist), Noori Chowdhury (Coordinator), Louise Vanasse (Coordinator), Mark Purdue (Research Officer). David Northrup (Manager for CATI). Centro per la Prevenzione Oncologia Torino, Piemonte (Italy): Roberto Zanetti (PI), Stefano Rosso (Data Manager), Carlotta Sacerdote (Coordinator). University of California, Irvine (USA): Hoda Anton-Culver (PI), Nancy Leighton (Coordinator), Maureen Gildea (Data Manager). University of Michigan, Ann Arbor (USA): Stephen Gruber (PI), Joe Bonner (Data Manager), Joanne Jeter (Coordinator). New Jersey Department of Health and Senior Services, Trenton (USA): Judith Klotz (PI), Homer Wilcox (Co-PI), Helen Weiss (Coordinator). University of North Carolina, Chapel Hill (USA): Robert Millikan (PI), Nancy Thomas (Co-Investigator), Dianne Mattingly (Coordinator), Jon Player (Laboratory Technician), Chiu-Kit Tse (Data Analyst). University of Pennsylvania, Philadelphia, PA (USA): Timothy Rebbeck (PI), Peter Kanetsky (Co-Investigator), Amy Walker (Laboratory Technician), Saarene Panossian (Laboratory Technician). Consultants: Harvey Mohrenweiser, University of California, Irvine, Irvine, CA (USA); Richard Setlow, Brookhaven National Laboratory, Upton, NY (USA).

Author details

[1]Division of Epidemiology and Biostatistics, European Institute of Oncology, Via Ramusio 1, Milan 20141, Italy. [2]Department of Occupational Health,

University of Milan, Milan, Italy. [3]Department of Dermatology, University of L'Aquila, L'Aquila, Italy. [4]Department of Statistics, University of Milan Bicocca, Milan, Italy. [5]Department of Biomedical Sciences and Biotechnologies, University of Brescia, Brescia, Italy. [6]Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, Germany. [7]Department of Dermatology, Instituto Valenciano de Oncologia, Valencia, Spain. [8]Department of Dermatology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [9]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [10]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA. [11]Department of Oncology and Pathology, Cancer Center, Karolinska Institutet, Stockholm, Sweden. [12]Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [13]Department of Internal Medicine and Medical Specialties, University of Genoa, Genoa, Italy. [14]Department of Dermatology, Leiden University Medical Center, Leiden, The Netherlands. [15]Murdoch Childrens Research Institute, Royal Children's Hospital, Victoria, Australia. [16]Menzies Research Institute Tasmania, University of Tasmania, Hobart, Australia. [17]Servicio de Dermatologia, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain. [18]Institute of Forensic Research, Krakow, Polandfv. [19]Department of Genetic and Pathology, Pomeranian Medical University, Polabska, Poland. [20]Department of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark. [21]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA. [22]Unit of Cancer Genetics, Istituto di Chimica Biomolecolare, CNR, Sassari, Italy. [23]Dptd. Oncologia medica y hematologia, Fundacion Investigation Hospital Clinico Universitario de Valencia- INCLIVA, Valencia, Spain. [24]Department of Dermatology, University of Athens, Andreas Sygros Hospital, Athens, Greece. [25]Division of Dermatology, Washington University, St. Louis, MO, USA. [26]Dermatological R&D Skin Research Department, POLA Chemical Industries, Yokohama, Japan. [27]Shibaura Institute of Technology, Tokyo, Japan. [28]Department of Dermatology, Oslo University Hospital, Oslo, Norway. [29]Department of Dermatology, University of Edinburgh, Edinburgh, UK. [30]International Prevention Research Institute, Lyon, France. [31]Department of Biochemistry and Molecular Biology, University of Murcia, Murcia, Spain. [32]Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada. [33]Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, University of Leeds, Leeds, UK. [34]UCL Institute of Child Health, London, UK. [35]Department of Forensic Molecular Biology, Erasmus MC University Medical Center, Rotterdam, The Netherlands. [36]Department of Dermatology, Erasmus MC University Medical Center, Rotterdam, The Netherlands.

Received: 19 April 2012 Accepted: 23 July 2012
Published: 3 August 2012

References

1. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D: **A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.** *Nature* 2001, **409:**928–933.
2. Chatzinasiou F, Lill CM, Kypreou K, Stefanaki I, Nicolaou V, Spyrou G, Evangelou E, Roehr JT, Kodela E, Katsambas A, Tsao H, Ioannidis JP, Bertram L, Stratigos AJ: **Comprehensive field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma.** *J Natl Cancer Inst* 2011, **103:**1227–1235.
3. Seminara D, Khoury MJ, O'Brien TR, Manolio T, Gwinn ML, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, Bray MS, Brenchley PE, Buffler PA, Casas JP, Chokkalingam AP, Danesh J, Davey Smith G, Dolan S, Duncan R, Gruis NA, Hashibe M, Hunter D, Jarvelin MR, Malmer B, Maraganore DM, Newton-Bishop JA, Riboli E, Salanti G, Taioli E, Timpson N, Uitterlinden AG, Vineis P, Wareham N, Winn DM, Zimmern R, Ioannidis JP: **The emergence of networks in human genome epidemiology: challenges and opportunities.** *Epidemiology* 2007, **18:**1–8.
4. Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR: **Commentary: meta-analysis of individual participants' data in genetic epidemiology.** *Am J Epidemiol* 2002, **156:**204–210.
5. Ioannidis JP, Rosenberg PS, Goedert JJ, Ashton LJ, Benfield TL, Buchbinder SP, Coutinho RA, Eugen-Olsen J, Gallart T, Katzenstein TL, Kostrikis LG, Kuipers H, Louie LG, Mallal SA, Margolick JB, Martinez OP, Meyer L, Michael NL, Operskalski E, Pantaleo G, Rizzardi GP, Schuitemaker H, Sheppard HW, Stewart GJ, Theodorou ID, Ullum H, Vicenzi E, Vlahov D, Wilkinson D, Workman C, Zagury JF, O'Brien TR: **Effects of CCR5-Delta32, CCR2-64I, and SDF-1 3'A alleles on HIV-1 disease progression: an international meta-analysis of individual-patient data.** *Ann Intern Med* 2001, **135:**782–795.
6. Raimondi S, Paracchini V, Autrup H, Barros-Dios JM, Benhamou S, Boffetta P, Cote ML, Dialyna IA, Dolzan V, Filiberti R, Garte S, Hirvonen A, Husgafvel-Pursiainen K, Imyanitov EN, Kalina I, Kang D, Kiyohara C, Kohno T, Kremers P, Lan Q, London S, Povey AC, Rannug A, Reszka E, Risch A, Romkes M, Schneider J, Seow A, Shields PG, Sobti RC, Sørensen M, Spinola M, Spitz MR, Strange RC, Stucker I, Sugimura H, To-Figueras J, Tokudome S, Yang P, Yuan J-M, Warholm M, Taioli E: **Meta- and pooled analysis of GSTT1 and lung cancer: a HuGE-GSEC review.** *Am J Epidemiol* 2006, **164:**1027–1042.
7. Paracchini V, Raimondi S, Gram IT, Kang D, Kocabas NA, Kristensen VN, Li D, Parl FF, Rylander-Rudqvist T, Soucek P, Zheng W, Wedren S, Taioli E: **Meta- and pooled analyses of the cytochrome P-450 1B1 Val432Leu polymorphism and breast cancer: a HuGE-GSEC review.** *Am J Epidemiol* 2007, **165:**115–125.
8. Klug SJ, Ressing M, Koenig J, Abba MC, Agorastos T, Brenna SM, Ciotti M, Das BR, Del Mistro A, Dybikowska A, Giuliano AR, Gudleviciene Z, Gyllensten U, Haws AL, Helland A, Herrington CS, Hildesheim A, Humbey O, Jee SH, Kim JW, Madeleine MM, Menczer J, Ngan HY, Nishikawa A, Niwa Y, Pegoraro R, Pillai MR, Ranzani G, Rezza G, Rosenthal AN, Roychoudhury S, Saranath D, Schmitt VM, Sengupta S, Settheetham-Ishida W, Shirasawa H, Snijders PJ, Stoler MH, Suárez-Rincón AE, Szarka K, Tachezy R, Ueda M, van der Zee AG, von Knebel Doeberitz M, Wu MT, Yamashita T, Zehbe I, Blettner M: **TP53 codon 72 polymorphism and cervical cancer: a pooled analysis of individual data from 49 studies.** *Lancet Oncol* 2009, **10:**772–784.
9. Bracci PM, Skibola CF, Conde L, Halperin E, Lightfoot T, Smith A, Paynter RA, Skibola DR, Agana L, Roman E, Kane E, Wiencke JK: **Chemokine polymorphisms and lymphoma: a pooled analysis.** *Leuk Lymphoma* 2010, **51:**497–506.
10. Truong T, Hung RJ, Amos CI, Wu X, Bickeböller H, Rosenberger A, Sauter W, Illig T, Wichmann HE, Risch A, Dienemann H, Kaaks R, Yang P, Jiang R, Wiencke JK, Wrensch M, Hansen H, Kelsey KT, Matsuo K, Tajima K, Schwartz AG, Wenzlaff A, Seow A, Ying C, Staratschek-Jox A, Nürnberg P, Stoelben E, Wolf J, Lazarus P, Muscat JE, Gallagher CJ, Zienolddiny S, Haugen A, van der Heijden HF, Kiemeney LA, Isla D, Mayordomo JI, Rafnar T, Stefansson K, Zhang ZF, Chang SC, Kim JH, Hong YC, Duell EJ, Andrew AS, Lejbkowicz F, Rennert G, Müller H, Brenner H, Le Marchand L, Benhamou S, Bouchardy C, Teare MD, Xue X, McLaughlin J, Liu G, McKay JD, Brennan P, Spitz MR: **Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium.** *J Natl Cancer Inst* 2010, **102:**959–971.
11. Lurie G, Wilkens LR, Thompson PJ, Shvetsov YB, Matsuno RK, Carney ME, Palmieri RT, Wu AH, Pike MC, Pearce CL, Menon U, Gentry-Maharaj A, Gayther SA, Ramus SJ, Whittemore AS, McGuire V, Sieh W, Pharoah PD, Song H, Gronwald J, Jakubowska A, Cybulski C, Lubinski J, Schildkraut JM, Berchuck A, Krüger Kjær S, Høgdall E, Fasching PA, Beckmann MW, Ekici AB, Hein A, Chenevix-Trench G, Webb PM, Beesley J, Australian Ovarian Cancer Study Group, Goodman MT, Ovarian Cancer Association Consortium: **Estrogen receptor beta rs1271572 polymorphism and invasive ovarian carcinoma risk: pooled analysis within the Ovarian Cancer Association Consortium.** *PLoS One* 2011, **6:**e20703.
12. Boffetta P, Armstrong B, Linet M, Kasten C, Cozen W, Hartge P: **Consortia in Cancer Epidemiology: lessons from InterLymph.** *Cancer Epidemiol Biomarkers Prev* 2007, **16:**197–199.
13. Smith-Warner SA, Spiegelman D, Ritz J, Albanes D, Beeson WL, Bernstein L, Berrino F, van den Brandt PA, Buring JE, Cho E, Colditz GA, Folsom AR, Freudenheim JL, Giovannucci E, Goldbohm RA, Graham S, Harnack L, Horn-Ross PL, Krogh V, Leitzmann MF, McCullough ML, Miller AB, Rodriguez C, Rohan TE, Schatzkin A, Shore R, Virtanen M, Willett WC, Wolk A, Zeleniuch-Jacquotte A, Zhang SM, Hunter DJ: **Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer.** *Am J Epidemiol* 2006, **163:**1053–1064.
14. Raimondi S, Sera F, Gandini S, Iodice S, Caini S, Maisonneuve P, Fargnoli MC: **MC1R variants, melanoma and red hair color phenotype: a meta-analysis.** *Int J Cancer* 2008, **122:**2753–2760.

15. Kollias N, Sayre RM, Zeise L, Chedekel MR: **Photoprotection by melanin.** *J Photochem Photobiol B* 1991, **9**:135–160.
16. Burchill SA, Ito S, Thody AJ: **Effects of melanocyte-stimulating hormone on tyrosinase expression and melanin synthesis in hair follicular melanocytes of the mouse.** *J Endocrinol* 1993, **137**:189–195.
17. Rouzaud F, Kadekaro AL, Abdel-Malek ZA, Hearing VJ: **MC1R and the response of melanocytes to ultraviolet radiation.** *Mutat Res* 2005, **571**:133–152.
18. Gerstenblith MR, Goldstein AM, Fargnoli MC, Peris K, Landi MT: **Comprehensive evaluation of allele frequency differences of MC1R variants across populations.** *Hum Mutat* 2007, **28**:495–505.
19. Kanetsky PA, Panossian S, Elder DE, Guerry D, Ming ME, Schuchter L, Rebbeck TR: **Does MC1R genotype convey information about melanoma risk beyond risk phenotypes?** *Cancer* 2010, **116**:2416–2428.
20. Williams PF, Olsen CM, Hayward NK, Whiteman DC: **Melanocortin 1 receptor and risk of cutaneous melanoma: a meta-analysis and estimates of population burden.** *Int J Cancer* 2011, **129**:1730–1740.
21. Gandini S, Sera F, Cattaruzza MS, Pasquini P, Abeni D, Boyle P, Melchi CF: **Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi.** *Eur J Cancer* 2005, **41**:28–44.
22. Gandini S, Sera F, Cattaruzza MS, Pasquini P, Zanetti R, Masini C, Boyle P, Melchi CF: **Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors.** *Eur J Cancer* 2005, **41**:2040–2059.
23. Anno S, Abe T, Sairyo K, Kudo S, Yamamoto T, Ogata K, Goel VK: **Interactions between SNP alleles at multiple Loci and variation in skin pigmentation in 122 Caucasians.** *Evol Bioinform Online* 2007, **3**:169–178.
24. Anno S, Abe T, Yamamoto T: **Interactions between SNP alleles at multiple loci contribute to skin color differences between caucasoid and mongoloid subjects.** *Int J Biol Sci* 2008, **4**:81–86.
25. Kennedy C, ter Huurne J, Berkhout M, Gruis N, Bastiaens M, Bergman W, Willemze R, Bavinck JN: **Melanocortin 1 receptor (MC1R) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color.** *J Invest Dermatol* 2001, **117**:294–300.
26. Bastiaens M, ter Huurne J, Gruis N, Bergman W, Westendorp R, Vermeer BJ, Bouwes Bavinck JN: **The melanocortin-1-receptor gene is the major freckle gene.** *Hum Mol Genet* 2001, **10**:1701–1708.
27. Bastiaens MT, ter Huurne JA, Kielich C, Gruis NA, Westendorp RG, Vermeer BJ, Bavinck JN, Leiden Skin Cancer Study Team: **Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair.** *Am J Hum Genet* 2001, **68**:884–894.
28. Council LM, Gardner JM, Helms C, Liu Y, Cornelius LA, Bowcock AM: **Contribution of genetic factors for melanoma susceptibility in sporadic US melanoma patients.** *Exp Dermatol* 2009, **18**:485–487.
29. Pastorino L, Cusano R, Bruno W, Lantieri F, Origone P, Barile M, Gliori S, Shepherd GA, Sturm RA, Bianchi-Scarra G: **Novel MC1R variants in Ligurian melanoma patients and controls.** *Hum Mutat* 2004, **24**:103.
30. Pastorino L, Bonelli L, Ghiorzo P, Queirolo P, Battistuzzi L, Balleari E, Nasti S, Gargiulo S, Gliori S, Savoia P, Abate Osella S, Bernengo MG, Bianchi Scarrà G: **CDKN2A mutations and MC1R variants in Italian patients with single or multiple primary melanoma.** *Pigment Cell Melanoma Res* 2008, **21**:700–709.
31. Ghiorzo P, Pastorino L, Pizzichetta MA, Bono R, Queirolo P, Talamini R, Annessi G, Bruno W, Nasti S, Gargiulo S, Battistuzzi L, Sini MC, Palmieri G, Scarrà GB: **CDKN2A and MC1R analysis in amelanotic and pigmented melanoma.** *Melanoma Res* 2009, **19**:142–145.
32. Dwyer T, Stankovich JM, Blizzard L, FitzGerald LM, Dickinson JL, Reilly A, Williamson J, Ashbolt R, Berwick M, Sale MM: **Does the addition of information on genotype improve prediction of the risk of melanoma and nonmelanoma skin cancer beyond that obtained from skin phenotype?** *Am J Epidemiol* 2004, **159**:826–833.
33. Branicki W, Brudnik U, Kupiec T, Wolanska-Nowak P, Wojas-Pelc A: **Determination of phenotype associated SNPs in the MC1R gene.** *J Forensic Sci* 2007, **52**:349–354.
34. Brudnik U, Branicki W, Wojas-Pelc A, Kanas P: **The contribution of melanocortin 1 receptor gene polymorphisms and the agouti signalling protein gene 8818A > G polymorphism to cutaneous melanoma and basal cell carcinoma in a Polish population.** *Exp Dermatol* 2009, **18**:167–174.
35. Branicki W, Brudnik U, Wojas-Pelc A: **Interactions between HERC2, OCA2 and MC1R may influence human pigmentation phenotype.** *Ann Hum Genet* 2009, **73**:160–170.
36. Debniak T, Scott R, Masojc B, Serrano-Fernandez P, Huzarski T, Byrski T, Debniak B, Gorski B, Cybulski C, Medrek K, Kurzawski G, van de Wetering T, Maleszka R, Kładny J, Lubinski J: **MC1R common variants, CDKN2A and their association with melanoma and breast cancer risk.** *Int J Cancer* 2006, **119**:2597–2602.
37. Gapska P, Scott RJ, Serrano-Fernandez P, Mirecka A, Rassoud I, Gorski B, Cybulski C, Huzarski T, Byrski T, Nagay L, Maleszka R, Sulikowski M, Lubinski J, Debniak T: **Vitamin D receptor variants and the malignant melanoma risk: a population-based study.** *Cancer Epidemiology* 2009, **33**:103–107.
38. Fargnoli MC, Chimenti S, Keller G, Höfler H, Peris K: **Identification of four novel melanocortin 1 receptor (MC1R) gene variants in a Mediterranean population.** *Hum Mutat* 2003, **21**:655.
39. Fargnoli MC, Spica T, Sera F, Pellacani G, Chiarugi A, Seidenari S, Carli P, Chimenti S, Peris K: **Re: MC1R, ASIP, and DNA repair in sporadic and familial melanoma in a Mediterranean population.** *J Natl Cancer Inst* 2006, **98**:144–145.
40. Fargnoli MC, Altobelli E, Keller G, Chimenti S, Höfler H, Peris K: **Contribution of melanocortin-1 receptor gene variants to sporadic cutaneous melanoma risk in a population in central Italy: a case–control study.** *Melanoma Res* 2006, **16**:175–182.
41. Fargnoli MC, Pike K, Pfeiffer RM, Tsang S, Rozenblum E, Munroe DJ, Golubeva Y, Calista D, Seidenari S, Massi D, Carli P, Bauer J, Elder DE, Bastian BC, Peris K, Landi MT: **MC1R variants increase risk of Melanomas harboring BRAF mutations.** *J Invest Dermatol* 2008, **128**:2485–2490.
42. Fernandez-de-Misa R, Hernández-Jimenez JG, Carretero Hernández G, Pérez-Méndez L, Aguirre-Jaime A, Flores C, Suárez Hernández J, Perera Molinero A, Claverie-Martín F: **The D84E variant of the alpha-MSH receptor 1 gene is associated with cutaneous malignant melanoma early onset.** *J Dermatol Sci* 2008, **52**:186–192.
43. Fernandez-de-Misa R, Pérez-Méndez LI, Hernández-Jiménez JG, Rodríguez Mdel C, Vilar MC, Suárez J, Claverie-Martín F: **Main pigmentary features and melanocortin 1 receptor (MC1R) gene polymorphisms in the population of the Canary Islands.** *Int J Dermatol* 2008, **47**:806–811.
44. Han J, Kraft P, Colditz GA, Wong J, Hunter DJ: **Melanocortin 1 receptor variants and skin cancer risk.** *Int J Cancer* 2006, **119**:1976–1984.
45. Nan H, Qureshi AA, Hunter DJ, Han J: **Interaction between p53 codon 72 polymorphism and melanocortin 1 receptor variants on suntan response and cutaneous melanoma risk.** *Br J Dermatol* 2008, **159**:314–321.
46. Nan H, Kraft P, Hunter DJ, Han J: **Genetic variants in pigmentation genes, pigmentary phenotypes, and risk of skin cancer in Caucasians.** *Int J Cancer* 2009, **125**:909–917.
47. Höiom V, Tuominen R, Käller M, Lindén D, Ahmadian A, Månsson-Brahme E, Egyhazi S, Sjöberg K, Lundeberg J, Hansson J: **MC1R variation and melanoma risk in the Swedish population in relation to clinical and pathological parameters.** *Pigment Cell Melanoma Res* 2009, **22**:196–204.
48. Kanetsky PA, Ge F, Najarian D, Swoyer J, Panossian S, Schuchter L, Holmes R, Guerry D, Rebbeck TR: **Assessment of polymorphic variants in the melanocortin-1 receptor gene with cutaneous pigmentation using an evolutionary approach.** *Cancer Epidemiol Biomarkers Prev* 2004, **13**:808–819.
49. Scherer D, Bermejo JL, Rudnai P, Gurzau E, Koppova K, Hemminki K, Kumar R: **MC1R variants associated susceptibility to basal cell carcinoma of skin: interaction with host factors and XRCC3 polymorphism.** *Int J Cancer* 2008, **122**:1787–1793.
50. Scherer D, Nagore E, Lorenzo Bermejo J, Figl A, Botella-Estrada R, Thirumaran RK, Angelini S, Hemminki K, Schadendorf D, Kumar R: **Melanocortin receptor 1 variants and melanoma risk: a study of two European populations (p NA).** *Int J Cancer* 2009, **125**:1868–1875.
51. Landi MT, Kanetsky PA, Tsang S, Gold B, Munroe D, Rebbeck T, Swoyer J, Ter-Minassian M, Hedayati M, Grossman L, Goldstein AM, Calista D, Pfeiffer RM: **MC1R, ASIP, and DNA repair in sporadic and familial melanoma in a Mediterranean population.** *J Natl Cancer Inst* 2005, **97**:998–1007.
52. Mengel-From J, Børsting C, Sanchez JJ, Eiberg H, Morling N: **Determination of cis/trans phase of variations in the MC1R gene with allele-specific PCR and single base extension.** *Electrophoresis* 2008, **29**:4780–4787.
53. Motokawa T, Kato T, Hashimoto Y, Katagiri T: **Effect of Val92Met and Arg163Gln variants of the MC1R gene on freckles and solar lentigines in Japanese.** *Pigment Cell Res* 2007, **20**:140–143.

54. Motokawa T, Kato T, Hashimoto Y, Takimoto H, Yamamoto H, Katagiri T: **Polymorphism patterns in the promoter region of the MC1R gene are associated with development of freckles and solar lentigines.** *J Invest Dermatol* 2008, **128:**1588–1591.
55. Fernandez L, Milne R, Bravo J, Lopez J, Aviles J, Longo M, Benitez J, Lazaro P, Ribas G: **MC1R: three novel variants identified in a malignant melanoma association study in the Spanish population.** *Carcinogenesis* 2007, **28:**1659–1664.
56. Stratigos AJ, Dimisianos G, Nikolaou V, Poulou M, Sypsa V, Stefanaki I, Papadopoulos O, Polydorou D, Plaka M, Christofidou E, Gogas H, Tsoutsos D, Kastana O, Antoniou C, Hatzakis A, Kanavakis E, Katsambas AD: **Melanocortin receptor-1 gene polymorphisms and the risk of cutaneous melanoma in a low-risk southern European population.** *J Invest Dermatol* 2006, **126:**1842–1849.
57. Stefanaki I, Stratigos AJ, Dimisianos G, Nikolaou V, Papadopoulos O, Polydorou D, Gogas H, Tsoutsos D, Panagiotou P, Kanavakis E, Antoniou C, Katsambas AD: **p53 codon 72 Pro homozygosity increases the risk of cutaneous melanoma in individuals with dark skin complexion and among noncarriers of melanocortin 1 receptor red hair variants.** *Br J Dermatol* 2007, **156:**357–362.
58. Dimisianos G, Stefanaki I, Nicolaou V, Sypsa V, Antoniou C, Poulou M, Papadopoulos O, Gogas H, Kanavakis E, Nicolaidou E, Katsambas AD, Stratigos AJ: **A study of a single variant allele (rs1426654) of the pigmentation-related gene SLC24A5 in Greek subjects.** *Exp Dermatol* 2009, **18:**175–177.
59. Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens AC, Kayser M: **Eye color and the prediction of complex phenotypes from genotypes.** *Curr Biol* 2009, **19:**R192–R193.
60. Wong TH, Jackson IJ, Rees JL: **The physiological and phenotypic determinants of human tanning measured as change in skin colour following a single dose of ultraviolet B radiation.** *Exp Dermatol* 2010, **19:**667–673.
61. Casula M, Muggiano A, Cossu A, Budroni M, Caracò C, Ascierto PA, Pagani E, Stanganelli I, Canzanella S, Sini MC, Palomba G, Palmieri G, The Italian. Melanoma Intergroup (IMI): **Role of key-regulator genes in melanoma susceptibility and pathogenesis among patients from South Italy.** *BMC Cancer* 2009, **9:**352.
62. Kanetsky PA, Rebbeck TR, Hummer AJ, Panossian S, Armstrong GK, Kricker A, Marrett LD, Millikan RC, Gruber SB, Culver HA, Zanetti R, Gallagher RP, Dwyer T, Busam K, From L, Mujumdar U, Wilcox H, Begg CB, Berwick M, GEM Study Group: **Population-based study of natural variation in the Melanocortin-1 Receptor Gene and Melanoma.** *Cancer Res* 2006, **66:**9330–9337.
63. Galore-Haskel G, Azizi E, Mohamdi H, Scope A, Chaudru V, Laitman Y, Barak F, Pavlotsky F, Demenais F, Friedman E: **MC1R variant alleles and malignant melanoma risk in Israel.** *Eur J Cancer* 2009, **45:**2015–2022.
64. Jiménez-Cervantes C, Germer S, González P, Sánchez J, Sánchez CO, García-Borrón JC: **Thr40 and Met122 are new partial loss-of-function natural mutations of the human melanocortin 1 receptor.** *FEBS Lett* 2001, **508:**44–48.
65. Rana BK, Hewett-Emmett D, Jin L, Chang BH, Sambuughin N, Lin M, Watkins S, Bamshad M, Jorde LB, Ramsay M, Jenkins T, Li WH: **High polymorphism at the human melanocortin 1 receptor locus.** *Genetics* 1999, **151:**1547–1557.
66. Akey JM, Wang H, Xiong M, Wu H, Liu W, Shriver MD, Jin L: **Interaction between the melanocortin-1 receptor and P genes contributes to inter-individual variation in skin pigmentation phenotypes in a Tibetan population.** *Hum Genet* 2001, **108:**516–520.
67. Valverde P, Healy E, Jackson I, Rees JL, Thody AJ: **Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans.** *Nat Genet* 1995, **11:**328–330.
68. Valverde P, Healy E, Sikkink S, Haldane F, Thody AJ, Carothers A, Jackson IJ, Rees JL: **The Asp84Glu variant of the melanocortin 1 receptor (MC1R) is associated with melanoma.** *Hum Mol Genet* 1996, **5:**1663–1666.
69. Smith R, Healy E, Siddiqui S, Flanagan N, Steijlen PM, Rosdahl I, Jacques JP, Rogers S, Turner R, Jackson IJ, Birch-Machin MA, Rees JL: **Melanocortin 1 receptor variants in an Irish population.** *J Invest Dermatol* 1998, **111:**119–122.
70. Healy E, Flannagan N, Ray A, Todd C, Jackson IJ, Matthews JN, Birch-Machin MA, Rees JL: **Melanocortin-1-receptor gene and sun sensitivity in individuals without red hair.** *Lancet* 2000, **355:**1072–1073.
71. Flanagan N, Healy E, Ray A, Philips S, Todd C, Jackson IJ, Birch-Machin MA, Rees JL: **Pleiotropic effects of the melanocortin 1 receptor (MC1R) gene on human pigmentation.** *Hum Mol Genet* 2000, **9:**2531–2537.
72. Naysmith L, Waterston K, Ha T, Flanagan N, Bisset Y, Ray A, Wakamatsu K, Ito S, Rees JL: **Quantitative measures of the effect of the melanocortin 1 receptor on human pigmentary status.** *J Invest Dermatol* 2004, **122:**423–428.
73. Mossner R, Anders N, Konig IR, Kruger U, Schmidt D, Berking C, Ziegler A, Brockmoller J, Kaiser R, Volkenandt M, Westphal GA, Reich K: **Variations of the melanocortin-1 receptor and the glutathione-S transferase T1 and M1 genes in cutaneous malignant melanoma.** *Arch Dermatol Res* 2007, **298:**371–379.
74. Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart CJ, McKeigue PM, Kittles RA: **Skin pigmentation, biogeographical ancestry and admixture mapping.** *Hum Genet* 2003, **112:**387–399.
75. Matichard E, Verpillat P, Meziani R, Gérard B, Descamps V, Legroux E, Burnouf M, Bertrand G, Bouscarat F, Archimbaud A, Picard C, Ollivaud L, Basset-Seguin N, Kerob D, Lanternier G, Lebbe C, Crickx B, Grandchamp B, Soufir N: **Melanocortin 1 receptor (MC1R) gene variants may increase the risk of melanoma in France independently of clinical risk factors and UV exposure.** *J Med Genet* 2004, **41:**e13.
76. Jannot AS, Meziani R, Bertrand G, Gérard B, Descamps V, Archimbaud A, Picard C, Ollivaud L, Basset-Seguin N, Kerob D, Lanternier G, Lebbe C, Saiag P, Crickx B, Clerget-Darpoux F, Grandchamp B, Soufir N, Melan Cohort: **Allele variations in the OCA2 gene (pink-eyed-dilution locus) are associated with genetic susceptibility to melanoma.** *Eur J Hum Genet* 2005, **13:**913–920.
77. Liboutet M, Portela M, Delestaing G, Vilmer C, Dupin N, Gorin I, Saiag P, Lebbe C, Kerob D, Dubertret L, Grandchamp B, Basset-Seguin N, Soufir N: **MC1R and PTCH gene polymorphism in French patients with basal cell carcinomas.** *J Invest Dermatol* 2006, **126:**1510–1517.
78. Guedj M, Bourillon A, Combadières C, Rodero M, Dieudé P, Descamps V, Dupin N, Wolkenstein P, Aegerter P, Lebbe C, Basset-Seguin N, Prum B, Saiag P, Grandchamp B, Soufir N, Melan Cohort Investigators: **Variants of the MATP/SLC45A2 gene are protective for melanoma in the French population.** *Hum Mutat* 2008, **29:**1154–1160.
79. Ichii-Jones F, Lear JT, Heagerty AH, Smith AG, Hutchinson PE, Osborne J, Bowers B, Jones PW, Davies E, Ollier WE, Thomson W, Yengi L, Bath J, Fryer AA, Strange RC: **Susceptibility to melanoma: influence of skin type and polymorphism in the melanocyte stimulating hormone receptor gene.** *J Invest Dermatol* 1998, **111:**218–221.
80. Clairmont A, Sies H, Ramachandran S, Lear JT, Smith AG, Bowers B, Jones PW, Fryer AA, Strange RC: **Association of NAD(P)H:quinone oxidoreductase (NQO1) null with numbers of basal cell carcinomas: use of a multivariate model to rank the relative importance of this polymorphism and those at other relevant loci.** *Carcinogenesis* 1999, **20:**1235–1240.
81. Strange RC, Ellison T, Ichii-Jones F, Bath J, Hoban P, Lear JT, Smith AG, Hutchinson PE, Osborne J, Bowers B, Jones PW, Fryer AA: **Cytochrome P450 CYP2D6 genotypes: association with hair colour, Breslow thickness and melanocyte stimulating hormone receptor alleles in patients with malignant melanoma.** *Pharmacogenetics* 1999, **9:**269–276.
82. Jones FI, Ramachandran S, Lear J, Smith A, Bowers B, Ollier WE, Jones P, Fryer AA, Strange RC: **The melanocyte stimulating hormone receptor polymorphism: association of the V92M and A294H alleles with basal cell carcinoma.** *Clin Chim Acta* 1999, **282:**125–134.
83. Ramachandran S, Fryer AA, Strange RC: **Genetic factors determining cutaneous basal cell carcinoma phenotype.** *Med Pediatr Oncol* 2001, **36:**559–563.
84. Box NF, Wyeth JR, O'Gorman LE, Martin NG, Sturm RA: **Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair.** *Hum Mol Genet* 1997, **6:**1891–1897.
85. Palmer JS, Duffy DL, Box NF, Aitken JF, O'Gorman LE, Green AC, Hayward NK, Martin NG, Sturm RA: **Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype?** *Am J Hum Genet* 2000, **66:**176–186.
86. Box NF, Duffy DL, Irving RE, Russell A, Chen W, Griffyths LR, Parsons PG, Green AC, Sturm RA: **Melanocortin-1 receptor genotype is a risk factor for basal and squamous cell carcinoma.** *J Invest Dermatol* 2001, **116:**224–229.

87. Voisey J, Box NF, van Daal A: **A polymorphism study of the human Agouti gene and its association with MC1R.** *Pigment Cell Res* 2001, **14:**264–267.
88. Duffy DL, Box NF, Chen W, Palmer JS, Montgomery GW, James MR, Hayward NK, Martin NG, Sturm RA: **Interactive effects of MC1R and OCA2 on melanoma risk phenotypes.** *Hum Mol Genet* 2004, **13:**447–461.
89. Beaumont KA, Shekar SN, Newton RA, James MR, Stow JL, Duffy DL, Sturm RA: **Receptor function, dominant negative activity and phenotype correlations for MC1R variant alleles.** *Hum Mol Genet* 2007, **16:**2249–2260.
90. Shekar SN, Duffy DL, Frudakis T, Montgomery GW, James MR, Sturm RA, Martin NG: **Spectrophotometric methods for quantifying pigmentation in human hair-influence of MC1R genotype and environment.** *Photochem Photobiol* 2008, **84:**719–726.
91. Duffy DL, Zhao ZZ, Sturm RC, Hayward KK, Martin NG, Montgomery GW: **Multiple pigmentation gene Polymorphisms account for a substantial proportion of risk of Cutaneous malignant Melanoma.** *J Investig Dermatol* 2010, **130:**520–528.
92. Binkley CJ, Beacham A, Neace W, Gregg RG, Liem EB, Sessler DI: **Genetic variations associated with red hair color and fear of dental pain, anxiety regarding dental care and avoidance of dental care.** *J Am Dent Assoc* 2009, **140:**896–905.
93. Hacker E, Hayward NK, Dumenil T, James MR, Whiteman DC: **The Association between MC1R Genotype and BRAF Mutation Status in Cutaneous Melanoma: findings from an Australian Population.** *J Investig Dermatol* 2010, **130:**241–248.
94. Latreille J, Ezzedine K, Elfakir A, Ambroisine L, Gardinier S, Galan P, Hercberg S, Gruber F, Rees J, Tschachler E, Guinot C: **MC1R Gene Polymorphism affects skin color and phenotypic features related to sun sensitivity in a population of French adult women.** *Photochem Photobiol* 2009, **85:**1451–1458.
95. Elfakir A, Ezzedine K, Latreille J, Ambroisine L, Jdid R, Galan P, Hercberg S, Gruber F, Malvy D, Tschachler E, Guinot C: **Functional MC1R-gene variants are associated with increased risk for severe photoaging of facial skin.** *J Invest Dermatol* 2010, **130:**1107–1115.
96. García-Closas M, Lubin JH: **Power and sample size calculations in case–control studies of gene-environment interactions: comments on different approaches.** *Am J Epidemiol* 1999, **149:**689–692.
97. Egger M, Davey Smith G, Schneider M, Minder C: **Bias in meta-analysis detected by a simple, graphical test.** *BMJ* 1997, **315:**629–634.
98. Miettinen OS: *Theoretical epidemiology*. New York, NY: John Wiley and Sons, Inc; 1985.
99. Huberman M, Langholz B: **Application of the missing-indicator method in matched case–control studies with incomplete data.** *Am J Epidemiol* 1999, **150:**1340–1345.
100. Stukel TA, Demidenko E, Dykes J, Karagas MR: **Two-stage methods for the analysis of pooled data.** *Stat Med* 2001, **20:**2115–2130.
101. Higgins JP, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** *Stat Med* 2002, **21:**1539–1558.
102. Breslow NE, Day NE: *The analysis of case–control studies. Statistical Methods in Cancer Research*, Volume 1. Lyons: IARC Scientific Publications; 1980.
103. Kooperberg C, Ruczinski I, LeBlanc ML, Hsu L: **Sequence analysis using logic regression.** *Genet Epidemiol* 2001, **21:**S626–S631.
104. Ruczinski I, Kooperberg C, LeBlanc M: **Logic regression.** *J Comput Graph Stat* 2003, **12:**475–511.
105. Kooperberg C, Ruczinski I: **Identifying interacting SNPs using Monte Carlo logic regression.** *Genet Epidemiol* 2005, **28:**157–170.
106. Schwender H, Ickstadt K: **Identification of SNP interactions using logic regression.** *Biostatistics* 2008, **9:**187–198.
107. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y, Miljkovic I, Rasmussen-Torvik L, Harris TB, Province MA, Borecki IB, Florez JC, Meigs JB, Cupples LA, Dupuis J: **Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP × environment regression coefficients.** *Genet Epidemiol* 2011, **35:**11–18.
108. Beaumont KA, Shekar SN, Cook AL, Duffy DL, Sturm RA: **Red Hair is the Null Phenotype of MC1R.** *Hum Mutat* 2008, **29:**E88–E94.
109. Excoffier L, Slatkin M: **Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population.** *Mol Biol Evol* 1995, **12:**921–927.
110. Hawley ME, Kidd KK: **HAPLO: a program using the EM algorithm to estimate frequencies of multi-site haplotypes.** *J Hered* 1995, **86:**409–411.
111. Davies JR, Randerson-Moor J, Kukalizch K, Harland M, Kumar R, Madhusudan S, Nagore E, Hansson J, Höiom V, Ghiorzo P, Gruis NA, Kanetsky PA, Wendt J, Pjanova D, Puig S, Saiag P, Schadendorf D, Soufir N, Okamoto I, Affleck P, García-Casado Z, Ogbah Z, Ozola A, Queirolo P, Sucker A, Barrett JH, van Doorn R, Bishop DT, Newton-Bishop J: **Inherited variants in the MC1R gene and survival from cutaneous melanoma: a BioGenoMEL study.** *Pig Cell and Melanoma Res* 2012, **25:**384–394.
112. Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG: **A multilevel model framework for meta-analysis of clinical trials with binary outcomes.** *Stat Med* 2000, **19:**3417–3432.
113. Fibrinogen Studies C, Jackson D, White I, Kostis JB, Wilson AC, Folsom AR, Wu K, Chambless L, Benderly M, Goldbourt U, Willeit J, Kiechl S, Yarnell JW, Sweetnam PM, Elwood PC, Cushman M, Psaty BM, Tracy RP, Tybjaerg-Hansen A, Haverkate F, de Maat MP, Thompson SG, Fowkes FG, Lee AJ, Smith FB, Salomaa V, Harald K, Rasi V, Vahtera E, Jousilahti P: **Systematically missing confounders in individual participant data meta-analysis of observational cohort studies.** *Stat Med* 2009, **28:**1218–1237.
114. Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, The Emerging Risk Factors Collaboration: **Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies.** *Int J Epidemiol* 2010, **39:**1345–1359.

doi:10.1186/1471-2288-12-116
Cite this article as: Raimondi *et al.*: Melanocortin-*1* receptor, skin cancer and phenotypic characteristics (M-SKIP) project: study design and methods for pooling results of genetic epidemiological studies. *BMC Medical Research Methodology* 2012 **12**:116.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central