UNIVERSITA' DEGLI STUDI DI TORINO

Department of Medical Sciences

PhD Program in Biomedical Sciences and Oncology XXXIV cycle

Human Genetic Curriculum

Academic Year 2021/2022

EPIGENOMICS OF MESOTHELIOMA

computational methods in cancer research

Giovanni Cugliari

Tutor: Prof. Giuseppe Matullo

PhD coordinator: Prof. Emilio Hirsch

ABSTRACT

Malignant pleural mesothelioma (MPM) is a rare and aggressive neoplasm with median survival of 12 months. Patients are usually diagnosed when current treatments have limited benefits, highlighting the need for noninvasive tests aimed at an MPM risk assessment tool that might improve life expectancy. Three hundred asbestos-exposed subjects (163 MPM cases and 137 cancer-free controls), from the same geographical region in Italy, were recruited. The evaluation of asbestos exposure was conducted considering the frequency, the duration and the intensity of occupational, environmental and domestic exposure. A genome-wide methylation array was performed to identify novel blood DNA methylation (DNAm) markers of MPM.

The first study "New DNA methylation signals for Malignant Pleural Mesothelioma risk assessment" investigated DNA methylation differences in easily accessible white blood cells (WBCs) between malignant pleural mesothelioma (MPM) cases and asbestos-exposed cancer-free controls. Epigenome-wide analysis (EWAS) revealed 12 single-CpGs associated with the disease and two of these showed high statistical power (99%) and effect size (>0.05) after false discovery rate (FDR) multiple comparison corrections: (i) cg03546163 in *FKBP5* significantly hypomethylated in cases, and (ii) cg06633438 in *MLLT1* statistically hypermethylated in cases. Based on the interaction analysis, asbestos exposure and epigenetic profile together may improve MPM risk assessment. Receiver Operation Characteristics (ROC) for Case-Control Discrimination showed a significant increase in MPM discrimination when DNAm information was added in the model (baseline model, BM: asbestos exposure, age, gender and white blood cells); area under the curve, AUC = 0.75; BM + cg03546163 at *FKBP5*. AUC = 0.89, 2.1×10^{-7}; BM + cg06633438 at *MLLT1*. AUC = 0.89, 6.3×10^{-8}. Validation and replication procedures, considering independent sample size and a different DNAm analysis technique, confirmed the observed associations. Our

results suggest the potential application of DNAm profiles in blood to develop noninvasive tests for MPM risk assessment in asbestos-exposed subjects.

The second study "DNA methylation of *FKBP5* as predictor of overall survival in Malignant Pleural Mesothelioma" is the first one to investigate DNA methylation changes in WBCs from easily accessible peripheral blood as MPM survival biomarkers. Considering 12 months as a cut-off for OS, epigenome-wide association study (EWAS) revealed statistically significant (p value = 7.7E-09) OS-related differential methylation of a single-CpG (cg03546163), located in the 5'UTR region of *FKBP5* gene. The Cox proportional hazards regression model highlighted that cg03546163 is an independent marker of prognosis in MPM patients with a better performance than traditional inflammation-based scores such as lymphocyte-to-monocyte ratio (LMR). Cases with DNAm < 0.45 at the cg03546163 had significantly poor survival compared with those showing DNAm ≥ 0.45 (mean: 243 versus 534 days; p value< 0.001). Epigenetic changes at the *FKBP5* gene were robustly associated with OS in MPM cases. Our results showed that blood DNA methylation levels could be promising and dynamic prognostic biomarkers in MPM.

The third study "Relationship between aging and Malignant Pleural Mesothelioma: epigenetic clock analyses" was focused on the investigations of the interaction between epigenetic age acceleration measures and asbestos exposure in easily accessible white blood cells (WBCs) between malignant pleural mesothelioma (MPM) cases and asbestos-exposed cancer-free controls. Age is one of the strongest predictors of chronic disease and mortality. Aging denotes a multitude of processes at the cellular level, and biological responses to aging differ among people, having thus an important role when considering the relationship with other disease-related covariates. Using two previously established methylation-based "clocks" (proposed by S. Horvath), namely intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA), we defined biological age acceleration for each of three hundred asbestos-exposed subjects. First, we explained the relationship between asbestos exposure and disease; EEAA

showed that biological age acceleration was statistically significantly associated with increased asbestos exposure (Estimate = 0.704, 95% CI: 0.067, 1.475, P= 0.043). Then, statistically significant differences between cases and controls were found. Considering interaction effects we showed the OR increase in the group with higher median values of asbestos exposure and DNAm-aging respect to the others. Finally, about 10% of the total effect of the asbestos exposure on MPM was independently mediated by biological aging variation (mediation effect). As a second outcome, we demonstrated that the mean of the number of total stochastic epigenetic mutations (hypo and hyper) was higher in cases with respect to controls. In particular, hypo-stochastic epigenetic mutations showed a mean difference between cases and controls about three-fold higher than hyper-stochastic epigenetic mutations. Moreover, mean stochastic epigenetic mutations increase in relation to asbestos exposure in cases but not in controls. Lastly, the trend of stochastic epigenetic mutations seems to be in the opposite direction stratifying by the two aging-related measures (EEAA and IEAA).

The fourth study "Epigenomics of Malignant Pleural Mesothelioma: a structural equation modeling" analyzed the overall MPM path including single-CpGs, asbestos exposure and epigenetic age acceleration measures in easily accessible white blood cells (WBCs) between malignant pleural mesothelioma (MPM) cases and asbestos-exposed cancer-free controls. No study has shown the simultaneous effect combining all these predictors with the aim to test the epigenomic pathway using a statistical approach. Structural equation modeling (SEM) is a largely confirmatory, rather than exploratory, technique; it is used to determine whether a model is valid than to find a suitable model. Asbestos exposure levels were extracted considering a quantitative measure; DNAm profiles have been used as single CpGs and to compute WBCs estimation and biological age measures. The SEM showed that all ten relationships (4 four associations and 6 covariances) included in the graph model were statistically significant.

The fifth study "*FKBP5*, a modulator of stress responses involved in Malignant Mesothelioma: the link between stress and cancer" reviewed the *FKBP5* gene involvement

in Malignant Pleural Mesothelioma (MPM) with the aim to investigate and explain the potential link between stress and cancer related outcomes. Stress is a well-established risk factor for a lot of disease phenotypes, including cancer. The risk for stress-related disorders is shaped by complex interactions among multiple environmental stressors and many genes with small individual effects on expressed phenotypes.

The results suggest the potential application of DNAm profiles in blood to develop noninvasive tests for MPM risk assessment in asbestos-exposed subjects. The identification of simple and valuable prognostic markers for MPM will enable clinicians to select patients who are most likely to benefit from aggressive therapies and avoid subjecting non-responder patients to ineffective treatment. Furthermore, aging-acceleration measures and stochastic epigenetic mutations should be considered as a proxy of stress-related environmental phenotype associations. Overall results explain the potential use of a suite of peripheral blood DNA methylation measures to better characterize the MPM biological path. Computational methods such as epigenome-wide association study (EWAS), mediation, interaction, structural equation modeling (SEM) are fundamental methodological approaches to better characterize causality in epigenetics. This investigation will increase the knowledge about the epigenetics contribution (in detail, blood DNm) in MPM focused on the development of non-invasive tests for asbestos-exposed subjects with the aim to monitor early detection indicators during the risk assessment and prognostic factors in MPM cases.

ABBREVIATIONS

AA: age acceleration

AUC: area under the curve

CpGs: cytosine-guanine dinucleotides

DNAm: DNA methylation

DNAm-age: DNA methylation age

EEAA: extrinsic epigenetic age acceleration

EWAS: epigenome-wide association study

FDR: false discovery rate

GLM: generalized regression models

GWAS: Genome-wide analysis

IEAA: intrinsic epigenetic age acceleration

LMR: lymphocyte-to-monocyte ratio

MPM: Malignant Pleural Mesothelioma

PCs: principal components

ROC curves: receiver operating characteristic curves

SEM: structural equation modeling

WBCs: white blood cells

CONTENTS

INTRODUCTION

Disease

Main characteristics

Malignant pleural mesothelioma (MPM) is an aggressive tumor. The disease usually develops after a long latency (20-40 years) following asbestos exposure [1]. The International Agency for Research on Cancer confirmed that all fibrous forms of asbestos are carcinogenic to humans, causing mainly mesothelioma, respiratory-tract tumors, mesothelioma, and cancer at other tissue sites [2].

Incidence and prevalence

Although MPM is considered a rare malignancy (prevalence 1–9/100,000), about 40,000 deaths have been estimated to occur each year globally [3, 4]. The World Health Organization estimates that 125 million people annually around the world are exposed to asbestos.

Risk factors

Asbestos is the principal carcinogen associated with MPM with an attributable risk of 80–90% [5]. In 1960 the first convincing epidemiologically evidence of a link between MPM and both occupational and environmental asbestos exposure was reported [6]. Later, many in vitro studies have demonstrated cytotoxic effects of asbestos fibers [7]. The use of asbestos began in the 1930s with the development of asbestos mines. Many developed industrialized countries have banned or restricted asbestos use in the past 30 years. Despite the ban on asbestos use, asbestos-containing construction materials still expose construction workers and civilians to asbestos, especially during demolition or reconstruction processes. At the same time, asbestos continues to be mined, exported and widely used in many nations including China, India and some Latin American countries. The disease usually develops after a long latency (20-40 years) following asbestos

exposure with risk increasing depending on duration and intensity of exposure [8]. Those at risk are the asbestos workers themselves and their family members. Male are more likely to get a mesothelioma than females [9]; cigarette smoking is not involved in mesothelioma development [10].

Pathogenesis

The rationale linked to the pathogenesis of MPM is that the asbestos fibers stick out from the lung surface and cause repeated cycles of scratching, damage, inflammation, and repair in the adjacent parietal mesothelial-cell layer [11, 12]. Inhaled asbestos (fibers longer than 5μm) is thought to arrive in the peritoneal cavity via the lymph nodes or entering by sputum or drinking water.

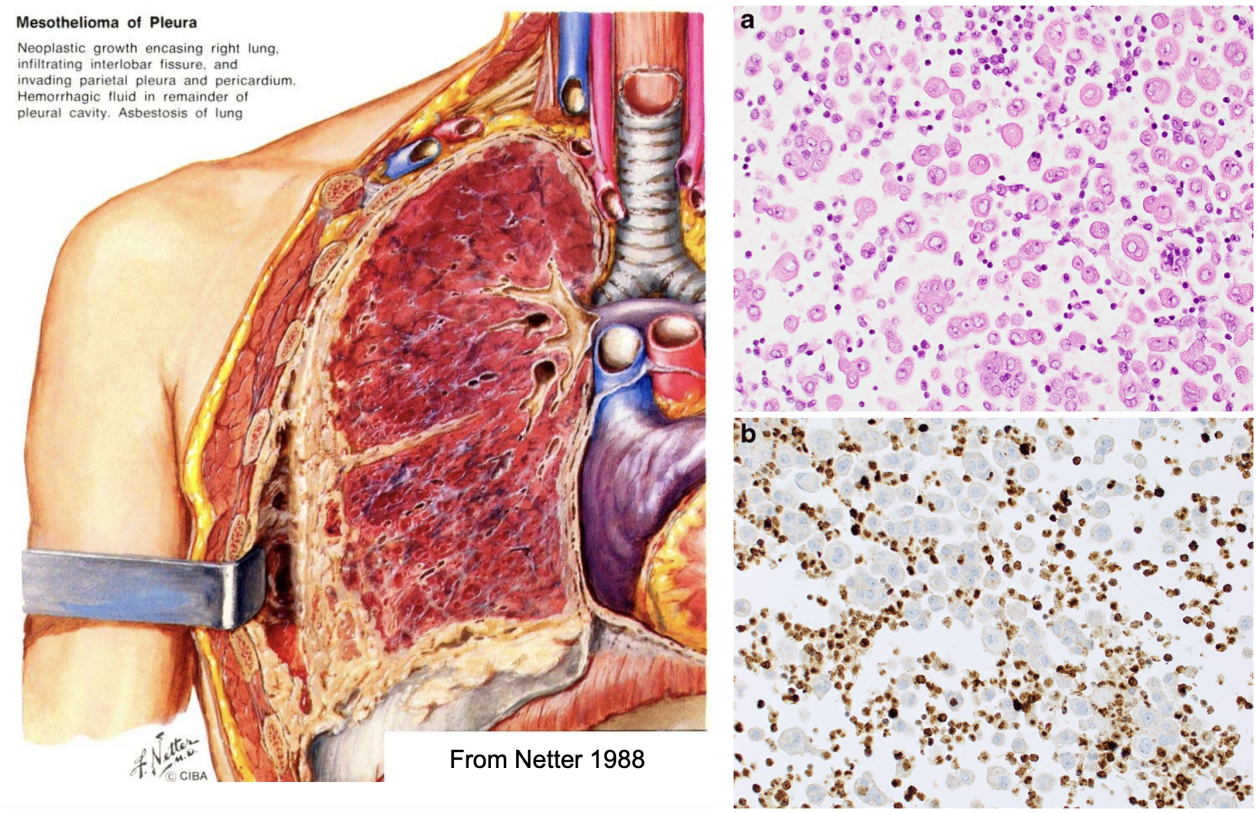

Figure 1. Netter, F. H. 1988. The CIBA Collection of Medical Illustrations, Volume 7: Respiratory System. CIBA-Geigy, Summit. (left). Malignant mesothelioma in pleural effusion sample. a Hypercellular specimen containing mostly mesothelial cells in clusters or 3D formations (H&E, × 20). b Loss of BAP1 by immunohistochemistry confirms a diagnosis of malignant mesothelioma (IHC, × 40) (right).

Phagocytosis of fibers by macrophages and oxidoreduction reactions on fiber surfaces are known to generate genotoxic reactive oxygen species that are capable of inducing DNA damage [13] and leading to genetic alterations in MPM [14] as for example the chromosomal aberrations [15, 16]. Recent research on the pathogenesis of MPM has indicated that both epigenetic and genetic alterations contribute to tumorigenesis. Asbestos affects mitochondrial DNA and functional electron transport resulting in mitochondrial-derived ROS, which has been shown to induce base oxidation, single-strand breaks, micronuclei, and apoptosis in lung alveolar epithelial cells [7, 18]. Lesions at sites of fiber deposition and alterations in gene expression are other relevant mechanisms in asbestos-induced neoplasia in lungs and other target organs [19].

Diagnosis

The combination of an unexplained pleural effusion and pleural pain should raise the suspicion of malignant mesothelioma, even if the initial cytologic findings are negative. Weight loss and fatigue are common later in the progression of pleural mesothelioma but are less so at presentation. Although a cytologic diagnosis can be made quickly, malignant mesothelioma is usually not diagnosed until two or three months after the onset of symptoms; delays of this length are especially frequent in centers in which the disease is uncommon. Mesothelioma is occasionally discovered incidentally on routine chest radiography. The most common presenting features in patients with peritoneal malignant mesothelioma are distention due to ascites, abdominal pain, and occasionally organ impairment, such as bowel obstruction.

Treatment

The patient with localized MPM was surgically cured and survived; however, the patient with diffuse MPM was ineligible for surgical treatment. Treatment ranges from chemotherapy or supportive care for advanced disease to aggressive surgery-based multimodality regimens for fit patients with limited disease. Available treatments have not proven their ability in significantly prolonging survival in comparison to supportive care. Currently, a randomized trial has demonstrated an increase of response rate and survival

when comparing cisplatin and pemetrexed versus cisplatin alone [20]. The first-line treatment is a combination of a multi targeted anti folate (pemetrexed or raltitrexed) and a platinum compound (cisplatin or carboplatin) [21, 22]. This regimen yields the best outcome in terms of median survival (12 months), median time to disease progression (6 months) and response rate (41%). There is no standard regimen for second line treatment, the chemotherapeutic agents used showed only marginal response rates [23-25]. Radiotherapy can be used to prevent tumor seeding along intervention sites, as adjuvant therapy after surgery or as palliative treatment. Other therapeutic strategies such as for example immunotherapy are promising but require further investigation and improvement. As other cancers, genesis of MPM is associated with genomic mutations but also epigenetic errors leading to modifications of gene expression [26, 27].

Prognosis

High mortality rate is associated with MPM. Although significant progress has been made in terms of etiology and pathogenesis, currently available therapeutic options have not significantly improved the survival outcome of patients on standard chemotherapeutic regimens [21]. The impact of chemotherapy on the outcome of patients with MPM is still controversial, the median survival being about 8-12 months [28-31]. Median survival duration post-diagnosis is less than 1 year. The roles of the new and old inflammatory indexes and markers in MPM prognosis were evaluated: increased age, stage 3–4 disease, the non-epithelial type, a low Karnofsky performance score, a high white blood cell count, and a low lymphocyte-to-monocyte ratio (LMR) were associated with a poor prognosis. The European Organization for Research and Treatment of Cancer (EORTC) and Cancer and Leukemia Group B (CALGB) devised two prognostic scoring systems for use in patients with mesothelioma and considered a combination of biological and clinical factors. However, these scoring systems are not routinely used for MPM prognosis because they are time-consuming and require costly equipment. Lastly, metastases are rarely the cause of death.

References

1. Sekido, Y. Molecular pathogenesis of malignant mesothelioma. *Carcinogenesis* 2013;34:1413-1419.
2. Straif, K; Benbrahim-Tallaa, L.; Baan, R.; et al. A review of human carcinogens--Part C: metals, arsenic, dusts, and fibers. *Lancet Oncol* 2009;10:453-454.
3. Rossini, M.; Rizzo, P; Bononi, I.; et al. New Perspectives on Diagnosis and Therapy of Malignant Pleural Mesothelioma. *Front Oncol* 2018;8:91.
4. Furuya, S.; Chimed-Ochir, O.; Takahashi, K.; et al. Global Asbestos Disaster. *Int J Environ Res Public Health* 2018;15.
5. McDonald, JC.; Armstrong, BG.; Edwards, CW.; Gibbs, AR.; Lloyd, HM.; Pooley, FD.; et al. Case-referent survey of young adults with mesothelioma: I. Lung fibre analyses. *Ann Occup Hyg*. 2001; 45: 513–518.
6. Wagner, JC.; Sleggs, CA.; Marchand, P. Diffuse pleural mesothelioma. *Br J Ind Med* 1960;17:260-71.
7. Kelsey, K.T.; et al. (1986) The in vitro genetic effects of fibrous erionite and crocidolite asbestos. *Br. J. Cancer*, 54, 107–114.
8. Bianchi, C.; Giarelli, L.; Grandi, G.; et al. Latency periods in asbestos-related mesothelioma of the pleura. *Eur J Cancer Prev* 1997;6(2):162–166.
9. Tischoff, I.; Neid, M.; Neumann, V.; Tannapfel A. Pathohistological diagnosis and differential diagnosis. *Recent Results Cancer Res* 2011;189:57–78.
10. Muscat, JE.; Wynder, EL. Cigarette smoking, asbestos exposure, and malignant mesothelioma. *Cancer Res* 1991;51(9):2263–7.
11. Xu, A.; Huang, X.; Lien, YC.; Bao, L.; Yu, Z.; Hei, TK. Genotoxic mechanisms of asbestos fibers: role of extranuclear targets. *Chem Res Toxicol* 2007, 20 (5): 724-733.
12. Choe, N.; Tanaka, S.; Kagan, E. Asbestos fibers and interleukin-1 upregulate the formation of reactive nitrogen species in rat pleural mesothelial cells. *Am J Respir Cell Mol Biol* 1998, 19 (2): 226-236.
13. Xu, A.; et al. Role of oxyradicals in mutagenicity and DNA damage induced by crocidolite asbestos in mammalian cells. *Cancer Res 1999* 59, 5922–5926.

14. Ault, JG.; Cole, RW.; Jensen, CG.; Jensen, LC.; Bachert, LA.; Rieder, CL. Behavior of crocidolite asbestos during mitosis in living vertebrate lung epithelial cells. *Cancer Res* 1995, 55 (4): 792-798.
15. Lee, WC.; Testa, JR. Somatic genetic alterations in human malignant mesothelioma. *Int J Oncol* 1999;14:181–8.
16. Murthy, SS.; Testa, JR. Asbestos, chromosomal deletions, and tumor suppressor gene alterations in human malignant mesothelioma. *J Cell Physiol* 1999;180:150–7.
17. Panduri, V.; Weitzman, SA.; Chandel, NS.; Kamp, DW. Mitochondrial-derived free radicals mediate asbestos-induced alveolar epithelial cell apoptosis. *Am J Physiol Lung Cell Mol Physiol* 2004, 286 (6): L1220-L1227.
18. Aljandali, A.; Pollack, H.; Yeldandi, A.; Li, Y.; Weitzman, SA.; Kamp, DW. Asbestos causes apoptosis in alveolar epithelial cells: role of iron-induced free radicals. *J Lab Clin Med* 2001, 137 (5): 330-339.
19. Gwinn, MR.; DeVoney, D.; Jarabek, AM.; Sonawane, B.; Wheeler, J.; Weissman, DN.; Masten, S.; Thompson, C. Meeting report: mode(s) of action of asbestos and related mineral fibers. *Environ Health Perspect* 2011, 119 (12): 1806-1810.
20. Netter, F. H. 1988. *The CIBA Collection of Medical Illustrations*, Volume 7: Respiratory System. CIBA-Geigy, Summit.
21. Wolf, AS.; Daniel, J.; Sugarbaker, DJ. Surgical techniques for multimodality treatment of malignant pleural mesothelioma: Extrapleural pneumonectomy and pleurectomy/decortication. *Semin Thorac Cardiovasc Surg* 2009;21(2):132–148.
22. Vogelzang, NJ.; Rusthoven, JJ.; Symanowski, J.; Denham, C.; Kaukel, E.; Ruffie, P.; et al. Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma. *J Clin Oncol* 2003;21(14):2636–44.
23. van Meerbeeck, JP.; Gaafar, R.; Manegold, C.; Van Klaveren, R.;, Van Marck, EA.; Vincent, M.; et al. Randomized phase III study of cisplatin with or without raltitrexed in patients with malignant pleural mesothelioma: an intergroup study of the European Organisation for Research and Treatment of Cancer Lung Cancer Group and the National Cancer Institute of Canada. *J Clin Oncol* 2005;23(28):6881–9.
24. Jassem, J.; Ramlau, R.; Santoro, A.; Schuette, W.; Chemaissani, A.; Hong, S.; et al. Phase IIItrial of pemetrexed plus best supportive care compared with best supportive care in

previously treated patients with advanced malignant pleural mesothelioma. *J Clin Oncol* 2008;26(10):1698–704.

25. Kindler, HL.; Millard, F.; Herndon, JE.; Vogelzang, NJ.; Suzuki, Y.; Green, MR. Gemcitabine for malignant mesothelioma: A phase II trial by the Cancer and Leukemia Group B. *Lung Cancer* 2001;31(2–3):311–7.
26. van Meerbeeck, JP.; Baas, P.; Debruyne, C.; Groen, HJ.; Manegold, C.; Ardizzoni, A.; et al. A Phase II study of gemcitabine in patients with malignant pleural mesothelioma. European Organization for Research and Treatment of Cancer Lung Cancer Cooperative Group. *Cancer* 1999;85(12):2577–82.
27. Gordon, GJ.; Jensen, RV.; Hsiao, LL.; Gullans, SR.; Blumenstock, JE.; Richards, WG.; et al. Using gene expression ratios to predict outcome among patients with mesothelioma. *J Natl Cancer Inst* 2003;95(8):598–605.
28. Gordon, GJ.; Rockwell, GN.; Jensen, RV.; Rheinwald, JG.; Glickman, JN.; Aronson, J.;, et al. Identification of novel candidate oncogenes and tumor suppressors in malignant pleural mesothelioma using largescale transcriptional profiling. *Am J Pathol* 2005;166(6):1827–40.
29. Pasello, G.; Ceresoli, GL.; Favaretto, A. An overview of neoadjuvant chemotherapy in the multimodality treatment of malignant pleural mesothelioma. *Cancer Treat Rev* 2012.
30. Tanrikulu, AC.; Abakay, A.; Kaplan, MA.; et al. A clinical, radiographic and laboratory evaluation of prognostic factors in 363 patients with malignant pleural mesothelioma. *Respiration* 2010;80:480–487.
31. Vogelzang, NJ. Malignant mesothelioma: diagnostic and management strategies for 1992. *Semin Oncol* 1992;19:64–71.
32. Schouwink, H.; Kors,e CM.; Bonfrer, JM.; et al. Prognostic value of the serum tumour markers Cyfra 21-1 and tissue polypeptide antigen in malignant mesothelioma. *Lung Cancer* 1999;25:25–32.

METHODOLOGICAL BACKGROUND

Epigenomics

Epigenomics focuses on genome-wide characterization of reversible modifications of DNA or DNA-associated proteins, such as DNA methylation or histone acetylation. Covalent modifications of DNA and histones are major regulators of gene transcription and subsequently of cellular fate.. Those modifications can be influenced both by genetic and environmental factors, can be long lasting, and are sometimes heritable.. While the role of epigenetic modifications as mediators of transgenerational environmental effects remains controversial, their importance in biological processes and disease development is evident from many epigenome-wide association studies that have been reported.

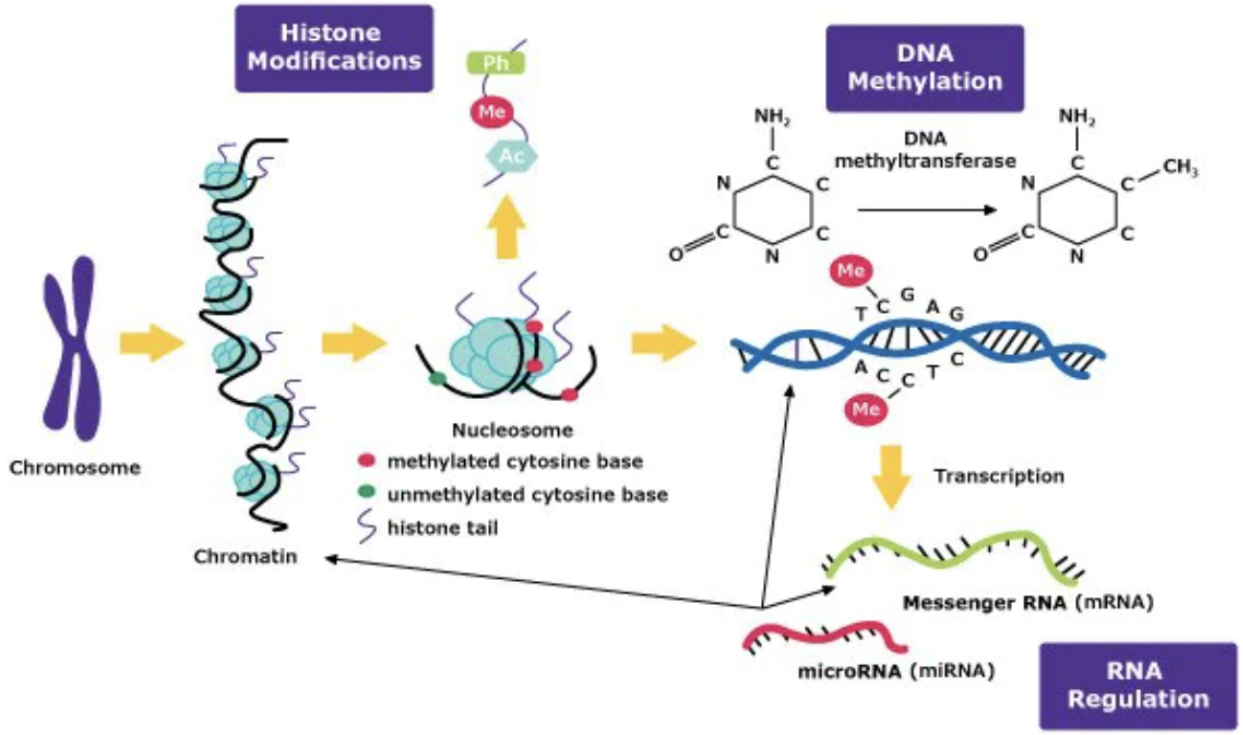

Figure 2. Epigenetics regulation.

Epigenetic signatures are often tissue-specific.. Thus, in addition to insight gained from identifying epigenetic modifications correlating with diseases, data generated by these studies has great potential to enhance our functional interpretation of genetic variants residing in those regions or of epigenetic markers associated with disease independently of genetic variation.

Presently, less information is available regarding mechanisms and clinical relevance of epigenetic derangements in cancer. Although genomic alterations are clearly associated with oncogenesis, more recent evidence shows that epigenetic modifications affect temporal and spatial control of gene activity required for homeostasis of complex organisms. By affecting gene activity, epigenetics also plays a major role during tumorigenesis.

DNA methylation (DNAm) is one of the epigenetic modifications with a key role in neoplasia and silencing repeat elements, which may also have an impact on disease etiology. DNAm patterns are specific to tissues and developmental stages, and they also change over time. The epigenome of a cell is highly dynamic, being governed by a complex interplay of genetic and environmental factors. Functionally, these perturbations may be involved, respectively, in tumor-suppressor genes silencing and genomic destabilization, crucial events in tumor development.

Distinguishing if epigenetic variation is causative or not of the disease process is not straightforward; this is a crucial step because it will help to elucidate the functional role of the disease-associated variation and its potential utility in terms of diagnosis or therapies.

Computational methods

Epigenome-wide association study (EWAS)

Generalized regression models (GLM) are very much like classical (Gaussian) linear models in most aspects, let aside the distribution of the error term and the expected value of the response, $\mu = E(Y)$, that is related to the linear predictor by means of a link function $\eta(\mu)$. Differential methylation (DMe) between cases and controls was tested by GLM at

single CpGs (EWAS, Epigenome-wide analysis) adjusted for the main predictors as gender, age, estimated WBCs, population stratification, and technical variability. For multiple comparisons tests, a Benjamini-Hochberg false discovery rate (FDR) p-value ≤ 0.05 was considered statistically significant. To ensure a power of the study greater than 99% (two-tailed test at 0.05 alpha error), only CpGs with mean difference (MD) of Beta-value between cases and controls ≥ |0.05|. Covariates were included step-by-step in sensitivity analysis to validate the association output considering effect size, standard error, 95% confidence interval and p-value variations. CpGs with MD between cases and controls ≥ |0.05| and Pfdr ≤ 0.05 were i) underwent gene set enrichment analysis to identify pathways potentially affected by MPM related methylation change; and ii) included in future statistical models to test for association. Then, we compared the area under the curve (AUC) for test receiver operating characteristic of two models by the DeLong test (R packages 'pROC' and 'ROCR'): this was done to test the potential improvement in disease discrimination when considering DMe information together with the main contributors information.

Parametric statistics make strong assumptions about the distribution: the distribution is normal, it's symmetric, the error term is independent of x (homoscedasticity), and there are no outliers. One way around this is to use non-parametric methods that make no assumptions about the underlying distribution. A common class of non-parametric methods are resampling methods, which allows the estimation of population variables via continual resampling of the empirical sample. Bootstrap is also an appropriate way to control and check the stability of the results. The number of resamples is the number of random samples with replacement from the original dataset. The sample size for each resample is equal to the sample size of the original dataset (for ex, n = 300). The number of resamples equals the number of observations on the histogram (for ex, B = 2000). The bias-corrected and accelerated (BCa) bootstrap interval was calculated with regard to single CpGs. The error of the estimates was also evaluated. Data were analyzed with R package 'bootstrap'. In the second step, we compared unsupervised clustering and conditional subgroups (quantile distribution) of single CpG. Gaussian finite mixture

model fitted by EM algorithm was implemented to highlight the classification components related to single-CpG (R package 'mclust').

Interaction

Most parametric tests require that residuals be normally distributed and that the residuals be homoscedastic. One approach when residuals fail to meet these conditions is to transform one or more variables to better follow a normal distribution. Often, just the dependent variable in a model will need to be transformed. For all the regression analyses, asbestos exposure doses (fibers/mL years) were rank transformed (R packages 'e1071' and 'GenABEL') to remove skewness. A two-sided test was used, with a 0.05 level of significance.

The statistical consideration of interaction effects in GLM, dichotomous choice models, is a well-developed tool in bioassay and epidemiology. Traditionally, the most frequent practice has been the analysis of simple main effects. Interactions are said to exist when a change in the level of one factor has different effects on the response variable, depending on the value of the other factor.

For example, The relationship between CpGs and asbestos exposure in MPM risk (odds ratio) was analyzed by logistic regression method adjusting by age, gender, SNP PCs, technical covariates and WBCs estimates. A binary classification was used both for asbestos exposure (above-median versus below-median) and for CpG methylation (above-median and below-median). MPM risk for a given CpG level and asbestos exposure was expressed by ORij, where the first index (i) indicated the asbestos exposure coded as below-median and above-median subjects; the second index (j) indicated the CpG coded as above-median and below-median; considering the direction of the effect, the same approach was used: regard to hypo-methylated CpGs above-median was the reference level; On the contrary, considering Hyper-methylated CpGs the reference level was below-median. Subjects with below-median asbestos exposure and reference level of CpG DNA methylation were considered as the baseline group, thus coding their MPM risk as OR00 = 1.

Interaction was analyzed in respect to both additive and multiplicative models based on the ORs obtained by logistic regression. Under the null hypothesis of no interaction under the additive model, OR (01,10,11) is not significantly different from 1. 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 reduces the effects of the other. Deviations from a multiplicative model were explored by multivariable logistic regression models including: asbestos exposure, CpG, the corresponding interaction term (CpG × exposure), adjusted by age, gender, SNP PCs, technical covariates and WBCs estimates. p-value < 0.05 was considered statistically significant. The same approach was considered using aging related measures with asbestos exposure to infer about the MPM OR.

Mediation

A mediator (M) is a variable that is on the causal path from an exposure (E) to an outcome variable (Y). Understanding mediation is useful for identifying potential modifiable risk factors lying between an exposure and an outcome including linear equations, structural equation models, marginal structural models and G-computation. Causal mediation analysis plays an essential role by helping to identify intermediate variables (or mediators) that lie in the causal pathway between the exposure and the outcome. Mediation (R package 'lavaan') aims to partition the total (causal) effect of E on Y into mediated effects (effects that operate by changing the mediator, M) and non-mediated effects.

In the MPM context, the rationale is: statistically significant CpGs (that show differential methylation between cases and controls in EWASs) are tested in subsequent analyses to investigate whether environmentally modified (asbestos exposure); age acceleration measures as mediators were also considered using the same path. Age, gender, SNP PCs, technical covariates and WBCs estimates were included in the statistical models to adjust for the potential confounding effect.

Structural equation modeling (SEM)

Path Diagrams play a fundamental role in structural equation modeling (SEM) as a device for showing which variables cause changes in other variables. Notice that, besides representing the linear equation relationships with arrows, the diagrams also contain some additional aspects. As the equation systems we examine become increasingly complicated, so do the covariance structures they imply. What causal modeling does allow us to do is examine the extent to which data fail to agree with one reasonably viable consequence of a model of causality. Although path diagrams can be used to represent causal flow in a system of variables, they need not imply such a causal flow. As such, they can convey linear relationships when no causal relations are assumed. SEM was implemented to simultaneously analyze any relationship previously performed (R package 'lavaan').

Survival analysis

The duration of survival and the median and mean event times with 95% confidence intervals were estimated according to the Kaplan–Meier method. The duration of survival was determined as the period between the time of diagnosis and the time of death (R package 'survival'). If patients were still alive, survival was defined as the period between the times of diagnosis until June 2016. The proportional hazards regression model was used for both the univariate and multivariate analyses (Cox proportional hazards model). Comparisons of overall survival were made using two-tailed log-rank tests with a 0.05 level of significance. Only variables with p values <0.1 in the univariate analysis were included in the final model for the multivariate analysis. In the Cox regression analysis, the backward conditional (stepwise-AIC, R package 'MASS') method was used. A p value <0.05 was considered to indicate statistical significance. LMR, CpG sites and age acceleration measures were considered as predictors in the regression model. Continued and categorical variables (quantile consideration) were also used.

Epigenomics biomarkers

Beta values

Genomic DNA was extracted from whole blood collected in EDTA by an on-column DNA purification method (QIAamp DNA Blood Mini Kit, QIAGEN GmbH, Germany), according to manufacturer's instructions. DNA integrity was checked by an electrophoretic run in standard TBE 0.5X buffer on a 1% low melting agarose gel (Sigma-Aldrich GmbH, Germany); DNA purity and concentration were assessed by a NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Five-hundred nanograms of genomic DNA for each sample were bisulfite treated (EZ-96 DNA Methylation-Gold Kit, Zymo Research Corporation, Irvine, CA, USA) to convert un-methylated cytosine to uracil. In downstream analyses the methylation status of a CpG site can thus be assessed by distinguishing between a C or T residue at the same position in the bisulfite treated DNA. The Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, CA, USA) was used to measure the methylation level of more than 485,000 individual CpG loci at a genome-wide resolution. Cases and controls were randomly and blindly distributed across conversion plates and methylation BeadChips. Twelve samples were analyzed on each BeadChip. As a "position effect" was reported for Illumina Methylation BeadChips, each sample position on the BeadChip was completely random as well. We further verified the randomization of the position on each BeadChip was effective by checking for a position effect, and we found no occurrence of position effect. BeadChips were processed according to manufacturer protocols. Briefly, the enzymatically fragmented bisulfite converted DNA is hybridized on-chip to locus specific DNA oligomers. A subsequent single base extension reaction with Biotin- or DNP-labeled ddNTPs allows the discrimination between methylated (unconverted) or un-methylated (converted) cytosines at each CpG site. After fluorescent staining, the BeadChips were imaged with a dedicated scanner and fluorescent signals recorded.

Data were inspected with the dedicated GenomeStudio software v2011.1 with Methylation module 1.9.0 (Illumina Inc., San Diego, CA), and quality checked according to the built in

quality controls, i.e. more than 200 control probes specifically designed to assess technical aspects of the bisulfite conversion process and of the BeadChip assay itself, including bisulfite Conversion, normalization (system background), staining efficiency, extension efficiency, hybridization efficiency, target removal (stripping after extension reaction), and specificity.

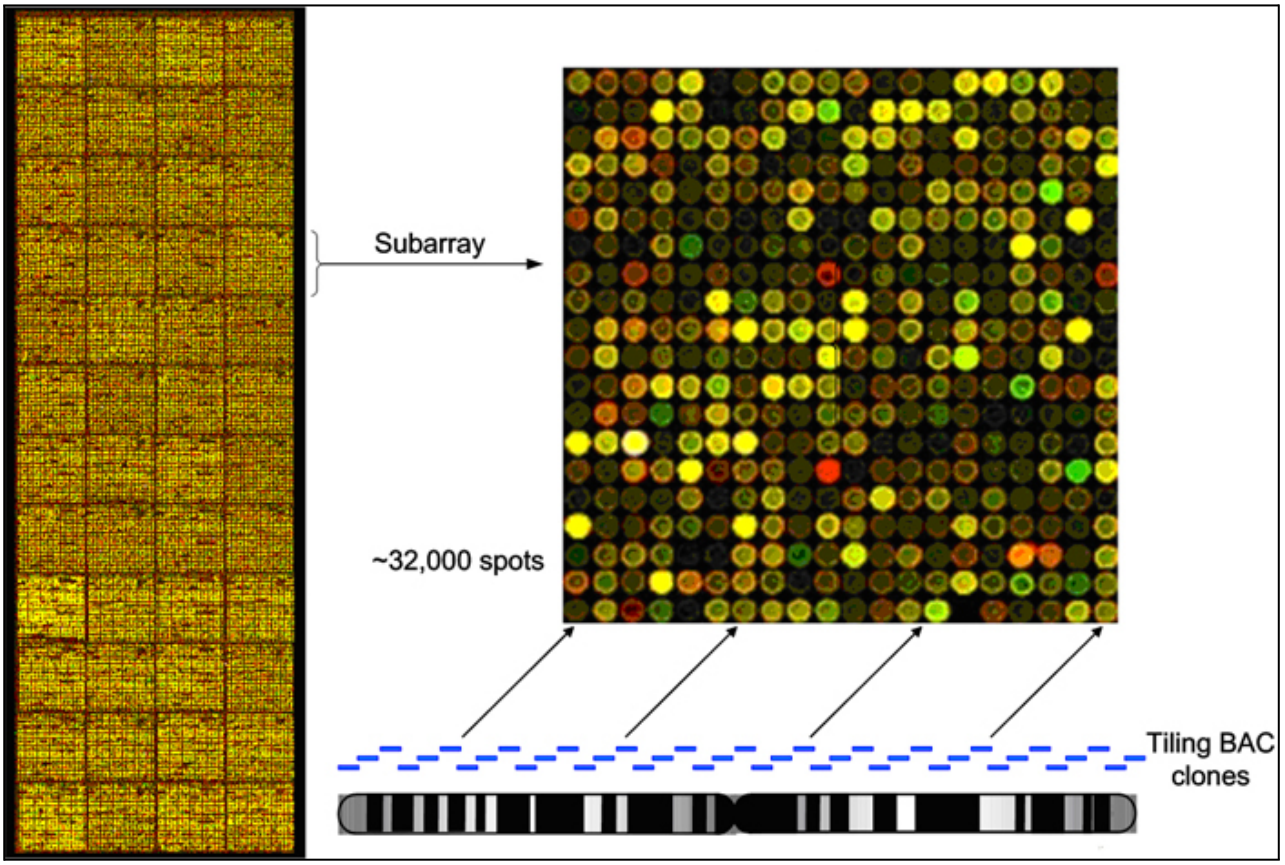

Figure 4. Whole genome microarray including 22 autosomes and 485,000 individual CpGs.

Raw data were analyzed with the R package 'methylumi'. The average methylation value at each locus, was computed as the ratio of the intensity of the methylated signal over the total signal (un-methylated + methylated): Beta-values represent the percentage of methylation at each individual CpG locus, ranging from 0 (no methylation) to 1 (full methylation). We excluded from the analyses: i) single Beta-values with detection p-value ≥ 0.01; ii) CpG loci with missing Beta-values in more than 20% of the assayed samples; iii)

CpG loci detected by probes containing SNPs with MAF ≥ 0.05 in the CEPH (Utah residents with ancestry from northern and western Europe, CEU) population; iv) samples with a global call rate ≤ 95%. Lastly, CpGs on chromosomes X and Y were excluded from the analysis.

To account for methylation assay variability and batch effects, we corrected all differential methylation analyses for "control probes" Principal Components (PCs). Using PCs assessed by principal a component analysis (PCA, R package 'stats') of the BeadChip's built-in control probes as a correction factor for statistical analyses of microarray data is a method that allows to consider the technical variability of several steps in the DNA-methylation analysis, from the bisulfite conversion to BeadChip processing. The first 10 PCs were thus included into the differential methylation analyses to correct for technical variability and batch effect.

White blood cells (WBCs)

WBC subtype percentages from genome-wide methylation data for each subject was used as a correction factor in regression models; in partially-adjusted regression models the first PC was included to correct for a substantial WBC variability.

Aging related measures

DNA methylation age (DNAm-age) of each subject was calculated by the algorithms proposed by Horvath which was derived from a range of tissues and cell types, using 353 probes targeted in the Illumina 27k and 450k methylation arrays. Aging eludes precise definition at the systemic level and denotes a multitude of processes at the cellular level. DNAm profiles of cytosine phosphate guanines (CpGs) have been used as indices of biological age. Notably, measures of epigenetic age in blood have been reported to be predictive of all-cause mortality after adjusting for chronological age and traditional risk factors. A positive residual of the DNAm-age estimate regressed on chronological age (AA) indicates that an individual is ahead of his or her chronological age, and a negative one suggests an individual is biologically "younger" than reflected by the chronological age. Two age AA indices, namely intrinsic epigenetic age acceleration (IEAA) and extrinsic

epigenetic age acceleration (EEAA), were calculated as well. EEAA is influenced by blood cell count contribution, whereas IEAA value is only weakly correlated with estimated measures of blood cell counts. EEAA is defined as the weighted average of DNAm age and imputed proportions of naïve CD8+ T cells, memory CD8+ T cells and plasmablasts.

Stochastic epigenetic mutations

The distribution and variability of methylation levels in our population were studied for each one of the 445,254 CpG sites using Box-and-whiskers plots in order to identify stochastic epigenetic mutations. For each probe, whenever the methylation level of one subject differed from the rest of the population we considered the outlier sample as epimutated for that locus. Thus, for each locus, epimutated subjects were identified as the extreme outliers with methylation level exceeding three times interquartile ranges Q1-(3 × IQR) and Q3+(3 × IQR). Finally, all epimutated loci were annotated in a new data matrix that allowed to calculate, for each subject, the total amount of epimutations and their genomic position. The Box-and-whiskers plot analysis was conducted using the boxplot function provided in the R car package and confirmed using the outlier function in the R outliers package. The analysis identified genes with an enriched number of epimutated probes (bonferroni corrected p-value < 0.05) confirming the presence of the epigenetic alterations previously reported in the medical report.

Lymphocyte-to-monocyte ratio (LMR)

Chronic inflammation plays a key role in the pathogenesis of cancer because of specific exposure. Several inflammation-based prognostic scores including the lymphocyte-to-monocyte ratio (LMR) reportedly predict survival in many malignancies. Several studies have shown that inflammation-based prognostic scores that include a combination of serum C-reactive protein (CRP) and albumin (ALB), a combination of neutrophil and lymphocyte counts (the neutrophil-to-lymphocyte ratio, NLR), a combination of platelet (PLT) and lymphocyte counts (the platelet-to-lymphocyte ratio, PLR) are associated with survival in patients with various cancers. Lymphocytes act as tumor suppressors by inducing cytotoxic cell death and inhibiting tumor cell proliferation

and migration. The important role of monocytes and macrophages in cancer has recently been uncovered.

Study design

Study population

Study subjects belong to a wider ongoing collaborative study on MPM, which is actively enrolling MPM cases and healthy subjects in the municipalities of Casale Monferrato (Piedmont region, Italy), an area with an exceptionally high incidence of mesothelioma caused by asbestos contamination at work and in the general environment from the asbestos-cement Eternit plant that was operational until 1986; additional MPM cases recruited in the main hospitals of the municipalities of Turin, Novara and Alessandria (Piedmont region, Italy).

The population-based case-control study included the incident cases of MPM diagnosed between 2000 and 2010 in the main hospitals of the reference centers after histological and/or cytological confirmation of MPM diagnosis; controls were selected (matched by date of birth and gender) during the same range of years from the local population, or among subjects not affected by neoplastic or respiratory conditions admitted to general medicine at the reference hospital. Subjects included in this study were selected according to the following criteria: i) availability of good quality DNA at the time of the analyses and ii) asbestos exposure above the background level. In total, 300 samples, 163 MPM cases of MPM and 137 non-MPM controls underwent DNAm analysis. Our study complies with the Declaration of Helsinki principles, and conforms to ethical requirements. All volunteers signed an informed consent form at enrollment. The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (formerly Human Genetics Foundation, Turin, Italy).

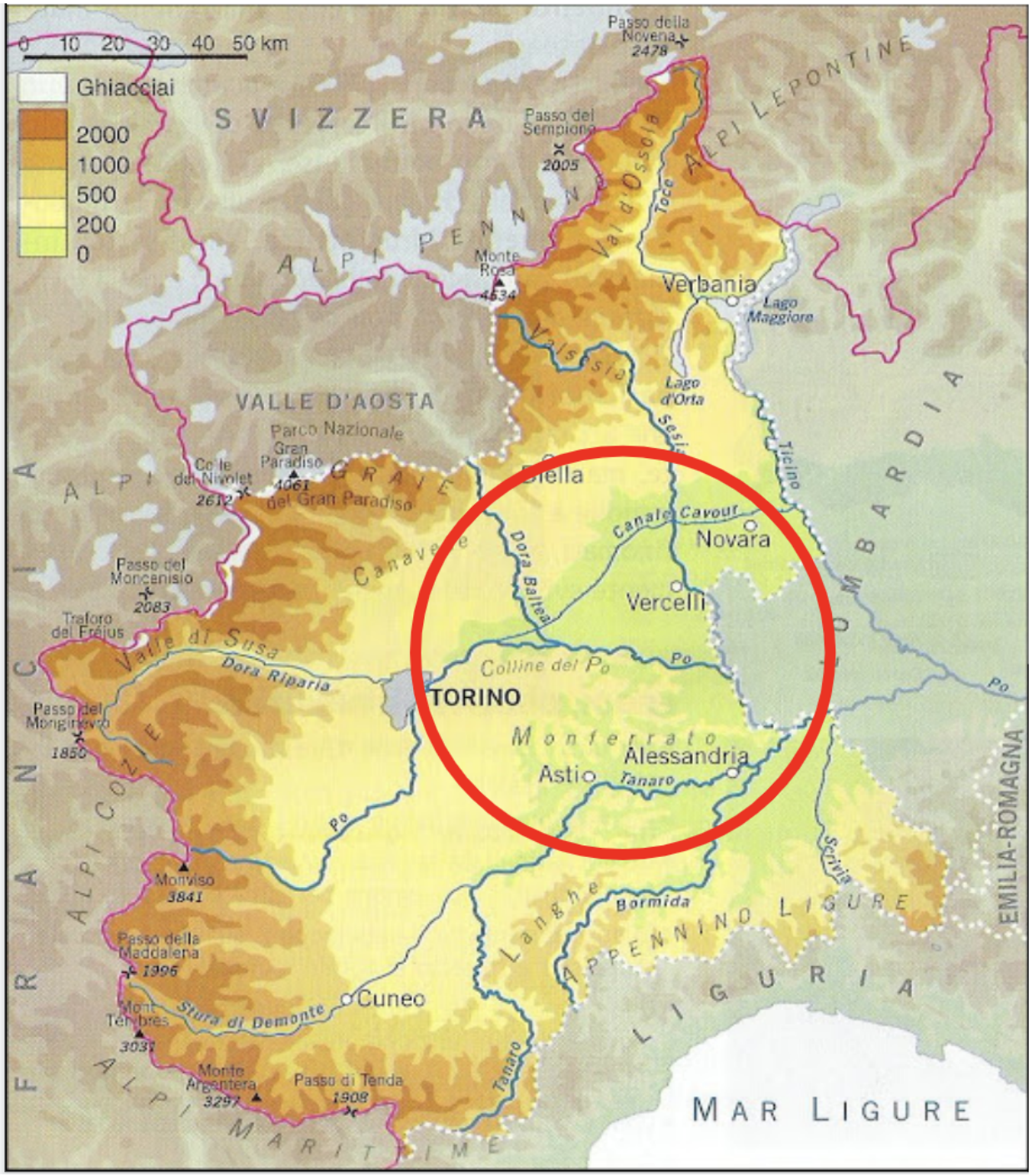

Figure 5. Population-based case-control study. Subjects were recruited in the municipalities of Casale Monferrato, Novara, Vercelli and Turin (Piedmont region, Italy).

Exposure assessment

For all subjects, occupational history and lifestyle habits information was collected through interviewer-administered questionnaires filled out at enrollment during a

face-to-face interview. Job titles were coded according to the International Standard Classification of Occupations, and trades according to the Statistical Classification of Economic Activities in the European Community. Frequency, duration and intensity of exposure were estimated, then a cumulative exposure index was computed. The evaluation of asbestos exposure (occupational, environmental, domestic) was conducted blindly to the case/control status by an experienced occupational epidemiologist: as regards occupational exposure, at least one and potentially many exposure patterns were assigned to every job held by a study subject; while regard to environmental exposures, exposures unrelated with work included living in proximity to industrial or natural sources of airborne asbestos were evaluated; then referring to domestic exposures, having asbestos containing materials installed at home or handling such materials during home repairs or leisure-time activities were considered. A score was assigned to the exposure determinants included in a fixed list, including active and passive exposures. Scores were not intended to directly drive exposure assessment, but their pattern drew exposure scenarios, documenting how the rater reconstructed each exposure circumstance. This step accomplished, exposure was assessed along the four axes of probability, frequency, intensity and duration.

Limitations

One potential limitation of this study is that this is a retrospective study. A large-scale prospective validation study is needed to confirm the results. Furthermore, EWAS associations can be causal as well as consequential for the phenotype in question, a difference from GWASs]. Reverse causation and confounding are particular issues for EWAS design. Although any EWAS association with disease is potentially an advance, being able to identify the direction of causality will greatly aid in determining the usefulness of epigenetic variation as, for example, a marker of disease progression, a target for reversal by treatment with drug or a measure of drug response by monitoring the kinetics of drug-induced epigenetic changes.

Study1

New DNA methylation signals for Malignant Pleural Mesothelioma risk assessment

Simple Summary: Our study investigated DNA methylation differences in easily accessible white blood cells (WBCs) between malignant pleural mesothelioma (MPM) cases and asbestos-exposed cancer-free controls. A multiple regression model highlighted that the methylation level of two single CpGs (cg03546163 in *FKBP5* and cg06633438 in *MLLT1*) are independent MPM markers. The epigenetic changes at the *FKBP5* and *MLLT1* genes were robustly associated with MPM in asbestos-exposed subjects. Interaction analyses showed that MPM cases and cancer-free controls showed DNAm differences which may be linked to asbestos exposure.

Abstract: Malignant pleural mesothelioma (MPM) is a rare and aggressive neoplasm. Patients are usually diagnosed when current treatments have limited benefits, highlighting the need for noninvasive tests aimed at an MPM risk assessment tool that might improve life expectancy. Three hundred asbestos-exposed subjects (163 MPM cases and 137 cancer-free controls), from the same geographical region in Italy, were recruited. The evaluation of asbestos exposure was conducted considering the frequency, the duration and the intensity of occupational, environmental and domestic exposure. A genome-wide methylation array was performed to identify novel blood DNA methylation (DNAm) markers of MPM. Multiple regression analyses adjusting for potential confounding factors and interaction between asbestos exposure and DNAm on the MPM odds ratio were applied. Epigenome-wide analysis (EWAS) revealed 12 single-CpGs associated with the disease. Two of these showed high statistical power (99%) and effect size (>0.05) after false discovery rate (FDR) multiple comparison corrections: (i) cg03546163 in *FKBP5*,

significantly hypomethylated in cases (Mean Difference in beta values (MD) = −0.09, 95% CI = −0.12|−0.06, $p = 1.2 \times 10^{-7}$), and (ii) cg06633438 in *MLLT1*, statistically hypermethylated in cases (MD = 0.07, 95% CI = 0.04|0.10, $p = 1.0 \times 10^{-6}$). Based on the interaction analysis, asbestos exposure and epigenetic profile together may improve MPM risk assessment. Above-median asbestos exposure and hypomethylation of cg03546163 in *FKBP5* (OR = 20.84, 95% CI = 8.71|53.96, $p = 5.5 \times 10^{-11}$) and hypermethylation of cg06633438 in *MLLT1* (OR = 11.71, 95% CI = 4.97|29.64, $p = 5.9 \times 10^{-8}$) genes compared to below-median asbestos exposure and hyper/hypomethylation of single-CpG DNAm, respectively. Receiver Operation Characteristics (ROC) for Case-Control Discrimination showed a significant increase in MPM discrimination when DNAm information was added in the model (baseline model, BM: asbestos exposure, age, gender and white blood cells); area under the curve, AUC = 0.75; BM + cg03546163 at *FKBP5*. AUC = 0.89, 2.1×10^{-7}; BM + cg06633438 at *MLLT1*. AUC = 0.89, 6.3×10^{-8}. Validation and replication procedures, considering independent sample size and a different DNAm analysis technique, confirmed the observed associations. Our results suggest the potential application of DNAm profiles in blood to develop noninvasive tests for MPM risk assessment in asbestos-exposed subjects.

Keywords: malignant pleural mesothelioma; asbestos exposure; DNA methylation; epigenome-wide analysis; interaction analysis

---

1. Introduction

Mesothelioma has a long latency period, usually emerging 20–40 years after asbestos exposure [1]. Malignant pleural mesothelioma (MPM) is rare (prevalence 1–9/100,000), but the corresponding annual death toll worldwide is still estimated at about 40,000 [2,3]. Each year, 125 million people are exposed to asbestos, according to a World Health Organization report [4]. The International Agency for Research on Cancer confirmed that all fibrous forms of asbestos are carcinogenic to humans. The main outcome of exposure is

mesothelioma, but cancer at other sites, such as respiratory-tract tumors, are moderately frequent [4]. Previous in vitro studies have demonstrated the cytotoxic effects of asbestos fibers [5,6].

A significant association between MPM and asbestos exposure has been reported, showing a clear, increasing trend in the odds ratio (OR) with increasing cumulative exposure among subjects exposed to over 10 fiber/mL-years [7]. Another study reported that the incidence of malignant mesothelioma (MM) was strongly associated with the proximity of one's residence to an asbestos exposure source [8].

DNA methylation (DNAm) is an epigenetic mechanism involved in gene expression regulation. In particular, dysregulation of promoter DNAm and histone modification are epigenetic mechanisms involved in human malignancies [9].

According to recent papers, both DNAm and genetic alterations may contribute to MPM tumorigenesis [10–15]. Whereas the genome remains consistent throughout one's lifetime, factors like ageing, lifestyle, environmental exposures and diseases can modify DNAm. The adaptive nature of DNAm means that it can be used to link environmental factors to the development of pathologic phenotypes such as cancers. Fasanelli et al. observed an association between exposure to tobacco and site-specific CpG methylation. They also used peripheral blood DNA to evaluate the importance of these epigenetic alterations in the aetiology of lung cancer [16].

There is less information on the mechanisms and clinical outcomes of epigenetic derangements in MPM [17–19]. Several studies have evaluated DNAm alterations in MM samples [20–22], but few of them focused on DNAm alteration in blood as a circulating marker. Fischer et al. examined serum DNAm of nine gene-specific promoters from MM cases [23]. A more recent paper identified hypomethylation of a single CpG in *FKBP5* in whole blood cells as a predictor of overall survival in MPM cases [13]. Guarrera et al. evaluated methylation levels in DNA from whole blood leukocytes as potential diagnostic markers for MPM and found a differential methylation between asbestos-exposed MPM cases and controls, mainly in genes related to the immune system [11]. The identification

of reliable DNAm biomarkers with high sensitivity and specificity for MPM risk assessment would be a major advancement.

This study was undertaken with the goal to identify new biomarkers for MPM risk assessment and to determine if peripheral blood DNAm profiles have any predictive value. The second goal was to evaluate the interaction effect of asbestos exposure with DNAm on MPM risk. Currently, there are no sensitive testing methods available for the screening of asbestos-exposed individuals who are at high risk of developing MPM. Thus, the identification of reliable MPM diagnostic biomarkers in peripheral blood might provide a tool for detecting the disease at an early stage.

2. Results

2.1. Epigenome-Wide Association Study (EWAS)

CpGs (445,254) passed quality control procedures and were considered for statistical analyses. EWAS revealed two statistically significant differentially methylated single-CpGs between case and control groups: cg03546163 in the *FKBP5* gene (Mean Difference in beta values (MD) = 0.09, 95% CI = −0.12|−0.06, $p = 1.2 \times 10^{-7}$, $p = 0.028$) and cg06633438 in the *MLLT1* gene (MD = 0.07, 95% CI = 0.04|0.10, $p = 1.0 \times 10^{-6}$, $p = 0.049$) after False Discovery Rate (FDR) post hoc correction (**Figure 1**; **Table 1**).

Another 10 CpGs showed hypo/hypermethylation in MPM considering FDR < 0.05 but not effect size (MD) cut off ≥ |0.05| (**Table 1**).

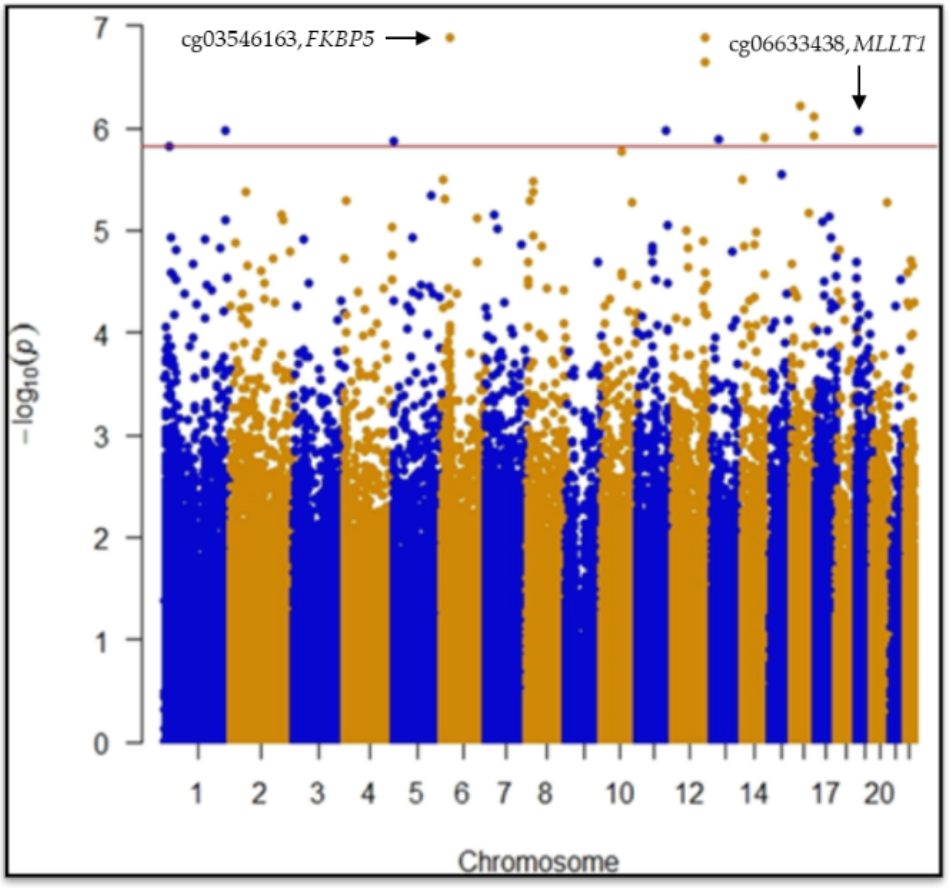

Figure 1. Manhattan plot for EWAS test on 450k single CpGs. Single-CpG DNAm was used as dependent variable adjusting for age, gender, White blood cells (WBCs: monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T) estimation, population stratification and technical variability. FDR post hoc line highlights statistically significant differences between cases and controls at single CpG level.

Bootstrap was computed to estimate measures of accuracy using random sampling methods. The bias-corrected and accelerated (BCa) bootstrap interval was calculated for cg03546163 in *FKBP5* (95% CIBCa = −0.16|−0.10, z0 = −0.008, a = 0.002) and cg06633438 in *MLLT1* (95% CIBCa = −0.06|−0.1, z0 = −0.011, a = 0.0004) genes, confirming the robustness of the results considering the sample under study.

Statistically significant differences in MD between cases and controls were found in the WBCs estimated (monocytes, $p = 6.0 \times 10^{-3}$; granulocytes, $p = 2.2 \times 10^{-16}$; B cells, $p = 1.1 \times$

10^{-12}; NK cells, $p = 3.6 \times 10^{-4}$; CD4+ T, $p = 2.2 \times 10^{-16}$; CD8+ T, $p = 6.8 \times 10^{-11}$; naïve CD4T, $p = 0.012$; naïve CD8T, $p = 7.0 \times 10^{-3}$).

In order to assess if smoking status, classified as current, former and never-smokers, could modify DNAm profiles, we performed a multivariate regression analysis with the same model used for the discovery phase. No evidence of methylation differences linked to different smoking levels was found for any of the twelve statistically significant CpGs.

Table 1. Differential DNAm analyses ordered by effect size. Information about single-CpGs, including location-related values and model outputs (effect size, standard error, *p* values).

Probe ID	Chr	Map Position	Gene Symbol	Ucsc Refgene Group	Snp Probe	Effect Size	Standard Error	*p* Value	Fdr	Sign
cg02869235	12	124726864			rs73223527	0.058	0.011	1.3×10^{-7}	0,028	*§
cg03546163	6	35654363	*FKBP5*	5′UTR		−0.089	0.016	1.3×10^{-7}	0,028	*§ⓗ
cg02353048	12	124718401				0.033	0.006	2.2×10^{-7}	0,032	*§
cg06633438	19	6272158	*MLLT1*	Body		0.069	0.014	1.0×10^{-6}	0,049	*§ⓗ
cg18860329	13	43354421	*C13orf30*	TSS1500		0.050	0.010	1.3×10^{-6}	0,049	*§
cg19782190	14	103487004	*CDC42BPB*	Body		0.043	0.009	1.2×10^{-6}	0,049	*§
cg06834916	5	95610				0.037	0.008	1.4×10^{-6}	0,049	*§
cg09479650	16	85578516			rs4843449	0.037	0.007	1.2×10^{-6}	0,049	*§
cg26680989	16	85560739			rs80332660	0.036	0.007	7.6×10^{-7}	0,049	*§
cg25409554	1	234871422				0.034	0.007	1.1×10^{-6}	0,049	*§
cg01201399	16	30793389	*ZNF629*	Body		0.030	0.006	6.1×10^{-7}	0,049	*§
cg17283266	11	111717611	*ALG9*	Body		−0.030	0.006	1.1×10^{-6}	0,049	*§

Control group was set as reference. Adjustment covariates: age, gender, population stratification, WBCs (monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T) estimation and technical variability. *: statistically significant at *p* value < 0.05; §: statistically significant at FDR post hoc adjustments. ⓗ: statistically significant at beta = 0.01.

2.2. Receiver Operation Characteristics (ROC) for Case-Control Discrimination

The baseline model (BM) including age, gender, asbestos exposure and WBCs was compared with BM adding the DNAm levels of cg03546163 or cg06633438. Receiver Operation Characteristics 8ROC9 curves showed a significant increase in MPM discrimination when DNAm information was added in the model (Table 2).

Table 2. Disease discrimination test considering (AUC) comparison between baseline model and models additionally including single-CpG.

Model	AUC	DeLong's Test
BM (asbestos exposure, age, gender and WBCs)	0.75	Reference
BM + cg03546163 (*FKBP5*)	0.89	2.1×10^{-7}
BM + cg06633438 (*MLLT1*)	0.89	6.3×10^{-8}

Models are shown as baseline model (BM) or BM + Single CpG DNAm. AUC Differences between the considered model and BM were estimated with DeLong's test.

2.3. Interaction Analysis

CpG sites and asbestos exposure were considered as predictors of MPM risk in the interaction model. Categorical variables (quantile information) were used considering median values.

We tested the interaction between asbestos exposure and DNAm levels at cg03546163 in *FKBP5* and cg06633438 in *MLLT1.*

Considering cg03546163 in *FKBP5,* DNA hypermethylation and low asbestos exposure levels were used as references, while for cg06633438 in *MLLT1,* DNA hypomethylation and low asbestos exposure levels were set as references (Table 3).

The OR was estimated as the relationship between the combination of single-CpGs DNAm levels and asbestos exposure quantile, and the reference (low median asbestos exposure and hypermethylation status for cg03546163, or hypomethylation status for cg06633438).

Age, gender, population stratification, and WBCs were included in the GLM (family = binomial) to adjust the interaction effect.

The relationship between asbestos exposures and single-CpG DNAm levels was evaluated. An increase of one unit of asbestos exposure (rank transformed fibers/mL years) was related to the *FKBP5* gene ($\beta = -0.016$, 95% CI = −0.031|−0.001, $p = 0.044$) and *MLLT1* gene ($\beta = -0.014$, 95% CI = 0.001|0.026, $p = 0.035$) methylation level variations.

Strong association between asbestos exposure and MPM risk, considering dichotomous distribution of asbestos exposure, was found (OR = 6.11, 95% CI = 3.73|10.20, $p = 1.8 \times 10^{-12}$). Quartile distribution of asbestos exposure was evaluated to estimate the potential incremental association with MPM risk (1st quartile: used as reference; 2nd quartile: OR = 1.83, 95% CI = 0.93|3.69, $p = 0.09$; 3rd quartile: OR = 6.63, 95% CI = 3.30|13.81, $p = 2.1 \times 10^{-7}$; 4th quartile: OR = 11.00, 95% CI = 5.26|24.30, $p = 7.3 \times 10^{-10}$).

Table 3. Interaction between asbestos exposure and single CpG DNAm on the MPM Odds ratios.

DNAm	Asbestos Exposure	OR	Std. Error	95% CI	*p* Value
cg03546163 (*FKBP5*)					
Hypo	Low	2.79	1.51	1.26|6.33	0.013
Hyper	High	7.21	1.54	3.17|17.27	4.6×10^{-6}
Hypo	High	20.84	1.59	8.71|53.96	5.5×10^{-11}
cg06633438 (*MLLT1*)					
Hyper	Low	1.29	1.63	0.70|3.81	0.258
Hypo	High	7.27	1.55	3.17|17.65	5.3×10^{-6}
Hyper	High	11.71	1.57	4.97|29.64	5.9×10^{-8}

Reference for cg03546163 in *FKBP5*: hypermethylation and low asbestos exposure levels; reference for cg06633438 in *MLLT1*: hypomethylation and low asbestos exposure levels.

2.4. Validation and Replication

For the replication and validation approaches, an independent sample of 140 MPM cases and 104 cancer-free asbestos-exposed controls from the same areas were considered, using a targeted DNAm analysis technique.

The direction and magnitude of the association was consistent for cg03546163 and cg06633438 DNAm. Patients showed significantly lower DNAm at cg03546163 (MD = −0.061, 95% CI = −0.087|−0.036, $p = 4.5 \times 10^{-6}$) and higher DNAm at cg06633438 (MD = 0.024, 95% CI = 0.061|0.013, $p = 4.0 \times 10^{-2}$) compared with controls. A multivariate analysis confirmed that DNAm at cg03546163 in *FKBP5* and cg06633438 in *MLLT1* were independently associated with MPM detection.

3. Discussion

In the present study, we used a whole genome microarray approach to investigate DNAm in WBCs from MPM cases and asbestos-exposed cancer-free controls from a region with a history of asbestos exposure (Piedmont, Italy) [10] in order to identify new noninvasive epigenetic markers related to MPM. The identification of reliable MPM diagnostic biomarkers in peripheral blood might improve risk assessment.

We observed hypomethylation of CpG cg03546163, located in the 5′ UTR region of *FKBP5* gene, in MPM cases compared to controls.

Epigenetic activation of the FKBP Prolyl Isomerase 5 (*FKBP5*) gene has been shown to be associated with increased stress sensitivity and the risk of psychiatric disorders [24]. *FKBP5* is an immunophilin and has an important role in immunoregulation and in protein folding and trafficking. It plays a role in transcriptional complexes and acts as a cotranscription factor, along with other proteins in the *FKBP* family [25]. The suggestion of a possible role of *FKBP5* in the development and progression of different types of cancer has stemmed from several studies. In particular, high protein expression has been linked to either suppression or promotion of tumour growth, depending on tumour type and microenvironment [26,27].

FKBP5 is involved in the NF-kB and AKT signaling pathways, both of which are implicated in tumorigenesis [28]. Notably, NF-kB appears to be frequently constitutively activated in malignant tumours and involved in the modulation of genes linked to cell motility, neoangiogenesis, proliferation and programmed cell death [29]. An epigenetic

upregulation of *FKBP5* could promote NF-kB activation [30]. STAT3-NFkB activity is involved in chemoresistance in MM cells [31], and NFkB was shown to be constitutively active as a result of asbestos-induced chronic inflammation [32].

CpG cg06633438 located in the body region of the *MLLT1* gene was hypermethylated in cases compared to controls.

The *MLLT1* gene encodes the ENL protein, a histone acetylation reader component of the super elongation complex (SEC), which promotes transcription at the elongation stage by suppressing transient pausing by the polymerase at multiple sites along the DNA. In acute myeloid leukemia, *MLLT1* regulates chromatin remodeling and gene expression of many important proto-oncogenes [31]. Yoshikawa and colleagues suggested that mesothelioma may be the consequence of the somatic inactivation of chromatin-remodeling complexes and/or histone modifiers, including *MLLT1* [30].

In mesothelioma patients with short-term recurrence after surgery, frequent 19p13.2 loss was reported. This region encompasses several putative tumor suppressors or oncogenes, including *MLLT1* [32].

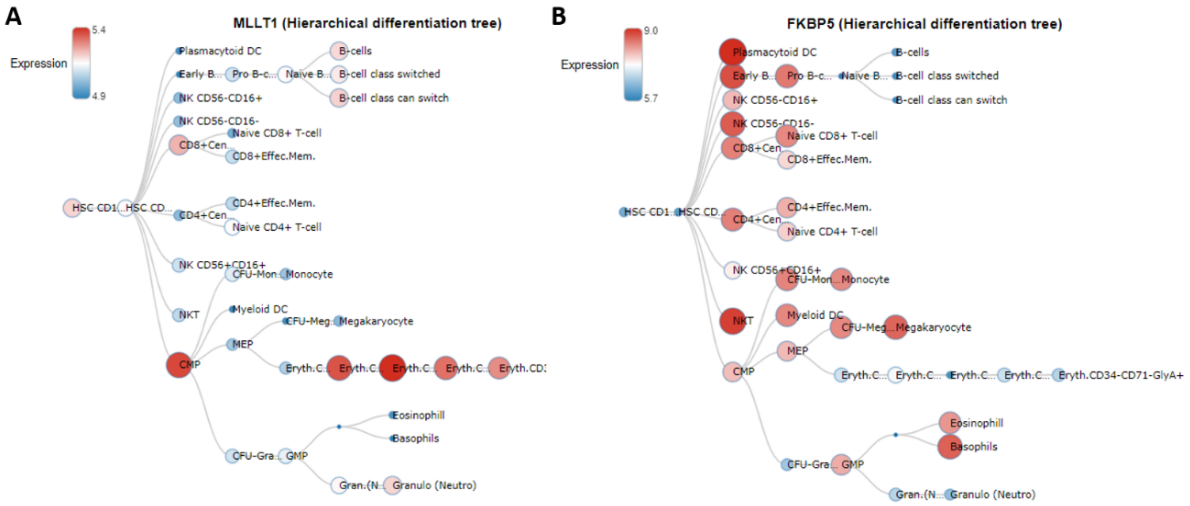

Figure 2. Expression profiles in normal human haematopoiesis. *MLLT1* (**A**) and *FKBP5* (**B**) expression profiles in normal human haematopoiesis as reported in the Blood Spot database (http://servers.binf.ku.dk/bloodspot/,accessed on 26 May 2021).

Interestingly, *MLLT1* and *FKBP5* showed opposite behavior, increasing and decreasing DNAm levels, respectively, in relation to MPM. This finding could reflect the opposite expression profiles of the two genes among all the different subtypes of white blood cells in normal human hematopoiesis, as reported in the Blood Spot database (http://servers.binf.ku.dk/bloodspot/,accessed on 26 May 2021) (Figure 2) [33].

Our interaction analysis showed that considering DNAm levels at *FKBP5* and *MLLT1* genes together with asbestos exposure levels may help to better define MPM risk for asbestos-exposed subjects.

Six single-CpGs showed differential methylation in patients, including those located in *C13orf30, CDC42BPB, ZNF629* and *ALG9* genes; the other six were not annotated to named genes. *ALG9* is a glycogene whose reduced expression has been described during the epithelial-to-mesenchymal transition, an essential process also involved in cancer progression [34]. The *CDC42BPB* gene is ubiquitously expressed in mammals and encodes a serine/threonine protein kinase, a member of the MRCK family [35]. The role of MRCKs in cytoskeletal reorganization during cell migration and invasion has been characterized [36]. The biological function of *C13orf30* and *ZNF629*, a DNA-binding transcription factor, is still to be established.

MPM cases and asbestos-exposed controls showed different proportions of estimated WBCs, which may denote the crucial implication of the immune system. It is known that in cancer, including mesothelioma, the immune system is affected [37], and there is evidence that asbestos directs antigen overstimulation, and that reactive oxygen species production induces functional changes in WBCs [38]. Indeed, in MPM cases, we showed a reduction of estimated CD4+ and CD8+ T lymphocytes, suggesting a weaker adaptive immune system [39]. This may reflect the possible occurrence of functional changes in WBC subtypes in MPM [40,41].

The need for reliable biomarkers is of extreme relevance for a disease such as MPM, which is characterized by the accumulation and persistence of asbestos fibers in the lungs, leading to a long latency period before clear clinical signs of the tumor are detectable. Several biomarkers for early MPM detection (e.g., mesothelin, osteopontin and fibulin-3)

have been proposed so far; however, some of them are still under investigation [42]. In this context, DNAm changes in easily-accessible WBCs may provide a useful tool to better assess MPM risk in asbestos-exposed subjects.

Our findings that DNAm levels in single-CpGs in *FKBP5* and *MLLT1* genes are independent markers of MPM in asbestos-exposed subjects suggest the potential use of blood DNAm analysis as a noninvasive test for MPM detection.

Some somatic gene alterations in lung cancer have been linked to tobacco smoke, but few data are available on the role of asbestos fibers: Andujar and colleagues investigate the mechanism of P16/CDKN2A alterations in lung cancer including asbestos-exposed patients. P16/CDKN2A gene inactivation in asbestos-exposed non-small-cell lung carcinoma (NSCLC) cases, a tumor independent of tobacco smoking but associated with asbestos exposure, mainly occurs via promoter hypermethylation, loss of heterozygosity and homozygous deletion, suggesting a possible relationship with an effect of asbestos fibers [43].We observed epigenetic deregulations in the blood of MPM patients compared to that of cancer-free controls, suggesting the potential use of DNAm for risk stratification among asbestos-exposed individuals.

If this observation can be verified in prospectively collected samples, it may be possible to use CpGs methylation to further improve MPM risk estimation for subjects with occupational and/or environmental asbestos exposure.

Limitation of the Study

Leukocyte DNAm may be a nonspecific marker related to a general, tumour-induced inflammatory status rather than a specific MPM biomarker. Further studies are therefore needed to support our findings.

One main limitation of the functional interpretation of our results is that all our cases had already developed MPM at recruitment: thus, our findings likely reflect disease status rather than being markers of the dynamic processes leading to MPM onset. The lack of MPM tissue from the same subjects also poses major constraints to the functional interpretation of our findings.

Notwithstanding the above limitations, the discrimination between MPM cases and asbestos-exposed cancer-free controls improved when DNAm levels were considered together with asbestos exposure levels.

4. Material and Methods

4.1. Study Population

Study subjects were part of a wider, ongoing collaborative study, which is actively enrolling MPM cases and cancer-free controls in the municipality of Casale Monferrato (Piedmont Region, Italy). This area was chosen due to its exceptionally high incidence of mesothelioma, caused by widespread occupational and environmental asbestos exposure originating from the Eternit asbestos-cement plant, which was operational until 1986 [44]. Additional MPM cases and cancer-free controls were recruited from other main hospitals of the Piedmont Region (in the municipalities of Turin, Novara and Alessandria). The ongoing collaborative study includes MPM cases diagnosed between incident MPM cases diagnosed between 2000 and 2010 after histological and/or cytological confirmation, and matched controls [45].

The present study included 159 MPM cases and 137 cancer-free controls from a larger case-control study, all of whom had genetic and blood DNAm data [46], good quality DNA at the time of the analyses, and information on asbestos exposure above the background level, as defined in Ferrante et al. [47]. MPM cases and asbestos-exposed cancer-free controls were matched by date of birth (±18 months) and gender. An additional 244 (140 MPM cases and 104 cancer-free controls) independent samples from the same case-control study were included for validation/replication analyses.

Table 4,5 shows the descriptive characteristics of controls and cases (Min, 1st Q, Median, Mean, 3rd Q and Max) that were considered in the statistical analysis (gender, age, asbestos exposure and WBC estimates: monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T). Asbestos exposure (occupational, environmental and domestic) was

normalized considering frequency, duration and intensity. Smoking status (current, former and never smokers) is also explained in Table 6.

Our study complied with the Declaration of Helsinki principles and conformed to ethical requirements. All volunteers signed an informed consent form at enrollment. The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

Table 4. Descriptive characteristics of cancer-free control group.

Variable	Controls (Male 100, Female 37)					
	Min	**1st Q**	**Median**	**Mean**	**3rd Q**	**Max**
Age	41.60	57.41	65.65	64.59	72.63	90.94
Asbestos exposure	−2.71	−0.97	−0.48	−0.44	0.09	1.73
Monocytes	0.00	0.05	0.06	0.07	0.08	0.26
Granulocytes	0.36	0.54	0.60	0.62	0.68	0.99
Natural Killer	0.00	0.04	0.07	0.08	0.11	0.29
B cells	0.00	0.07	0.09	0.09	0.11	0.19
CD4+ T	0.00	0.10	0.14	0.14	0.19	0.35
CD8+ T	0.00	0.03	0.06	0.07	0.10	0.23

Minimum (Min), First Quartile (1st Q), Median, Mean, Third Quartile (3rt Q) and Maximum (Max) of variables related to cancer-free controls.

Table 5. Descriptive characteristics of MPM group.

Variable	Cases (Male 113, Female 50)					
	Min	**1st Q**	**Median**	**Mean**	**3rd Q**	**Max**
Age	33.90	61.19	68.68	67.59	75.17	90.80
Asbestos exposure	−2.71	−0.21	0.39	0.37	0.98	2.94
Monocytes	0.00	0.05	0.07	0.08	0.10	0.20
Granulocytes	0.37	0.67	0.74	0.74	0.81	1.03
Natural Killer	0.00	0.02	0.05	0.06	0.08	0.23
B cells	0.00	0.05	0.06	0.06	0.08	0.16
CD4+ T	0.00	0.03	0.07	0.08	0.11	0.22
CD8+ T	0.00	0.00	0.02	0.03	0.04	0.22

Minimum (Min), First Quartile (1st Q), Median, Mean, Third Quartile (3rt Q) and Maximum (Max) of variables related to MPM cases.

Table 6. Descriptive characteristics of smoking status stratified by disease.

Smoking Habits	Cases (163)		Controls (137)	
	n	%	*n*	%
Current smokers	29	17.79	30	21.90
Former smokers	54	33.13	60	43.80
Never smokers	75	46.01	47	34.31

n and % of the three levels of smoking status stratified by disease.

4.2. Exposure Assessment

Information on occupational history and lifestyle habits were collected from all subjects through interviewer-administered questionnaires, which were completed during face-to-face interviews at enrollment. Job titles were coded in two ways according to the International Standard Classification of Occupations [47] and the Statistical Classification of Economic Activities in the European Community.

A cumulative exposure index was computed considering frequency, duration and intensity of asbestos exposure. Occupational, environmental and domestic asbestos exposure were evaluated by an experienced occupational epidemiologist [47], and exposure doses (fibers/mL years) were rank-transformed to remove skewness.

4.3. Blood DNAm Analysis and Beta-Value Extraction

DNAm levels were measured in DNA from whole blood collected at enrollment using the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA). For blood DNAm analysis (including quality control) please refer to the previous work of the same group [11].

We used the R package 'methylumi' to analyze DNAm data. The average methylation value at each locus was computed as the ratio of the intensity of the methylated signal over the total signal (unmethylated + methylated) [48]. Beta-values ranging from 0 (no methylation) to 1 (full methylation) represent the percentage of methylation at each individual CpG locus.

We excluded the following from the analyses: (i) single beta-values with a *p*-value for detection ≥ 0.01; (ii) CpG loci that had missing beta-values in more than 20% of the assayed samples; (iii) CpG loci detected by probes containing single nucleotide polymorphisms (SNPs) with MAF ≥ 0.05 in the CEPH (Utah residents with ancestry from northern and western Europe, CEU) population; and iv) samples with a global call rate ≤ 95%. We also excluded CpGs on chromosomes X and Y.

4.4. Batch Effect, Population Stratification and White Blood Cells Estimations

All differential methylation analyses were corrected for "control probes" Principal Components (PCs) to account for variability and batch effects in methylation assays. We used PCs assessed by principal component analysis of the BeadChip's built-in control probes as a correction factor for statistical analyses of microarray data. This method allows researchers to account for the technical variability in the different steps in DNAm analysis, from bisulfite conversion to BeadChip processing [49].

An individual's geographic origins may influence DNAm profiles, which could potentially introduce bias. To take this into consideration, we took advantage of the available data from our previous study, which includes a genome-wide genotyping dataset from the same study subjects [50]. When genome-wide genotyping was used to calculate the first PCs, they were shown to correlate with different geographic origins [51].

For each subject, we extracted WBC subtype percentages, estimated based on genome-wide methylation data. This method provides quantification of the composition of leukocytes that can be achieved by simple histological or flow cytometric assessments, with an admissible range of variability [52].

4.5. Statistical Analyses

4.5.1. Epigenome-Wide Association Study

An association test was used to analyze the mean differences (MD) in single-CpG methylation between MPM cases and asbestos-exposed cancer-free controls. We performed multiple regression analysis adjusted for age, gender, estimated WBCs

(monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T), population stratification (first 2 PCs) and technical variability (first 10 PCs). For multiple comparison tests, a FDR p value ≤ 0.05 was considered statistically significant.

Bootstrapping was performed using random sampling methods to estimate the measures of accuracy defined in terms of bias, variance, confidence intervals and prediction error. Bootstrapping can also be applied to control and check the results for stability. The bias-corrected and accelerated (BCa) bootstrap interval was calculated with regard to single CpGs.

ROC for Case-Control Discrimination was implemented, and the AUC metric was applied to estimate the predictive performance of a binary classification (cases/controls). The baseline model (BM) included age, gender, asbestos exposure and WBCs, and was compared with the BM after adding the DNAm levels of statistically significant, single-CpGs at EWAS. AUC differences between BMs before and after the addition of DNAm levels were estimated with DeLong's test.

4.5.2. Statistical Power

To ensure a study power greater than 99% (two-tailed test at $\alpha = 0.05$ and $\beta = 0.01$), only CpGs with a MD between cases and controls ≥ |0.05| were selected.

Covariates were included step-by-step in a sensitivity analysis to validate the association output considering effect size, standard error, 95% confidence interval and p value variations.

Gene set enrichment analyses were carried out on CpGs with a False Discovery Rate p value (P_{FDR}) ≤ 0.05 to identify pathways that may be affected by MPM-related changes in methylation.

All statistical analyses were conducted using the open source software R (4.0.2).

4.5.3. Interaction Analysis

Logistic regression was used to analyze the relationship between CpGs and asbestos exposure in MPM risk (odds ratio), adjusting for age, gender, SNP PCs and WBCs

estimates. Asbestos exposure was classified as above-median or below-median, and CpG methylation was categorized as above-median or below-median.

MPM risk for a given CpG level and asbestos exposure was expressed by ORij, where i indicates the asbestos exposure (below-median or above-median) and j indicates the CpG (above-median or below-median). Considering the direction of the effect, the same approach was used: for hypomethylated CpGs, above-median was used as the reference level, whereas below-median was used for hypermethylated CpGs.

Subjects with below-median asbestos exposure and reference-level CpG DNAm were considered the baseline group, and their MPM risk was coded as OR00 = 1. Interaction was analyzed with respect to both additive and multiplicative models based on the ORs obtained by logistic regression.

Synergistic interaction (positive interaction) implies that the combined action of two factors in an additive model is greater than the sum of their individual effects. Antagonistic interaction, on the other hand, means that when two factors are present in an additive model, the action of one reduces the effect of the other.

Multivariable logistic regression models were used to explore any deviations from a multiplicative model, including asbestos exposure, CpG and the corresponding interaction term (CpG × exposure). All models were adjusted for age, gender, SNP PCs, technical covariates and WBCs estimates. p-values < 0.05 were considered statistically significant.

4.6. Validation and Replication

DNAm signal validation and replication was done by the EpiTYPER MassARRAY assay (Agena Bioscience). This assay uses a MALDI-TOF mass spectrometry-based method to quantitatively assess the DNA methylation state of the CpG sites of interest [53]. DNA (500 ng) was bisulfite-converted as indicated in Section 4.3.

PCR amplification, treatment with SAP solution and Transcription/RNase A cocktails were performed according to the manufacturer's instructions, and the mass spectra were analyzed by an EpiTYPER analyzer. The MassARRAY assay cannot discriminate between CpGs located in close proximity in the sequence, so instead, the close neighboring CpGs

are analyzed as "Units", i.e., the measured methylation level is the average of the methylation levels of the CpGs cumulatively analyzed within the Unit. In the case of cg03546163, the measured methylation level is the average between two CpG sites located in very close proximity (Figure S1). For cg06633438, the two adjacent signals were considered, since the results for the model did not differ for effect size, standard error, 95% CI or *p* value (Figure S2).

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Location of cg03546163 investigated by EpiTYPER MassARRAY, Figure S2: Location of cg06633438 investigated by EpiTYPER MassARRAY.

cg03546163

TTTTTGTTTAGGATGAATTAGTTTGGAGAGGAAAGAATATTTTGTTTGTATATTTGT
AGAAGAAGTTATTGTTGTTAAAAGTTAAGTTGGTTTTTTGATGTTTAGTGTGGATAG
GTTGAATAATAATTTATTTTTA**CG(CpG1)**TT**CG(CpG2)**GAGGGTTTATTTTATGTA
AATAGTTGAAAGGAATTGGATAAGATTGTAGTTTTT

Figure S1. Location of cg03546163 investigated by EpiTYPER MassARRAY. Locations of cg03546163 (CpG2, in yellow) and a second CpG site very close (CpG1, in red), investigated by EpiTYPER MassARRAY. Flanking primers are underlined. The CpG sites could not be tested individually due to MassARRAY technology constraints, but had to be tested jointly as a single unit: the methylation level is the cumulative value of all the sites within the CpG unit.

cg06633438

TTTTTTTAGTTTTGGGATTTTGAGATATTAGGATAG**CG(CpG1)**TTATTAATTTAGG
AGTAGTTTTTGTGTTTGTGTGTGTA**CG(CpG2)**TGTG**CG(CpG3)**TTTGTGTGTAGGT
TGGATTTTGATTTTAAAAGGGTTAGGATATGAAAAGAGA**CG(CpG4)**AG**CG(CpG5)**T
TTTAGATTTGAGGAAATATA**CG(CpG6)CG(CpG7)**TGGAAGA**CG(CpG8)**TG**CG(CpG**
9)CG(CpG10)GGGTT**CG(CpG11)**GGTTTTTTTTTGAGGGAGTGATAAGT**CG(CpG12**
)TGGAAAGTTTTAATATGG**CG(CpG13)**GG**CG(CpG14)**GGGGTGGGGG**CG(CpG15)**A
TAGGATTTGTTTGTTTTAG**CG(CpG16)**GTTTTAT**CG(CpG17)CG(CpG18)**TGATAA
TTGAGAGTTTAGAGGGGGA**CG(CpG19)**GAGAGGGAGGGAAGT**CG(CpG20)**GTTAGG
CG(CpG21)T**CG(CpG22)**GGGTATTTAGTTAGGTGG**CG(CpG23)**GGGTTTGAA**CG(C**
pG24)GGGTTT**CG(CpG25)**AAATGTTAAAGAAGAGATAGGTATTGGAGGGATTGTA
GTTATTGGGATATATTTGGTTAGTTGT**CG(CpG26)**AG**CG(CpG27)**ATGG**CG(CpG2**
8)TGGAAAGG**CG(CpG29)**TAGATGGGAGTTGTTTGGTTTGGGGTT**CG(CpG30)**GGA
GGG**CG(CpG31)**GTTTTGTTAGGGGAGGGTGTGGGATTTAGGTTATGGTTTGGTTGG
TTT

Figure S2. Location of cg06633438 investigated by EpiTYPER MassARRAY. Locations of cg06633438 (CpG10, in yellow) and other CpG sites very close (CpGn, in red), investigated by EpiTYPER MassARRAY. Flanking primers are underlined. The CpG sites could not be tested individually due to MassARRAY technology constraints, but had to be tested jointly as a single unit: the methylation level is the cumulative value of all the sites within the CpG unit.

Institutional Review Board Statement: Ethics approval and consent to participate: The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

References

1. Sekido, Y. Molecular pathogenesis of malignant mesothelioma. *Carcinogenesis* 2013, *34*, 1413–1419.
2. Rossini, M.; Rizzo, P.; Bononi, I.; Clementz, A.; Ferrari, R.; Martini, F.; Tognon, M.G. New Perspectives on Diagnosis and Therapy of Malignant Pleural Mesothelioma. *Front. Oncol.* 2018, *8*, 91.
3. Furuya, S.; Chimed-Ochir, O.; Takahashi, K.; David, A.; Takala, J. Global Asbestos Disaster. Int. J. Environ. Res. Public Health 2018, 15, 1000.

4. Straif, K.; Benbrahim-Tallaa, L.; Baan, R.; Grosse, Y.; Secretan, B.; El Ghissassi, F.; Bouvard, V.; Guha, N.; Freeman, C.; Galichet, L.; et al. A review of human carcinogens—Part C: Metals, arsenic, dusts, and fibres. Lancet Oncol. 2009, 10, 453–454.
5. Jaurand, M.C. Mechanisms of fiber-induced genotoxicity. Environ. Health Perspect. 1997, 105, 1073–1084.
6. Kelsey, K.T.; Yano, E.; Liber, H.L.; Little, J.B. The in vitro genetic effects of fibrous erionite and crocidolite asbestos. Br. J. Cancer 1986, 54, 107–114.
7. Iwatsubo, Y.; Pairon, J.C.; Boutin, C.; Ménard, O.; Massin, N.; Caillaud, D.; Orlowski, E.; Galateau-Salle, F.; Bignon, J.; Brochard, P. Pleural mesothelioma: Dose-response relation at low levels of asbestos exposure in a French population-based case-control study. Am. J. Epidemiol. 1998, 148, 133–142.
8. Howel, D.; Arblaster, L.; Swinburne, L.; Schweiger, M.; Renvoize, E.; Hatton, P. Routes of asbestos exposure and the development of mesothelioma in an English region. Occup. Environ. Med. 1997, 54, 403–409.
9. Kanherkar, R.R.; Bhatia-Dey, N.; Csoka, A.B. Epigenetics across the human lifespan. Front. Cell Dev. Biol. 2014, 2, 49.
10. Ferrante, D.; Bertolotti, M.; Todesco, A.; Mirabelli, D.; Terracini, B.; Magnani, C. Cancer mortality and incidence of mesothelioma in a cohort of wives of asbestos workers in Casale Monferrato, Italy. Environ. Health Perspect. 2007, 115, 1401–1405.
11. Guarrera, S.; Viberti, C.; Cugliari, G.; Allione, A.; Casalone, E.; Betti, M.; Ferrante, D.; Aspesi, A.; Casadio, C.; Grosso, F.; et al. Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. J. Thorac. Oncol. 2019, 14, 527–539.
12. Matullo, G.; Guarrera, S.; Betti, M.; Fiorito, G.; Ferrante, D.; Voglino, F.; Cadby, G.; Di Gaetano, C.; Rosa, F.; Russo, A.; et al. Genetic variants associated with increased risk of malignant pleural mesothelioma: A genome-wide association study. PLoS ONE 2013, 8, e61253.
13. Cugliari, G.; Catalano, C.; Guarrera, S.; Allione, A.; Casalone, E.; Russo, A.; Grosso, F.; Ferrante, D.; Viberti, C.; Aspesi, A.; et al. DNA Methylation of FKBP5 as Predictor of Overall Survival in Malignant Pleural Mesothelioma. Cancers 2020, 12, 3470.
14. Betti, M.; Aspesi, A.; Sculco, M.; Matullo, G.; Magnani, C.; Dianzani, I. Genetic predisposition for malignant mesothelioma: A concise review. Mutat. Res. 2019, 781, 1–10.

15. Guo, G.; Chmielecki, J.; Goparaju, C.; Heguy, A.; Dolgalev, I.; Carbone, M.; Seepo, S.; Meyerson, M.; Pass, H.I.. Whole-exome sequencing reveals frequent genetic alterations in *BAP1*, *NF2*, *CDKN2A*, and *CUL1* in malignant pleural mesothelioma. Cancer Res. 2015, 75, 264–269.
16. Fasanelli, F.; Baglietto, L.; Ponzi, E.; Guida, F.; Campanella, G.; Johansson, M.; Grankvist, K.; Johansson, M.; Assumma, M.B.; Naccarati, A.; et al. Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. Nat. Commun. 2015, 15, 10192.
17. Moore, L.D.; Le, T.; Fan, G. DNA methylation and its basic function. Neuropsychopharmacology 2013, 38, 23–38.
18. Bononi, A.; Napolitano, A.; Pass, H.I.; Yang, H.; Carbone, M. Latest developments in our understanding of the pathogenesis of mesothelioma and the design of targeted therapies. Expert. Rev. Respir. Med. 2015, 9, 633–654.
19. Vandermeers, F.; Neelature, S.; Costa, C.; Hubaux, R.; Cosse, J.P.; Willems, L. The role of epigenetics in malignant pleural mesothelioma. Lung Cancer 2013, 81, 311–318.
20. Zhang, X.; Tang, N.; Rishi, A.K.; Pass, H.I.; Wali, A. Methylation profile landscape in mesothelioma: Possible implications in early detection, disease progression, and therapeutic options. Methods Mol. Biol. 2015, 1238, 235–247.
21. Goto, Y.; Shinjo, K.; Kondo, Y.; Shen, L.; Toyota, M.; Suzuki, H.; Gao, W.; An, N.; Fujii, M.; Murakami, H.; et al. Epigenetic profiles distinguish malignant pleural mesothelioma from lung adenocarcinoma. Cancer Res. 2009, 69, 9073–9082.
22. Christensen, B.C.; Houseman, E.A.; Poage, G.M.; Godleski, J.J.; Bueno, R.; Sugarbaker, D.J.; Wiencke, J.K.; Nelson, H.H.; Marsit, C.J.; Kelsey, K.T. Integrated profiling reveals a global correlation between epigenetic and genetic alterations in mesothelioma. Cancer Res. 2010, 70, 5686–5694.
23. Fischer, J.R.; Ohnmacht, U.; Rieger, N.; Zemaitis, M.; Stoffregen, C.; Kostrzewa, M.; Buchholz, E.; Manegold, C.; Lahm, H. Promoter methylation of *RASSF1A*, *RARbeta* and *DAPK* predict poor prognosis of patients with malignant mesothelioma. Lung Cancer 2006, 54, 109–116.
24. Matosin, N.; Halldorsdottir, T.; Binder, E.B. Understanding the molecular mechanisms underpinning gene by environment interactions in psychiatric disorders: The *FKBP5* model. Biol. Psychiatry 2018, 83, 821–830.

25. Kang, C.B.; Hong, Y.; Dhe-Paganon, S.; Sup Yoon, H. FKBP family proteins: Immunophilins with versatile biological functions. Neurosignals 2008, 16, 318–325.
26. Li, L.; Lou, Z.; Wang, L. The role of *FKBP5* in cancer aetiology and chemoresistance. Br. J. Cancer 2011, 104, 19–23.
27. Romano, S.; D'Angelillo, A.; Romano, M.F. Pleiotropic roles in cancer biology for multifaceted proteins FKBPs. Biochim. Biophys. Acta 2015, 1850, 2061–2068.
28. Staibano, S.; Mascolo, M.; Ilardi, G.; Siano, M.; De Rosa, G. Immunohistochemical analysis of *FKBP51* in human cancers. Curr. Opin. Pharmacol. 2011, 11, 338–347.
29. Zannas, A.S.; Jia, M.; Hafner, K.; Baumert, J.; Wiechmann, T.; Pape, J.C.; Arloth, J.; Ködel, M.; Martinelli, S.; Roitman, M.; et al. Epigenetic upregulation of *FKBP5* by aging and stress contributes to NF-kappaB-driven inflammation and cardiovascular risk. Proc. Natl. Acad. Sci. USA 2019, 116, 11370–11379.
30. Yoshikawa, Y.; Sato, A. Biallelic germline and somatic mutations in malignant mesothelioma: Multiple mutations in transcription regulators including *mSWI/SNF* genes. Int. J. Cancer 2015, 136, 560–571.
31. Zhou, J.; Ng, Y.; Chng, W.J. ENL: Structure, function, and roles in hematopoiesis and acute myeloid leukemia. Cell Mol. Life Sci. 2018, 75, 3931–3941.
32. Canino, C.; Luo, Y.; Marcato, P.; Blandino, G.; Pass, H.I.; Cioce, M. *STAT3-NFkB/DDIT3/CEBPbeta* axis modulates *ALDH1A3* expression in chemoresistant cell subpopulations. Oncotarget 2015, 6, 12637–12653.
33. Bagger, F.O.; Kinalis, S.; Rapin, N. BloodSpot: A database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. Nucleic Acids Res. 2019, 47, 881–885.
34. Schins, R.P.; Donaldson, K. Nuclear Factor Kappa-B Activation by Particles and Fibers. Inhal. Toxicol. 2000, 12, 317–326.
35. Tan, Z.; Lu, W.; Li, X.; Yang, G.; Guo, J.; Yu, H.; Li, Z.; Guan, F. Altered N-Glycan expression profile in epithelial-to-mesenchymal transition of NMuMG cells revealed by an integrated strategy using mass spectrometry and glycogene and lectin microarray analysis. J. Proteome Res. 2014, 13, 2783–2795.

36. Kale, V.P.; Hengst, J.A.; Desai, D.H.; Dick, T.E.; Choe, K.N.; Colledge, A.L.; Takahashi, Y.; Sung, S.S.; Amin, S.G.; Yun, J.K. A novel selective multikinase inhibitor of ROCK and MRCK effectively blocks cancer cell migration and invasion. Cancer Lett. 2014, 354, 299–310.
37. Tanrikulu, A.C.; Abakay, A.; Komek, H.; Abakay, O. Prognostic value of the lymphocyte-to-monocyte ratio and other inflammatory markers in malignant pleural mesothelioma. Environ. Health Prev. Med. 2016, 21, 304–311.
38. Zhao, Z.S.; Manser, E. PAK and other Rho-associated kinases—Effectors with surprisingly diverse mechanisms of regulation. Biochem. J. 2005, 386 Pt 2, 201–214.
39. Nishimura, Y.; Kumagai-Takei, N.; Matsuzaki, H.; Lee, S.; Maeda, M.; Kishimoto, T.; Fukuoka, K.; Nakano, T.; Otsuki, T. Functional alteration of natural killer cells and cytotoxic T lymphocytes upon asbestos exposure and in malignant mesothelioma patients. Biomed. Res. Int. 2015, 238431.
40. Maeda, M.; Nishimura, Y.; Kumagai, N.; Hayashi, H.; Hatayama, T.; Katoh, M.; Miyahara, N.; Yamamoto, S.; Hirastuka, J.; Otsuki, T. Dysregulation of the immune system caused by silica and asbestos. J. Immunotoxicol. 2010, 7, 268–278.
41. Miura, Y.; Nishimura, Y.; Katsuyama, H.; Maeda, M.; Hayashi, H.; Dong, M.; Hyodoh, F.; Tomita, M.; Matsuo, Y.; Uesaka, A.; et al. Involvement of IL-10 and Bcl-2 in resistance against an asbestos-induced apoptosis of T cells. Apoptosis 2006, 11, 1825–1835.
42. Cristaudo, A.; Bonotti, A.; Guglielmi, G.; Fallahi, P.; Foddis, R. Serum mesothelin and other biomarkers: What have we learned in the last decade? J. Thorac. Dis. 2018, 10, 353–359.
43. Andujar, P.; Wang, J.; Descatha, A.; Galateau-Sallé, F.; Abd-Alsamad, I.; Billon-Galland, M.A.; Blons, H.; Clin, B.; Danel, C.; Housset, B.; et al. p16INK4A inactivation mechanisms in non-small-cell lung cancer patients occupationally exposed to asbestos. Lung Cancer 2010, 67, 23–30.
44. Dianzani, I.; Gibello, L.; Biava, A.; Giordano, M.; Bertolotti, M.; Betti, M.; Ferrante, D.; Guarrera, S.; Betta, G.P.; Mirabelli, D.; et al. Polymorphisms in DNA repair genes as risk factors for asbestos-related malignant mesothelioma in a general population study. Mutat. Res. 2006, 599, 124–134.
45. Betti, M.; Ferrante, D.; Padoan, M.; Guarrera, S.; Giordano, M.; Aspesi, A.; Mirabelli, D.; Casadio, C.; Ardissone, F.; Ruffini, E.; et al. *XRCC1* and *ERCC1* variants modify malignant mesothelioma risk: A case-control study. Mutat. Res. 2011, 708, 11–20.

46. Betti, M.; Casalone, E.; Ferrante, D.; Aspesi, A.; Morleo, G.; Biasi, A.; Sculco, M.; Mancuso, G.; Guarrera, S.; Righi, L.; et al. Germline mutations in DNA repair genes predispose asbestos-exposed patients to malignant pleural mesothelioma. Cancer Lett. 2017, 405, 38–45.
47. Ferrante, D.; Mirabelli, D.; Tunesi, S.; Terracini, B.; Magnani, C. Pleural mesothelioma and occupational and non-occupational asbestos exposure: A case-control study with quantitative risk assessment. Occup. Environ. Med. 2016, 73, 147–153.
48. Du, P.; Zhang, X.; Huang, C.C.; Jafari, N.; Kibbe, W.A.; Hou, L.; Lin, S.M. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinform. 2010, 11, 587.
49. Lehne, B.; Drong, A.W.; Loh, M.; Zhang, W.; Scott, W.R.; Tan, S.T.; Afzal, U.; Scott, J.; Jarvelin, M.R.; Elliott, P.; et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015, 16, 37.
50. Campanella, G.; Polidoro, S.; Di Gaetano, C.; Fiorito, F.; Guarrera, S.; Krogh, V.; Palli, D.; Panico, S.; Sacerdote, C.; Tumino, R.; et al. Epigenetic signatures of internal migration in Italy. Int. J. Epidemiol. 2015, 44, 1442–1449.
51. Di Gaetano, C.; Voglino, F.; Guarrera, S.; Fiorito, G.; Rosa, F.; Di Blasio, A.M.; Manzini, P.; Dianzani, I.; Betti, M.; Cusi, D.; et al. An overview of the genetic structure within the Italian population from genome-wide data. PLoS ONE 2012, 7, e43759.
52. Houseman, E.A.; Accomando, W.P.; Koestler, D.C.; Christensen, B.C.; Marsit, C.J.; Nelson, H.H.; Wiencke, J.K.; Kelsey, K.T. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinform. 2012, 13, 86.
53. Ehrich, M.; Nelson, M.R.; Stanssens, P.; Zabeau, M.; Liloglou, T.; Xinarianos, G.; Cantor, C.R.; Field, J.K.; van den Boom, D. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc. Natl. Acad. Sci. USA 2005, 102, 15785–15790.

Study2

DNA methylation of *FKBP5* as predictor of overall survival in Malignant Pleural Mesothelioma

Simple Summary: Our study is the first one to investigate DNA methylation changes in WBCs from easily accessible peripheral blood as MPM survival biomarkers. The Cox proportional hazards regression model highlighted that cg03546163 is an independent marker of prognosis in MPM patients with a better performance than traditional inflammation-based scores such as lymphocyte-to-monocyte ratio (LMR). Biological validation and replication showed that epigenetic changes at the *FKBP5* gene were robustly associated with OS in MPM cases. The identification of simple and valuable prognostic markers for MPM will enable clinicians to select patients who are most likely to benefit from aggressive therapies and avoid subjecting non-responder patients to ineffective treatment.

Abstract: Malignant pleural mesothelioma (MPM) is an aggressive tumor with median survival of 12 months and limited effective treatments. The scope of this study was to study the relationship between blood DNA methylation (DNAm) and overall survival (OS) aiming at a non-invasive prognostic test. We investigated a cohort of 159 incident asbestos exposed MPM cases enrolled in an Italian area with high incidence of mesothelioma. Considering 12 months as a cut-off for OS, epigenome-wide association study (EWAS) revealed statistically significant (p value = 7.7E-09) OS-related differential methylation of a single-CpG (cg03546163), located in the 5'UTR region of *FKBP5* gene. This is an independent marker of prognosis in MPM patients with a better performance than traditional inflammation-based scores such as lymphocyte-to-monocyte ratio (LMR). Cases with DNAm < 0.45 at the cg03546163 had significantly poor survival compared with those

showing DNAm ≥ 0.45 (mean: 243 versus 534 days; p value< 0.001). Epigenetic changes at the *FKBP5* gene were robustly associated with OS in MPM cases. Our results showed that blood DNA methylation levels could be promising and dynamic prognostic biomarkers in MPM.

Keywords: malignant pleural mesothelioma; asbestos exposure; DNA methylation; lymphocyte-to-monocyte ratio; epigenome-wide analysis; survival analysis

---

1. Introduction

Malignant pleural mesothelioma (MPM) is an aggressive tumor. The disease usually develops after a long latency (20-40 years) following asbestos exposure [1]. Although MPM is considered a rare malignancy (prevalence 1–9/100,000), about 40,000 deaths have been estimated to occur each year globally [2, 3]. The World Health Organization estimates that 125 million people annually around the world are exposed to asbestos. The International Agency for Research on Cancer confirmed that all fibrous forms of asbestos are carcinogenic to humans, causing mainly mesothelioma, respiratory-tract tumors, mesothelioma, and cancer at other tissue sites [4].

The prognosis of MPM is poor with a median survival of about 12 months from the diagnosis [5].

Generally, the first-line treatment is a combination of a multi-targeted anti folate (pemetrexed or raltitrexed) drug and a platinum compound (cisplatin or carboplatin) [6]. Currently, only a single randomized trial demonstrated an increase in survival time when comparing cisplatin and pemetrexed versus cisplatin alone [7]; unfortunately, most patients became resistant to this treatment and relapsed rapidly. No oncogenic driver has been identified and molecular pathways leading to MPM have also not yet been clearly determined. Other therapeutic strategies such as immunotherapy are promising but require further investigation and improvement [8].

Recent research on the pathogenesis of MPM indicated that i) both genetic and epigenetic alterations contribute to asbestos-induced tumorigenesis [9, 10], ii) inflammation-based prognostic scores that include lymphocyte counts are associated with survival [11].

MPM has a low frequency of protein-altering mutations (~25 mutations per tumor), compared to many other tumors [12]. Moreover, germline mutations in different genes mainly involved in DNA damage repair confer moderate-to-high genetic risk of MPM development [13]. The BAP1-tumor predisposition syndrome is the most studied genetic condition associated with MPM development and is caused by mutations in the BRCA1-associated protein 1 (BAP1) gene [13].

In the last 10 years, epigenetic markers, such as DNA methylation (DNAm) and microRNAs (miRNAs), have gained popularity as possible early diagnostic and prognostic biomarkers in cancer research, including MPM. While genetic markers may differ from case to case in most cancer patients (i.e., each patient may carry a different mutation within the same gene), different subjects show variable levels of epigenetic biomarkers in specific target regions and different tissues depending on disease status [14].

DNA methylation is one of the epigenetic factors that can be altered in cancer tissues [15]. However, regarding mechanisms and clinical outcome of epigenetic derangements in MPM, less information is available [16]. Although DNAm is stable, it can be modified throughout life by several factors such as ageing, lifestyle, environmental exposures, and diseases. It thus represents an adaptive phenomenon linking environmental factors and the development of pathologic phenotypes such as cancers. DNAm changes are considered to possibly play a role in MPM progression, and have therefore been suggested as a potential tool for prognosis [18].

The fact that epigenetic modifications, unlike genetic changes, are potentially reversible, may open new perspectives for patient clustering and novel therapeutic options. A reliable prognostic biomarker that offers high sensitivity and specificity would be a major advancement for MPM. Blood-based biomarkers that have been explored in MPM include megakaryocyte potentiating factor (an alternative cleavage product of the mesothelin

precursor protein) [19]; fibulin 3 is also found in pleural fluid, and high levels appear to correlate with advanced disease [20].

Considering clinical end-point, low pleural fluid glucose and high C-reactive protein and pleural thickening represent the main prognosis factors [21]. Recent studies confirm that a combination of epigenetic alterations is superior with respect to an only genetic approach on overall survival [17].

This study was undertaken with the goal of better characterizing the MPM overall survival (OS) evaluating the potential predictive value of peripheral blood DNAm profiles. The second goal was the comparison of the DNAm prognostic performance with the broadly used lymphocyte-to-monocyte ratio (LMR) method.

2. Results

2.1. Epigenome-wide association study

EWAS revealed a statistically significant hypo-methylated single-CpG (cg03546163) in *FKBP5* gene in the low survival group after Bonferroni post-hoc correction (**Figure 1**).

Bootstrap was computed to estimate the measures of accuracy, using random sampling methods.

Other 5 CpGs in *FKBP5* gene showed hypomethylation in poor MPM survivors, with un-adjusted p value < 0.05 (**Table 2**); instead, no CpGs in *FKBP5* gene showed statistically significant hypermethylation in poor MPM survivors.

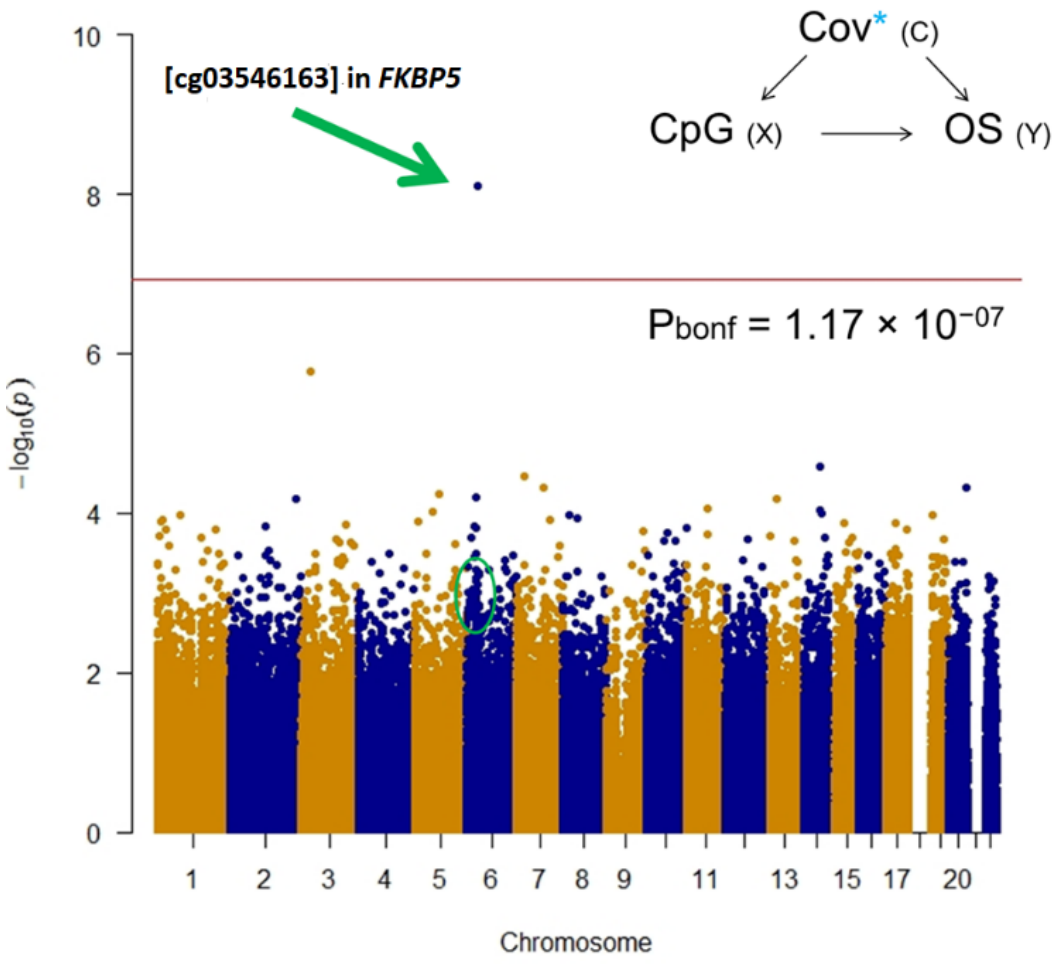

Figure 1. Manhattan plot for EWAS test on 450k single CpGs. Overall survival was used as a dependent variable considering 12 months as cut-off adjusting for age, gender, histological subtype, asbestos exposure, WBCs estimation, population stratification and technical variability. Bonferroni post hoc line highlights statistically significant differences on OS at single CpG level.

Table 2. Differential DNAm analyses at *FKBP5* gene ordered by effect size (low survival group was used as reference). Information about single CpGs including location-related values and model outputs (effect size, standard error, p values).

TargetID	CHR	UCSC_REFGENE_GROUP	ENHANCER	Probe_start	Probe_end	Closest_TSS	Dist_closest_TSS	Gene_name	Effect_size	SE	p_value	Bonf	Sign
cg03546163	6	5'UTR;5'UTR;5'UTR;5'UTR	NA	35654313	35654363	35656691	2329	*FKBP5*	0.12	0.02	7.71E-09	0.0032804	*§
cg00052684	6	5'UTR	TRUE	35694195	35694245	35696396	2152	*FKBP5*	0.04	0.02	0.0145890	1	*
cg00130530	6	5'UTR;TSS1500;TSS1500;TSS1500	NA	35657152	35657202	35656718	-483	*FKBP5*	0.03	0.01	0.0014908	1	*
cg19226017	6	TSS1500;Body	NA	35697185	35697235	35696396	-788	*FKBP5*	0.03	0.01	0.0216391	1	*
cg08915438	6	TSS1500;Body	NA	35697709	35697759	35696396	-1362	*FKBP5*	0.02	0.01	0.0507796	1	*
cg14642437	6	5'UTR;5'UTR;5'UTR;5'UTR	NA	35652471	35652521	35656691	4171	*FKBP5*	0.02	0.01	0.0307181	1	*
cg25114611	6	TSS1500;Body	NA	35696820	35696870	35696396	-473	*FKBP5*	0.02	0.01	0.0804351	1	
cg16052510	6	Body;Body;Body;Body	TRUE	35603093	35603143	35656691	53549	*FKBP5*	0.01	0.01	0.2017837	1	
cg03591753	6	5'UTR	NA	35659141	35659191	35656718	-2422	*FKBP5*	0.01	0.01	0.0712878	1	
cg23416081	6	5'UTR	TRUE	35693573	35693623	35696396	2824	*FKBP5*	0.01	0.01	0.3001815	1	
cg19014730	6	5'UTR;5'UTR;5'UTR;5'UTR	TRUE	35635985	35636035	35656691	20707	*FKBP5*	0.01	0.01	0.5109240	1	
cg20813374	6	5'UTR;TSS1500;TSS1500;TSS1500	NA	35657130	35657180	35656718	-461	*FKBP5*	0.01	0.01	0.5386224	1	

cg07061368	6	5'UTR;5'UTR;5'UTR;5'UTR	TRUE	35631736	35631786	35656691	24956	*FKBP5*	0.00	0.01	0.4407199	1
cg08636224	6	5'UTR;TSS1500;TSS1500;TSS1500	NA	35657871	35657921	35656718	-1202	*FKBP5*	0.00	0.00	0.18248273	1
cg01294490	6	TSS200;TSS200;5'UTR;TSS1500	NA	35656906	35656956	35656718	-187	*FKBP5*	0.00	0.01	0.4213002	1
cg07485685	6	5'UTR;Body	NA	35696060	35696110	35696396	336	*FKBP5*	0.00	0.00	0.8479419	1
cg14284211	6	Body;Body;Body;Body	TRUE	35570224	35570274	35656691	86468	*FKBP5*	0.00	0.01	0.9743447	1
cg17030679	6	5'UTR;Body;1stExon	NA	35696300	35696350	35696396	97	*FKBP5*	0.00	0.00	0.9557194	1
cg00862770	6	5'UTR;5'UTR;5'UTR;5'UTR	NA	35655764	35655814	35656691	928	*FKBP5*	0.00	0.00	0.9399041	1
cg00140191	6	5'UTR;5'UTR;5'UTR;5'UTR	NA	35656193	35656243	35656691	450	*FKBP5*	0.00	0.00	0.8823881	1
cg00610228	6	5'UTR;Body	NA	35695934	35695984	35696396	463	*FKBP5*	0.00	0.00	0.873762	1
cg07633853	6	Body;Body;Body;Body	TRUE	35569421	35569471	35656691	87221	*FKBP5*	0.00	0.01	0.9654276	1
cg10300814	6	Body;Body;Body;Body	TRUE	35565066	35565116	35480646	-84469	*TULP1*	0.00	0.00	0.6206779	1
cg16012111	6	TSS200;TSS200;TSS200;5'UTR	NA	35656758	35656808	35656718	-39	*FKBP5*	0.00	0.00	0.5190471	1
cg06937024	6	5'UTR;Body	NA	35695440	35695490	35696396	908	*FKBP5*	0.00	0.00	0.1350045	1
cg08586216	6	5'UTR;5'UTR;5'UTR;5'UTR	TRUE	35612301	35612351	35656691	44341	*FKBP5*	0.00	0.00	0.1056313	1
cg17085721	6	5'UTR;5'UTR;5'UTR;5'UTR	TRUE	35645291	35645341	35656691	11351	*FKBP5*	0.00	0.00	0.2115825	1
cg02665568	6	Body;Body;Body	NA	35544468	35544518	35480646	-63821	*TULP1*	-0.01	0.01	0.2947576	1

cg15929276	6	5'UTR	TRUE	35687456	35687506	35696396	8940	*FKBP5*	-0.01	0.01	0.4559690	1
cg06087101	6	Body;3'UTR;Body;Body	NA	35551882	35551932	35480646	-71285	*TULP1*	-0.02	0.02	0.2037838	1

Low survival group was set as reference. Adjustment covariates: age, gender, asbestos exposure, histological subtype, smoke, population stratification, WBCs estimation and technical variability. *: statistically significant at p value< 0.05; §: statistically significant at Bonferroni and FDR post hoc adjustments.

2.2. Survival analysis

CpG sites and LMR were considered as predictors in the regression model. Categorical variables (quantile information) were used.

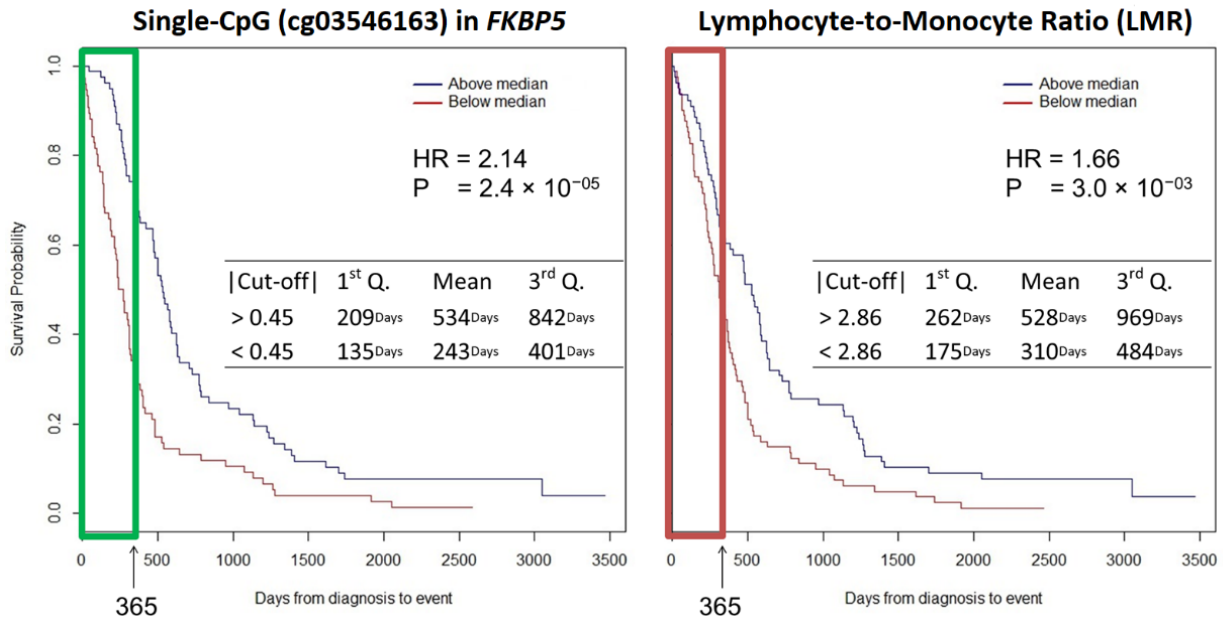

Figure 2. K-M survival curves show a) cg03546163: patients with a DNAm < 0.45 had significantly poor survival compared with a DNAm ≥ 0.45 (mean, 243 versus 534 days; p value< 0.001); and b) LMR: patients with values < 2.86 had significantly poor survival compared with patients with values ≥ 2.86 (mean, 310 versus 528 days; p value< 0.001). cg03546163 is an independent marker of prognosis in patients with MPM and performs better than LMR ($HR_{cg03546163}$= 2.14 vs HR_{LMR}= 1.66).

Cox model was computed considering the same list of covariates included in the EWAS. Patients with DNAm < 0.45 at the cg03546163 had significantly poorer survival compared with subjects with DNAm ≥ 0.45 (mean, 243 versus 534 days; p value < 0.001). Survival at the 1st and the 3rd Quartiles was 135 versus 209 days and 401 versus 842 days, respectively, comparing patients with single CpG DNAm < 0.45 with those with single CpG DNAm ≥ 0.45. The Multivariate analysis showed that cg03546163 DNAm at *FKBP5* was independently associated with OS. Kaplan–Meier curves revealed that a decrease of

methylation at cg03546163 (< 0.45) was significantly associated with worse OS (HR = 2.14 p value< 0.0001) (**Figure 2a**).

Patients with LMR < 2.86 had significantly poorer survival compared with patients with LMR ≥ 2.86 (mean, 310 versus 528 days; p value< 0.001). Survival at 1st Quartile was 175 versus 262 days whereas at 3rd Quartile was 484 versus 969 days comparing patients with LMR < 2.86 with those with LMR > 2.86. LMR was independently associated with OS: Kaplan–Meier curves showed that decreased LMR (< 2.86) was significantly associated with decreased OS (HR= 1.66; p value<0.01) (**Figure 2b**).

Histological subtype (epithelioid versus non-epithelioid), smoking status (current, never and former) and asbestos exposure showed no statistically significant results on survival.

2.3. Validation and replication

The statistically significant association between cg03546163 DNAm and OS was confirmed in an independent sample of patients (replication) and using a different targeted DNAm analysis technique (validation). A sample of 133 MPM cases (58 low survivors and 75 high survivors) was recruited and stratified in low and high OS considering the same cut-off (365 days).

The same model used for the discovery phase was performed. Patients with below median OS had significantly lower DNAm at the cg03546163 compared with those with above median OS (mean, 188 versus 786 days; p value< 0.001). The 1st Quartile was 113 versus 482 days and the 3rd Quartile was 262 versus 862 days comparing patients with DNAm difference (reference above median OS, MD: -0.04, 95%CI: -0.07|-0.01, p value: 0.04) at the cg03546163. The Multivariate analysis confirmed that cg03546163 DNAm at *FKBP5* was independently associated with OS.

3. Discussion

A growing number of studies reported on the identification of epigenetic prognostic biomarkers in several cancers [17, 35].

This study focused on the exploration of epigenetic factors related to MPM survival in MPM incident cases from Piedmont (Italy), a region with a well documented history of asbestos exposure [36].

More than 450k methylation sites were evaluated in DNA from whole blood looking for new insights related to overall survival in MPM. The main result was the hypomethylation of a single CpG (cg03546163) in the 5' UTR region of *FKBP5* gene in patients with poorer survival compared to patients with longer survival; it also showed to be an independent marker of prognosis in MPM patients. This result has been replicated in a different series of patients belonging to the same cohort using the Sequenom Quantitative DNAm analysis.

In general, a combination of epigenetic and clinical factors is under investigation in clinical prognosis and survival, including tumor histology, gender, hemoglobin level, platelet and white blood cell count, and lactate dehydrogenase level [37].

Recently, due to the important role of inflammation in the development of MPM, several studies investigated the effect of inflammation-based biomarkers on the prognosis [11, 35]. We selected the LMR for the comparison because its performance was previously reported to be higher than other inflammation-based markers in MPM [38].

To validate the prognostic value of the observed CpG methylation site, we compared our result with the LMR score.

Kaplan–Meier survival curves for MPM patients highlighted cg03546163 methylation at *FKBP5* gene as a prognostic factor superior to the LMR score.

The FKBP Prolyl Isomerase 5 (*FKBP5*), also known as *FK506* binding protein 51 (*FKBP51*) is a member of immunophilin protein family, which contributes to the immunoregulation and to the basic cellular processes involving protein folding and trafficking. Together with other members of the FKBPs family, this protein participates in transcriptional complexes and acts as a co-transcription factor.

Although no studies have investigated the methylation of *FKBP5* as a prognostic factor in MPM, a growing number of whole-blood studies investigated its DNA methylation levels

in order to explain the impact of environmental stress in the aetiology and treatment of several diseases [39]. Interestingly, in a recent study on the Behcet's disease (BD) it was demonstrated that hypomethylation in the 5'UTR region (including cg03546163) of *FKBP5* characterized cases and it was strongly associated with high gene expression, suggesting a possible role of DNA methylation in the pathogenesis [40].

Other five single CpGs at *FKBP5* showed hypomethylation in poor survivors: this evidence supports the potential overall contribution of *FKBP5* on the patient differentiation by OS.

In several human cancer tissues, a relevant role for *FKBP5* in sustaining cancer cell growth and aggressiveness has been documented. In particular, for glioma [41], prostate cancer and melanoma [42], a strict correlation between protein abundance and aggressiveness has been demonstrated.

Probably, the relationship between *FKBP5* and tumor progression and aggressiveness, is represented by its implication in NF-kB and AKT signaling pathways, with key roles in tumorigenesis and response to antineoplastic chemotherapy [43].

Moreover, a well characterized anti-apoptotic effect is mediated by NF-κB transcription factors and *FKBP5* has documented anti-apoptotic effects: recent studies hypothesized that *FKBP5* could promote inflammation, by activating the master immune regulator NF-kB, after an epigenetically upregulation due to aging and stress [44, 45].

Previous studies conducted on various cancer types, showed that upregulation of *FKBP5* gene expression is associated with drug resistance [46]. In a study on ovarian cancer cell lines, the up-regulation of *FKBP5* increased the resistance to chemotherapeutic agents, whereas the gene silencing sensitized ovarian cancer cells to taxol [47]. In the present study we could not evaluate *FKBP5* gene expression due to the lack of available RNA, which was not collected in the study. However, this should be further addressed and verified in future studies.

One study demonstrated that overexpression of *FKBP5* increased the chemosensitivity through the AKT pathway [44]. Similar research supported this observation making

FKBP5 an effective biomarker for sensitivity to chemotherapy; patient responses to chemotherapy may be determined by the variation in *FKBP5* levels [48].

Limitation of the study

Being able to identify the direction of causality will greatly aid in determining the usefulness of epigenetic variation.

Leukocyte DNA methylation could mainly represent a nonspecific marker related to a general inflammatory status due to the presence of a tumor rather than a specific MPM biomarker and further studies should be carried out to support our findings.

As an additional limitation, we have therapy information only for a small subset of patients and we could not test treatment-specific OS differences in relation to *FKBP5* methylation levels.

4. Material and methods

4.1. Study population

Study subjects belong to a wider ongoing collaborative study on MPM, which is actively enrolling MPM cases in the municipalities of Casale Monferrato (Piedmont region, Italy), an area with an exceptionally high incidence of mesothelioma caused by widespread asbestos exposure for locals, both occupational and environmental, due to the asbestos-cement Eternit plant that was operational until 1986 [22]. Additional MPM cases were recruited in the main hospitals of the municipalities of Turin, Novara and Alessandria (Piedmont region, Italy). The study included incident MPM cases diagnosed between 2000 and 2010 after histological and/or cytological confirmation of MPM diagnosis [23, 24].

No peritoneal cases were considered with the aim to better identify epigenetics characteristics of MPM.

In the present study, 159 MPM cases belonging to a larger case-control study with genetic [10, 25] and blood DNAm dat [9] were selected according to the following criteria: i) availability of good quality DNA at the time of the analyses and ii) asbestos exposure above the background level, as defined in [26]. Additional 133 independent samples from the same cohort have been included for the validation/replication analyses.

Descriptive information of MPM patients are shown in **Table 1**. Median survival (365 days) was used as a cut-off value to stratify patients in high and low survivors.

No differences in categorical (centre, gender, smoke, histotype) and continuous (Asbestos exposure, WBCs composition) variables among low and high survivors were found.

Our study complies with the Declaration of Helsinki principles and conforms to ethical requirements. All volunteers signed an informed consent form at enrollment. The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

Table 1. Descriptive information of MPM patients. Median survival (365 days) was used as a cut-off value to stratify patients in high and low survivors.

Categorical variable	**level**	**low OS (n = 79)**		**high OS (n = 80)**	
		N	%	N	%
Centre	Casale	50	63.3	46	57.5
	Torino	29	36.7	34	42.5
Gender	Males	59	74.7	50	62.5
	Females	20	25.3	30	37.5
Smoke	Current	20	26.3	8	10.3
	Former	24	31.6	29	37.2
	Never	32	42.1	41	52.6
Histotype	Epithelioid	44	55.7	61	76.3
	Sarcomatoid	14	17.7	2	2.5
	Biphasic	17	21.5	11	13.8
	Undefined	2	2.5	1	1.3
	Not known	2	2.5	5	6.3

Continuous variable	level	low OS		high OS	
		Mean	SD	Mean	SD
Overall Survival (days)	-	198.7	101.6	957.8	698.7
Age (years)	-	67.7	12.4	67.5	9.6
Asbestos Exp. (norm)	-	1.4	1.5	1.5	1.9
CD8T (%)	-	2.9	4.5	3	3.4
CD4T (%)	-	6.8	5.3	8.8	5.4
Natural Killer (%)	-	4.9	4.9	6.3	4.1
B cell (%)	-	6.1	2.8	6.4	2.7
Monocytes (%)	-	8.1	4.1	7.6	4.4
Granulocytes (%)	-	75	13	72	10

Asbestos Exposure (occupational, environmental and domestic) was normalized considering frequency, duration and intensity.

4.2. Exposure assessment

For all subjects, occupational history and lifestyle habits information were collected through interviewer-administered questionnaires filled out at enrollment during a face-to-face interview. Job titles were coded according to the International Standard Classification of Occupation [26] and according to the Statistical Classification of Economic Activities in the European Community.

Frequency, duration and intensity of exposure were estimated, then a cumulative exposure index was computed. The evaluation of asbestos exposure (occupational, environmental and domestic) was conducted by an experienced occupational epidemiologist. For the selection criteria and descriptive evaluation, asbestos exposure doses (fibers/mL years) were rank transformed to remove skewness.

4.3. Blood DNAm analysis

Genomic DNA was extracted from whole blood collected in EDTA by an on-column DNA purification method (QIAamp DNA Blood Mini Kit, QIAGEN GmbH, Germany),

according to manufacturer's instructions. DNA integrity was checked by an electrophoretic run in standard TBE 0.5X buffer on a 1% low melting agarose gel (Sigma-Aldrich GmbH, Germany); DNA purity and concentration were assessed by a NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Five hundred nanograms of genomic DNA for each sample were bisulfite treated (EZ-96 DNA Methylation-Gold Kit, Zymo Research Corporation, Irvine, CA, USA) to convert un-methylated cytosine to uracil. Cases were randomly and blindly distributed across conversion plates.

The Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, CA, USA) was used to measure the methylation level of more than 485,000 individual CpG loci at a genome-wide resolution [27].

Twelve samples were analyzed on each BeadChip. As a "position effect" was reported for Illumina Methylation BeadChips, each sample position on the BeadChip was completely random as well. We further verified the randomization of the position on each BeadChip was effective by checking for a position effect, and we found no occurrence of it. BeadChips were processed according to manufacturer protocols. Data were inspected with the dedicated GenomeStudio software v2011.1 with Methylation module 1.9.0 (Illumina Inc., San Diego, CA), and quality checked as previously described [28].

4.4. Beta-value extraction

Raw DNAm data were analyzed with the R package *'methylumi'*. The average methylation value at each locus was computed as the ratio of the intensity of the methylated signal over the total signal (un-methylated + methylated) [29]. Beta-values represent the percentage of methylation at each individual CpG locus, ranging from 0 (no methylation) to 1 (full methylation).

We excluded from the analyses: i) single Beta-values with detection p-value ≥ 0.01; ii) CpG loci with missing Beta-values in more than 20% of the assayed samples; iii) CpG loci detected by probes containing SNPs with MAF ≥ 0.05 in the CEPH (Utah residents with

ancestry from northern and western Europe, CEU) population; iv) samples with a global call rate ≤ 95%. Lastly, CpGs on chromosomes X and Y were excluded from the analysis.

4.5. Batch effect, population stratification and White Blood Cells estimations

To account for methylation assay variability and batch effects, we corrected all differential methylation analyses for "control probes" Principal Components (PCs). Using PCs assessed by principal component analysis of the BeadChip's built-in control probes as a correction factor for statistical analyses of microarray data is a method that allows to account for the technical variability of several steps in the DNAm analysis, from the bisulfite conversion to BeadChip processing [30].

Geographic origins of subjects may influence DNAm profiles. To consider this source of potential bias, we took advantage of the whole genome genotyping dataset from the same subjects from our previous study [10]. The first PCs calculated based on genome-wide genotyping were shown to correlate with different geographic origins of people [31, 32].

White Blood Cells (WBC) subtype percentages calculated based on genome-wide methylation data [33]. for each subject were extracted. This method quantifies the normally mixed composition of leukocytes beyond what is possible by simple histological or flow cytometric assessments. In a diverse array of diseases and following numerous immune-toxic exposures, leukocyte composition will critically inform the underlying immune-biology to most chronic medical conditions. Then, it is necessary to extract and control for the percentage of involved WBCs with the aim to infer a functional biological pathway.

LMR score was calculated from the DNAm-estimated WBCs by dividing the total lymphocyte count by the monocyte count.

4.6. Epigenome-wide association study (EWAS)

Association test was used to analyze the mean differences (MD) at single-CpG methylation between low and high survival. Multiple regression analysis adjusted for age, gender, histological subtype, asbestos exposure, smoke, estimated WBCs, population stratification

(first 2 PCs) and technical variability (first 10 PCs) was implemented. For multiple comparisons tests, Bonferroni p value≤ 0.05 was considered statistically significant.

Using random sampling methods, bootstrap was implemented to estimate the measures of accuracy defined in terms of bias, variance, confidence intervals and prediction error. Bootstrap is also an appropriate way to control and check the stability of the results. The bias-corrected and accelerated (BCa) bootstrap interval was calculated with regard to single CpGs.

4.7. Survival analysis

The survival time was determined as the time between the date of diagnosis and the date of death. If patients were still alive at the last follow-up (2016), survival was defined as the time from the date of diagnosis until June 2016. The time and the median event times with 95% confidence intervals were estimated according to the Kaplan–Meier method. The proportional hazards regression model was used for both the univariate and multivariate analyses (Cox model).

Comparison of OS curves was performed using two-tailed log-rank tests with a 0.05 level of significance. Only variables with p value < 0.1 in the univariate analysis were included in the final model for the multivariate analysis. In the Cox regression analysis, the backward conditional method (stepwise-AIC) was used. LMR and CpG sites were considered as predictors in the regression model.

4.8. Statistical Power

To ensure a power of the study greater than 80% (two-tailed test at 0.05 alpha error), only CpGs with mean difference (MD) of Beta-value between low and high survival of ≥ |0.035| were selected. Covariates were included step-by-step in sensitivity analysis to validate the association output considering effect size, standard error, 95% confidence interval and p value variations.

CpGs with Bonferroni p value (P_{FDR}) ≤ 0.05 underwent gene set enrichment analysis to identify pathways potentially affected by MPM related methylation changes.

All statistical analyses were conducted using the open source software R (4.0.2).

4.9. Validation and replication

Sequenom MassARRAY for the DNAm signal validation and replication was used. In detail, the EpiTYPER assay (Sequenom) uses a MALDI-TOF mass spectrometry-based method to quantitatively assess the DNA methylation state of CpG sites of interest [34]. DNA (500 ng) was bisulfite-converted using the EZ-96 DNA Methylation Kit (Zymo Research) with the following modifications: incubation in CT buffer for 21 cycles of 15 minutes at 55°C and 30 seconds at 95°C, elution of bisulfite-treated DNA in 100 μl of water. The treatment converts unmethylated Cytosine into Uracil, leaving methylated Cytosine unchanged. In this way, variations in the sequence are produced depending on DNA methylation status of the original DNA molecule.

PCR amplification, treatment with SAP solution and Transcription/RNase A cocktails were performed according to the protocol provided by Sequenom and the mass spectra were analyzed by EpiTYPER analyzer (Sequenom, San Diego, CA). As the MassARRAY assay is unable to discriminate between CpGs located at close vicinity to each other in the sequence, the close neighboring CpGs are analyzed as "Units", i.e. the measured methylation level is the average of the methylation levels of the CpGs cumulatively analyzed within the Unit. In the case of cg03546163 the measured methylation level is the average between two CpG sites located very close (**Figure S1**).

The amplicon for cg03546163 (chr6:35,654,364) encompasses 196bp (chr6:35,654,222-chr6:35,654,418 (GRCh37/hg19)) and PCR was performed on 10 ng of converted DNA using the following primers:

-cg03546163_10FW: aggaagagagTTTTTGTTTAGGATGAATTAGTTTGG;

-cg03546163_T7RV:cagtaatacgactcactatagggagaaggctAAAAACTACAATCTTATCCAATTC CTTT.

5. Conclusions

Our results suggest the potential use of DNAm analysis in blood to develop non-invasive tests for prognostic evaluation in MPM; our study is the first to demonstrate that a single CpG in *FKBP5* gene is an independent marker of prognosis in patients with MPM and is superior to the LMR inflammation-based prognostic score. The identification of simple and valuable prognostic markers for MPM will enable clinicians to select patients who are most likely to benefit from aggressive therapies and avoid subjecting non-responder patients to ineffective treatment. Moreover, epigenetic modifications such as DNAm are potentially reversible and can open new perspectives for epigenetic therapies in MPM. Knowledge of epigenetic changes has provided new therapeutic opportunities against cancer. To allow better approach of cancer cell inhibitory strategies, may be useful the understanding of molecular mechanisms that underlie cellular DNA epigenetic alterations. In this context, we reported epigenetic deregulations in blood samples from MPM patients in relation to OS, paving the road to both patients' stratification and the possible discovery of new combined therapeutic options in MPM.

Variables such as histological subtype, in previous literature have been associated with prognostic outcome [49-52]. In this study, to achieve the highest level of statistical power, it was not possible to stratify the analyses by histotype due the non-equal proportion of MPM cases, but we considered the histotype for all adjustments. Studies of large populations are needed to investigate the relationship between prognostic markers and treatment regimens focusing on specific histological subtypes.

MPM is an aggressive tumor associated with poor outcomes, yet the best treatment options remain controversial, in particular with regards to the role of surgery in treatment of this disease. To address this topic, some population-based studies demonstrated the effect of surgery on improved survival in the epithelioid group but not in biphasic or sarcomatoid groups; Although cancer-directed surgery was used more often in epithelioid and biphasic patients compared with sarcomatoid patients [53]. In this analysis, the authors did not observe statistically significant differences in overall survival between

stages for any histologic subtype. This was similar to results reported in the IASLC-IMIG study, which did not find differences in median survival between stages for patients who underwent any type of cancer-directed surgery [54].

The usage of Methylation alterations in clinical specimens as biomarkers could be recognized. Non-invasively obtained, methylation-based biomarkers detected in blood cells from cancer patients offer significant practical advantages, being promising and dynamic prognostic markers.

Appendices: Supplementary:

cg03546163

TTTTTGTTTAGGATGAATTAGTTTGGAGAGGAAAGAATATTTTGTTTGTATATTTGT
AGAAGAAGTTATTGTTGTTAAAAGTTAAGTTGGTTTTTTGATGTTTAGTGTGGATAG
GTTGAATAATAATTTATTTTTA**CG**(CpG1)TT**CG**(CpG2)GAGGGTTTATTTTATGTA
AATAGTTGAAAGGAATTGGATAAGATTGTAGTTTTT

Figure S1. Locations of cg03546163 (CpG2, in bold) and a second CpG site very close (CpG1, in red), investigated by Sequenom MassARRAY. Flanking primers are underlined. The CpG sites could not be tested individually due to MassARRAY technology constraints, but had to be tested jointly as a single unit: the methylation level is the cumulative value of all the sites within the CpG unit.

Institutional Review Board Statement: Ethics approval and consent to participate: The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

References

1. Sekido, Y. Molecular pathogenesis of malignant mesothelioma. *Carcinogenesis* 2013;34:1413-1419.

2. Rossini, M.; Rizzo, P.; Bononi, I.; et al. New Perspectives on Diagnosis and Therapy of Malignant Pleural Mesothelioma. *Front Oncol* 2018;8:91.
3. Furuya, S.; Chimed-Ochir, O.; Takahashi, K.; et al. Global Asbestos Disaster. *Int J Environ Res Public Health* 2018;15.
4. Straif, K.; Benbrahim-Tallaa, L.; Baan, R.; et al. A review of human carcinogens--Part C: metals, arsenic, dusts, and fibers. *Lancet Oncol* 2009;10:453-454.
5. Curran, D.; Sahmoud, T.; Therasse, P.; et al. Prognostic factors in patients with pleural mesothelioma: the European Organization for Research and Treatment of Cancer experience. *J Clin Oncol* 1998;16:145-152.
6. Jassem, J.; Ramlau, R.; Santoro, A.; et al. Phase III trial of pemetrexed plus best supportive care compared with best supportive care in previously treated patients with advanced malignant pleural mesothelioma. *J Clin Oncol* 2008;26:1698-1704.
7. Vogelzang, NJ.; Rusthoven, JJ.; Symanowski, J.; et al. Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma. *J Clin Oncol* 2003;21:2636-2644.
8. Yap, TA.; Aerts, JG.; Popat, S.; et al. Novel insights into mesothelioma biology and implications for therapy. *Nat Rev Cancer* 2017;17:475-488.
9. Guarrera, S.; Viberti, C.; Cugliari, G.; et al. Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. *J Thorac Oncol* 2019;14:527-539.
10. Matullo, G.; Guarrera, S.; Betti, M.; et al. Genetic variants associated with increased risk of malignant pleural mesothelioma: a genome-wide association study. *PLoS One* 2013;8:e61253.
11. Tanrikulu, AC.; Abakay, A.; Komek, H.; et al. Prognostic value of the lymphocyte-to-monocyte ratio and other inflammatory markers in malignant pleural mesothelioma. *Environ Health Prev Med* 2016;21:304-311.
12. Guo, G.; Chmielecki, J.; Goparaju, C.; et al. Whole-exome sequencing reveals frequent genetic alterations in BAP1, NF2, CDKN2A, and CUL1 in malignant pleural mesothelioma. *Cancer Res* 2015;75:264-269.
13. Betti, M.; Aspesi, A.; Sculco M, et al. Genetic predisposition for malignant mesothelioma: A concise review. *Mutat Res* 2019;781:1-10.

14. Gai, W.; Sun, K. Epigenetic Biomarkers in Cell-Free DNA and Applications in Liquid Biopsy. *Genes (Basel)* 2019;10.
15. Moore, LD.; Le, T.; Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* 2013;38:23-38.
16. Bononi, A.; Napolitano, A.; Pass, HI.; et al. Latest developments in our understanding of the pathogenesis of mesothelioma and the design of targeted therapies. *Expert Rev Respir Med* 2015;9:633-654.
17. Vandermeers, F.; Neelature Sriramareddy, S.; Costa, C.; et al. The role of epigenetics in malignant pleural mesothelioma. *Lung Cancer* 2013;81:311-318.
18. Ferrari, L.; Carugno, M.; Mensi, C.; et al. Circulating Epigenetic Biomarkers in Malignant Pleural Mesothelioma: State of the Art and critical Evaluation. *Front Oncol* 2020;10:445.
19. Creaney, J.; Yeoman, D.; Demelker, Y.; et al. Comparison of osteopontin, megakaryocyte potentiating factor, and mesothelin proteins as markers in the serum of patients with malignant mesothelioma. *J Thorac Oncol* 2008;3:851-857.
20. Creaney, J.; Dick, IM.; Meniawy, TM.; et al. Comparison of fibulin-3 and mesothelin as markers in malignant mesothelioma. *Thorax* 2014;69:895-902.
21. Tanrikulu, AC.; Abakay, A.; Kaplan, MA.; et al. A clinical, radiographic and laboratory evaluation of prognostic factors in 363 patients with malignant pleural mesothelioma. *Respiration* 2010;80:480-487.
22. Ferrante, D.; Bertolotti, M.; Todesco, A.; et al. Cancer mortality and incidence of mesothelioma in a cohort of wives of asbestos workers in Casale Monferrato, Italy. *Environ Health Perspect* 2007;115:1401-1405.
23. Dianzani, I.; Gibello, L.; Biava, A.; et al. Polymorphisms in DNA repair genes as risk factors for asbestos-related malignant mesothelioma in a general population study. *Mutat Res* 2006;599:124-134.
24. Betti, M.; Ferrante, D.; Padoan, M.; et al. XRCC1 and ERCC1 variants modify malignant mesothelioma risk: a case-control study. *Mutat Res* 2011;708:11-20.
25. Betti, M.; Casalone, E.; Ferrante, D.; et al. Germline mutations in DNA repair genes predispose asbestos-exposed patients to malignant pleural mesothelioma. *Cancer Lett* 2017;405:38-45.

26. Ferrante, D.; Mirabelli, D.; Tunesi, S.; et al. Pleural mesothelioma and occupational and non-occupational asbestos exposure: a case-control study with quantitative risk assessment. *Occup Environ Med* 2016;73:147-153.
27. Bibikova, M.; Barnes, B.; Tsan, C.; et al. High density DNA methylation array with single CpG site resolution. *Genomics* 2011;98:288-295.
28. Bibikova, M.; Lin, Z.; Zhou, L.; et al. High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 2006;16:383-393.
29. Du, P.; Zhang, X.; Huang, CC.; et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010;11:587.
30. Lehne, B.; Drong, AW.; Loh, M.; et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol* 2015;16:37.
31. Campanella, G.; Polidoro, S.; Di Gaetano, C.; et al. Epigenetic signatures of internal migration in Italy. *Int J Epidemiol* 2015;44:1442-1449.
32. Di Gaetano, C.; Voglino, F.; Guarrera, S.; et al. An overview of the genetic structure within the Italian population from genome-wide data. *PLoS One* 2012;7:e43759.
33. Houseman, EA.; Accomando, WP.; Koestler, DC.; et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;13:86.
34. Ehrich, M.; Nelson, MR.; Stanssens, P.; et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* 2005;102:15785-15790.
35. Costa-Pinheiro, P.; Montezuma, D.; Henrique, R.; et al. Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics* 2015;7:1003-1015.
36. Comba, P.; D'Angelo, M.; Fazzo, L.; et al. Mesothelioma in Italy: the Casale Monferrato model to a national epidemiological surveillance system. *Ann Ist Super Sanita* 2018;54:139-148.
37. Pass, H.; Giroux, D.; Kennedy, C.; et al. The IASLC Mesothelioma Staging Project: Improving Staging of a Rare Disease Through International Participation. *J Thorac Oncol* 2016;11:2082-2088.

38. Yamagishi, T.; Fujimoto, N.; Nishi, H.; et al. Prognostic significance of the lymphocyte-to-monocyte ratio in patients with malignant pleural mesothelioma. *Lung Cancer* 2015;90:111-117.
39. Argentieri, MA.; Nagarajan, S.; Seddighzadeh, B.; et al. Epigenetic Pathways in Human Disease: The Impact of DNA Methylation on Stress-Related Pathogenesis and Current Challenges in Biomarker Development. *EBioMedicine* 2017;18:327-350.
40. Yu, H.; Du, L.; Yi, S.; et al. Epigenome-wide association study identifies Behcet's disease-associated methylation loci in Han Chinese. *Rheumatology (Oxford)* 2019;58:1574-1584.
41. Yamaguchi, I.; Nakajima, K.; Shono, K.; et al. Downregulation of PD-L1 via *FKBP5* by celecoxib augments antitumor effects of PD-1 blockade in a malignant glioma model. *Neurooncol Adv* 2020;2:vdz058.
42. Romano, S.; Simeone, E.; D'Angelillo, A.; et al. *FKBP51*s signature in peripheral blood mononuclear cells of melanoma patients as a possible predictive factor for immunotherapy. *Cancer Immunol Immunother* 2017;66:1143-1151.
43. Li, L.; Lou, Z.; Wang, L. The role of *FKBP5* in cancer aetiology and chemoresistance. *Br J Cancer* 2011;104:19-23.
44. Staibano, S.; Mascolo, M.; Ilardi, G.; et al. Immunohistochemical analysis of *FKBP51* in human cancers. *Curr Opin Pharmacol* 2011;11:338-347.
45. Zannas, AS.; Jia, M.; Hafner, K.; et al. Epigenetic upregulation of *FKBP5* by aging and stress contributes to NF-kappaB-driven inflammation and cardiovascular risk. *Proc Natl Acad Sci U S A* 2019;116:11370-11379.
46. Romano, MF.; Avellino, R.; Petrella, A.; et al. Rapamycin inhibits doxorubicin-induced NF-kappaB/Rel nuclear activity and enhances the apoptosis of melanoma cells. *Eur J Cancer* 2004;40:2829-2836.
47. Sun, NK.; Huang, SL.; Chang, PY.; et al. Transcriptomic profiling of taxol-resistant ovarian cancer cells identifies *FKBP5* and the androgen receptor as critical markers of chemotherapeutic response. *Oncotarget* 2014;5:11939-11956.
48. Romano, S.; D'Angelillo, A.; Romano MF. Pleiotropic roles in cancer biology for multifaceted proteins FKBPs. *Biochim Biophys Acta* 2015;1850:2061-2068.

49. van der Bij, S.; Koffijberg, H.; Burgers, JA.; et al. Prognosis and prognostic factors of patients with mesothelioma: a population-based study. *Br J Cancer* 2012;107:161-4.
50. Nojiri, S.; Gemba, K.; Aoe, K.; et al. Survival and prognostic factors in malignant pleural mesothelioma: a retrospective study of 314 patients in the west part of Japan. *Jpn J Clin Oncol* 2011;41:32-9.
51. Temel, JS.; Greer, JA.; Muzikansky, A.; et al. Early palliative care for patients with metastatic non-small-cell lung cancer. *N Engl J Med* 2010;363:733-42.
52. Neumann, V.; Rütten, A.; Scharmach, M.; et al. Factors influencing long-term survival in mesothelioma patients—results of the German mesothelioma register. *Int Arch Occup Environ Health* 2004;77:191-9.
53. Meyerhoff, RR.; Yang, CF.; Speicher, PJ; et al.. Impact of mesothelioma histologic subtype on outcomes in the Surveillance, Epidemiology, and End Results database. *J Surg Res*. 2015 Jun 1;196(1):23-32.
54. Rusch, VW.; Giroux, D.; Kennedy, C.; et al. Initial analysis of the international association for the study of lung cancer mesothelioma database. *J Thorac Oncol*. 2012;7:1631.

Study3

Relationship between aging and Malignant Pleural Mesothelioma: epigenetic clock analyses

Simple Summary: The study was focused on the investigations of the interaction between epigenetic age acceleration measures and asbestos exposure in easily accessible white blood cells (WBCs) between malignant pleural mesothelioma (MPM) cases and asbestos-exposed cancer-free controls.

Abstract: Age is one of the strongest predictors of chronic disease and mortality. Aging denotes a multitude of processes at the cellular level, and biological responses to aging differ among people, having thus an important role when considering the relationship with other disease-related covariates. Malignant Pleural Mesothelioma (MPM) patients are characterized by a short overall survival time, highlighting the need for non-invasive tests aiming at an MPM risk assessment. Three hundred asbestos-exposed subjects (163 MPM cases and 137 cancer-free controls), from the same geographical region in Italy, were recruited. Using two previously established methylation-based "clocks" (proposed by S. Horvath), namely intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA), we defined biological age acceleration for each of three hundred asbestos-exposed subjects. Asbestos exposure-related variation in age acceleration indices and 95% confidence intervals (CIs) were estimated using multiple linear regression models. First, we explained the relationship between asbestos exposure and disease; EEAA showed that biological age acceleration was statistically significantly associated with increased asbestos exposure (Estimate = 0.704, 95% CI: 0.067, 1.475, P= 0.043). Then, statistically significant differences between cases and controls were found; considering

interaction effect we showed the OR increase in the group with higher median values of asbestos exposure and DNAm-aging respect to the others; finally, about 10% of the total effect of the asbestos exposure on MPM was independently mediated by biological aging variation (mediation effect). As a second outcome, we demonstrated that the mean of the number of total stochastic epigenetic mutations (hypo and hyper) was higher in cases with respect to controls. In particular, hypo-stochastic epigenetic mutations showed a mean difference between cases and controls about three-fold higher than hyper-stochastic epigenetic mutations. Moreover, mean stochastic epigenetic mutations increase in relation to asbestos exposure in cases but not in controls. Lastly, the trend of stochastic epigenetic mutations seems to be in the opposite direction stratifying by the two aging-related measures (EEAA and IEAA).

Our results suggest the potential use of age acceleration measures from DNAm profiles in blood as a proxy of asbestos exposure assessment. This will allow the development of non-invasive tests for asbestos-exposed subjects with the aim to best characterize and monitor early detection indicators in MPM.

Keywords: asbestos exposure; DNA methylation; age acceleration, malignant pleural mesothelioma.

---

1. Introduction

Malignant pleural mesothelioma (MPM) is a rare (prevalence 1–9/100,000) [2] tumor characterized by a long latency period of 20-40 years after first asbestos exposure [1,5,6].
A significant association between MPM and asbestos exposure has been reported, showing a clear, increasing trend in the odds ratio (OR) with increasing cumulative exposure among subjects exposed to over 10 fiber/mL-years [1].

According to recent papers, both DNAm and genetic alterations may contribute to MPM tumorigenesis [2-7]. Whereas the genome remains consistent throughout one's lifetime, factors like ageing, lifestyle, environmental exposures, and diseases can modify DNAm. The adaptive nature of DNAm means that it can be used to link environmental factors to the development of pathologic phenotypes such as cancers.

There is less information on the mechanisms and clinical outcomes of epigenetic derangements in MPM [8-10]. Several studies have evaluated DNAm alterations in MM samples [11-13], but few of them focused on DNAm alteration in blood as a circulating marker. Fischer et al. examined serum DNAm of nine gene-specific promoters from MM cases [14]. The identification of reliable DNAm based scores with high sensitivity and specificity for MPM risk assessment would be a major advancement.

This study was undertaken with the goal to indagate new biomarkers for MPM risk assessment and to determine if peripheral blood DNAm profiles, asbestos exposure values and epigenetic age measures simultaneously analyzed have any predictive value.

The idea was to evaluate the interaction effect of asbestos exposure with epigenetic ages measures on MPM risk. Currently, there are only one sensitive testing methods available for the screening of asbestos-exposed individuals, who are at high risk of developing MPM. Thus, the identification of reliable age related MPM diagnostic biomarkers in peripheral blood might increase the biological path of MPM development and the tools performance for detecting the disease at an early stage.

Stochastic epimutations (SEMs) that are not shared among subjects and that only partially affect the mean methylation level of the group remain unexplored. Another aspect of this study is to investigate the relationship between stochastic epigenetic mutations (SEMs) and MPM with the aim to better characterize the burden between MPM cases and controls. Interaction between asbestos exposure and SEM was evaluated to infer on the MPM odds ratio (OR).

2. Results

DNA methylation (DNAm) profiles have been used to compute biological age. Using two previously established methylation-based "clocks" (Horvath), namely intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA), we defined biological age acceleration for each of three hundred asbestos-exposed subjects.

Strong association between asbestos exposure and MPM risk, considering dichotomous distribution of asbestos exposure, was found (OR = 6.11, 95% CI = 3.73 ⁞ 10.20, P = 1.8×10^{-12}); quartile distribution of asbestos exposure was evaluated to estimate the potential incremental association with MPM risk (1^{st} quartile: used as reference; 2^{nd} quartile: OR = 1.83, 95% CI = 0.93 ⁞ 3.69, P = 0.09; 3^{rd} quartile: OR = 6.63, 95% CI = 3.30 ⁞ 13.81, P = 2.1×10^{-07}; 4^{th} quartile: OR = 11.00, 95% CI = 5.26 ⁞ 24.30, P = 7.3×10^{-10}).

DNAm-age was estimated for each subject from methylation levels in blood cells and showed a strong correlation with chronological age (Spearman's rank correlation rho = 0.80, p < 2.2E-16).

EEAA, which does not account for WBC composition, was significantly different between cases and controls: EEAAcases = 0.51 [-3.35–5.75], EEAAcontrols = -0.42 [-5.04–2.57], Wilcoxon p = 0.007, GLM effect size(se) = 2.27(0.86), p = 0.009.

EEAA showed that biological age acceleration was statistically significantly associated with increased asbestos exposure (Estimate = 0.704, 95% CI: 0.067, 1.475, P= 0.043).

As a second outcome, we demonstrated that the mean of the number of total stochastic epigenetic mutations (hypo and hyper) was higher in cases with respect to controls.

In particular, hypo-stochastic epigenetic mutations showed a mean difference between cases and controls about three-fold higher than hyper-stochastic epigenetic mutations. Moreover, mean stochastic epigenetic mutations increase in relation to asbestos exposure in cases but not in controls. Lastly, the trend of stochastic epigenetic mutations seems to be in the opposite direction stratifying by the two aging-related measures (EEAA and IEAA).

2.1. Interaction Analysis

Asbestos exposure-related variation in age acceleration indices and 95% confidence intervals (CIs) were estimated using multiple linear regression models. First, we explained the relationship between asbestos exposure and disease (**Figure 1**);

EEAA showed that biological age acceleration was statistically significantly associated with increased asbestos exposure (Estimate = 0.704, 95% CI: 0.067, 1.475, P= 0.043) (**Figure 2**);

than, statistically significant differences between cases and controls were found (**Figure 3**);

considering interaction effect we showed the OR increase in the group with higher median values of asbestos exposure and DNAm-aging respect to the others (**Figure 4**);

finally, about 10% of the total effect of the asbestos exposure on MPM was independently mediated by biological aging variation (mediation effect).

EEAA and asbestos exposure were considered as predictors of MPM risk in the interaction model. Categorical variables (quantile information) were used considering median values.

We tested the interaction between EEAA and asbestos exposure.

The OR was estimated as the relationship between the combination of EEAA levels and asbestos exposure quantile, and the reference (low median EEAA levels and low median asbestos exposure levels). Age, gender, population stratification, and WBCs (monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T) were included in the GLM (family=Binomial) to adjust the interaction effect.

As shown in **Table 1** high (>median) levels of EEAA and asbestos exposure refer to an OR of 9 respect to low levels.

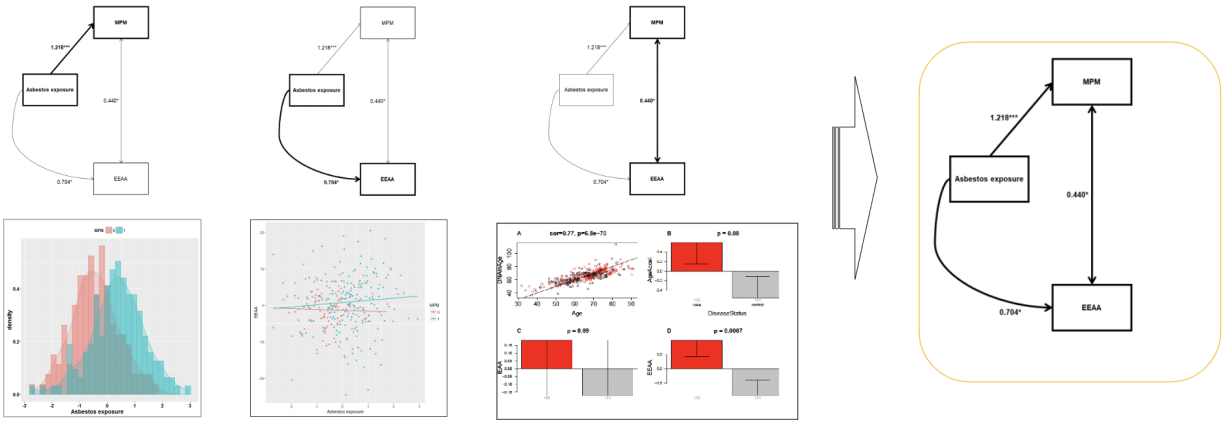

Figure 1-4. Interaction chart. Asbestos exposure, EEAA and MPM were included in the path.

Table 1. Interaction between aging (EEAA) and asbestos exposure on the MPM Odds ratios.

Aging	Asbestos exposure	OR	95% CI	P Value
(EEAA)				
HIgh	Low	1	-1 ⁞ 2	NS
Low	HIgh	4	2 ⁞ 8	1×10^{-04}
High	High	9	4 ⁞19	1×10^{-09}

Reference: hyper-methylation and low asbestos exposure levels.

3. Discussion

In the present study, we used DNAm values to investigate aging in WBCs from MPM cases and asbestos-exposed cancer-free controls from a region with a history of asbestos exposure (Piedmont, Italy) [10] in order to improve the assessment using new non-invasive epigenetic markers related to MPM. The identification of reliable MPM age related biomarkers in peripheral blood might improve risk assessment. DNAm profiles have been used as indices of biological age and the roles of epigenetic age acceleration measures were investigated.

Our results suggest the potential application of DNAm profiles, considering age acceleration measures, in blood to develop non-invasive tests for asbestos-exposed subjects. In detail, epigenetic age acceleration may play a role as a proxy of internal dose of asbestos exposure with the aim to better estimate individual risk and possibly for early detection of MPM.

The stochastic epigenetic mutations approach can add information at the level of epigenetic evaluation in the context of MPM. stochastic epigenetic mutations can be used as outcome or mediator in association models in order to better understand its contribution to MPM development. Considering stochastic epigenetic mutations occurrence and asbestos exposure levels may allow clinicians to better evaluate MPM risk.

3.1. Limitation of the Study

EEAA values may be a nonspecific marker related to a general, tumour-induced inflammatory status rather than a specific MPM biomarker; further studies are therefore needed to support our findings.

One main limitation of the functional interpretation of our results is that all our cases had already developed MPM at recruitment: thus, our findings likely reflect disease status rather than being markers of the dynamic processes leading to MPM onset. The lack of MPM tissue from the same subjects also poses major constraints to the functional interpretation of our findings.

Notwithstanding the above limitations, the discrimination between MPM cases and asbestos-exposed cancer-free controls improved when DNAm levels were considered together with asbestos exposure levels.

4. Material and Methods

4.1. Study Population

Study subjects are part of a wider, ongoing collaborative study, which is actively enrolling MPM cases and cancer-free controls in the municipality of Casale Monferrato (Piedmont

Region, Italy). This area was chosen due to its exceptionally high incidence of mesothelioma, caused by widespread occupational and environmental asbestos exposure originating from the Eternit asbestos-cement plant, which was operational until 1986 [15]. Additional MPM cases and cancer-free controls were recruited from other main hospitals of the Piedmont Region (in the municipalities of Turin, Novara and Alessandria). The ongoing collaborative study includes MPM cases diagnosed between incident MPM cases diagnosed between 2000 and 2010 after histological and/or cytological confirmation, and matched controls [16].

The present study included 159 MPM cases and 137 cancer-free controls from a larger case-control study, all of whom had genetic and blood DNAm data [17], good quality DNA at the time of the analyses, and information on asbestos exposure above the background level, as defined in Ferrante et al. [18]. MPM cases and asbestos-exposed cancer-free controls were matched by date of birth (±18 months) and gender. An additional 244 (140 MPM cases and 104 cancer-free controls) independent samples from the same case-control study were included for validation/replication analyses.

Our study complies with the Declaration of Helsinki principles and conforms to ethical requirements. All volunteers signed an informed consent form at enrollment. The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

4.2. Exposure Assessment

Information on occupational history and lifestyle habits were collected from all subjects through interviewer-administered questionnaires, which were completed during face-to-face interviews at enrollment. Job titles were coded in two ways: according to the International Standard Classification of Occupations [18] and the Statistical Classification of Economic Activities in the European Community.

A cumulative exposure index was computed, considering frequency, duration and intensity of asbestos exposure. Occupational, environmental and domestic asbestos

exposure was evaluated by an experienced occupational epidemiologist [18], and exposure doses (fibers/mL years) were rank-transformed to remove skewness.

4.3. Batch Effect, Population Stratification and White Blood Cells estimations

An individual's geographic origins may influence DNAm profiles, which could potentially introduce bias. To take this into consideration, we took advantage of the available data from our previous study, which includes a genome-wide genotyping dataset from the same study subjects [50]. When genome-wide genotyping was used to calculate the first PCs, they were shown to correlate with different geographic origins [19]. For each subject, we extracted WBC subtype percentages, estimated based on genome-wide methylation data; this method gives a quantification of the composition of leukocytes than can be achieved by simple histological or flow cytometric assessments, with an admissible range of variability [20].

4.4. Aging measures calculation

DNAm-age was estimated for each subject from DNAm levels on the overall sample according to the method proposed by Horvath et al. Estimated DNAm-age was compared with chronological age by the Spearman correlation test. Age acceleration (AA) was calculated as the residuals of the DNAm-age estimate regressed on chronological age in both MPM cases and cancer-free controls: positive AA suggests a biological age older than chronological age, whereas a negative AA suggests that the individual is biologically younger.

Two additional AA indices, intrinsic epigenetic AA (IEAA) and extrinsic epigenetic AA (EEAA), were calculated as well. EEAA is influenced by blood cell count contribution, whereas IEAA value is only weakly correlated with estimated measures of blood cell counts. The two indices, estimated from DNAm data, are thus indicators of two different cellular ageing processes, one (EEAA) dependent on and the other (IEAA) independent from WBC count.

4.5. Statistical Analysis

An association test was used to analyze the mean differences (MD) in EEAA and asbestos exposure levels between MPM cases and asbestos-exposed cancer-free controls.

Multiple regression analysis adjusted for age, gender, estimated WBCs (monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T), population stratification (first 2 PCs) and technical variability (first 10 PCs) was performed. For multiple comparison tests, a FDR p value ≤ 0.05 was considered statistically significant.

Logistic regression was used to analyze the relationship between EEAA and asbestos exposure in MPM risk (odds ratio), adjusting for age, gender, SNP PCs and WBCs (monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T) estimates. Asbestos exposure and EEAA were classified as above-median or below-median.

MPM risk for a given EEAA level and asbestos exposure was expressed by ORij, where i indicates the asbestos exposure (below-median or above-median) and j indicates the EEAA (above-median or below-median).

Subjects with below-median asbestos exposure and EEAA were considered the baseline group, and their MPM risk was coded as OR00 = 1. Interaction was analyzed in respect to both additive and multiplicative models based on the ORs obtained by logistic regression.

Synergistic interaction (positive interaction) implies that the combined action of two factors in an additive model is greater than the sum of their individual effects. Antagonistic interaction, on the other hand, means that when two factors are present in an additive model, the action of one reduces the effect of the other.

Multivariable logistic regression models were used to explore any deviations from a multiplicative model, including asbestos exposure, EEAA, and the corresponding interaction term (EEAA × exposure). All models were adjusted for age, gender, SNP PCs, technical covariates and WBCs (monocytes, granulocytes, natural killer, B cells, CD4+ T and CD8+ T) estimates. p-values < 0.05 were considered statistically significant.

All statistical analyses were conducted using the open source software R (4.0.2).

4.5. Stochastic epigenetic mutations

The distribution and variability of methylation levels in our population were studied for each one of the 445,254 CpG sites using Box-and-whiskers plots in order to identify stochastic epigenetic mutations [21]. For each probe, whenever the methylation level of one subject differed from the rest of the population we considered the outlier sample as epimutated for that locus. Thus, for each locus, epimutated subjects were identified as the extreme outliers with methylation level exceeding three times interquartile ranges Q1-(3 × IQR) and Q3+(3 × IQR). Finally, all epimutated loci were annotated in a new data matrix that allowed to calculate, for each subject, the total amount of epimutations and their genomic position. The Box-and-whiskers plot analysis was conducted using the boxplot function provided in the R car package and confirmed using the outlier function in the R outliers package. The analysis identified genes with an enriched number of epimutated probes (bonferroni corrected p-value < 0.05) confirming the presence of the epigenetic alterations previously reported in the medical report.

Institutional Review Board Statement: Ethics approval and consent to participate: The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

References

1. Iwatsubo, Y.; Pairon, J.C.; Boutin, C.; Ménard, O.; Massin, N.; Caillaud, D.; Orlowski, E.; Galateau-Salle, F.; Bignon, J.; Brochard, P. Pleural mesothelioma: dose-response relation at low levels of asbestos exposure in a French population-based case-control study. *Am J Epidemiol.* 1998, 148, 133-42.
2. Ferrante, D.; Bertolotti, M.; Todesco, A.; Mirabelli, D.; Terracini, B.; Magnani, C. Cancer mortality and incidence of mesothelioma in a cohort of wives of asbestos workers in Casale Monferrato, Italy. *Environ Health Perspect*. 2007, 115, 1401-5.
3. Guarrera, S.; Viberti, C.; Cugliari, G.; Allione, A.; Casalone, E.; Betti, M.; Ferrante, D.; Aspesi, A.; Casadio, C.; Grosso, F.; et al. Peripheral Blood DNA Methylation as Potential

Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. *J Thorac Oncol.* 2019, 14, 527-539.

4. Matullo, G.; Guarrera, S.; Betti, M.; Fiorito, G.; Ferrante, D.; Voglino, F.; Cadby, G.; Di Gaetano, C.; Rosa, F.; Russo, A.; et al. Genetic variants associated with increased risk of malignant pleural mesothelioma: a genome-wide association study. *PLoS One.* 2013, 8:e61253.
5. Cugliari, G.; Catalano, C.; Guarrera, S.; Allione, A.; Casalone, E.; Russo, A.; Grosso, F.; Ferrante, D.; Viberti, C.; Aspesi, A.; et al. DNA Methylation of FKBP5 as Predictor of Overall Survival in Malignant Pleural Mesothelioma. *Cancers.* 2020, 21, 12(11), 3470.
6. Betti, M.; Aspesi, A.; Sculco, M.; Matullo, G.; Magnani, C.; Dianzani, I. Genetic predisposition for malignant mesothelioma: A concise review. *Mutat Res.* 2019, 781, 1-10.
7. Guo, G.; Chmielecki, J.; Goparaju, C.; Heguy, A.; Dolgalev, I.; Carbone, M.; Seepo, S.; Meyerson, M.; Pass, H.I.. Whole-exome sequencing reveals frequent genetic alterations in BAP1, NF2, CDKN2A, and CUL1 in malignant pleural mesothelioma. *Cancer Res.* 2015, 75, 264-269.
8. Moore, L.D.; Le, T.; Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology.* 2013, 38, 23-38.
9. Bononi, A.; Napolitano, A.; Pass, H.I.; Yang, H.; Carbone, M. Latest developments in our understanding of the pathogenesis of mesothelioma and the design of targeted therapies. *Expert Rev Respir Med.* 2015, 9, 633-654.
10. Vandermeers, F.; Neelature, S.; Costa, C.; Hubaux, R.; Cosse, JP.; Willems, L. The role of epigenetics in malignant pleural mesothelioma. *Lung Cancer.* 2013, 81, 311-318.
11. Zhang, X.; Tang, N.; Rishi.; Pass, H.I.; Wali, A. Methylation profile landscape in mesothelioma: possible implications in early detection, disease progression, and therapeutic options. *Methods Mol Biol.* 2015, 1238, 235-47.
12. Goto, Y.; Shinjo, K.; Kondo, Y.; Shen, L.; Toyota, M.; Suzuki, H.; Gao, W.; An, N.; Fujii, M.; Murakami, H.; et al. Epigenetic profiles distinguish malignant pleural mesothelioma from lung adenocarcinoma. *Cancer Res.* 2009, 69(23), 9073-82.
13. Christensen, B.C.; Houseman, E.A.; Poage, G.M.; Godleski, J.J.; Bueno, R.; Sugarbaker, D.J.; Wiencke, J.K.; Nelson, H.H.; Marsit, C.J.; Kelsey, K.T. Integrated profiling reveals a

global correlation between epigenetic and genetic alterations in mesothelioma. *Cancer Res.* 2010, 70(14), 5686-94.

14. Fischer, J.R.; Ohnmacht, U.; Rieger, N.; Zemaitis, M.; Stoffregen, C.; Kostrzewa, M.; Buchholz, E.; Manegold, C.; Lahm, H. Promoter methylation of RASSF1A, RARbeta and DAPK predict poor prognosis of patients with malignant mesothelioma. *Lung Cancer.* 2006, 54, 109-16.
15. Dianzani, I.; Gibello, L.; Biava, A.; Giordano, M.; Bertolotti, M.; Betti, M.; Ferrante, D.; Guarrera, S.; Betta, G.P.; Mirabelli, D.; et al. Polymorphisms in DNA repair genes as risk factors for asbestos-related malignant mesothelioma in a general population study. *Mutat Res.* 2006, 599, 124-134.
16. Betti, M.; Ferrante, D.; Padoan, M.; Guarrera, S.; Giordano, M.; Aspesi, A.; Mirabelli, D.; Casadio, C.; Ardissone, F.; Ruffini, E.; et al. XRCC1 and ERCC1 variants modify malignant mesothelioma risk: a case-control study. *Mutat Res.* 2011, 708, 11-20.
17. Betti, M.; Casalone, E.; Ferrante, D.; Aspesi, A.; Morleo, G.; Biasi, A.; Sculco, M.; Mancuso, G.; Guarrera, S.; Righi, L.; et al. Germline mutations in DNA repair genes predispose asbestos-exposed patients to malignant pleural mesothelioma. *Cancer Lett.* 2017, 405, 38-45.
18. Ferrante, D.; Mirabelli, D.; Tunesi, S.; Terracini, B.; Magnani, C. Pleural mesothelioma and occupational and non-occupational asbestos exposure: a case-control study with quantitative risk assessment. *Occup Environ Med.* 2016, 73, 147-153.
19. Di Gaetano, C.; Voglino, F.; Guarrera, S.; Fiorito, G.; Rosa, F.; Di Blasio, AM.; Manzini, P.; Dianzani, I.; Betti, M.; Cusi, D.; et al. An overview of the genetic structure within the Italian population from genome-wide data. *PLoS One.* 2012, 7, e43759.
20. Houseman, E.A.; Accomando, W.P.; Koestler, D.C.; Christensen, B.C.; Marsit, C.J.; Nelson, H.H.; Wiencke, J.K.; Kelsey, K.T. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012, 13, 86.
21. Gentilini, D.; Garagnani, P.; Pisoni, S.; Bacalini, MG.; Calzari, L.; Mari, D.; Vitale, G.; Franceschi, C.; Di, Blasio AM. Stochastic epigenetic mutations (DNA methylation) increase exponentially in human aging and correlate with X chromosome inactivation skewing in females. *Aging* (Albany NY). 2015 Aug;7(8):568-78.

Study4

Epigenomics of Malignant Pleural Mesothelioma: a structural equation modeling

Simple Summary: The study was focused on the investigations of the overall MPM path including single-CpGs, asbestos exposure, epigenetic age acceleration measures in easily accessible white blood cells (WBCs) between malignant pleural mesothelioma (MPM) cases and asbestos-exposed cancer-free controls.

Abstract: Malignant Pleural Mesothelioma (MPM) patients are characterized by a short overall survival time, highlighting the need for non-invasive tests aiming at an MPM risk assessment. Three hundred asbestos-exposed subjects (163 MPM cases and 137 cancer-free controls), from the same geographical region in Italy, were recruited. Recently, asbestos exposure, aging, single CpGs DNA methylation (DNAm) and white blood cells (WBCs) composition have been individually associated with Malignant Pleural Mesothelioma (MPM). No study has shown the simultaneous effect combining all these predictors with the aim to test the epigenomic pathway using a statistical approach. Structural equation modeling (SEM) is a largely confirmatory, rather than exploratory, technique; It is used to determine whether a model is valid than to find a suitable model. Asbestos exposure levels were extracted considering a quantitative measure; DNAm profiles have been used as single CpGs and to compute WBCs estimation and biological age measures. The SEM showed that all ten relationships (4 four associations and 6 covariances) included in the graph model were statistically significant. Our results suggest the potential use of a suite of peripheral blood DNA methylation measures to better characterize the MPM biological path. This will allow to increase the knowledge about the epigenetics contribution in MPM

and more in detail to develop non-invasive tests for asbestos-exposed subjects with the aim to monitor early detection indicators during the risk assessment.

Keywords: asbestos exposure; DNA methylation; SEM, Malignant Pleural Mesothelioma.

---

1. Introduction

Malignant Pleural Mesothelioma (MPM) patients are characterized by a short overall survival time, highlighting the need for non-invasive tests aiming at an MPM risk assessment. Three hundred asbestos-exposed subjects (163 MPM cases and 137 cancer-free controls), from the same geographical region in Italy, were recruited.

Recently, asbestos exposure, aging related measures, single CpGs DNA methylation (DNAm) and white blood cells (WBCs) composition have been individually associated with Malignant Pleural Mesothelioma (MPM). No study has shown the simultaneous effect combining all these predictors with the aim to test the epigenomic pathway using a statistical approach.

Structural equation modeling (SEM) is a versatile framework that enables modeling of complex multivariate data and multiple predictors in unison [1]. Furthermore, SEMs allow for the evaluation of total, path-specific, direct, and indirect effects between all specified variables [2]. SEM is a largely confirmatory, rather than exploratory, technique; it is used to determine whether a model is valid rather than to find a suitable model. Asbestos exposure levels were extracted considering a quantitative measure; DNAm profiles have been used as single CpGs and to compute WBCs estimation and biological age measures. Structural equation modeling (SEM) is a powerful, multivariate technique found increasingly in scientific investigations to test and evaluate multivariate causal relationships. SEMs differ from other modeling approaches as they test the direct and indirect effects on pre-assumed causal relationships.

SEM is a combination of two statistical methods: confirmatory factor analysis and path analysis. Confirmatory factor analysis should be the best approach to infer causality in medicine. Omics data can be simultaneously used with phenotypes with the goal to extract a multitude of relationships including regressions and covariates. Traditional approaches differ from the SEM approach in several areas: First, SEM is a highly flexible and comprehensive methodology; Second, traditional methods specify a default model whereas SEM requires formal specification of a model to be estimated and tested. SEM offers no default model and places few limitations on what types of relations can be specified. SEM model specification requires researchers to support hypotheses with theory or research and specify relations a priori; Third, SEM is a multivariate technique incorporating observed (measured) and unobserved variables (latent constructs) while traditional techniques analyze only measured variables. Multiple, related equations are solved simultaneously to determine parameter estimates with SEM methodology; Fourth, SEM allows researchers to recognize the imperfect nature of their measures. SEM explicitly specifies error while traditional methods assume measurement occurs without error; Fifth, traditional analysis provides straightforward significance tests to determine group differences, relationships between variables, or the amount of variance explained. SEM provides no straightforward tests to determine model fit. Instead, the best strategy for evaluating model fit is to examine multiple tests; Sixth, SEM resolves problems of multicollinearity. Multiple measures are required to describe a latent construct (unobserved variable). Multicollinearity cannot occur because unobserved variables represent distinct latent constructs; Finally, a graphical language provides a convenient and powerful way to present complex relationships in SEM. Model specification involves formulating statements about a set of variables. A diagram, a pictorial representation of a model, is transformed into a set of equations. The set of equations are solved simultaneously to test model fit and estimate parameters.

2. Results

The result of single path analysis determines if the relationships are unidirectional or bidirectional, by using the study design characteristics; in this context, the retrospective direction is able to adopt the specifications related to any single row. After this phase we move to the outline of the model by determining the number and relationships of measured and latent variables. A path diagram depicting the structural and measurement models will guide the second phase when the focus will be on the model identification and test, as described next.

The SEM showed that all ten relationships (**Table 1**) (4 four associations and six covariances) included in the graph model were statistically significant.

Table 1. SEM analyses of MPM path.

Lhs	Op	Rhs	Estimate	SE	P value	95%CI L	95%CI U
MPM	~	Asbestos exposure	0,19732	0,02686	2,05E-13	0,14467	0,24998
cg03546163	~	Asbestos exposure	-0,02973	0,00805	2,20E-04	-0,04551	-0,01396
PC-WBC	~	Asbestos exposure	0,00244	0,00084	3,49E-03	0,0008	0,00408
EEAA	~	Asbestos exposure	0,70416	0,39329	4,34E-02	0,06668	1,47500
cg03546163	~~	MPM	-0,02669	0,00401	2,67E-11	-0,03454	-0,01884
MPM	~~	PC-WBC	0,00282	0,00042	1,37E-11	0,002	0,00364
cg03546163	~~	PC-WBC	-0,00084	0,00012	2,06E-11	-0,00108	-0,00059
EEAA	~~	PC-WBC	0,02373	0,00578	4,03E-05	0,01241	0,03506
MPM	~~	EEAA	0,43976	0,18192	1,56E-02	0,0832	0,79632
cg03546163	~~	EEAA	-0,11628	0,05438	3,25E-02	-0,22287	-0,00968

SEM including all selected covariates. Stratification by the relationship as regression or covariance.

3. Discussion

Our insights suggest the potential use of a suite of peripheral blood DNA methylation measures to better characterize the MPM biological path. We simultaneously analyzed i) asbestos exposure (continuous, log-transformed to optimize the normality of the distribution), ii) MPM risk (binomial distribution, cases-controls), iii) single-CpG DNAm (cg03546163 in *FKBP5*, strongly associated with MPM risk), PC-WBCs (1st principal component, calculated using the 6 common WBCs) and EEAA (extrinsic epigenetic age acceleration, one of the most aging-related epigenetic measures). Conceptually, due the study design charachteristics, asbestos exposure is the only one variable that should be considered as a predictor in all regression between the available information. SEM results depict the effect of asbestos exposure on MPM risk, single-CpG DNAm, PC-WBCs and EEAA. Moving to covariance considerations, MPM risk shows correlation with single-CpG DNAm, PC-WBCs and EEAA. Single-CpG also shows significant direct negative correlation with PC-WBCs and EEAA; lastly, PC-WBCs show direct positive correlation with EEAA. As per my knowledge, this study is the first to show causal relationship between all the included information focusing on the MPM epigenetics. SEM estimated all the coefficients in the model to be the best fit in determining the causality through direct and indirect relationships. The outcome of the study is shown in the illustrated diagrammatic representation (**Figure 1**). This study addresses gaps in knowledge pertaining to epigenetic factors affected by asbestos exposure and its rate of change. However, SEMs only test the associations of the assumed causal structure as specified by the investigator. The presence or absence of arrows must be specified a priori based on expert subject matter knowledge and assumptions. This study will increase the knowledge about the epigenetics contribution in MPM and more in detail to develop non-invasive tests for asbestos-exposed subjects with the aim to monitor early detection indicators during the risk assessment.

4. Material and Methods

Path diagram (**Figure 1**), on the other hand, aimed to find the causal relationship among variables by creating a path diagram.

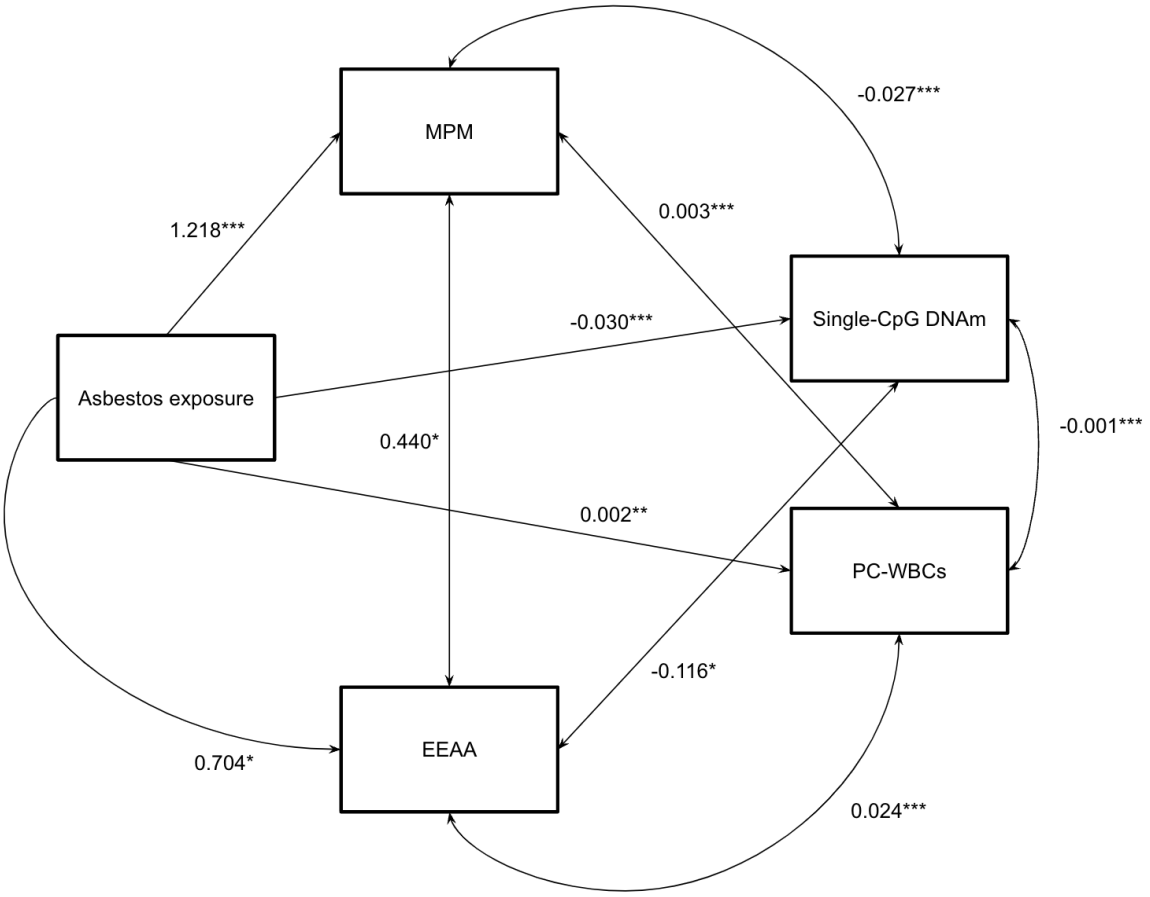

Figure 1. Path diagram. Asbestos exposure, EEAA, MPM, Single-CpG DNAm and PC-WBCs were included in the path. A single-headed arrow shows the regression while a double-headed arrow shows the covariance between the two variables.

Thanks to the previous works, I identified the MPM Research Problem developing hypotheses about the relationships among variables that are based on theory, previous empirical findings or both. These relationships may be direct or indirect whereby intervening variables may mediate the effect of one variable on another.

The Two main goals in SEM are: i) to understand the patterns of correlation/covariance among a set of variables and ii) to explain as much of their variance as possible with the

model specified. Structural Equation modeling relies on several statistical tests to determine the adequacy of model fit to the data.

The chi-square test indicates the amount of difference between expected and observed covariance matrices. A chi-square value close to zero indicates little difference between the expected and observed covariance matrices. In addition, the probability level must be greater than 0.05 when chi-square is close to zero.

The Comparative Fit Index (CFI) is equal to the discrepancy function adjusted for sample size. CFI ranges from 0 to 1 with a larger value indicating better model fit. Acceptable model fit is indicated by a CFI value of 0.90 or greater.

Root Mean Square Error of Approximation (RMSEA) is related to residuals in the model. RMSEA values range from 0 to 1 with a smaller RMSEA value indicating better model fit. In general, acceptable model fit is indicated by an RMSEA value of 0.06 or less.

SEM was implemented to simultaneously analyze any relationship previously performed (R package 'lavaan') [3].

Institutional Review Board Statement: Ethics approval and consent to participate: The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (IIGM, Candiolo, Italy).

References

1. Sanchez, BN.; Houseman, EA.; Ryan, LM. Residual-based diagnostics for structural equation models. *Biometrics*. 2009; 65(1):104-115.

2. VanderWeele, TJ. Invited commentary: structural equation models and epidemiologic analysis. *Am J Epidemiol*. 2012;176(7):608-612.

3. Rosseel, Y. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*. 2012, 48(2), 1-36.

Study5

FKBP5, a modulator of stress responses involved in Malignant Mesothelioma: the link between stress and cancer

Simple Summary: The study reviewed the *FKBP5* gene involvement in Malignant Pleural Mesothelioma (MPM) with the aim to investigate and explain the potential link between stress and cancer related outcomes.

Abstract: Malignant pleural mesothelioma (MPM) is a rare and aggressive neoplasm. Stress is a well-established risk factor for a lot of disease phenotypes, including cancer. The risk for stress-related disorders is shaped by complex interactions among multiple environmental stressors and many genes with small individual effects on expressed phenotypes. Asbestos exposure-related epigenetics modifications can be considered as risk mediator and driver for MPM prognosis evaluation. Furthermore, aging-acceleration measures should be considered as a proxy of stress-related environmental phenotype associations. Some recent research showed that *FKBP5* may be a key factor in the stress response and that transcriptomic data can provide insight into stress-related pathophysiology. Epigenetic activation of the *FKBP5* gene has increased stress sensitivity and the MPM risk of disease or prognosis outcomes and may be studied for potential treatment evaluation.

Keywords: Mesothelioma, stress, *FKBP5*, risk assessment, survival, asbestos exposure, aging, white blood cells

---

Malignant pleural mesothelioma (MPM) is a rare tumour characterized by a long latency period after asbestos exposure and poor survival. Due to the complexity of risk assessment of exposed subjects, the studies focused on biological mechanisms need to be improved. Stress related measures are involved in a multitude of disease phenotypes, including cancer.

An important modulator of stress responses is FK506-binding protein 51 (*FKBP5*/*FKBP51*), which, among other functions, acts as a co-chaperone that modulates glucocorticoid receptor (GR) activity. These results suggest that *FKBP5* may be a key factor in the stress response and that transcriptomic data can provide insight into stress-related pathophysiology [1]. A recent list of reports indicates a strong association of *FKBP5* functions with several neurological diseases, including posttraumatic stress disorder (PTSD) [2-4]. Epigenetic activation of the *FKBP5* gene has increased stress sensitivity and the risk of psychiatric disorders [5].

By modulating GR signaling, *FKBP5* has the potential to modulate the actions of glucocorticoids, hormones with pleiotropic effects that can affect essentially every body tissue [6]. Although in psychiatry and neuroscience *FKBP5* is most discussed as a modulator of glucocorticoid signaling, it is important to highlight that it also interacts with a host of other molecular partners, affecting several cellular processes.

However, none of the GWAS meta-analyses showed strong associated signals for this genetic locus yet. More consistent are reports of *FKBP5* × specific environmental stress interactions altering the risk for psychiatric disorders. Furthermore, *FKBP5* functions have also been correlated with multiple other diseases and processes, including type 2 diabetes, adipogenesis, fatty acid metabolism, and cancers [7].

In several cancers, a strong negative correlation of *FKBP5* expression with severity of disease has been observed [8-11]. Epigenetics can represent one concrete possibility to improve the mechanical characterization of the disease with the goal of early detection and prognosis stratification.

DNA methylation differences in white blood cells (WBCs) between MPM cases and asbestos-exposed cancer-free controls highlighted some interesting differences [12], including asbestos exposure-related hypo-methylation of *FKBP5* in the top marker of risk assessment; interaction analysis showed that considering DNAm levels at *FKBP5* together with asbestos exposure levels may help to better define MPM risk for asbestos-exposed subjects [13]; another recent paper identified hypomethylation of the same CpG in *FKBP5* as a predictor of overall survival in MPM cases with better performance than traditional inflammation-based scores such as lymphocyte-to-monocyte ratio (LMR) [14].

FKBP5 is an immunophilin and has an important role in immunoregulation and protein folding and trafficking. It plays a role in transcriptional complexes and acts as a co-transcription factor, along with other proteins in the FKBP family [15]. During the last few years, The hypothesis of a possible role of *FKBP5* in the development and progression of different types of cancer has stemmed from several studies. High protein expression has been linked to either suppression or promotion of tumor growth, depending on tumor type and microenvironment [16-17].

FKBP5 is involved in the NF-kB and AKT signaling pathways, both of which are implicated in tumorigenesis [18]. Notably, NF-kB appears to be frequently constitutively activated in malignant tumors and involved in the modulation of genes linked to cell motility, neoangiogenesis, proliferation, and programmed cell death [19]. Epigenetic upregulation of *FKBP5* could promote NF-kB activation [20]. STAT3-NFkB activity is involved in chemoresistance in MM cells, and NFkB was shown to be constitutively active because of asbestos-induced chronic inflammation [21]. Previous studies conducted on various cancer types showed that upregulation of *FKBP5* gene expression is associated with drug resistance [22]. The tissue- and context-specific *FKBP5* expression should be considered when examining the consequences of *FKBP5* dysregulation and when considering *FKBP5* as a candidate therapeutic target. A similar study supported this observation by making *FKBP5* an effective biomarker for sensitivity to chemotherapy; patient responses to chemotherapy may be determined by the variation in *FKBP5* levels

[17]. One study on ovarian cancer cells lines denoted that the upregulation of *FKBP5* may increase the resistance to chemotherapeutic agents, whereas the gene silencing sensitized ovarian cancer cells to taxol [23].

Lastly, the risk for stress-related disorders is shaped by complex interactions among multiple environmental stressors and many genes with small individual effects on expressed phenotypes. Elucidating these complex interactions at a systems level is a challenging task but may contribute to improving the holistic understanding of stress-related disorders. Furthermore, aging-related epigenetics measures should show interesting associations between stress-related phenotypes and disease to better characterize clusters of exposed subjects. This review further supports the notion that stress can affect cancer outcomes in exposed subjects, perhaps by interfering with neural mechanisms involved in controlling the oncogenesis pathway for early detection, prognosis, and treatment.

References

1. Kwon, J.; Kim, Y.J.; Choi, K. et al. 2019. Identification of stress resilience module by weighted gene co-expression network analysis in Fkbp5-deficient mice. *Mol Brain* 12, 99.
2. Duman, R.S.; Girgenti, MJ. 2019. Molecular and cellular studies of PTSD: postmortem transcriptome analysis and novel therapeutic targets. *J Neurosci Res* 97:292–299.
3. Jaworska-Andryszewska, P.; Rybakowski, J.K. 2019. Childhood trauma in mood disorders: neurobiological mechanisms and implications for treatment. *Pharmacol Rep* 71:112–120.
4. Talarowska, M. 2020. Epigenetic mechanisms in the neurodevelopmental theory of depression. *Depress Res Treat* 2020:6357873.
5. Madison, N.; Halldorsdottir, T.; Binder, E.B. Understanding the molecular mechanisms underpinning gene by environment interactions in psychiatric disorders: The *FKBP5* model. Biol. *Psychiatry* 2018, 83, 821–830.
6. Nicolaides, N.C.; Charmandari, E.; Chrousos, G.P.; Kino, T. 2014. Recent advances in the molecular mechanisms determining tissue sensitivity to glucocorticoids: novel mutations, circadian rhythm, and ligand-induced repression of the human glucocorticoid receptor. *BMC Endocr Disord* 14: 71.

7. Hausl, A.S.; Balsevich, G.; Gassen, N.C.; Schmidt, M.V. 2019. Focus on *FKBP51*: a molecular link between stress and metabolic disorders. *Mol Metab* 29:170–181.
8. Pei, H.; Li, L.; Fridley, B.L.; Jenkins, G.D.; Kalari, K.R.; Lingle, W.; Petersen, G.; Lou, Z.; Wang, L. 2009. *FKBP51* affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell* 16:259–266.
9. Yu, J.; Qin, B.; Wu, F.; Qin, S.; Nowsheen, S.; Shan, S.; Zayas, J.; Pei, H.; Lou, Z.; Wang, L. 2017. Regulation of serine-threonine kinase Akt activation by NAD+-dependent deacetylase SIRT7. *Cell Rep* 18:1229–1240.
10. Daneri-Becerra, C.; Zgajnar, N.R.; Lotufo, C.M.; Ramos, Hryb A.B.; Piwien-Pilipuk, G.; Galigniana, M.D. 2019. Regulation of *FKBP51* and *FKBP52* functions by post-translational modifications. *Biochem Soc Trans* 47:1815–1831.
11. Zgajnar, N.R.; De Leo, S.A.; Lotufo, C.M.; Erlejman, A.G.; Piwien-Pilipuk, G.; Galigniana, M.D. 2019. Biological actions of the Hsp90-binding immunophilins *FKBP51* and *FKBP52*. *Biomolecules* 9:52.
12. Guarrera, S.; Viberti, C.; Cugliari, G.; Allione, A.; Casalone, E.; Betti, M.; Ferrante, D.; Aspesi, A.; Casadio, C.; Grosso, F.; et al. Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. *J. Thorac. Oncol.* 2019, 14, 527–539.
13. Cugliari, G.; Allione, A.; Russo, A.; Catalano, C.; Casalone, E.; Guarrera, S.; Grosso, F.; Ferrante, D.; Sculco, M.; La Vecchia, M.; et al. New DNA Methylation Signals for Malignant Pleural Mesothelioma Risk Assessment. *Cancers* 2021, 13, 2636.
14. Cugliari, G.; Catalano, C.; Guarrera, S.; Allione, A.; Casalone, E.; Russo, A.; Grosso, F.; Ferrante, D.; Viberti, C.; Aspesi, A.; et al. DNA Methylation of *FKBP5* as Predictor of Overall Survival in Malignant Pleural Mesothelioma. *Cancers* 2020, 12, 3470.
15. Kang, C.B.; Hong, Y.; The-Paganon, S.; Sup Yoon, H. FKBP family proteins: Immunophilins with versatile biological functions. *Neurosignals* 2008, 16, 318–325.
16. Li, L.; Lou, Z.; Wang, L. The role of *FKBP5* in cancer aetiology and chemoresistance. Br. J. Cancer 2011, 104, 19–23.
17. Romano, S.; D'Angelillo, A.; Romano, M.F. Pleiotropic roles in cancer biology for multifaceted proteins FKBPs. Biochim. *Biophys. Acta* 2015, 1850, 2061–2068.

18. Staibano, S.; Mascolo, M.; Ilardi, G.; Siano, M.; De Rosa, G. Immunohistochemical analysis of *FKBP51* in human cancers. Curr. Opin. *Pharmacol*. 2011, 11, 338–347.
19. Zannas, A.S.; Jia, M.; Hafner, K.; Baumert, J.; Wiechmann, T.; Pape, J.C.; Arloth, J.; Ködel, M.; Martinelli, S.; Roitman, M.; et al. Epigenetic upregulation of *FKBP5* by aging and stress contributes to NF-kappaB-driven inflammation and cardiovascular risk. *Proc. Natl. Acad. Sci. USA* 2019, 116, 11370–11379.
20. Yoshikawa, Y.; Sato, A. Biallelic germline and somatic mutations in malignant mesothelioma: Multiple mutations in transcription regulators including mSWI/SNF genes. *Int. J. Cancer* 2015, 136, 560–571.
21. Zhou, J.; Ng, Y.; Chng, W.J. ENL: Structure, function, and roles in hematopoiesis and acute myeloid leukemia. *Cell Mol. Life Sci.* 2018, 75, 3931–3941.
22. Romano, M.F.; Avellino, R.; Petrella, A.; Bisogni, R.; Romano, S.; Venuta, S. Rapamycin inhibits doxorubicin-induced NF-kappaB/Rel nuclear activity and enhances the apoptosis of melanoma cells. *Eur. J. Cancer* 2004, 40, 2829–2836.
23. Sun, N.K.; Huang, S.L.; Chang, P.Y.; Lu, H.P.; Chao, C.C.K. Transcriptomic profiling of taxol-resistant ovarian cancer cells identifies *FKBP5* and the androgen receptor as critical markers of chemotherapeutic response. *Oncotarget* 2014, 5, 11939–11956.

CONCLUSION

Although the mechanisms of carcinogenesis in MPM are incompletely understood, chronic inflammation is critically involved in the pathogenesis of MPM. It is well known that systemic inflammatory response plays an important role in cancer progression. Generally cases showed higher values for Granulocytes and lower values for Monocytes, B cells, Natural Killer cells CD4+ T, CD8+ T, Naïve CD4T and Naïve CD8T compared to controls. Moreover, functional changes in WBCs were related to asbestos-induced immune system deregulation and autoimmune response asbestos related antigenic overstimulation, putatively linked to MPM occurrence. Indeed, in our study, the observed reduction of estimated CD4+ and CD8+ T lymphocytes in MPM cases suggests a weaker adaptive immune system and is compatible with the possible occurrence of functional changes in cellular subpopulations in MPM.

This population-based case-control study was conducted to improve knowledge on the relationship between MPM and quantitative exposure to asbestos, based on the evidence arising from the area of Casale Monferrato (Italy). MPM strongly denoted association with asbestos. Only a limited number of studies on MPM had previously been carried out in general populations and comparison of results was made difficult by differences in the assessment of exposure and in criteria of analysis.

In recent years, there is a growing interest in DNAm profile changes as possible diagnostic biomarkers in cancer research. Altered DNA methylation is frequently observed in cancer, and DNAm profiles of specific genes were already proposed as potential tools for cancer detection, risk prediction, and prognosis. In this context, we used novel insights and statistical methodologies with the aim to improve the knowledge related to the early detection in MPM. Thus, the identification of reliable MPM diagnostic biomarkers in peripheral blood might provide the tool for detecting the disease at a more treatable stage.

In the present project, we used a whole genome microarray approach to investigate DNA-methylation in white blood cells (WBCs) from MPM cases and healthy controls, aiming at the identification of a panel of differentially methylated CpG markers indicative of MPM in asbestos exposed subjects. Cases denoted hypomethylation detected in *FKBP5* gene. Adding this statistically significant single-CpG DNAm to only asbestos exposure risk factor, operating characteristic (ROC) curves showed a significant increase in MPM discrimination. If confirmed by functional studies, our findings of DMe in WBCs suggest that DNAm changes in the immune system components may play a role in the oncogenic process triggered by asbestos exposure.

Exposure to asbestos fibers has been shown to be associated with aberrant patterns of DNA methylation at single CpG in *FKBP5* in peripheral blood of exposed individuals. Concerning the relationship of epigenetic age acceleration measures and MPM, cases show high levels compared to controls in extrinsic epigenetic age acceleration (EEAA), while no statistically significant differences on intrinsic epigenetic age acceleration (IEAA) were found. Based on the interaction analysis, it appears that the effect of asbestos exposure on MPM depends on age acceleration measures. The mediation analysis identified that part of the total effect was mediated by differential single-CpG methylation in the *FKBP5*; one potential explanation for these findings, restricting to the level of DNA methylation at same single CpG in *FKBP5*, is that the association between DNA methylation and MPM might just reflect the known causal effect of asbestos exposure on MPM, as DNA methylation is a strong biomarker for asbestos exposure. Lastly structural equation modeling (SEM) was implemented to simultaneously analyze any prior relationship. DNA methylation plays an important role in the etiology of cancer, through the use of mediation analysis.

Although several treatment options have been delivered to patients with MPM, the median survival time is approximately 12 months. It is well known that systemic inflammatory response plays an important role in cancer progression. Chronic inflammation is critically involved in the pathogenesis of MPM as a potential result of

asbestos exposure. Systemic inflammation can predict clinically meaningful outcomes, such as overall survival (OS) and response to systemic treatment. Considering WBCs, lymphocytes act as tumor suppressors by inducing cytotoxic cell death and inhibiting tumor cell proliferation and migration, then the important role of monocytes in cancer has recently been uncovered. More recently, some studies demonstrated that lymphocyte-to-monocyte ratio (LMR) is associated with prognosis in several cancers. In our cohort, the multivariate analysis showed that the single CpG DNA methylation at *FKBP5* gene and LMR were independently associated with OS.

Our results suggest the potential use of DNA methylation profiles, considering age acceleration measures and LMR, in blood to develop non-invasive tests for MPM detection in asbestos-exposed subjects. Using advanced statistical modeling, as for example the SEM, can improve the knowledge about complex system biology of MPM.

High-throughput technologies have revolutionized medical research. The advent of genotyping arrays enabled large-scale genome-wide association studies and methods for examining global transcript levels, which gave rise to the field of "integrative genetics". Other omics technologies, such as proteomics and metabolomics, are now often incorporated into the methodology of biological researchers. As compared to studies of a single omics type, multi-omics offers the opportunity to understand the flow of information that underlies disease. Such complementary effects and synergistic interactions between omic layers in the life course can only be captured by integrative study of multiple molecular layers. Building upon the success in single-omics discovery research, population studies started adopting the multi-omics approach to better understanding the molecular function and disease etiology. Multi-omics approaches integrate data obtained from different omic levels to understand their interrelation and combined influence on the disease processes.

The identification of simple and valuable prognostic markers for MPM will enable clinicians to select patients who are most likely to benefit from intensive therapy, and avoid subjecting unsuitable candidates to futile treatment. In addition to its involvement

in the pathogenesis of cancer-related cachexia, systemic inflammation can predict clinically meaningful outcomes, such as overall survival and response to systemic treatment. Knowledge of epigenetics has provided new therapeutic opportunities against cancer. Here, we summarize the current knowledge pertaining to epigenetic deregulations (restricting to peripheral blood DNA methylation) in MPM and present different useful models for estimating the occurrence of malignant mesothelioma (MM) after asbestos exposure and predicting potential treatment options.

SCIENTIFIC REPORTS

PUBLICATIONS

EWAS (focus of the thesis)

1. Blood cell DNA methylation biomarkers in preclinical malignant pleural mesothelioma: The EPIC prospective cohort. *International Journal of Cancer*, Online ahead of print. Alessandra Allione et al. (2022).
2. New DNA Methylation Signals for Malignant Pleural Mesothelioma Risk Assessment. *Cancers* 13, 2636. Giovanni Cugliari et al. (2021).
3. DNA Methylation of *FKBP5* as Predictor of Overall Survival in Malignant Pleural Mesothelioma. *Cancers* 12 (11), 3470. Giovanni Cugliari et al. (2020).
4. Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. *Journal of Thoracic Oncology* 14 (3), 527-539. Simonetta Guarrera et al. (2019).

Gene expression

5. Interleukin-3-Receptor-α in Triple-Negative Breast Cancer (TNBC): An Additional Novel Biomarker of TNBC Aggressiveness and a Therapeutic Target. *Cancers*, 14(16), 3918. Malvina Koni et al. (2022).
6. DNA repair and cancer in colon and rectum: Novel players in genetic susceptibility. *International journal of cancer* 146 (2), 363-372. Barbara Pardini et al. (2020).

GWAS and rare variants

7. Functional and clinical implications of genetic structure in 1686 Italian exomes. *Human Mutation* 42 (3), 272-289. Giovanni Birolo et al. (2021).
8. Advances in the Genetics of Hypertension: The Effect of Rare Variants. *International journal of molecular sciences* 19 (3), 688. Alessia Russo et al. (2018).

RoH and IBD

9. Associations of autozygosity with a broad range of human phenotypes. *Nature Communication* 10(1), 1-17. David W Clark et al. (2019).

Multi-Omics

10. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nature Genetics* 53, 1311-1321. Josine L Min et al. (2021).

Mendelian Randomization

11. Design and quality control of large-scale two-sample Mendelian randomisation studies. *MedRxiv*. Philip Haycock et al. (2021).
12. The association between genetically elevated polyunsaturated fatty acids and risk of cancer. *The Lancet Oncology.* Philip Haycock et al. (2022).

Meta-analysis

13. DNA methylation signature of chronic low-grade inflammation and its role in cardio-respiratory diseases. *Nature Communication* 13(1), 1-14. Matthias Wielscher et al. (2022).
14. Association of maternal prenatal smoking GFI1-locus and cardio-metabolic phenotypes in 18,212 adults. *EBioMedicine* 38, 206-216. Priyanka Parmar et al. (2018).

Artificial Intelligence

15. Analysis of the effects of orthodontic treatment on Instagram posts using artificial intelligence. *Journal of Aligner Orthodontics*, 9-16. Simone Parrini et al. (2022).
16. Improving the prediction of cardiovascular risk with machine-learning and DNA methylation data. *IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology (CIBCB)*. Giovanni Cugliari et al. (2019).

Neuroscience

17. Data Science in Neuroscience: A Review of The EEG Analytical Workflow. *Research & Investigations in Sports Medicine* 6 (2), 486-488. Giovanni Cugliari et al. (2020).
18. Electrocortical activity during stretch reflex in athletes. *Sport Sciences for Health*, 14(3): 625-631. Marco Ivaldi et al. (2018).
19. Delta and alpha rhythms are modulated by the physical movement knowledge in acrobatic gymnastics: an EEG study in visual context. *Sport Sciences for Health*, 13(3): 563-569. Marco Ivaldi et al. (2018).

Orthodontics

20. Lower Molar Distalization Using Clear Aligners: Bodily Movement or Uprighting? A Preliminary Study. *Appl. Sci*, 12, 7123. Elisa Rota et al. (2022).
21. Analysis of Class II Intermaxillary Elastics Applied Forces: An in-vitro Study. *Front. Dent. Med*, 18 January 2022 Sec. Dental Materials. Tommaso Castroflorioet al. (2022).
22. Micro computed tomography evaluation of Invisalign aligner thickness homogeneity. *Angle Orthod*. May; 91(3): 343–348. Edoardo Mantovani et al. (2021).
23. Short term dentoskeletal effects of mandibular advancement clear aligners in Class II growing patients. A prospective controlled study according to STROBE Guidelines. *European Journal of Paediatric Dentistry* vol. 22/2. SErene Ravera et al. (2021).
24. Interceptive treatment effects of EF performed appliance in pre-pubertal and pubertal skeletal Class II growing patients: A retrospective controlled study. *European journal of pediatric dentistry*, 21:123. Serena Ravera et al. (2020).

25. Scanning electron microscopy analysis of aligner fitting on anchorage attachments. *J Orofac Orthop*, 80(2); 79-87. Edoardo Mantovani et al. (2019).
26. Postural changes in orthodontic patients treated with clear aligners: a rasterstereographic study. *Journal of Electromyography and Kinesiology* 38; 44-48. Simone Parrini et al. (2018).
27. Does occlusal splint affect posture? A randomized controlled trial. *Cranio*. Ilaria De Giorgi et al. (2018).
28. Scanning electron microscopy evaluation of aligner fit on teeth. *The Angle Orthodontist,* 88(5): 596-601. Edoardo Mantovani et al. (2018).

Sleep Bruxism

29. Effects on Sleep Bruxism Activity of Three Different Oral Appliances: One Year Longitudinal Cohort Study. *Curr Drug Deliv*, May 19. Andrea Bargellini et al. (2022).
30. Use of clear aligners in the orthodontic treatment of sleep bruxism patients: a randomised controlled trial. *Journal of Aligner Orthodontics*, 293-300. Andrea Bargellini et al. (2020).
31. Effects of clear aligners on sleep bruxism: randomized controlled trial. *Journal of Biological Regulators and Homeostatic Agents*, ;32(2 Suppl. 2):21-29. Tommaso Castroflorio et al. (2018).

Nutrition and supplementation

32. Effects of nutritional intake on performance in master athletes during an extreme ultra-trail. *Journal of Human Sport and Exercise*, 5(4). Felicina Biorci et al. (2019).
33. Relationship of chlorophyll supplement and platelet-related measures in endurance athletes: a randomized, double-blind, placebo-controlled study. *Sport Sciences for Health*, 14(2): 449-454. Giovanni Cugliari et al. (2018).

Athletic performance

34. Variations in athletic performance with occlusal splint in track and field athletes: a randomized clinical trial. *The Journal of Sports Medicine and Physical Fitness*, 62(3):375-381. Simone Parrini et al. (2021).
35. Heart rate analysis during very deep dives in flooded caves: single case study. *International Journal of Physical Education, Fitness and Sports*, 9(3): 9-17. Marco Ivaldi et al. (2020).

CONFERENCES

1. Epigenomics of Malignant Pleural Mesothelioma: a structural equation modeling. *European Journal of Human Genetics*. Giovanni Cugliari et al. (2022).
2. The association between genetically elevated polyunsaturated fatty acids and risk of cancer. *ASHG*. Philip Haycock et al. (2022).
3. Comparison of Polygenic Risk Scores for Coronary Artery Disease in an Italian prospective cohort: the EPICOR study. *European Journal of Human Genetics*. Alessandro De Gregorio et al. (2022).
4. Interaction between asbestos exposure and stochastic epigenetic mutations in Malignant Pleural Mesothelioma. *European Journal of Human Genetics*. Giovanni Cugliari et al. (2021).
5. Relationship Between Asbestos Exposure and Epigenetic Age Acceleration in Malignant Pleural Mesothelioma. *International Conference of the International Mesothelioma Interest Group*. Giovanni Cugliari et al. (2021).
6. Relationship Between Stochastic Epigenetic Mutations and Asbestos Exposure in Malignant Pleural Mesothelioma. *International Conference of the International Mesothelioma Interest Group*. Giovanni Cugliari et al. (2021).
7. DNA methylation biomarker as predictor of survival in patients with malignant pleural mesothelioma. *International Conference of the International Mesothelioma Interest Group*. Giovanni Cugliari et al. (2021).
8. Asbestos exposure and epigenetic clock. Epigenomics of Common Diseases. Giovanni Cugliari et al. (2021).
9. Stochastic epigenetic mutations in Malignant Pleural Mesothelioma. *Epigenomics of Common Diseases.* Giovanni Cugliari et al. (2021).

10. Asbestos exposure is associated with epigenetic signs of aging. *Food, health and factors promoting longevity*. Giovanni Cugliari et al. (2021).
11. Association between asbestos exposure and epigenetic age acceleration. *European Journal of Human Genetics.* Giovanni Cugliari et al. (2020).
12. Peripheral blood DNA methylation as prognostic tool in Malignant Pleural Mesothelioma. *European Journal of Human Genetics.* Giovanni Cugliari et al. (2019).
13. Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. *European Journal of Human Genetics.* Simonetta Guarrera et al. (2019).
14. Meta-Analysis of Maternal Prenatal Smoking GFI1-Locus and Cardio-Metabolic Phenotypes in Adults. *Preprint Lancet.* Priyanka Parmar et al. (2018).
15. Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects. International Conference of the *International Mesothelioma Interest Group.* Simonetta Guarrera et al. (2018).

EDITORIAL REVIEWS

MDPI Topical Editor: Editorial of the special issue of the International Journal of Molecular Sciences (Molecular Mechanisms and Therapies of Malignant Mesothelioma, ISSN 1422-0067). This special issue belongs to the section "Molecular Oncology".

More than twenty reviews were performed during the PhD program including:

1. Cancers
2. Vaccines
3. IJMS
4. Cells
5. Life
6. Ijerph
7. Genes
8. Biomedicines
9. Current Oncology
10. Bioengineering

COURSES

1. International School in Data Science: Machine Learning and its Application in Cancer Genomics. IUSS - Scuola Universitaria Superiore Pavia. (2021).
2. International School in Deep Learning and Reinforcement Learning. Machine Learning Department at CMU. (2021).
3. International School in Quantum Computing Hard and Software. EPFL, École polytechnique fédérale de Lausanne. (2021).
4. International School in Discovering the unknown: AI in healthcare. Università degli studi di Pavia. (2021).
5. International School in Resource-aware Machine Learning. TU Dortmund University. (2020).