Gene-asbestos interaction in malignant pleural mesothelioma susceptibility

TITLE: Gene-asbestos interaction in malignant pleural mesothelioma susceptibility.

Running: Gene-asbestos interaction in pleural mesothelioma.

Sara Tunesi[1,2]*, Daniela Ferrante[1], Dario Mirabelli[2,3], Silvano Andorno[1], Marta Betti[4], Giovanni Fiorito[5,6], Simonetta Guarrera[5,6], Elisabetta Casalone[5,6], Monica Neri[7], Donatella Ugolini[8], Stefano Bonassi[7], Giuseppe Matullo[§5,6], Irma Dianzani[§4], Corrado Magnani[§1,3].

[1] Department of Translational Medicine - Unit of Medical Statistics and Cancer Epidemiology, CPO Piemonte and University of Piemonte Orientale, Novara, Italy.
[2] Center for Cancer Epidemiology and Prevention, City of Health and Science Hospital, Torino, Italy
[3] Interdepartmental Center "G. Scansetti", University of Turin, Italy.
[4] Department of Health Sciences, University of Piemonte Orientale, Novara, Italy.
[5] Genomic Variation in Human Population and Complex Disease Unit, Human Genetics Foundation, Turin, Italy.
[6] Department of Medical Sciences, University of Turin, Turin, Italy.
[7] Unit of Clinical and Molecular Epidemiology, IRCCS San Raffaele Pisana, Rome, Italy.
[8] Department of Internal Medicine, University of Genoa and IRCSS AOU San Martino-IST-Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy.

* To whom correspondence should be addressed.
Department of Translational Medicine
University of Piemonte Orientale
Via Solaroli,17
28100 Novara, Italy
Tel: +39 03213732183; Fax: +39 03213733112; Email: sara.tunesi@med.unipmn.it

[§]*these authors jointly directed this work.*

Running title: GWAS interaction analysis in mesothelioma risk

The authors declare no conflict of interests.

Abbreviations: CI, confidence intervals; CVC, Cross Validation Consistency; GAI gene-asbestos interaction; GMDR, Generalized Multifactor Dimensionality Reduction; GWAS, genome-wide association study; HWE Hardy-Weinberg equilibrium; LRT likelihood ratio test; MPM, malignant pleural mesothelioma; OR, odds ratio; RERI, Relative Excess Risk due to Interaction; SI, Synergy Index; SNP, Single Nucleotide Polymorphism; TBA, Testing Balanced Accuracy;

Keywords: mesothelioma, asbestos, gene-environment interaction, GWAS

Abstract (MAX 250)

Asbestos exposure is the main risk factor for malignant pleural mesothelioma (MPM), a rare aggressive tumor. Nevertheless, on average less than 10% of subjects highly exposed to asbestos develop MPM, suggesting the possible involvement of other risk factors. To identify the genetic factors that may modulate the risk of MPM, we conducted a gene-environment interaction analysis including asbestos exposure and fifteen Single Nucleotide Polymorphisms (SNPs) previously identified through a genome-wide association study (GWAS; 370,000 genotyped SNPs, 5 million imputed SNPs) on Italian subjects. In the present study, we assessed gene-asbestos interaction (GAI) on MPM risk using Relative Excess Risk due to Interaction (RERI) and Synergy Index (SI) for additive interaction, and V index for multiplicative interaction. Generalized Multifactor Dimensionality Reduction (GMDR) analyses were also performed. Positive deviation from additive GAI was found for six SNPs (rs1508805, rs2501618, rs4701085, rs4290865, rs10519201, rs763271), and four of them (rs1508805, rs2501618, rs4701085, rs10519201) deviated also from multiplicative models. However, after Bonferroni correction only deviation form multiplicative model were still significant for rs1508805 and rs4701085. GMDR analysis showed a strong MPM risk due to asbestos exposure and suggested a possible synergistic effect between asbestos exposure and rs1508805, rs2501618, and rs5756444. Our results suggested that GAI may play an additional role on MPM susceptibility, given that asbestos exposure appears as the main risk factor.

Summary: our gene-wide interaction study of asbestos and mesothelioma, based on 392 cases and 367 controls, suggested that additive and multiplicative gene-environment interaction might be associated with an increased risk, in addition to asbestos exposure only.

Introduction

Malignant pleural mesothelioma (MPM) is a rare, aggressive tumor, characterized by treatment resistance and poor prognosis (1). Although rare in the past, MPM frequency increased, in relation to asbestos use (1). The only clearly established risk factors for MPM are exposure to asbestos and other asbestiform minerals such as erionite and, to a lesser extent, ionizing radiation for medical purposes (1,2).

A genetic component in the etiology of the disease (3) might in part explain the relative rarity of MPM also in heavily exposed cohorts (2), the reports of familial clustering (4–6) and the results of candidate-gene association studies (3,7).

Matullo *et al.* (8) identified 15 and 5 imputed genotyped SNPs associated to MPM in a GWAS on an Italian study sample of 407 MPM cases and 389 healthy controls, and concluded that genetic risk factors may play an additional role in the development of MPM (8). Cadby *et al.* (9) in a companion GWAS study on MPM observed other associated SNPs but failed to replicate results from Matullo *et al.* (8). They hypothesized the lack of replication might be explained by differences in population genetic structure, type of asbestos exposure, or different asbestos exposure assessment.

The present study further investigates the interactions between candidate SNPs (8) and asbestos exposure, and their effects in modulating MPM risk in Italian population.

Methods

Ethics statement

All MPM cases and controls included in the present paper gave written informed consent. This study was performed according to the principles of the Declaration of Helsinki and in agreement with ethical requirements. Approval was obtained from the Istituto Nazionale per la Ricerca sul Cancro Ethics Committee for the studies in Genoa and La Spezia, and from the Human Genetics Foundation (HuGeF) Ethics Committee for the studies in Casale Monferrato and Turin.

Study sample

We presented statistical analysis of interaction in the ~~study sample was the same as in the~~ GWAS study by Matullo *et al.* (8). The study sample ~~It~~ was composed of MPM cases and controls from cities

located in Northern Italy: Casale Monferrato and Turin in the Piedmont Region, and Genoa and La Spezia in the Liguria Region. Casale Monferrato sample was a population-based MPM case-control study (10), and included 241 MPM patients and 252 population controls. Turin sample was a hospital-based MPM case-control study (7), and consisted of 91 MPM patients and 56 controls (non-neoplastic and non-respiratory conditions). The hospital-based study in Genoa and La Spezia included 75 incident MPM cases (11), and 81 controls ~~including~~ (healthy subjects or patients hospitalized for non-neoplastic and non-respiratory conditions). All study subjects were of Italian origin and caucasican ethnicity. Criteria of eligibility also included: be residing in the study area at the time of diagnosis and pathological confirmation of the diagnosis (based on histology or cytology with immunohistochemical staining). For practical reasons, the study in Turin was limited to cases and controls admitted to the main metropolitan hospitals. Cases and controls were individually matched by age and gender.

After reviewing the individual occupational histories, collected during questionnaire-based personal interviews, asbestos exposure was classified by an expert (D.M.) as ''absent/unlikely'' (no acknowledged occupational or environmental exposure), ''low'' (low exposure probability, or definite exposure at low level), and ''high'' (definite and high exposure, corresponding in principle to asbestos-cement and asbestos-textile workers, insulators, shipyard workers and dockers and similar activities). Further details on the exposure assessment process were given in Magnani *et al.* (12).

SNP genotyping and genotyping quality controls

Whole-genome genotyping was performed ~~by~~ using a Human CNV370 - Quad Bead Chip (Illumina Inc., San Diego, CA, USA) for 716 samples. The remaining 80 samples were tested on a Human610-Quad (which includes 100% of the HumanCNV370 Bead Chip SNPs) as the Human CNV370-Quad had been discontinued. Genotypes assignment was done by Genome Studio V 2011.1 (Illumina Inc., San Diego, CA). Five SNPs (rs2236304, rs7632718, rs9833191, rs10815216, rs73034881) were assessed by 5'nuclease assay (TaqMan Assay, Life Technologies Inc). We tested the compatibility of the two platforms measuring the minor allele frequencies in our controls and no significant differences were observed (data not shown).

Quality controls were conducted in the main GWAS (8) and are only summarized here. A cut-off a genotyping call rate of 0.98 was set, leading to the exclusion of 18 study subjects. SIdentity By Descent (IBD) estimation using the Identity By State (IBS) distance was used to check genotypic

Commento [ST1]: Capire questo pezzo se è corretto

identity or relatedness among subjects (13). Subjects with IBD≥0.05 (n = 16) were considered consanguineous and excluded from further analyses. We additionally excluded three samples with an X chromosome inbreeding homozygosity estimate of about 0.5. In total, thirty-seven subjects (4.64%) were removed from the analysis, leaving 759 subjects (392 cases and 367 controls). SNPs with minor allele frequency, <1% (n = 15,252), those having > 0.05 missing genotypes (n = 11,535) and those deviating from Hardy-Weinberg equilibrium (HWE) in the control population (P ≤0.001, n = 1,157) were excluded from the analysis, for a final study data-set of 330,879 SNPs, which were analyzed for their potential association with MPM.

Population structure and SNPs selection

Fifteen SNPs (rs2236304, rs742109, rs1508805, rs2501618, rs4701085, rs4290865, rs9536579, rs7632718, rs9833191, rs3801094, rs7841347, rs10519201, rs5756444, rs10815216, rs73034881) resulted from the GWAS analyses (8). These SNPs were selected by logistic regression analysis on per allele additive model adjusting by age, gender, PCA cluster, center of recruitment and asbestos exposure level (8). Only genotyped SNPs were considered for the present study.

Statistical analysis

The relationship between SNP and asbestos exposure in MPM causation was analyzed by logistic regression method adjusting by age, gender, PCA cluster and center. A binary classification was used both for asbestos exposure (exposed vs non-exposed) and for genotypes. Subjects with exposure coded as "'absent/unlikely exposure" were considered as non-exposed, while subjects coded as "low" or "high" were considered as exposed.

~~Generally speaking,~~ MPM risk for a given SNP and asbestos exposure was expressed by $OR_{i,j}$ where the first index (i) indicated the asbestos exposure coded as 0 for non-exposed and 1 for exposed subjects; the second index (j) indicated the SNP variant coded as 0 for the major allele and 1 for subjects bearing one or two copies of the minor allele. Subjects non-exposed and homozygous for the major allele were considered as reference group, thus coding their MPM risk as $OR_{00}=1$.

Interaction was analyzed in respect to both additive and multiplicative models. Deviation from an additive model was explored as the Relative Excess Risk due to Interaction (RERI) and the Synergy Index (SI) (14). Confidence intervals (CI) were calculated by the delta method (15). RERI was defined

as $[OR_{11}-OR_{01}-OR_{10}+1]$ and SI as $[OR_{11}-1]/[(OR_{01}-1) + (OR_{10}-1)]$ where the $OR_{i,j}$ represents the odds ratio estimated using logistic regression adjusted by age, gender, PCA cluster, and center of recruitment.

Under the null hypothesis of no interaction, RERI is not significantly different from 0, whereas SI is not significantly different from 1. RERI > 1 indicates positive interaction and RERI < 0 negative interaction. SI > 1 means positive interaction and SI < 1 negative interaction. Synergistic interaction (positive interaction) implies that the combined action between two factors in an additive model is greater than the sum of individual effects. On the contrary, antagonistic interaction means that in the presence of two factors in an additive model, the action of one exposure variable reduces the effects of the other (14).

Deviations from a multiplicative model were explored by multivariable logistic regression models including: asbestos exposure, one SNP at time and the corresponding interaction term (SNP × exposure); models were adjusted for age, gender, PCA cluster and center of recruitment. P-values for multiplicative interaction were calculated by comparing the full model including a multiplicative interaction term to a reduced model without it, using the likelihood ratio test (LRT) (16).

The Multiplicativity index $V=OR_{11}/(OR_{01} \times OR_{10})$ (17) was also calculated. A value of 1 for V suggested a multiplicative joint effect, whereas values greater or lower than 1 indicated an interaction that is more or less than multiplicative, respectively.

The generalized multifactor dimensionally reduction (GMDR) (18) method was applied to analyze high order SNP-exposure interactions (GMDR v7, software obtained fromhttp://www.ssg.uab.edu/gmdr/). GMDR is considered a model-free method. Potential confounders (age, gender, PCA cluster, and center of recruitment) were included as covariates. The data were divided in 10 sets: 9 for training and 1 for testing. N (from 1 to 5) factors were selected from the training set and their combinations were represented in n-dimensional space. The GMDR classified each combination (multifactorial class) as ‘high risk’ or ‘low risk’, thus reduced the n-dimensional space to one dimensional with two levels. For each possible model size (one-factor, two-factor, etc.), the model with the lowest misclassification error was selected. “Leave one-out-cross-validation” was used to calculate the prediction error and to evaluate the predictive ability of the model to fit the test set. The result was a set of models, and the Testing Balanced Accuracy (TBA) and Cross Validation Consistency (CVC) indexes were used to

determinate the overall best model. TBA was calculated as (sensitivity+specificity)/2. ORs were computed by 'low risk' vs 'high risk' combinations. P values were determined by sign test, a robust non parametric test implemented in GMDR software (19). The model with higher TBA and CVC and P-value less of equal to 0.05 derived from sign test was considered as the best one.

HWE was reassessed for the 15 SNPs comparing the observed genotype frequencies with the expected frequencies using Chi-squared test with significance level at P<0.05.
Statistical significance was set at P<0.05, confidence intervals (CIs) were estimated at 95%.

Results

After standard GWAS quality control procedure, as reported in (8), we considered eligible for the statistical analyses 759 subjects, 367 MPM cases and 392 controls. The general characteristics of the sample are reported in Table I. For each SNP we performed the gene-environment interaction analysis considering only subjects with known asbestos exposure and genotype.
All the polimorfisms were in HWE at P<0.05, except rs2236304 (P=0.022) and rs9536579 (P=0.034) and rs10815216 (P=0.004). This may be due to the small size of the sample or to chance.

The baseline association of covariate with MPM onset showed statistically significant differences for age (P<0.001), PCA cluster (P=0.040) and centre (vs P>0.001) and multivariate analyses were adjusted for (Tab. II).

Stratified analysis

The joint effect of exposure and each of the candidate SNPs assessed by multivariate logistic regressions, controlling by possible confounders (age, PCA cluster, centre and sex), is shown in Table III.

For example, considering the SNP rs2236304 the risk of mesothelioma was increased (OR=3.34 95%CI 1.00 – 11.09) in non-exposed subjects carrying the minor allele compared to non-exposed and homozygous for the major allele (reference category), and it was also increased in exposed individuals homozygous for the major allele (OR=16.00 95%CI 5.45 – 46.98). The effect further increased (OR=28.87 95%CI 9.95 – 83.78) when both factors (exposure and minor allele) were present.

Evidence of synergistic interaction between the minor allele of this polymorphism and asbestos exposure is further evaluated by indicators presented in Table IV (see later).

Considering “non exposed-non carrying at risk alleles” as the reference group, ORs among “exposed-non carrying” subjects ranged from 4.56 (rs7632718) to 35.83 (rs5756444); among “exposed-carrying” subjects it ranged from 3.92 (rs10815216) to 28.87 (rs223604) (Table III).

Rs1508805, rs2501618, rs4701085, rs4290865, rs10519201 (SHC4) and rs7632718 (SLC74A14) showed a similar pattern: null effect or protection among non-exposed and risk increase among the exposed.

For rs5756444, in the absence of exposure, the minor allele triplicated the risk (OR 3.37, borderline statistical significance) while it showed no effect among the exposed (OR from 35.83 to 21.34 with overlapping CIs).

Analysis without adjusting for confounding (age, sex, PCA cluster and center) showed similar results (Supplementary Table S1).

Additive and multiplicative interaction

RERI, SI, V and the statistical significance of the interaction term in the multiplicative logistic regression model are reported in Table IV, for each SNP.

In respect to deviation from additivity, significant positive interaction between SNP and exposure was found for rs1508805, rs2501618 (*CEP350* gene), rs4701085, rs4290865, rs10519201 (*SHC4*), rs7632718 (*SLC74A14*) according to RERI and SI indexes. Rs73034881 (SDK1/FOXK1) showed borderline RERI and statistically significant SI, with negative interaction. Significant negative interaction between SNP and exposure was also found according to SI index for rs9536579, rs5756444, rs10815216, rs9833191 (THRB) and rs7303881 (SDK1/FOXK1). After accounting for multiple comparisons using Bonferroni correction, SI index was still statistically significant a part for eight SNPs (rs1508805, rs2501618, rs4701085, rs4290865, rs9536579, rs10519201, rs5756444, s9833191) but RERI was no longer statically significant for any SNPs.

Statistically significant deviation from the multiplicative model was observed for: rs1508805, rs2501618 (*CEP350*), rs4701085, rs10519201 (*SHC4*), rs575644, and rs10815216. Except for

rs575644 (V=0.18), all these deviation from multiplicative model indicated a more than multiplicative interaction (V>1) between SNP and exposure. After accounting for multiple comparison using Bonferroni correction of P= 0.003 (0.05/15) the interaction remaned statistically significant in multiplicative scale, for rs150885 and rs471085.

Rs1508805 (Table III) in absence of exposure the minor allele conferred protection (OR=0.19), whereas exposure doubled the risk (from 5.22 to 10.33).

~~Rs1508805, rs2501618, rs4701085, rs4290865, rs10519201 (SHC4) and rs7632718 (SLC74A14) showed a similar pattern: null effect or protection among non-exposed and risk increase among the exposed.~~

~~For rs5756444, in the absence of exposure, the minor allele triplicated the risk (OR 3.37, borderline statistical significance) while it showed no effect among the exposed (OR from 35.83 to 21.34 with overlapping CIs).~~

In summary (Table IV) six SNPs (rs1508805, rs2510618, rs4701085, rs10519201 (SHC4), rs5765444, rs10815216) showed deviation from multiplicative model. SNPs rs575644 and rs10815216 had a significant deviation from multiplicativity while the deviation from additivity was observed only according to SI index. Similar trends were observed in unadjusted association analyses (Supplementary Table S2).

GMDR

Table V shows the result obtained from the GMDR analysis for one to five-factors models adjusted by covariates; the best models for different sizes were reported. According to GMDR selection model (19), the best model is the one with maximum TBA, maximum CVC and P-value derived from sign test less or equal to 0.05. There was no single model with all of these characteristics. The model with maximum CVC (10/10), included exposure only, had the second highest TBA (63.39%), OR= 9.71 (95% CI 4.02 **–** 22.44) and sign test P-value <0.001. The model including SNP rs2501618, rs1508805, rs5756444 and exposure had the third CVC (5/10) and the first maximum TBA (64.45%), sign test P-value <0.001 and OR=6.64 (95% CI 3.60 **–** 11.30) (Table V).

~~We verified that the~~ ORs estimated by GMDR using a classification ('high risk' vs 'low risk') corresponded, as expected, to the ORs estimated by logistic regression using the same classification (results not shown).

Discussion

This is the first study systematically examining interactions between asbestos exposure and a set of candidate SNPs emerging from a GWAS on MPM. We considered for each SNP both additive and multiplicative interactions with asbestos exposure and some significant results were found.

The age differences between cases and control observed in Casale Monferrato and Genova studies were due to different participation of cases and controls invited to the study. We observed a lower participation of controls in older ages, in particular among women (7). The main analyses were always adjusted for age, gender, PCA cluster and center.

Interaction analysis is dependent on the selection of the joint effect of interest. In the absence of "a priori" knowledge and of theoretical reason for choosing either, both additive and multiplicative models were tested. Interaction on the additive scale is present when the joint effect of the two risk factors is different from the sum of the individual effects. Interaction on the multiplicative scale is characterized by joint effect of the two risk factors different from the product of the individual effects (14). In this study, deviation from additive model was assessed by RERI and SI indexes. All of the selected SNPs presented SI index with significant values, suggesting deviation from additivity. The RERI index, on the contrary, was more restrictive. Indeed, as noted by Assmann (20), SI is generally statistically more unstable than RERI, when estimated using ORs instead of relative risks, as in the present study.

When multiple comparison were considered using Bonferroni correction SI remained statistically significant for eight SNPs (rs1508805, rs2501618, rs4701085, rs4290865, rs9536579, rs10519201, rs5756444, s9833191) and multiplicative interaction for two SNPs (rs1508805 and rs10519201 (*SHC4*)). However, we do not entirely agree on the use of Bonferroni correction here, because all the SNPs were already selected under consideration of multiple comparison in the GWAS analysis (8) and therefore repeating the correction for the present interaction analysis is too conservative .

The GMDR analysis suggested that the major contribution to the development of MPM was due to asbestos exposure, even after consideration of the selected SNPs. In fact, the model with the highest

TBA value (a four-way model including: exposure, rs1508805, rs2501618, rs5756444) had a low value of CVC (5/10); the model including only asbestos exposure had an only slightly lover TBA value (64.45% vs 63.39%) but a much better CVC value (10/10), so it is the best one. Without adjusting for confounding variables the GMDR selected the same two models and including the same SNPs that were selected by the model adjusted by confounders (Supplementary Table S3)

The most of the selected SNPs from GMDR analysis were also selected by additive or multiplicative interaction analysis; however the results we obtained from the GMDR analysis indicated the preeminent role of asbestos exposure and offer limited support for an interaction between asbestos exposure and some variant alleles.

In the present study, we selected the significant SNPs from our published GWAS (8), as no other evidence was previously reported in the literature. In a recent publication, Cadby *et al*. (9) also investigated MPM risk with a genome-wide association study, but their findings were not replicated in our Italian sample, apart some evidence of replication in the *SDK1* gene region. Because the SNPs they detected were not included in our most significant fifteen SNPs, they were not considered in these analyses.

SNPs included in the present study have a limited 'a priori' association with MPM risk or other asbestos health effects.

The rs2501618, located in*CEP350* gene, and selected as deviating from both additive and multiplicative models, was found associated to atopy in a previous paper (21) studying potential candidate genes for asthma or atopy. In our work, rs2501618 reduced the MPM risk in non-exposed subjects but increased the risk in exposed subjects with a synergistic interaction between asbestos exposure and the minor allele.

Deviation from additive interaction was found for SNP rs10519201, located in *SHC4* gene. SI additive interaction index remind significant after Bonferroni correction. This SNP showed association with psychiatric illness (eating-disorder) in Boraska *et al*. (22).

Rs7632718 is located in *SLC7A14* (solute carrier family 7 member 14), which lies in 3q26.2 regions, that was one of the replicating regions in Cadby *et al*. (9). Although no link with MPM had been previously reported for *SLC7A14*, a chromosomal gain of this region has been described in MPM (23),

suggesting a possible involvement of other genes. Cadby et al (9) indicated *SDK1* and the region around this gene as most consistently associated with MPM risk in both Australian and Italian studies.

In the present analysis, although not statistically significant for all the interaction indexes, the rs73034881, located in *SDK1/FOXK1* region, is suggestive of negative (protective) additive interaction between variant allele and asbestos exposure. It is interesting to note that FOXK1 is an interactor of BAP1, whose deleterious mutations are responsible for a cancer prone syndrome that included mesothelioma in its phenotype.

In these analyses we found a possible interaction, both additive and multiplicative, between asbestos exposure and both rs2501618 (CEP350) and rs10519201 (SHC4). Interaction is also suggested by four-way GMDR interaction analysis that included exposure, rs2501618 (CEP350), rs10519201 (SHC4), and rs5756444. Although these genes have not been directly associated with MPM, their involvement in several cancer types has been described. CEP350 interacts directly with FGFR1 oncogene partner (FOP), a critical protein in the myeloproliferative disorders (24) and SHC4 has been reported to activate both Ras-dependent and Ras-independent migratory pathways in melanomas (25) Their involvement in cancer suggests a possible role in MPM pathogenesis, interacting with asbestos exposure.

We did not find association for other genes associated to MPM in literature results. We performed an additional analysis with the Variant Effect Predictor software (http://www.ensembl.org/info/docs/tools/vep/index.html) to determine the effect of the SNPs. Four variants (rs742109, rs9833191, rs10519201, rs5756444) resulted to be localized in regulatory regions suggesting putative functional consequences.

THRB and *MMP14* are reported as to be dysregulated in MPM (26,27). *THRB* encodes for thyroid hormone receptor beta (TRb), which could function as a tumor suppressor and *MMP14* (matrix metallopeptides 14) has been reported to influence overall survival in MPM cases (27) but we did not find any significant interaction with asbestos exposure in relation to MPM risk. *PVT1* (Pvt1 oncogenic (non-protein coding)) gene is involved in several types of cancer (28,29) but no significant interaction between asbestos and *PVT1*-rs7841347 was found.

None of the SNPs deviated from HWE showed a clear interaction with asbestos exposure in MPM development; although rs9536579 and rs2236304 showed deviation from the additive model for SI index, and rs10815216 showed borderline deviation from both the multiplicative model and the additive model for SI index, moreover after Bonferroni correction neither SI index nor V index were stratistically significant for rs1081521 and both rs2236304 showed SI not statistically significant.

Several limitation of the current study should be acknowledged. The statistical power is limited: the sample size is critical in general for all gene-environment interaction studies, and in particular for rare diseases such as MPM. The number of cases (with minor allele variant and not) exposed to asbestos is very limited, as the rule in MPM studies. This may influence the estimated ORs and their confidence intervals and increase the variability of interaction indexes. In order to better investigate the relationship between gene and asbestos exposure a larger sample size would be required. ~~As we selected significant SNPs from our GWAS, no further multiple comparison correction for GWAS threshold was performed.~~ Only SNPs previously selected by association study (8) were considered in this gene-environment analysis. We cannot exclude that other SNPs with weaker effect in MPM risk might interact with asbestos-exposure, but for the through investigation it will be necessary a replication study with huger numbers.

Finally, it is also possible that rare variants could contribute to gene-asbestos interaction but our GWAS did not take into account rare variants. The availability of methods for complete genome sequencing will allow to circumvent the problem linked to the identification of rare variants, whose involvement should be better investigated in future studies, in order to avoid missing potential associations.

Commento [ST2]: chiarire

In conclusion, our results give some suggestions on the combined contribution of genetic background in asbestos-related carcinogenesis of the pleura, indicating that genetic factors may interact with asbestos exposure on MPM risk with different modes of interaction. Asbestos however remains the major risk factor for MPM.

Although the interpretation of observed results is still unclear because of the limited *apriori* evidence of genetic factors modulating the effects of asbestos exposure, an independent replication of our results, together with functional data, could contribute to understand MPM physiopathology, and to better define the MPM risk profile of subjects at high level of asbestos' exposure.

Funding source

The present study was supported by Italian National Institute of Health - ISS, “Progetto Amianto” (to C.M.). Research on Malignant Mesothelioma, including the GWAS study, was supported by the Regione Piemonte, Ricerca Sanitaria Finalizzata 2007, 2008, 2009 (to I.D.); Regione Piemonte, Bando Scienze umane e sociali, Progetto RoPHS (to ID); Regione Piemonte, Ricerca Scientifica Applicata 2003 (to D.M., C.M.); Fondazione Buzzi Unicem Onlus 2007 (to I.D.); CIPE (to I.D.); AIRC (to I.D.) and Human Genetics Foundation - HuGeF (to G.M.). The Casale case-control study was supported by a grant from Regione Piemonte, Ricerca Sanitaria Finalizzata 2004 (to C.M.).

Acknowledgements

We wish to thank all of the patients and healthy controls that generously participated in the case control studies on MM.

Conflict of Interest Statement: D.M and C.M. acted as expert witnesses for the public prosecutor in criminal trials involving the liability of industrial managers for asbestos-related cancers among their employees.

References

1. Pinto, C. *et al.* (2013) Second Italian consensus conference on malignant pleural mesothelioma: state of the art and recommendations. *Cancer Treat. Rev.*, **39**, 328–39.

2. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans (2011) A review of Human carcinogens: *IARC Monogr. Eval. Carcinog. Risks to Humans Twelfth Ed. 2011 2012*, **ed. Lyon:**,

3. Neri, M. *et al.* (2008) Genetic susceptibility to malignant pleural mesothelioma and other asbestos-associated diseases. *Mutat. Res.*, **659**, 126–36.

4. Ascoli, V. *et al.* (2007) Mesothelioma in blood related subjects: report of 11 clusters among 1954 Italy cases and review of the literature. *Am. J. Ind. Med.*, **50**, 357–69.

5. Ugolini, D. *et al.* (2008) Genetic susceptibility to malignant mesothelioma and exposure to asbestos: the influence of the familial factor. *Mutat. Res.*, **658**, 162–71.

6. De Klerk, N. *et al.* (2013) Familial aggregation of malignant mesothelioma in former workers and residents of Wittenoom, Western Australia. *Int. J. Cancer*, **132**, 1423–8.

7. Betti, M. *et al.* (2011) XRCC1 and ERCC1 variants modify malignant mesothelioma risk: a case-control study. *Mutat. Res.*, **708**, 11–20.

8. Matullo, G. *et al.* (2013) Genetic variants associated with increased risk of malignant pleural mesothelioma: a genome-wide association study. *PLoS One*, **8**, e61253.

9. Cadby, G. *et al.* (2013) A genome-wide association study for malignant mesothelioma risk. *Lung Cancer*, **82**, 1–8.

10. Dianzani, I. *et al.* (2006) Polymorphisms in DNA repair genes as risk factors for asbestos-related malignant mesothelioma in a general population study. *Mutat. Res.*, **599**, 124–34.

11. Ugolini, D. *et al.* (2008) The CREST biorepository: a tool for molecular epidemiology and translational studies on malignant mesothelioma, lung cancer, and other respiratory tract diseases. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 3013–9.

12. Magnani, C. *et al.* (2001) Increased risk of malignant mesothelioma of the pleura after residential or domestic exposure to asbestos: a case-control study in Casale Monferrato, Italy. *Environ. Health Perspect.*, **109**, 915–9.

13. Purcell, S. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.

14. Rothman KJGS (1998) *Modern epidemiology*, Lippincott Williams and Wilkins.

15. Hosmer, D.W. *et al.* (1992) Confidence interval estimation of interaction. *Epidemiology*, **3**, 452–6.

16. Hosmer, D. *et al.* (1989) *Applied logistic regression*, John Wiley & Sons.

17. Lee, P.N. (2001) Relation between exposure to asbestos and smoking jointly and the risk of lung cancer. *Occup. Environ. Med.*, **58**, 145–53.

18. Lou, X.-Y. *et al.* (2008) A combinatorial approach to detecting gene-gene and gene-environment interactions in family studies. *Am. J. Hum. Genet.*, **83**, 457–67.

19. Lou, X.-Y. *et al.* (2007) A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am. J. Hum. Genet.*, **80**, 1125–37.

20. Assmann, S.F. *et al.* (1996) Confidence intervals for measures of interaction. *Epidemiology*, **7**, 286–90.

21. Castro-Giner, F. *et al.* (2009) A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS). *BMC Med. Genet.*, **10**, 128.

22. Boraska, V. *et al.* (2012) Genome-wide association analysis of eating disorder-related symptoms, behaviors, and personality traits. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.*, **159B**, 803–11.

23. Gray, S.G. *et al.* (2009) In arrayed ranks: array technology in the study of mesothelioma. *J. Thorac. Oncol.*, **4**, 411–25.

24. Yan, X. *et al.* (2006) A complex of two centrosomal proteins, CAP350 and FOP, cooperates with EB1 in microtubule anchoring. *Mol. Biol. Cell*, **17**, 634–44.

25. Pasini, L. *et al.* (2009) Melanoma: targeting signaling pathways and RaLP. *Expert Opin. Ther. Targets*, **13**, 93–104.

26. Melaiu, O. *et al.* A review of transcriptome studies combined with data mining reveals novel potential markers of malignant pleural mesothelioma. *Mutat. Res.*, **750**, 132–40.

27. Crispi, S. *et al.* (2009) Global gene expression profiling of human pleural mesotheliomas: identification of matrix metalloproteinase 14 (MMP-14) as potential tumour target. *PLoS One*, **4**, e7016.

28. Zeidler, R. *et al.* (1994) Breakpoints of Burkitt's lymphoma t(8;22) translocations map within a distance of 300 kb downstream of MYC. *Genes. Chromosomes Cancer*, **9**, 282–7.

29. Storlazzi, C.T. *et al.* (2004) Identification of a commonly amplified 4.3 Mb region with overexpression of C8FW, but not MYC in MYC-containing double minutes in myeloid malignancies. *Hum. Mol. Genet.*, **13**, 1479–85.