UNIVERSITY OF SALERNO

DEPARTMENT OF MEDICINE, SURGERY AND DENTISTRY

"SCUOLA MEDICA SALERNITANA"

PhD COURSE - XXXIII CYCLE

TRANSLATIONAL MEDICINE OF DEVELOPMENT AND ACTIVE AGING

***Curriculum* INNOVATIVE TECHNOLOGIES IN TRANSLATIONAL MEDICINE**

PhD DISSERTATION

Integrated approach in Human Biomonitoring: epigenetic profile for environment-health study correlation.

PhD Course Coordinator:

Ch.mo Prof. **Palmiero Monteleone**

Tutor:

Prof.ssa **Roberta Tarallo**

Candidate:

Biancamaria Pierri

Number 8800900024

ACADEMIC YEAR 2019/2020

UNIVERSITÀ DEGLI STUDI DI SALERNO

DIPARTIMENTO DI MEDICINA, CHIRURGIA E ODONTOIATRIA

"SCUOLA MEDICA SALERNITANA"

Corso di Dottorato di Ricerca - Ciclo XXXIII

MEDICINA TRASLAZIONALE DELLO SVILUPPO E DELL'INVECCHIAMENTO ATTIVO

***Curriculum* TECNOLOGIE INNOVATIVE IN MEDICINA TRASLAZIONALE**

TESI DI DOTTORATO

Integrated approach in Human Biomonitoring: epigenetic profile for environment-health study correlation.

Coordinatore:

Ch.mo Prof. **Palmiero Monteleone**

Tutor:

Prof.ssa **Roberta Tarallo**

Candidata:

Biancamaria Pierri

Matricola 8800900024

ANNO ACCADEMICO 2019/2020

Alla mia amata famiglia.

"[…]L'onestà e la sincerità ti rendono vulnerabile…non importa, sii onesto e sincero.
Quello che hai costruito per anni può essere distrutto in una notte…non importa, costruisci. […]"

Madre Teresa di Calcutta

To my beloved family.

"[…]Honesty and transparency make you vulnerable…be honest and transparent anyway.
What you spend years creating other could destroy overnight…create anyway. […]"

Mother Teresa

ABSTRACT

Human Biomonitoring (HB) represents a useful tool to investigate the complex relationship existing between environment and health.

Assess human exposure to external stressors means to deepen observe, with a holistic approach, all factors contributing to generate the exposure itself.

Considering the "Campania Region (Italy)" case study, the present PhD dissertation introduce and describe an innovative HB protocol that try to make up to the lack of data integration from complementary science fields. The biomonitoring model is based on a whole systemic analytical evaluation of the environmental context. The paradigm of the protocol considers three elements: identification of pollution sources, of pollutants migration ways and of effects on target organisms. It pursues its aims enrolling healthy human cohorts, in order to identify potential risks of exposure to pollutants and to potentially predict correlated clinical outcomes. Monitoring of environmental matrices, detection of exposure and effect biomarkers in human specimens, together with epidemiological evaluations, are integrated in a multilevel analysis in order to depict a framework of the regional state of contamination.

A particular focus in the dissertation is dedicated to epigenetic profiles obtained with DNA methylation array analysis, the most widely molecular mark observed in environmental epigenetic studies. The modulation of gene expression programs in response to environmental exposure is a key to understand gene-environment interactions and to deep insight into possible phenotypic and clinical effects on susceptible populations, as well as into the etiology of cancers and chronic-degenerative diseases. This means the possibility of translating the acquired knowledge into public health interventions and developing prevention strategies. Indeed, the long-term objective of the study, over the current PhD dissertation, is to create an integrated, dynamic map of environmental contamination in the Campania Region, in order to support the implementation of public health policy and provide a scientific reference model for the evaluation of exposure risk assessment.

According to a "One Health" perspective, the model aspires to be translated in different contexts and for various applications, and to become a useful tool for comparative, multi-criteria and multi-disciplinary analysis.

ABSTRACT

Il Biomonitoraggio Umano (HB) rappresenta un utile strumento per indagare la complessa relazione esistente tra ambiente e salute.

Valutare l'esposizione umana a stressors ambientali significa osservare profondamente, con un approccio olistico, tutti i fattori che contribuiscono a generare l'esposizione stessa.

Considerando il caso studio "Regione Campania (Italia)", la presente dissertazione di Dottorato di Ricerca introduce e descrive un innovativo protocollo di HB, che tenta di sopperire alla carenza di integrazione di dati tra discipline scientifiche complementari.

Il modello di biomonitoraggio si basa su una valutazione analitica e sistemica del contesto ambientale. Il paradigma del protocollo considera tre elementi: l'individuazione delle sorgenti di contaminazione, delle vie di migrazione dei contaminanti e dell'effetto sugli organismi target. Persegue tale scopo arruolando coorti di popolazione sana, al fine di identificare potenziali rischi di esposizione ai contaminanti e predire possibili conseguenze cliniche correlate. Un'analisi multilivello integra insieme il monitoraggio delle matrici ambientali, l'individuazione di biomarcatori di esposizione e di effetto in fluidi biologici umani, le valutazioni epidemiologiche, al fine di delineare una fotografia dello stato di contaminazione sul territorio regionale.

Un particolare focus nella dissertazione è dedicato ai profili epigenetici ottenuti dall'analisi degli array di metilazione del DNA, il marcatore molecolare più ampiamente utilizzato in studi di epigenetica ambientale. La modulazione dell'espressione genica in risposta all'esposizione ambientale rappresenta una chiave di comprensione delle interazioni gene-ambiente, utile per approfondire possibili effetti fenotipici e clinici sulle popolazioni suscettibili, così come l'eziologia delle neoplasie e delle patologie cronico-degenerative. Ciò permette di traslare le conoscenze acquisite in azioni di sanità pubblica e sviluppare strategie di prevenzione. Al di là della presente dissertazione di dottorato, difatti, l'obiettivo a lungo termine dello studio è di creare una mappa dinamica ed integrata dei livelli di contaminazione ambientale campani, al fine di supportare l'implementazione di interventi di sanità pubblica e fornire un modello scientifico di riferimento per la valutazione del rischio da esposizione.

In accordo con una prospettiva di tipo "One Health", il modello aspira ad essere traslato in differenti contesti e per varie applicazioni, e a diventare uno strumento utile per l'analisi comparativa, multi-criteriale e multidisciplinare.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Environment and Health

The environment, meaning as the whole physical, chemical, and biological factors external to a person (Pruss-Ustun et al., 2016), is an actual determinant for human life.

According to this omni-comprehensive definition, a fine and breakable balance governs the relationships between its biotic and abiotic compartments, and any perturbation affecting the system could be able to move this balance towards new arrangements. A combination of actions and reactions tries to restore the system balance, moved away from its "optimum", triggering mechanisms of positive (enhancing) or negative (regulatory) feedback of the phenomena. This type of regulation results in different condition of healthiness or risk for the unsettled ecosystem and its components.

Humans, as part of the environment and of the biotic compartment, are influenced by the succession of these events and by regulatory mechanisms during their lifespan. They are exposed to environmental events and contribute through their activities to the changes occurring in the environment itself. The main difference between natural and human-induced ecosystem changes is in the spatial and time scale of events.

Human activities, including population growth, industrialization, economic development, and urbanization, cause disturbances over time scales of decades to a few hundreds of years. Such environmental impacts and pollution processes could occur at smaller (local) spatial scale (e.g. point-source pollution associated with a hazardous waste site) or at a much larger scale, encompassing entire regions, until the entire planet (Maximilian et al., 2019).

It is well-known that, as dark side of the anthropization, the widespread of contaminants in the environmental matrices define a heavy threat on ecosystems health and particularly on target organisms, highlighting the environmental pollution as a growing concern for global health, and representing it as one of the great existential challenges of the Anthropocene epoch (Landrigan et al., 2018).

As stated by the World Health Organization (WHO), in 2012 approximately 24% of deaths recorded worldwide have been linked to environment, which is roughly 13.7 million deaths a year (this data includes infectious diseases related death). Moreover, the Global Burden of Disease (GBD) study estimates that the pollution-related disease was responsible for 9 million premature deaths in 2015 (16% of total global mortality) (Forouzanfar et al., 2016). Especially air, water, and soil contamination, together with electromagnetic pollution, industrial activities, in-door and occupational exposure, represent factors potentially correlated to chronic and degenerative diseases onset (Pruss-Ustun et al., 2016).

Over the past decades scientific research has supposed a link between exposure to environmental stressors and human diseases (Ziech et al., 2010). On the basis of these solid

scientific evidences a lot of substances derived from pollution phenomena were classified as carcinogenic or mutagenic due to their dangerousness (https://monographs.iarc.fr/agents-classified-by-the-iarc/). Most harmful compounds have been identified in occupational health studies as characteristics of professional exposure. For example, one of the most studied occupational carcinogens, the asbestos, had a terrible history of proved correlations between professional exposure and adverse clinical outcomes (Siemiatycki 2007; Frank et al., 2014).

Nevertheless, the sources of pollution are increasingly widespread and heterogeneous and, above all, follow diverse time of action in terms of exposure. This means that the phenomena do not affect only workers or cohorts of people with acute events, but the whole general population and, in a holistic point of view, affect the entire biotic compartment. Hence, humans are exposed to many environmental stressors during life, but the real balance is often represented by the addition of a series of environmental factors not linked to pollution, called confounders, which contribute to generate the final consequence in term of health or diseases onset (e.g. individual life or food habits).

The interpretations as to the extent of environmental or confounders remains unclear own to the multifactorial nature of diseases. Some observations suggest the minor role of environmental pollutants in chronic diseases incidence rather than lifestyle factors, others underline contaminants as the major contributing factors (Belpomme et al., 2007a; Belpomme et al., 2007 b; Irigaray et al., 2007; Clapp et al., 2008; Miller et al., 2012, Joseph et al., 2013).

In 2018, The Lancet Commission on pollution and health has developed the concept of "Pollutome", as the totally of all forms of pollution that have the potential to harm human health causing illnesses (Fig. 1A). In the opinion of the Commission it could be viewed as a fully contained (nested) subset of the "Exposome", the term coined in 2005 by Wild to identify every exposure to which an individual is subjected from conception to death. The Exposome complements the genome by providing a comprehensive description of lifelong exposure history (Wild 2005 and 2011), and its multidisciplinary measurement could be considered a useful way to improve prevention strategies from a public health perspective (Siroux et al., 2016) (Fig. 1B).

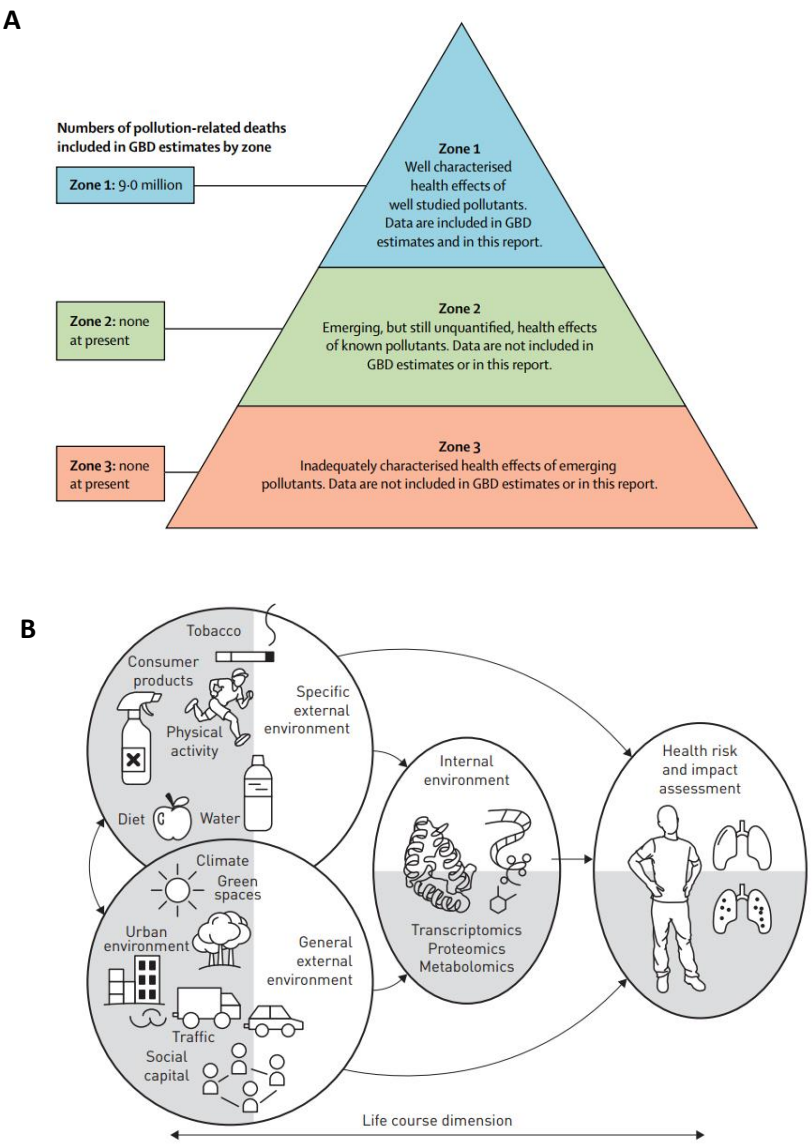


Figure 1. **(A)** The Pollutome. The totally of all forms of pollution that have the potential to harm human health. The Lancet Commission on pollution and health divides the Pollutome into three zones. Zone 1: well-established pollution-disease pairs, with robust estimate contribution to the global burden of disease; zone 2: emerging effects of known pollutants, where evidence of causation is building but associations are not fully characterized; zone 3: new and emerging pollutants, whose effects are unknown and not yet quantified (source Landrigan et al., 2018). **(B)** The Exposome. Effects and interactions between the three different domains of the Exposome: the general external, the specific external and the internal. The internal exposome comprises processes as metabolism, inflammation, aging etc.…The specific external includes for example lifestyle and food habits. The general external is the wider climatic, social, economic and psychological influences of the environment (source Siroux et al., 2016).

Up to this short overview, environmental pollution problems could be considered as a multilevel sum of past and present events, which could or not affect people at the same time, and nor to the same intensity or way (Trevors, 2010). Notwithstanding in the last decades there was a raise attention and a higher awareness for this hot issue, the complexity

and the interconnection of the problems require likewise complex and interconnected solutions.

Only through effective methods and models it could be possible to identify the variables and the causal links that feed the phenomena and influence each other's parts. Observational studies focusing on these phenomena, on ecosystem alterations and on related biological modifications, allow to go back to the source of the alteration itself and to identify mechanisms of cause-effect. Therefore, the state of well-being or disease will be determined by the correlations between all the variables that contribute to generate the exposure. Actually, the "exposure" represents the key concept that could give, by its measurement, the chance to define the relationships between environment and health.

1.2 Environmental Risk Assessment and Human Biomonitoring

The complex investigation of the causal link between environment and health could be interpreted through a multi-level analysis that starts from the definition of environmental exposure, through its evaluation, up to the assessment of the associated risk for health.

Thus, exposure measurement becomes a useful tool in understanding, characterizing, and reducing associated health risks (U.S. EPA 2019).

In a general meaning, the "Risk" could be defined as an expectation determined by the observation of phenomena and variables, with a percentage of uncertainty depending on the field of application. From a mathematical point of view, the risk is defined as a combination of the probability of an event occurring and the severity of its damage, therefore the extent of the consequences determined. Accordingly, the risk becomes a two-dimensional assessment expressed by the two variables: probability and magnitude. Thanks to these two factors, it permits to identify, predict or prevent, a series of phenomena whose effects could be deleterious. The system investigated must be screened, analyzed, deepened, in its entire structure, through protocols and methods that allow to generate a predictive definition and verification model. This model should be a tool at the service of authorized decision-making levels.

To measure the exposure means to assess the "Environmental Exposure Risk".

In general, it is carried out to examine the effects of an agent on humans (Health Risk Assessment) and on ecosystems (Ecological Risk Assessment) (Dinis et al., 2009).

The European Environment Agency in 1998 described a series of key task in general Risk Assessment to carry out especially for the environmental ones as guideline:

i. Problem formulation: pollutants characteristics and chemistry;
ii. Hazard identification: presence and toxicity of chemicals;
iii. Release assessment: sources and rate releases;

iv. Exposure assessment: target of exposure and type of exposure;
v. Consequence or effect assessment: how is it toxic and at what exposure levels and what is the effect on the receptors;
vi. Risk estimation and characterization: what does the risk assessment tell us about the situation and what are the risks, quantitative or qualitative;
vii. Risk evaluation: how important is the risk to those affected, those who create it, and those who control it.

It follows that the risk characterization and/or risk evaluation are used as input for risk management in order to come up with an answer to which actions should be taken and how should the remaining risks be handled (Fairman et al., 1998).

The scientific and methodological approach aimed to estimate exposure importance and extension is the Human Biomonitoring (HB).

According to Centre of Disease Control and Prevention (CDC) definition, HB is a method for assessing human exposure to pollutants or their effects, by measuring these chemicals, their metabolites or reaction products in human specimens (CDC, 2005).This means a repeated and controlled quantification of these markers, on susceptible subjects (or exposed in the past) in a general environment or in a workplace (Zielhuis and Henderson 1986; Manno et al. 2010).

Exposure to environmental pollutants occurs through different routes, such as inhalation, ingestion, and dermal absorption and the amount of pollutant uptake is named as the "absorbed dose." Thus, the body burden of a specific pollutant is determined by factors, such as the pollutant's concentration in a specific environmental medium, its physical and chemical properties, and timing of exposure, as well as individual factors, such as uptake, metabolism and excretion rates. HB considers all these factors by measuring the concentrations of a chemical or its metabolites in human matrices (WHO, 2015).

It was firstly used in occupational medicine and then transferred to life habits or territorial emergencies.

Concerning occupational medicine, the human biomonitoring allows to figure out and acquire data about health consequences after exposure to toxic compounds, thanks to its measured time and known type of contact to single elements or to a mixture. Examples are petroleum distribution workers (Heibati et al., 2018) or incinerator workers (Mauriello et al., 2017) exposure assessment.

Different analytical methods to estimate the uptake of occupational chemicals in human body fluids were developed in the early 20th century (e.g. lead and a few other compounds and relative metabolites) (Angerer et al., 2007). Many years later ('80s), HB was included in exposure and health surveillance programs as part of regulatory requirements, for instance for chemical industries. Indeed, when HB is compared to environmental evidence,

it shows itself to be an effective tool to improve occupational risk assessment at individual or group level, supporting the same health surveillance plans (Manno et al., 2010). During the health surveillance each worker is usually subjected to imaging, instrument, and biochemical testing to detect any clinically relevant occupation-dependent change. The use of biomarkers and of HB approach represents not only the method to estimate the internal exposure or to predict potential health effects, but also the chance to assess a causal relationship between health impairment and chemical exposure, when a change is firstly detected in exposed workers.

Therefore, in the field of occupational medicine, the "occupational risk assessment" could be defined as a qualitative and quantitative characterization of an occupational risk; it consists of three fundamental tools: environmental monitoring, health surveillance and biological monitoring. It becomes clear that the advantage of using biomarkers and HB approach in occupational surveillance and study is recognizable when the toxicological significance of the compound object of the analysis is sufficiently known. Some examples are below. Zidek et al in 2017 reported in a review the use of HB data for regulatory risk assessment in "Canada Chemicals Management Program". DEMOCOPHES study (Schwedler et al., 2017) was developed and carried out in Europe, with an attempt of harmonize HB studies on European scale, in order to support regulation. Same purposes for Human Biomonitoring for Europe (HB4EU) project (www.hbm4eu.eu), which joints efforts of many countries and environmental agency, coordinating and advancing HB in Europe, to provide scientific evidence for chemical policy making.

Nowadays the application of HB is beyond the limits of the occupational surveillance and is an attractive approach for research activities (Manno et al., 2014).

Despite the impressive potential of the HB, particularly in occupational risk assessment, it is not always correctly used or interpreted. Some ethical issues become evident. As early as in 2002, The Code of Ethics of the International Commission on Occupational Health (ICOH, 2002), highlighted some questions:

1. The aims of the study: if HB is performed for research interest or for health surveillance.
2. The definition of an appropriate study design: depending on the aims, the right choice of biomarkers, of matrices, of sample size (WHO's guidelines suggest a minimum of 120 subjects randomly selected), of reference group, of inclusion and exclusion criteria of enrollment.
3. The scientific approach and methodology.
4. The biomarkers validation and their predictive value.
5. The operative aspects: samples collections and analysis.
6. Interpretation of biomarkers data and management of the results.
7. Data protection and communication.

Beyond the occupational risk assessment, when HB is combined with health data, they together create an epidemiological study, aimed to observe the body burden of pollutants and their health effects. In addition, when HB is surrounded by environmental evidence it gives the possibility to obtain a systemic and full view of the condition investigated. In fact, HB alone usually does not reveal exposure sources and routes. Thus, environmental monitoring remains crucial for the development of targeted policy actions (WHO, 2015).

Monitoring and analyses on environmental matrices (soil, water, air, food, biosentinel models), gives the overview of territorial context, a state of art of critical issues and of potential contaminated areas, mapping the spatial dimension of compounds and geolocalizing the putative sources of pollution (Balassone et al. 2016; Ducci et al. 2017; Ducci et al. 2019).

Actually, in the systemic overview, the multidimensionality and the complexity of the observations are what gives the chance to divide the phenomenon in several levels of analysis, which, in the final intersection, could provide the global reading and interpretation.

Despite the great potential of the approach, the main problem in the majority of environmental assessment protocols and human biomonitoring studies seems to be exactly the lack of data integration and methodological sharing between different science fields, as well as the lack of a holistic approach through which acquire a real snapshot of the overall degree of contamination. Frequently also the epidemiological analyses are twisted by a unilateral view, led by the observations on mortality and diseases incidence, linked with a single source of pollution, without any consideration for the sum of variables contributing to the condition.

1.3 Exposure, Effect and Genetic Susceptibility Biomarkers

HB protocols consider the determination of exposure, effects and susceptibility biomarkers in order to define the internal dose of pollutants and their direct effects on biological and cellular system (Migliore 2018) (Fig.2).

Biomarkers of Exposure are the results of exposure itself and estimate chemicals or metabolites, in body fluids or in tissues, reflecting the internal dose and the level of pollutants absorbed. Inorganic and organic compounds, such as heavy metals, dioxins, dioxins-like and polychlorinated biphenyl (PCBs), and some of their derivatives, are among them. Biomarkers of Effect represent the biochemical or molecular reaction of the system to the xenobiotic compounds. They identify early and generally reversible biological effects: the measure of DNA damages, chromosome aberrations, sister chromatid exchange, DNA adducts and proteins, inflammatory cytokines, epigenetic changes.

Biomarkers of Susceptibility reflect the individuality, that is the predisposition to be susceptible to adverse effects related to exposure. For example, polymorphisms of relevant xenobiotic metabolizing enzymes.

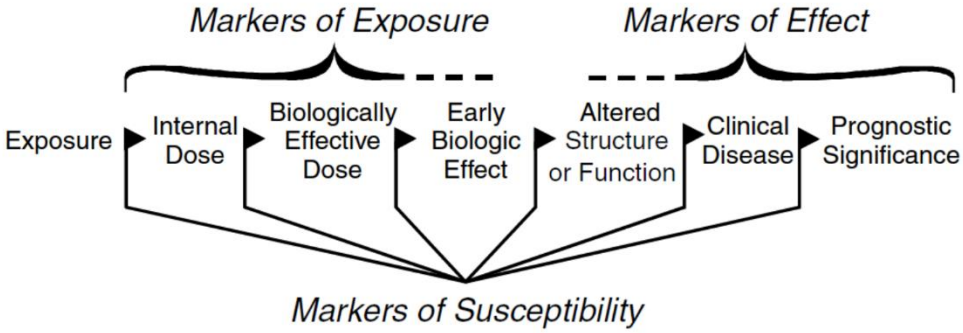


Figure 2. The continuum from exposure to disease. During different life stages biomarkers allow to evaluate the way and the magnitude of chemical interferences on biological processes, up to the alteration of structure and function, which means disease onset (source De Bord et al., 2015).

An important issue, in a correct HB study design, is the appropriate choice of the sample matrix to analyze. It could depend on the study design, on the aims and on the type of exposure investigated in the HB. In the last decades, the improvement in technologies and methodological approaches allow to have very low limits of quantification (LOQ) and, in consequence, have expanded the use of non-invasive matrices even if with low concentration of xenobiotics investigated.

As reported in Table 1, blood is one of the most suitable matrices for a series of investigations, given the half-life of its cellular components and the balance at systemic levels with other organs and tissues, even if it is limited by the invasiveness of the collection procedures (Smolders et al. 2009). A lot of heavy metals, such as lead and cadmium, are measured in blood. Dioxins and dioxins-like compounds, given their lipophilic nature, were originally evaluated in adipose tissue, but from ‘80s they are successfully estimated in serum, even if its low percentage of fats (WHO, 2015). Over the recent years, because of the great amount of serum required for high resolution analysis methods for these organic compounds (Brasseur et al., 2014), a lot of HB considers alternative specimens. For instance, breast milk of new mums is as a good reference matrix for this type of measurement given its lipophilic nature and given the level of information obtained on mother and child at the same time (Fang et al., 2013; Schuhmacher et al., 2019).

Urine is probably the most frequently used matrix in HB to quantify the degree of environmental or occupational exposure to pollutants, especially for substances with short biological half-life. The collection and analysis of urine samples carries no associated risk, and large volumes can at once be gathered per individual (Esteban et al., 2009). Urinary

samples are useful for the analysis of inorganic compounds and metabolites of organic ones, for instance volatile organic compounds, like vinyl chloride, or inorganic mercury, which reflects a recent exposure to this pollutant.

Other less common matrices, like hair, nails, saliva and teeth are practical for HB on children cohorts, given their non-invasiveness. These alternatives, with derivative of the most common ones, are generally used for in-depth analysis and particular tests. For example erythrocytes, a blood derivative, are the more appropriate for the Hexavalent Chromium Cr(VI) speciation; in fact Cr(VI) ions, after uptake by inhalation or percutaneously are carried in the blood plasma and penetrate, depending on the concentration, into the erythrocytes, which represent the golden standard target for its quantitative determination (Lewalter et al., 1985). Similar consideration for hair samples, used for Mercury (Hg) speciation. The growth rate of hair (1 cm per month) and the tendency of Hg to accumulate in hair make it possible to estimate long-term exposure. In addition, tracing the dynamic changes along hair growing direction provides an insight into average exposure on timescale ranging from weeks to months (Li et al., 2019a). Generally, more than the 80% of Hg in hair is as Metyl-Hg, which is taken up by hair follicles as MeHg-cysteine complexes (Pino et al., 2018). The semen is useful for the evaluation of Lead exposure; because of the evidence of structural alteration of spermatozoa in workers exposed to the element (Naha et al., 2005), there is a growing use of this specimen for the evaluation of exposure and the correlation with male infertility (Fatima et al., 2010 and 2015; Onul et al., 2018).

Table 1. Biological matrices used in HB studies (source WHO, 2015).

Matrix	Population	Advantages	Limitations	Compounds measured in the matrix
Blood, serum, plasma	General	In equilibrium with all organs and tissues. Well established standard operating procedures (SOPs) for sampling.	Invasive; trained staff and special materials required. Volume limitation. Special conditions for transport and shipment.	POPs, metals/trace elements, organic compounds, tobacco smoke. e.g.: alkylphenols, mercury, lead, BFRs, dioxins, water disinfection byproducts, fluorinated compounds, organochlorine pesticides, organophosphate pesticides, phthalates, PCBs, dioxins.
Urine	General	Non-invasive, easy collection, no volume limitation. Allows analysis of metabolite	Composition of urine varies over time.	Metals/trace elements, organic compounds, tobacco smoke. Metabolites of environmental pollutants. e.g.: mercury, cadmium, arsenic, organochlorine compounds, BPA, organophosphate pesticides, parabens, phthalates, PAHs, benzene.
Hair	General, with few exceptions (i.e. neonates)	Non-invasive; minimum training required for sampling. No special requirements for transport and storage. Information about cumulative exposure during previous months. Segmental analysis is possible.	Hair is exposed to the environment and can be contaminated. Potential variations with subject's hair colour, hair care or race.	Metals/trace elements, POPs e.g.: total mercury, methylmercury, arsenic, cadmium, parabens, organochlorine compounds
Cord blood	Specific	Non-invasive; provides information about mother and child. Well defined SOP for peripheral blood can be used for cord blood.	Only available at birth in maternity ward settings. Ethical constraints. Special conditions for transport and storage.	POPs and other organic compounds, metals/trace elements, tobacco smoke, e.g.: alkylphenols, mercury, lead, BFRs, dioxins, water disinfection byproducts, fluorinated compounds, organochlorine compounds, organophosphate pesticides, phthalates, PCBs.
Breast milk	Specific	Provides information about mother and child. Enriched with lipophilic compounds.	Somewhat invasive. Restricted period of availability. Depuration of chemicals during lactation should be considered.	POPs, metals/trace elements, organic compounds, tobacco, e.g.: alkylphenols, BPA, dioxins, BFRs, fluorinated compounds, PCBs, organochlorine pesticides, lead, cadmium, mercury, phthalates
Amniotic fluid	Specific	Invasive	Limited to women undergoing amniocentesis or Caesarean section.	POPs, organic compounds, metals/trace elements, e.g.: phthalates, mercury, organochlorine pesticides
Placenta	Specific	Non-invasive	Restricted to certain period of life; requires homogenization	Metals/trace elements, POPs, organic compounds, e.g.: mercury, cadmium, lead, organochlorine pesticides, BPA, BFRs, dioxins, PCBs, PAHs, phthalates

Matrix	Population	Advantages	Limitations	Compounds measured in the matrix
Meconium	Specific	Non-invasive, easy collection. Reflects prenatal exposure.	Only available at birth.	Metals/trace elements, POPs, organic compounds, tobacco, e.g.: mercury, cadmium, organochlorine pesticides, PCBs, phthalates
Nails	General	Non-invasive, easy collection. No special storage or transport requirements. Provide information about short and long term exposure.	Exposed to the environment and can be contaminated (toenails are less exposed). Less documented for HBM applications.	Metals/trace elements, tobacco, e.g.: arsenic, mercury, cadmium, lead
Deciduous teeth	Specific	Non-invasive. No special requirements for transport and storage.	Low availability; restricted to a certain period of life. Less documented for HBM applications.	Metals/trace elements, organic compounds, tobacco, e.g.: lead, cadmium
Sweat	General	Non-invasive	Difficult collection. Less documented for HBM applications.	Metals/trace elements, organic compounds, e.g.: lead, cadmium

Matrix	Population	Advantages	Limitations	Compounds measured in the matrix
Semen	Specific	Mainly used to measure biomarkers of effect.	Invasive. Only available from males. Less documented for HBM applications	Metals/trace elements, POPs, organic compounds, e.g.: mercury, phthalates, dioxins, PCBs, organochlorine pesticides
Exhaled air	General	Non-invasive. Direct assessment of exposure through air	Limited to volatile chemicals. Difficult sampling, transport and storage.	Metals, disinfection byproducts, e.g.: lead, cadmium, trihalomethanes
Saliva	General	Non-invasive, easy collection	Lower concentrations of analytes than in blood; requires sensitive analytical techniques. Variation in flow rate and composition. The use of stimulant or absorbent pads can interfere with analysis. Less documented for HBM applications.	Metals/trace elements, organic compounds, POPs, tobacco, e.g.: cadmium, phthalates, BPA, PCBs, dioxins

1.4 Omics Sciences and Translational Knowledge for Human Biomonitoring

Beyond exposure, effect and susceptibility biomarkers, the WHO's guidelines underline the role over the last decades of the "omics sciences" for HB, which offered the opportunity to enhance the knowledge of the exposure-response continuum in health risk assessment.
According to McHale et al., (2010), the perturbations causing the onset of human diseases can be assessed by measuring the components of the "responsome", by means toxicogenomics technologies. Especially they refer to trascriptome, proteome, miRNome and methylome. In point of fact, the study of genome structure and function and its response to external stressors could integrate some fundamental aspects in HB assessment by toxicogenomic perspective, in order to understand molecular mechanisms subtended to toxicity (National Research Council, 2007; Vineis et al. 2013). A deep insight into gene-environment interaction, through pathways, genetic, protein and metabolic alterations, pointed to predict phenotypic and clinical effects on susceptible populations (Albertini et al. 2000; Krishanu 2011). Noteworthy, these observations have also underlined the involvement of endocrine pathways, immune modulation, inflammatory response, as silent or symptomatic mediators of biological system perturbation associated to environmental exposure (Chighizola et al. 2012; Gilbert et al. 2016; Almeida et al. 2019; Hassan et al. 2019; Macchi et al. 2019; Zhao et al. 2019).
The toxicogenomics analysis proceeds by two different approaches: one focused mainly on identifying mechanisms of toxicity in animal studies or *in vitro*, another carrying out analysis on well-characterized exposed cohorts of people, in order to directly determine biomarkers of exposure and early effects and assess dose-response (McHale et al., 2010).
The omics techniques (adductomics, transcriptomics, proteomics, epigenomics and metabolomics), provide a molecular "fingerprint" of the exposure or early effects, comparable with profiles generated from already known toxicants.
A series of study have described, with an omics perspective, the relationship between exposure and molecular answer to external stressors (Forrest et al., 2005; Zhai et al., 2005; van Leeuwen et al., 2008; McHale et al., 2009). Reading these studies, it is clear the need of an accurate experimental design and a detailed data collection. Moreover, in the analysis of the advancing in risk assessment (National Research Council, 2008), authors suggest, as recommendation for this approach, the harmonization of cancer and non-cancer risk observations, to correct the deficiencies in the treatment of uncertainty and variability for this type of quantitative risk assessment. Human toxicogenomics data, if unbiased, could potentially generate biomarkers and enrich the knowledge on the mechanisms underlying several human diseases. In fact, toxicogenomics endpoints not only describe gene-environment interaction and, with an appropriate sample size, could potentially give information on human heterogeneity.

Most human chronic diseases are result of the interactions between inherited genetic factors and modifiable environmental factors (Hunter, 2005). Likewise, most cancer cases are consequence of the interaction of genetic variants and environmental factors. In 1997 Perera estimated that only the 5% of cancers' development is led by genetic factors. For example, genetic susceptibility contributes only for 5-10% on the total incidence of breast cancer cases. The development of sporadic breast cancers might be indeed related to the interaction of genetic variation with environmental exposures (e.g. radiation, organic compounds, carcinogens, xenoestrogens) (Balmain et al., 2003). Molecular epidemiological studies have revealed that many environmental factors may interact with genetic variants to affect the risk of cancer development (Ning et al., 2014). In this context, toxicogenomics has become a fundamental tool for HB, with the primary goal to understand the correlations between environmental stress and human disease susceptibility, exploring the molecular mechanisms of environmental mutagens and/or carcinogens, and identifying potential biomarkers of disease and toxicity, including mutagenicity and carcinogenicity (Waters et al., 2004). Generally, there are two ways to perform toxicogenomics studies: (i) starting with conventional toxicological research and then focusing on omics approaches to detect systematic biological effects, and (ii) starting with the omics study followed by conventional toxicological research to interrogate molecular mechanisms. In this effort, it becomes a challenge to determine the mechanisms of toxicity induced by mixtures containing many chemical components (Ning et al., 2014).

Understanding the interactions between genetic and environmental factors will provide insights into the etiology of cancers and chronic-degenerative diseases. This means the possibility of developing prevention strategy and translate the acquired knowledge into public health actions. Ning et al., supposed that in the next years, powerful infrastructures will be developed to support the storage, access, analysis and management of the huge amount and diversity of data derived from HB and toxicogenomics approach. Powerful bioinformatics tools will be developed for identification of genetic variants and genetic-environmental interactions that contribute to cancer risk. This will hopefully lead to better diagnosis, treatments, and ultimately prevention.

1.5 Environmental epigenetics

Environmental factors can affect the expression of the genome also through alterations in the epigenome, which is intended as the whole epigenetic modifications present along the DNA sequences in a particular cell type (Wild et al., 2013). The term "epigenetic", literally "above the genetics", was firstly used by Conrad Waddington in the 40s to explain "[…] the interactions of genes with their environment which bring the phenotype into being". Epigenetic modifications do not entail change in the genome code, but only in the gene expression. Hence, it provides the link between genes and phenotype. Thanks to the advancing in this research field, it was suggested the role of epigenetic mechanisms in modulation of gene expression programs in response to environmental exposure. Actually, epigenome functions can be considered an interface between the genome and the environment (Hou et al., 2012). Despite the dynamism, due to its reversible changes, an alteration in epigenome may represents a more stable signature of environmental exposure than transcriptomic ones (Wild et al., 2013). Some authors identified epigenetic marks as early changes in the biological system investigated in reaction to exposure to some external stressors. Among them, the tobacco smoking (Vaissere et al., 2009; Wan et al., 2012), the benzene (Bollati et al., 2007), the particulate matters and air pollution (Tarantini et al., 2009; Bollati et al., 2010), the arsenic (Bailey et al., 2013). It is clear that epigenetic plays a mechanistic role in the consequent disease processes. Since epigenetic regulates cellular plasticity, that is the ability of a cell to rapidly respond to its environment, and is a key regulator of critical cell processes including genome stability and gene transcription, it offers an attractive biological mechanism to explain how environmental exposures can cause disease (Ladd-Acosta and Fallin, 2016).

Epigenetic modifications include DNA methylation, histone modifications and long non-coding RNA activity. They are physiologically involved in the regulation of many cellular processes, including DNA-protein interactions, suppression of transposable elements mobility, cellular differentiation, embryogenesis, X-chromosome inactivation and genomic imprinting, regulating chromatin remodeling and spatial and temporal gene expression (Meissner et al., 2008; Portela and Esteller, 2010).

1.5.1 DNA Methylation mechanisms

DNA methylation is the most deepened epigenetic mechanism in relation to environment stressors. It occurs when a methyl group ($-CH_3$) is covalently bonded to the carbon 5 (5C) position of a cytosine and this happens exclusively in the context of CpG dinucleotides (Clark et al., 1995). The CpG dinucleotides tend to cluster in some regions, even if they are distributed in the whole human genome, representing about 1% (Varriale and Bernardi, 2010). High-density regions called "CpG islands" have more than 500 bases, with more

than 55% GC (guanidine and cytosine) content and an expected/observed CpG ratio of > 0.65. Approximately 40-60% of human gene promoters contains CpG islands and are usually unmethylated in physiological conditions; about 6% of them become methylated in a spatial- (tissue) and temporal- (lifespan) specific manner (Fig. 3a). Up to 2 Kb (kilobases) upstream the CpG islands, are located the "CpG shores", with more than 75% of tissue-specific differentially methylated regions. Methylation in shores shows higher correlation with gene expression than CpG islands (Fig. 3b). Finally, "CpG shelves" are located about at 4 Kb from islands. Occasionally some CpG clusters are located within the body of the gene (Fig. 3c), or even in the UTR region (untranslated region), where they are normally more susceptible to methylation (Esteller, 2002). Considerable amount of CpG are also found in repetitive DNA elements such as transposons and retrotransposon-like elements. Highly methylated CpG in these regions are necessary to prevent transposition, chromosome and genome instability (Ehrlich et al., 1982) (Fig. 3d).

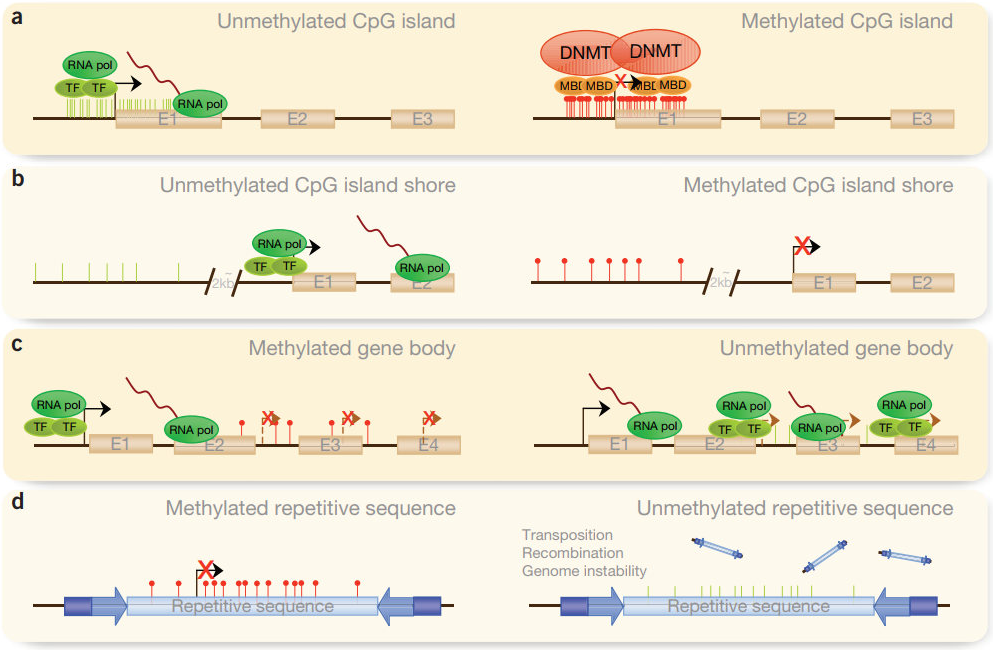


Figure 3. DNA methylation patterns. The figure describes on the left column a normal condition, on the right the altered one. (a) CpG islands when are located at promoters of genes are normally unmethylated, allowing transcription. When they are methylated cause transcriptional inactivation. (b) Same mechanism characterizes CpG shore. (c) Methylation along the gene body in a normal scenario is closely related to transcriptional activation, elongation efficiency and prevention of spurious initiation of transcription. It is positively associated to gene expression. In altered scenario the loss of methylated sites along the gene body could promote origin of transcription at incorrect points. (d) Methylation at repetitive sequences allows to prevent chromosome instability, translocations and gene disruption. Unmethylated scenario is associated to diseases onset (source Portela and Esteller, 2010).

DNA methylation is mediated by DNMT (DNA methyltransferase) enzymes, which allow to transfer a methyl group from S-adenosyl methionine (SAM) to the 5C of DNA cytosine.

The members of the DNMT family, which effectively possess the methyltransferase activity, are classified in "*de novo* DNMTs" (DNMT3A and DNMT3B) and "maintenance DNMTs" (DNMT1). During the embryonic development DNMT3A and DNMT3B are responsible of the methylation pattern setting. DNMT1, the maintenance enzyme, is able to preserve the methylation pattern already existing, because of its affinity for emi-methylated DNA sequences. It has the role of catalyze methylation on a DNA strand associated to a methylated CpG dinucleotides during semi-conservative DNA replication. Indeed, it is the most abundant DNMT in the cell, and is expressed mostly during S phase of the cell cycle.

The 5C of the cytosine could also be modified in 5- hydroxymethylcytosine (5hmC), the oxidative product of 5-metylcytosine, catalyzed by TET (ten-eleven translocation) family enzymes. This epigenetic mark plays a role in embryonic stem cells and blood cells differentiation, in neuronal development, through a mechanism of DNA demethylation able to regulate gene expression (Shock et al., 2011; Pastor et al.,2013).

Altered methylation pattern were highlighted in some pathogenic processes. Neoplastic conditions are generally characterized by gene-specific hypermethylation with a global hypomethylation scenario. While the global hypomethylation leads to chromosome instability, aberrant translocations and mutations, the hypermethylation that occur in CpG islands associated to gene promoters, cause transcriptional inactivation. When the gene is involved in the cell cycle control, DNA repair, Tp53 signaling, apoptosis, then the carcinogenesis is triggered (Portela and Esteller, 2010).

Aberrant DNA hypermethylation has been reported in the promoter regions silencing tumor-suppressor genes such as CDKN2A (cyclin-dependent kinase inhibitor 2A), CDKN2B (cyclin-dependent kinase inhibitor 2B), TP73 (tumor protein p73), MLH1 (mutL homolog 1), APC (adenomatosis polyposis coli), BRCA1 (breast cancer 1), MGMT (O-6-methylguanineDNA methyltransferase), VHL (von Hippel-Lindau tumor-suppressor), GSTP1(glutathione S-transferase pi 1), CDH1 (cadherin 1) and DAPK1 (death-associated protein kinase 1) (Sharma et al., 2010). For this reason, hypermethylated promoters are considered cancer biomarkers, with great diagnostic and prognostic potential (Li et al., 2009). Otherwise, the global hypomethylation could also be caused by dysregulation in DNMTs expression. Furthermore, epigenetic alterations are involved in neurodevelopmental, neurodegenerative and autoimmune diseases (Javierre et al., 2008; Urdinguio et al., 2009).

1.5.2 The role of DNA methylation in environmental epigenetics

DNA methylation is the most widely epigenetic mark observed in the environmental epigenetic studies. The exposure to external stressors has the potential to modify gene expression and adult disease susceptibility in different ways through changes in the

methylation patterns. These epigenetic changes can be inherited in somatic cells during mitosis, providing a potential mechanism by which the environmental impact on the epigenome can affect gene expression in long-term.

In addition, considering that the methylation pattern is also inherited in germinal cells division, the hypothesis of a transgenerational epigenetic inheritance represents a growing research field, especially referred to environmental exposure inducing epigenetic changes (Jirtle and Skinner, 2007; Feil and Fraga, 2012; Ladd Acosta and Fallin, 2015). Evidence from animal studies indicates that prenatal and post-natal environmental factors exposure (during the life-stages when a series of demethylation and methylation programming occur), can result in altered epigenetic programming and subsequent change in the risk of developing diseases. Likewise, mice germlines show that environmental influences could be linked to disease phenotypes in adulthood through the inherited epigenome modifications (Morgan et al., 1999; Lane et al., 2003; Anway et al., 2005; Anway and Skinner, 2006; Pembrey et al., 2006).

A series of studies on human cohorts were conducted to investigate the relationship between exposure to heavy metals, organic compounds and changes in DNA methylation patterns. Some evidence highlighted by HB studies are reported in Table 2, adapted from Migliore 2018 (Chapter 18).

Table 2. Human Biomonitoring using epigenetic biomarkers (adapted from Migliore, 2018 – Chapter 18).

Exposure	Sample size	Biological matrix	Epigenetic effect reported	Bibliography
Benzene	213 adults	Whole Peripherical Blood	DNA hypomethylation in LINE-1 and AluI	Bollati et al., 2007
PAH - urban smog (Polycyclic Aromatic Hydrocarbons)	56 children	Leukocytes from Cord Blood	DNA hypermethylation in CpG island associated with ACSL3 gene	Perera et al., 2009
POPs (Persistent Organic Pollutants)	70 adults	Whole Peripherical Blood	Global DNA hypomethylation	Rusiecki et al., 2008
Particulate Matters	718 adults	Leukocytes from whole peripherical blood	Global DNA hypomethylation	Baccarelli et al., 2009
Arsenic by drinking water	16 women	Leukocytes from whole peripherical blood	DNA methylation changes for breast cancer and diabetes correlated genes	Bailey et al., 2013

Endocrine Disruptor	196 pregnant women	Placenta	DNA hypomethylation in DMR (Differentially Methylated Region) IGF2 and H19	La Rocca et al., 2014

Arsenic exposure from drinking water was associated with pathways involved with breast cancer and diabetes onset by Bailey et al. in 2013. Also, Ameer et al. in 2017 associated arsenic exposure from drinking water with decreased gene expression and DNA hypermethylation in peripheral blood. The pathway analysis performed in this study suggests the involvement of genes related to cell death and cancer.

Benzene low-dose exposure was related, in HB on workers cohort as well in 2007 by Bollati et al., to a significant reduction in LINE-1 and AluI methylation, the repeated sequences representative of global DNA methylation. Results obtained were consistent with global methylation observed in malignant cells and particularly in acute myelogenous leukemia (AML). In addition, they verified hypermethylation in P15 promoter (which normally shows low or not methylation), and hypomethylation in MAGE-1 gene (with high methylation in normal tissues). Even these two observations were consistent with leukemia condition. Likewise, POPs and Particulate Matters from urban smog cause global DNA hypomethylation (Rusiecki et al., 2008; Baccarelli et al., 2009).

A systematic review published in 2015 by Ruiz-Hernandez et al., reported a synthetic overview of the studies evaluating the association of some environmental chemicals and DNA methylation. The Figure 4, adapted from the review, describes possible mechanism of action elucidating these interactions.

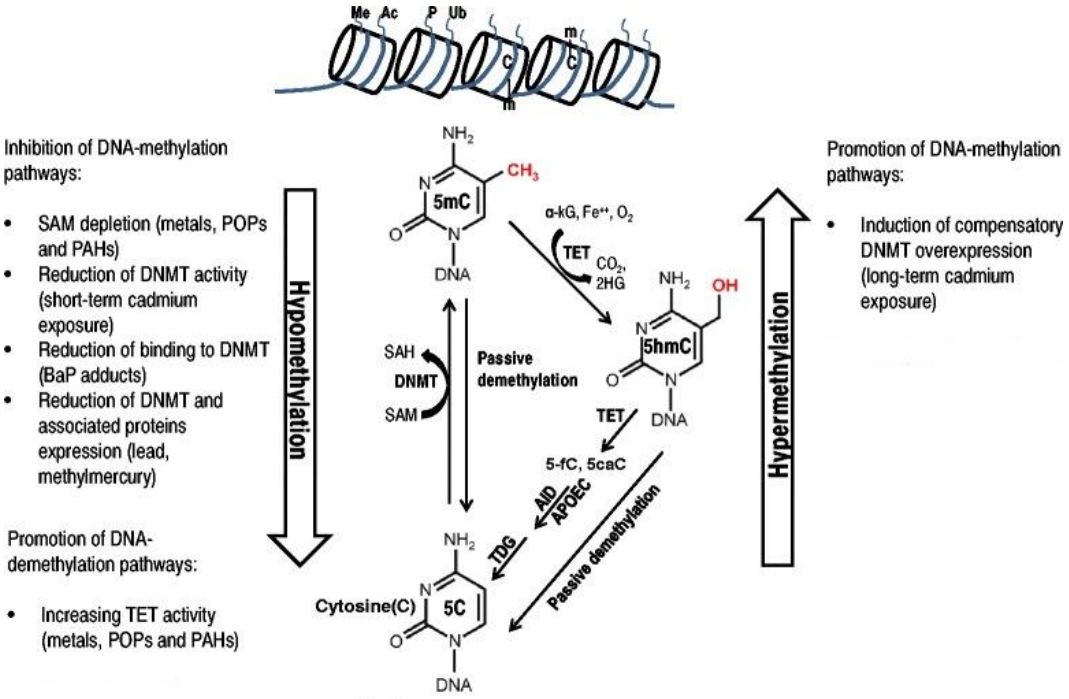


Figure 4. Possible mechanisms of action for environmental chemicals on DNA methylation. Metals, POPs and PAH tend to increase reactive oxygen species (ROS) formation. This alters methionine synthesis pathways, leading to a reduced synthesis of S-adenosylmethionine (SAM), substrate for DNMT enzymes. SAM depletion potentially inhibits DNA methylation, resulting in global DNA hypomethylation. Oxidative stress is proposed to stimulate the alpha-ketoglutarate (α-KG) production, which activated TET enzyme that, in presence of iron and oxygen, triggers demethylation pathway.

Short-term cadmium, PAH, lead and mercury exposure direct reduce DNMT concentration and activity. Conversely, long-term cadmium exposure induced compensatory DNMT overexpression, which increase global DNA methylation (adapted from Ruiz-Hernandez et al, 2015).

1.6 Environmental Context Overview: the Campania Region case study

Scientific literature documents some biomonitoring and epidemiological analyses about national areas of particular interest. Seveso in the Northern of Italy and Taranto in the South are two examples (Hay, 1980; Brambilla et al., 1998; Signorini et al., 2000; Warner et al., 2013; Vimercati et al., 2016; Buononato et al., 2016; Ingelido et al., 2017; Lovreglio et al., 2018). Among them, one of the most famous italian environmental contexts, considered in the last decades as a polluted land, is Campania Region. For several years, some areas of Campania have been scenario of irregular waste disposal and illegal dumping, defining a potential danger for spreading of persistent pollutants in the environment. In addition, due to a series of uncontrolled burning waste events, some of these areas gained the name of "Land of Fires" (LOF) (Legambiente Reports, 2013 and 2015). Over the years, public opinion has focused its attention on potential environmental damage arising from the LOF and its supposed correlations with adverse pathological outcomes, generating alarm. Probably because of an intense and inaccurate media coverage, and probably because of a complete unawareness of people on the true situation, the attention was brought on the Campania products, triggering a widespread distrust for what came from the territory, with a huge damage for agri-food sector. This emergency prompted special intervention by the Italian government, which set up an inter-ministerial task force (Law 6/2014) charged with guaranteeing the safety of agricultural production.

Many substances, chemical products and derivatives of burning and waste disposal are known to be harmful (especially those featured in the IARC ranking, in which substances are classified as carcinogenic on the basis of solid scientific evidences - https://monographs.iarc.fr/agents-classified-by-the-iarc/). Consequently, those associated with practices and events responsible of the LOF phenomenon, have been characterized from epidemiological and analytical point of view in various institutional and research contexts. This has made Campania one of the most thoroughly investigated Regions in Italy regarding the environmental pollution.

Several studies have been carried out (Mazza et al., 2018); some have focused on the measurement of exposure in humans or animals (Basile et al., 2009; Rivezzi et al., 2013; De Felip et al., 2014; Esposito et al., 2014; De Roma et al., 2017), others on the epidemiological evaluation of exposure and health effects (Senior and Mazza, 2004; Fazzo et al., 2011; Pirastu et al., 2013; Comba et al., 2014; Triassi et al., 2015). One series of monitoring activities on environmental and biological matrices showed levels of pollutants consistent with those expected for burning waste, as dioxins, dioxin-like and polychlorinatedbiphenyls compounds (Rivezzi et al., 2013). Likewise, compounds originating from urban areas and from rural and industrial activities have been detected

with specific distribution in the provinces (Pacini et al., 2013; Ducci et al., 2017; Thiombane et al., 2018). By contrast, several studies have revealed levels of pollutants in Campania comparable to or lower than the values registered in other European countries or in the industrialized areas of Northern Italy (Ulaszewska et al., 2011; De Felip et al., 2014). In 2014, Esposito et al. did not detect significantly higher levels of organic compounds (dioxins, dioxin-like, polychlorinated biphenyls congeners) in blood serum from people living in urban Naples and in "triangle of death" area, than those recorded in other European countries and Italian Regions. This was probably due to the lower presence of industrial plants and other recognized sources. In 2013, Pacini et al. reported that the main bodies of water in several areas of the Campania Region were in poor condition, owing to their relatively severe sediment contamination. Nevertheless, in the fish samples analyzed, they did not find high levels of the compounds investigated.

Concerning epidemiological evaluation, data from scientific literature actually reveal a decreased trend for oncological mortality from 1988 to 2009 in Campania, at the regional level, and, to a lesser extent, in the Provinces of Naples and Caserta (where LOF municipalities are concentrated), although in these same areas it was registered a significant deviation, from regional and national trend, with mortality excesses for some neoplastic and chronic conditions (Crispo et al., 2013). Furthermore, in 2014-2015, the SENTIERI report documented cancer incidence and mortality rates in towns in the LOF areas that were comparable to those recorded in areas with a history of industrial pollution in Northern and Southern Italy.

With regard to disease epidemiology, in 2014 Pirastu et al. reported higher rates of all-cause mortality and mortality due to neoplastic and chronic diseases in both sexes in two Campania areas, "Litorale Domizio-Flegreo and Agro-Aversano" and "Area Litorale Vesuviano", in comparison with the regional population; these data were later confirmed by Zona et al. in a 2019 update of the fifth SENTIERI report. However, only some of the medical conditions investigated can be correlated '*a priori*' with environmental causes. While the above authors hypothesize a correlation between disease onset and mortality rates with unauthorized waste disposal or occupational exposure, they also strongly recommend improving biomonitoring, epidemiologic evaluation and primary health interventions.

1.7 Integrated Approach for Human Biomonitoring

Diverse scientific publications on Campania Region case study and HB have stoked an epidemiological controversy (Cantoni, 2016). It is well known that the genesis of neoplastic disease is complex and multifactorial, therefore its investigation requires the integration of different evaluations on determining factors. In fact, the epidemiological evidence needs to take in account some other putative determinants of disease, including lifestyle, lack of a culture of prevention, social deprivation, and even the governance of the Regional Health System. Various epidemiological, clinical, geographical models have shown statistically significant correlations between several kinds of chronic or neoplastic diseases and exposure to diverse chemicals. However, much of the evidence available have been limited by the lack of fundamental information, such as environmental characterization in areas of analysis or the failure to investigate biomarkers or to adopt a molecular approach (Barba et al., 2011). None of the above-mentioned epidemiological evaluations have considered all the features of the phenomenon in its entirety, nor was able to analyze a homogeneous data-set: integration of data on the territory (soils, waters and air), food, animals, humans, or full-blown clinical outcomes (including confounding factors). Moreover, none of them have assessed the complex system of variables, environmental matrices and toxicokinetic which contribute to exposure and the putative risk for citizens' health. The real critical point in this type of analysis is the complexity of the environmental system, which includes biotic and abiotic compartments that coexist in a delicate balance.

Hence, to support a complete understanding of the phenomena, the entire system needs to be integrated into a single multilevel analysis, in order to define causal relationships throughout the network of factors involved.

For example, in HB studies, a fundamental contribute should be provided by the monitoring of the environmental contexts, especially by pollutants data in matrices (soil, water and air), in order to acquire a synthetic indicator of the contamination levels in the system analyzed. Additionally, in order to assess the contribution of diet and lifestyle, HB participants should be subjected to a validated questionnaire which collect these evidence (e.g. EPIC – European Prospective Investigation into Cancer and Nutrition studies); they also should undergo a complete medical examination to assess clinical parameters and anamnestic data. Finally, the sample size and the choice of matrices of analysis should fit with the primary endpoints of the study, with the possibility of integrating exposure, effects and susceptibility biomarkers (Pierri et al., 2020).

Concerning LOF and Campania Region case study, since data collected in different HB have so far been unable to demonstrate a causal link between human exposure to pollutants from dumps or other sources and increased mortality due to chronic and degenerative

diseases in some municipalities, this implicates an enhanced and compelling requirement of a methodological breakthrough.

As mentioned above, some assessments have not considered a set of specific variables covering environmental features, context observations and confounding aspects (lifestyle and dietary habits, social and economic deprivation, lack of a culture of prevention).

This underscores the pressing need of applying a new biomonitoring model, a new multidisciplinary and systemic approach with an analytical design capable of depicting the real state of contamination, and then to look in depth for the molecular mechanisms and the potential phenotypic consequences, with a view of assessment the exposure risk for residents and susceptible individuals.

2. AIMS OF THE STUDY

The purpose of the PhD dissertation is to introduce and describe an innovative HB protocol, carried out in the complex environmental context of the Campania Region, including the "Land of Fire" area. The dissertation is especially focused on innovations in HB study design and procedures and on biomarkers of exposure and effect considered in the health risk assessment for people residing in the regional territory.

The biomonitoring model is based on a whole systemic analytical evaluation of the environmental context. The paradigm of the protocol considers three elements: identification of pollution sources, of pollutants migration ways and of effects on target organisms, and pursues its aims enrolling healthy human cohorts, in order to identify potential risks of exposure to pollutants and to potentially predict correlated clinical outcomes.

The following chapters will describe:

1) Innovative study design and procedures, with environmental background characterization.
2) Exposure and effect biomarkers aimed to have a direct quantification of the environmental impact on the general population.
3) Exploring the environmental epigenetic marks, as DNA methylation, through an "omics" approach.
4) Data integration from several levels of analysis: monitoring of environmental matrices, human biomonitoring, toxicogenomic analysis and epidemiological evaluation, in order to obtain a framework of the regional condition, giving a spatial and geographical reference to pollutants, their sources and their migration ways. This could make up for the lack of methodological sharing and data merging among complementary scientific fields.

The long-term objective of the study, over the current PhD dissertation, is to create an integrated, dynamic map of environmental contamination in the Campania Region, in order to support the implementation of public health interventions and provide a reference model for the evaluation of exposure risk assessment.

Given its holistic nature, the model aspires to be translated in different contexts and for various applications, and to become a useful tool for comparative and multi-criteria analysis elsewhere.

3. MATERIALS AND METHODS

3.1 Study Design

SPES "Exposure Study on Susceptible People" is an epidemiological observational cohort study promoted by the Istituto Zooprofilattico Sperimentale del Mezzogiorno (IZSM) of Portici (Naples) in collaboration with the National Tumor Institute IRCCS "G. Pascale" of Naples. The protocol of the study was approved by the ethics committee of the IRCCS "G. Pascale" through Commissioner Deliberation n.590 of 03/08/2016.
The PhD project was carried out in partnership between Laboratory of Molecular Medicine and Genomics (Lab. Med. Mol. Ge.), Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, and the "National Reference Center for Analysis and Correlation Study of Environment, Animals and Humans", IZSM.

The study design consists of two parts:

1. A context overview, in order to individuate source of contamination (variables) and the migration pathways of pollutants, and to define, by means of a multi-criteria analysis, geographical areas of interest as "clusters", which are then ranked according to a computed environmental "Pressure Index".
2. A transversal biomonitoring of people living in these clusters, in order to estimate "exposure" by quantifying pollutants in biological samples and investigating its biomolecular effects through in-depth "omics" analysis.

3.1.1 Environmental characterization of the Campania Region

Data from the scientific literature, the ARPAC (environmental regional agencies) database, Potential Hazard Maps (Minolfi et al., 2016), local monitoring projects on the Land of Fire and on the Campania Region, and preliminary results from the Campania Trasparente project (www.campaniatrasparente.it) (Fig.5) allowed us to have an overview of the territorial context and potentially contaminated areas, by mapping the spatial distribution of chemical compounds and geo-localizing the putative sources of pollution.

Figure 5. Campania Trasparente Environmental Monitoring. The environmental background of SPES biomonitoring. The search for pollutants in bioindicator organisms and environmental matrices yields a guideline for identifying impact areas and Risk Indexes in the municipalities observed.

3.1.2 Pressure Index computing for cluster generation and impact areas definition

We adopted an innovative approach in order to construct a "Comparative Risk Assessment Model" based on a multi-criteria analysis and combinatorial calculation of the percentage weight assigned to variables that contribute to determining the exposure.

The paradigm, shown in Figure 6, defines the evaluation chart model, from the selection of sources responsible for contamination, through the identification of migration pathways of pollutants, up to the final targets of exposure.

This approach provides a better understanding of the environmental impact determined by the sum of variables, the concerted action which generates the exposure, or rather creates part of the "exposome" (Wild, 2005 and 2012; Dennis et al., 2017), which could condition human health and putative targets.

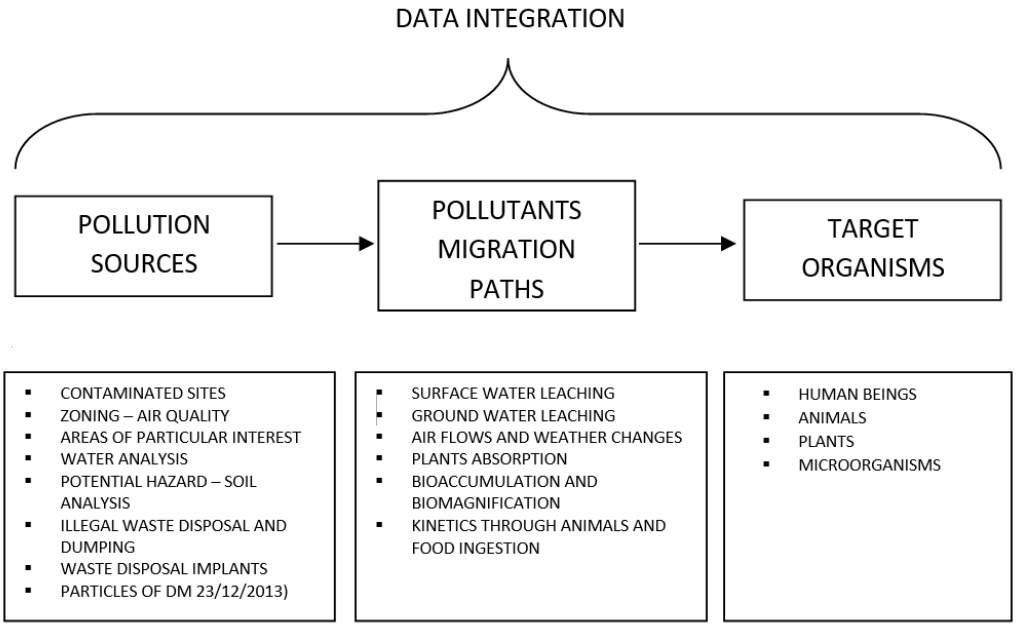


Figure 6. Paradigm of the method, from the selection of variables (sources of pollution) for AHP model, through identification of migration pathways of pollutants, to final target organisms. Migration pathways of pollutants represent the junction between the origin of pollutants and the final targets.

The variables shown in Table 3 were selected in such a way as to include all the factors involved in the pollution process, i.e. sources of potential contamination, analytical data on environmental matrices (soil, water, air), migration pathways from contamination sources to potential targets, and the targets themselves. Each variable was inserted into a rank system which attributes a score to each single factor, depending on the weight of the factor in generating the exposure; this operation was done by means of the multi-criteria decision support technique Analytic Hierarchy Process (AHP) (Saaty, 2008).

Table 3. Variables selected for Multi-criteria Analysis by the Analytical Hierarchy Process (AHP).

	VARIABLES – SOURCES OF POLLUTION	Environmental Parameters for IRc computation
1	Contaminated sites	Sites according to Legislative Decree 152/2006; sites of classes 3, 4 and 5 according to Decrees n.56 of 09/03/2015 and n.191 of 08/19/2015 related to the "Land of Fires".
2	Zoning	Use of land for agriculture, industry and urban purposes; housing density and evaluation of air quality.
3	Areas of particular interest	Sites of national and local interest present in the regional territory; illegal and pending landfills and potentially contaminated sites surveyed since at least one value above the CSCs was found.
4	State of groundwater and surface water	Water Analysis data from Campania Trasparente and ARPAC monitoring.
5	Potential Hazard	Soil analysis data from Minolfi et al., 2016, Campania Trasparente and ARPAC monitoring.
6	Illegal waste disposal and dumping	Data monitoring by SMA Campania (https://www.smacampania.info/progetto-terra-dei-fuochi/).
7	Waste disposal plants	Composting plants, waste-to-energy plants, STIR plants, controlled landfills
8	Particles covered "Land of Fire" Decree	DM 23/12/2013 - TdF 2a, 2b and 5, 4 and 3 of class A particles.

The AHP enables several alternatives to be compared with regard to a plurality of criteria, and yields an overall evaluation of each variable, which then can be ordered, for instance, according to a preference axis. Thus, all variables were geo-referenced and ordered in a geographic database (using Geographic Information System - GIS) and their percentage weight distributions were defined by means of AHP (Fig. 7).

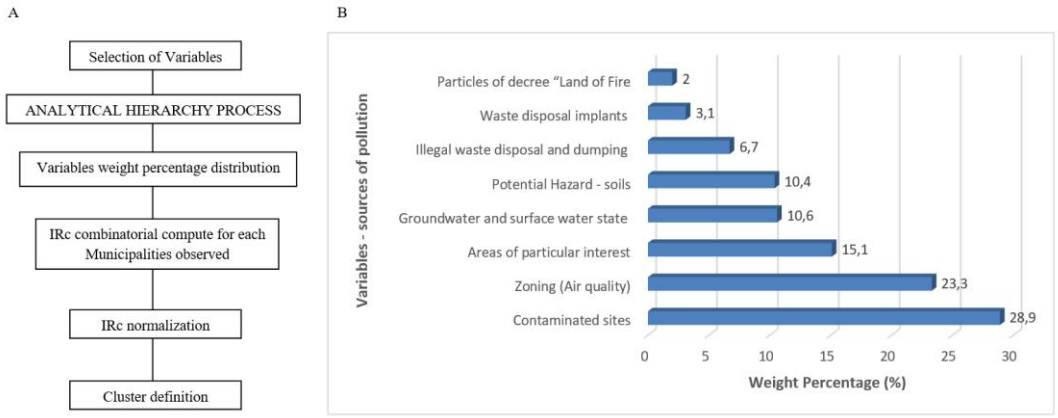


Figure 7. (A) Flowchart for the definition of "Municipal Risk Index" (IRc) and cluster generation.
(B) Percentage distribution of weights of variables resulting from Analytical Hierarchy Process. For each source of pollution, the AHP model defines the importance of criteria by means of pairwise comparison. Comparisons create a calculation matrix from which the generation of a priority vector gives the final percentage distribution of the weight of each variable.

This enabled us to compute a "Municipal Risk Index" (IRc): our "Pressure Index", given by the sum of the weights of the variables characterizing the territory of the municipality in analysis. This value ensures diversification of the pressures exerted by each source and reproduces different levels of risk; to sum up, this index provides a synthetic representation of the level of pressure exerted by the contamination sources on the municipal territory. The IRcs have been subsequently normalized (by a relativization method with variation range), and divided into homogeneous classes using the Natural Breaks Algorhitm (Dent 1999; Slocum 1999): 21 clusters comprising several municipalities whose score of risk index is included in a given range (Table 4). A completely different approach from the epidemiological computed indexes of Martuzzi et al., 2009 and Musmeci et al., 2010.

Table 4. Computed Municipal Index Risk IRc: range and clustering.

N. CLUSTER	CLUSTER	IRc	Normalized IRc
1	HIGH_1	833-1015	16-19
2	HIGH_2	1051.7-1216.10	21-23
3	HIGH_3	1240,8 - 1375,9	24-26
4	HIGH_4	1408.4 - 1514	27-29
5	HIGH_5	1563 - 1687.5	30- 32
6	HIGH_6	1692.3 - 1876.2	33 - 36
7	HIGH_7	1916.4 - 2049.7	37-40
8	HIGH_8	2101.3-2286.7	41-44
9	HIGH_9	2429.6 - 2650.1	47-51
10	HIGH_10	2729.6 - 3005.7	53-58
11	HIGH_11	3251.3 - 3347.8	63-65
12	HIGH_12	4125.9 - 5108.1	81-100
13	MEDIUM_1	791.7 - 1200.8	15-23
14	MEDIUM_2	1544.2 - 1749.4	30-34
15	MEDIUM_3	1847.5 - 3373.1	36 -66
16	LOW_1	265.6 - 499	4 - 9
17	LOW_2	525.4 - 815.6	9 - 15
18	LOW_3	812 - 1087.3	15 - 21
19	SABATO	1100.6 - 2077.9	21 - 40
20	IRNO_1	687 - 1201.3	13 - 23
21	IRNO_2	1711.1	33

The spatial distribution of environmental pressures, as yielded by the variables and by the computed IRc, provided an exhaustive representation of the territory, characterized mainly by three increasing levels of impact: high impact, medium impact and low impact (Fig. 8A). The high-impact area (Fig. 8B) included most of the territory of the Naples and Caserta provinces; this area is densely populated, has poor-quality groundwater and exceeds “contamination threshold concentrations” (CSCs) in soil samples. Numerous contaminated or potentially contaminated sites are located in this area, according to Legislative Decree 152/2006, denoting a high level of environmental pressure. The medium-impact area (Fig. 8C) included municipalities in the Naples, Salerno and Avellino provinces, which have poor-quality groundwater and exceed CSCs in soils. In addition, some municipalities show groundwater concentrations of organic halogenated compounds far above CSC values. The “Irno Valley” and “Sabato Valley”, which are situated in this area, contribute to this contamination through the presence of large industrial facilities and the orographic conformation of the sites. The low-impact area (Fig. 8D), covering the province of Salerno, the coastline (Cilento) and inland area (Sele-Tanagro), is characterized by a good state of groundwater, a low presence of CSCs and a low anthropization rate.

On the basis of the migration pathways of pollutants, an intervention area was identified for each level.

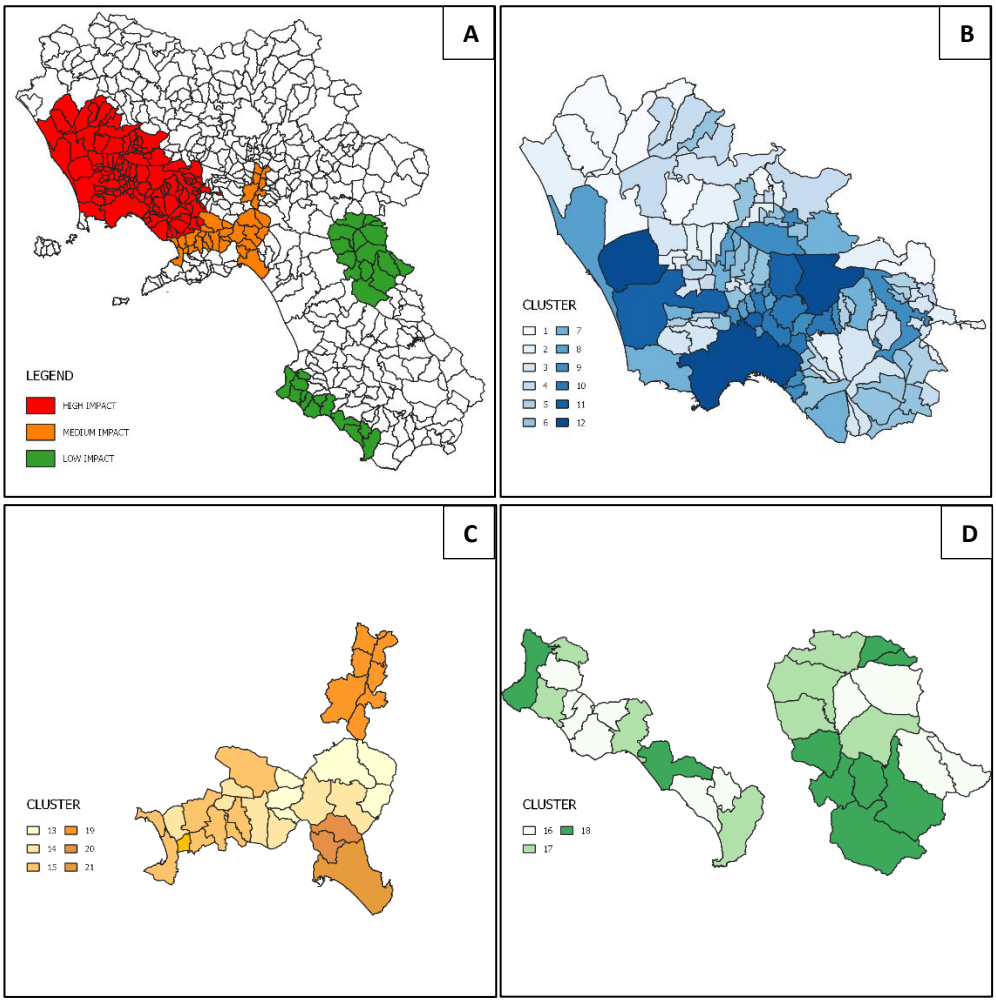


Figure 8. Clusters and Impact Areas. (A) Map of Campania Region showing areas of impact and clustering: (B) high impact, (C) medium impact, (D) low impact.

Based on the Comparative Risk Assessment Model and on the clustering of the municipalities in areas linked to different levels of human exposure, a target population in the HB protocol was identified in a ratio of 4:2:1 in high-, medium- and low-impact areas, respectively: 12 clusters were selected in the high-impact area, 6 in the medium-impact area and 3 in the low-impact area.

3.1.3 Study Population– SPES "Exposure Study on Susceptible People"

In order to integrate epidemiological observations, biomonitoring and toxicogenomic approaches, the SPES study design envisioned the recruitment of healthy residents in the Campania Region, the aim being to further investigate the state of exposure and its putative pathological consequences.

The scientific protocol involved the evaluation of a set of biomarkers able to quantify exposure and to detect individual genetic and molecular responses to the potential damage suffered.

In detail, exposure biomarkers comprised the presence of organic and inorganic compounds; effect biomarkers included sentinels of morphological, structural and functional alteration of cell systems; individual biomarkers of genetic susceptibility involved subjective features of predisposition or resistance to disease (Fig. 9).

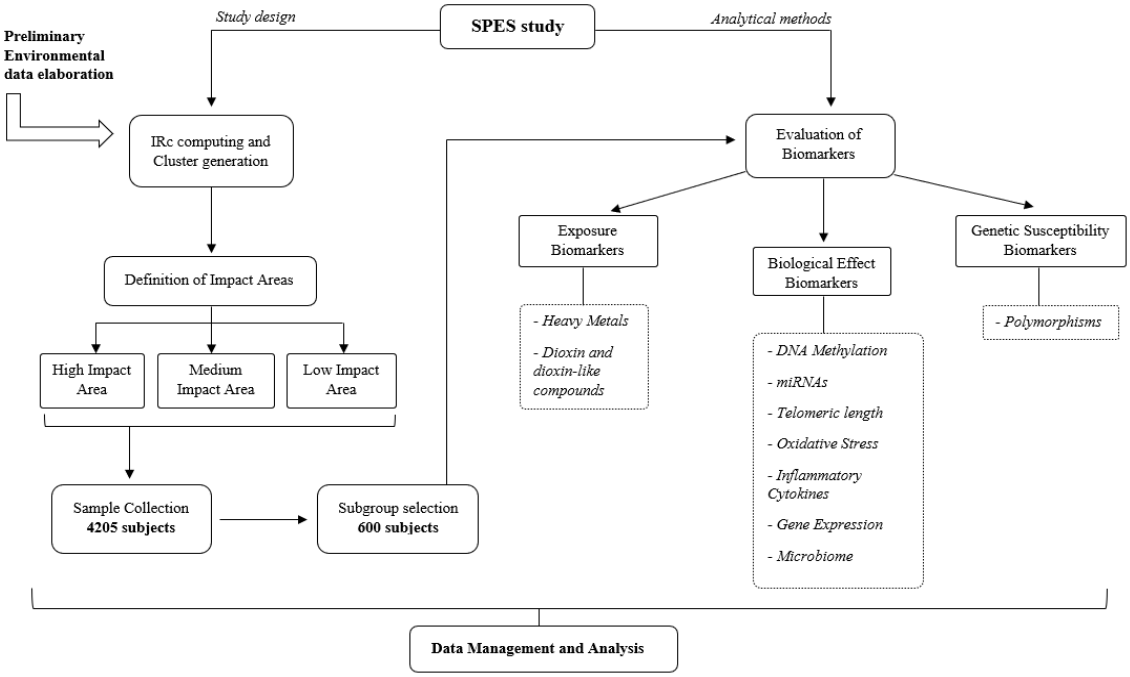


Figure 9. SPES Study Design. Selection of target population and analytical methods.

The study design considers an integrated set of assessments on healthy subjects living in several regional areas with different environmental "Pressure Indexes", computed by analyzing environmental data and literature references, as explained in previous paragraph. A total of 4205 subjects were enrolled; the sample size was fixed in such a way as to guarantee optimal power on statistical tests (with a significance level of $\alpha = 0.06$, the power is 90%); subjects were selected from the various impact areas in a ratio of 4:2:1, in cohorts of 200 people for each cluster.

The inclusion criteria were:

- good health (no severe or neoplastic conditions),
- age range 20-50,
- both sexes,
- no viral infectious (HIV, HCV, HBV),
- residence in the municipalities under study.

Among the inclusion criteria, it is important to underline that the age-range chosen (20-50 years) allows a better correlation between lifestyle and exposure to work-related risks; indeed, in younger or older age-groups, this correlation may be underestimated or

overestimated, respectively. In this way, we tried to reduce the contribution of some confounding factors. In addition, the choice of the age-range guarantees greater adherence to the project and a higher probability of being able to conduct long-term follow-up. Furthermore, in addition to having a negative history of HIV, HCV, HBV (to ensure the safety of operators and avoid bias) and providing informed consent, the participants must have lived for at least 5 years in these municipalities.

3.1.4 People cohorts enrollment and sample collection

The viability of the study was ensured through precise planning of the operating procedures; this guaranteed the ethical and scientific validity of the methods and actions in each phase and the rapid achievement of the final endpoints.

A Contract Research Organization (CRO) supervised all procedures during protocol performing.

The operational steps of the study were divided into five main areas of action:

(i) volunteers recruiting;
(ii) enrollment and medical examination;
(iii) samples collection;
(iv) experimental phase and reporting of results;
(v) data management and elaboration.

The selection procedures were implemented through random extrapolation of the population sample by means of an algorithm (Linear Congruential Generator - LCG) that randomly selects potential participants from municipal registries on the basis of the criteria indicated. The alignment of this random sample with the provincial phonebooks permitted to obtain the telephone contacts that the dedicated call center can make use of.

The activity of the call center was accompanied in parallel by the activity of recruitment units throughout the territory; these geo-located the potential participants in the municipalities of interest, identified stakeholders and promoted scientific communication. Indeed, scientific and health communication was one of the secondary endpoints of the study and was aimed at promoting awareness and transferring knowledge of health and prevention issues through specific campaigns tailored on the basis of social deprivation indexes.

On recruitment, participants signed a document in which they consented to the confidential treatment of their personal information, in accordance with current legislation on privacy (EU 679/2016). Eligible subjects, who were willing to participate, were registered in a database (www.openspes.campaniatrasparente.it), in which all the associated information, documents and reports are stored and managed. With the submission of an informed consent the volunteers become enrolled and, in order to ensure the protection of privacy, everyone is associate to an univocal ID code, as well as applied data encryption and policies

to restrict access to information. All interactions with the system take place via secure channels using internationally recognized security protocols (for instance https).

Synergic teams of recruiters, doctors and nurses completed the entire recruitment, enrollment, medical examination and sample collection steps in the different impact areas in the same time period (november 2016 – august 2017).

Participants first underwent medical examination, completed a questionnaire on lifestyle and eating habits (EPIC questionnaire – European Prospective Investigation into Cancer and Nutrition - validated by WHO – World Health Organization; Riboli, 1992; Margetts et al., 1997), and filled in an anamnestic (professional, familial, clinical and pathological) case report form (CRF).

Finally, samples of blood (about 100 ml in several aliquots) urine and stool were collected. These biological samples constituted the biological matrices in which biomarkers of exposure, effect and genetic susceptibility were measured, in order to further investigate any causal link between environmental contamination and health.

Every sample collected was identified by a unique barcode, associate with the ID of the subject.

Samples were transported with secondary and thirdly boxes and GPS- Data Logger Device, in order to guarantee transport temperature control and traceability.

Aliquots not used for immediate analysis have been stored in the IZSM Biobank dedicated to the study (www.biobancaizsm.it).

3.1.5 Subgroup selection and analytical methods

In addition to routine biochemical-clinical investigations and virologic investigations (HBV, HCV, HIV, HEV), which define an entrance profile of the individual, the analytical step has provided the search in the biological samples of exposure, effect and susceptible biomarkers.

Regarding biomarkers of exposure, we have considered Potentially Toxic Metals (MPT), Dioxins, Furans and PCBs linked to exposure to different sources of pollution, able to persist in the environment and in living organisms for a long time (WHO 2015). They were detected by chemical and biological methods: Inductively Coupled Plasma - Mass Spectrometry methods (ICP-MS) for MPT; High Resolution - Gas Chromatography (HR-GC) (adapted from US EPA Methods 1613 e 1668) and CALUX Assay (Warner et al. 2005; Sakthivel et al. 2016) for organic pollutants, dioxins, furans, dioxin-like, PCBs, PAH and other metabolites.

About Effect biomarkers, SPES protocol have considered: evaluation of the telomeric length (De Felice et al. 2012; Borghini et al. 2016), of the inflammatory cytokines (Gruzieva et al., 2017), of the oxidative stress (Rossner et al. 2013; Brucker et al. 2013), total plasmatic protein, as well as insights on epigenomic mechanisms such as the DNA

methylation (by pyrosequencing and microarray) (Plusquin et al. 2017; Demanelis et al. 2019), circulating miRNA (La Rocca et al. 2016; Vrijens et al. 2016), panel of gene expression profile (Ameer et al. 2017; Espin-Perez et al. 2018), microbiome (Claus et al. 2016) and metabolomic panel (Saberi et al. 2013; Li et al. 2019b).

Finally, biomarkers of genetic susceptibility. Details are shown in table 5.

Table 5. Scheme of Exposure, Effects and Susceptible Biomarkers.

BIOMARKER	BIOLOGICAL MATRIX	N. SAMPLES TESTED
➢ EXPOSURE BIOMARKERS:		
- Heavy Metals (20)	Serum	4205
- Organic - Dioxins, Furans, dioxin -like PCBs, non-dioxin-like PCBs	serum	600
- Bioassay CALUX for determination of organic pollutants	plasma	600
- Trace Elements	serum	600
- Speciation Cr VI (hexavalent chromium)	erythrocytes	600
- Speciation As (Arsenic)	urine	180
➢ EFFECT BIOMARKERS:		
- Telomeric length	DNA from whole blood	600
- Inflammatory cytokines (Luminex assay panel)	plasma	600
- Redox status	erythrocytes e mRNA from whole blood	600
- miRNA (panel of 18 miRNA)	miRNA from whole blood	600
- DNA Methylation (Pyrosequencing)	DNA from whole blood	600
- DNA Methylation (Array Infinium Methylation 850k Illumina)	DNA from whole blood	600
- Plasmatic markers (AOPP, MDA, homocystein, thiol, total protein)	plasma	600
- Copy Number Variation mitochondrial DNA	DNA from whole blood	600
- Gene Expression – Nanostring Panel	RNA from whole blood	600
- Microbiome	stool	600
- Metabolome	serum	600
➢ GENETIC SUSCEPTIBILITY BIOMARKERS:		
- Polymorphisms (hOGG1, XRCC3, XRCC1, CYPA1)	DNA from whole blood	600

In order to preliminary test all the biomarkers and identify the set of specific and trustable ones (especially in view of optimizing funding resources), a subgroup of 600 subjects has been selected between the 4205 enrolled.

The subgroup represents a specific part of SPES population, eligible for the following inclusion criteria:

- No smoking;
- No alcohol abuse;
- No use of drugs for chronic diseases (in order to avoid iatrogenic alterations);

and balance criteria:

- Ratio of 2:2:1 between high, medium and low impact areas, in such way as to maximizing the difference between clusters;
- Stratified by body mass index (no obesity of II, III and IV classes);

and, finally, with Professional Identification.

Details in Table 6.

Table 6. Subgroup stratification per gender, age, clusters and impact areas.

		male			female			total per	
	Clusters	**Age range 20-29**	**Age range 30-39**	**Age range 40-49**	**Age range 20-29**	**Age range 30-39**	**Age range 40-49**	**Cluster**	**Impact Areas**
HIGH	9	**10**	**10**	**10**	**10**	**10**	**10**	60	
	10	**10**	**10**	**10**	**10**	**10**	**10**	60	**240**
	11	**10**	**10**	**10**	**10**	**10**	**10**	60	
	12	**10**	**10**	**10**	**10**	**10**	**10**	60	
MEDIUM	13	**10**	**10**	**10**	**10**	**10**	**10**	60	
	14	**10**	**10**	**10**	**10**	**10**	**10**	60	**240**
	19	**10**	**10**	**10**	**10**	**10**	**10**	60	
	20+21	**10**	**10**	**10**	**10**	**10**	**10**	60	
LOW	16	**10**	**10**	**10**	**10**	**10**	**10**	60	
	17	**10**	**10**	**10**	**10**	**10**	**10**	60	**120**
		100	100	100	100	100	100		
		male total		**300**	***female total***		**300**		
						TOTAL			**600**

3.1.6 Sample processing and Biobanking

Biological specimens collected were blood, urine and stool.

Blood was collected in tubes vacutainer as follow:

- n. 8 vacutainers for serum (BD REF 366468 SST II Advance Tube);
- n. 6 vacutainers for plasma and whole blood (BD REF 367864 K_2EDTA Tube);
- n. 2 vacutainers for plasma (BD REF 368886 LH Lithium Heparin);
- n. 2 Tempus Blood RNA (REF 4342792);

for a total amount of about 100 ml.

Fresh samples were allocated to biochemical and virological analysis, within 3 hours from venous sampling (1 vacutainer for serum and 2 vacutainers with K_2EDTA for each subject). The remaining part was processed as follow:

- Serum was separated by centrifugation at +4°C, 10 minutes at 2000 g. Aliquots obtained were stored at -80°C in cryovials until analysis.
- Plasma from K_2EDTA and Lithium Heparin vacutainers was obtained after centrifugation at +4°C, 10 minutes at 2500 g. Then separated and stored at -80°C in cryovials until analysis.
- Whole blood aliquots (K_2EDTA vacutainers) were inverted several times and then transferred to cryovials for storage at -80°C until analysis.
- Tempus Blood RNA were vortexed and then stored at -80°C as manufacturer's instructions.
- Urine, collected in sterile urine container, was aliquoted in cryovials and stored at -20°C until analysis.
- Stool, collected in sterile container, was directly stored at -80°C until DNA extraction for microbiome analysis.

The biobank dedicated to the study (www.biobancaizsm.it) has a system of remote monitoring of temperature. Each aliquot is registered on the platform by a sample code and position coordinates in the fridge.

3.2 Biomarkers of Exposure

3.2.1 Heavy Metals

The determinations were performed by IZSM's Department of Chemistry.

All study participants samples (4205) were analyzed for 19 metals (arsenic [As], beryllium [Be], cadmium [Cd], cobalt [Co], chromium [Cr], copper [Cu], iron [Fe56 and Fe57], mercury [Hg], lithium [Li], manganese [Mn], molybdenum [Mo], nickel [Ni], lead [Pb], antimony [Sb], selenium [Se], strontium [Sr], thallium [Tl], vanadium [V], and zinc [Zn].

The measurement was carried out in serum by Inductively Coupled Plasma Mass Spectrometry (ICP-MS) NexION 350 (PerkinElmer, Waltham, USA) equipped with a concentric nebulizer (Meinhard Associates, Golden, USA), a cyclonic spray chamber and a quartz torch with a quartz injector tube (2 mm internal diameter). Argon gas (99.9999%) and Helium 6.0 were supplied from Rivoira (Italy).

For each analyte, accuracy and precision were assessed by internal quality controls and using certified reference materials.

Calibration standard solutions and internal standards were prepared by successive dilution of a high purity multi-element Standard Solution at 10 mg/L, obtained from Perkin Elmer

(Norwalk, CT). Internal standard solution containing yttrium (Y), rhodium (Rh), and lutetium (Lu) (each at 1000 mg/L) were acquired from Perkin Elmer. Superpure grade nitric acid 68% (v/v) was obtained from VWR International (Belgium).

Serum was thawed at room temperature and mixed by gentle rotation. Then, 500 µl of sample were diluted 1:10 (v/v) with nitric acid 67-69 %.

Trace element concentrations were determined in serum by ICP-MS and the signal intensities were compared to a standard calibration curve for each trace element.

The following isotopes were selected: 75As, 111Cd, 9Be, 59Co, 52Cr, 63Cu, 57Fe, 202Hg, 7Li, 55Mn, 98Mo, 60Ni, 208Pb, 121Sb, 78Se, 88Sr, 205Tl, 51V and 66Zn.

The correlation coefficient (R2) of calibration curves for all the trace elements was always greater than 0.99, showing a good linear relationship throughout the selected ranges of concentration. Each sample was analyzed in duplicate and the mean concentration was used in all statistical analyses.

3.2.2 Dioxins, Furans and PCBs congeners by High Resolution Gas Chromatography – Mass Spectrometry

The determinations were performed by IZSM's Department of Chemistry in partnership with University of Liegi (Belgium). The analysis was carried out on the 600 subgroup samples. Analytical method used was adapted from Brasseur et al. (2014).

This protocol allowed the determination of congeners belonging to classes of compounds as Dioxins PCDD (polychlorinated dibenzo-p-dioxin), Furans PCDF (polychlorinated dibenzofuran), DL-PCB (Dioxin like - polychlorinated biphenyl) and NDL-PCB (Non-Dioxin like - polychlorinated biphenyl) were analyzed.

For the purpose of the present dissertation will be considered only the following variables:

- Sum of PCDD/PCDF pg WHO-TEQ/g lipids.
- Sum of DL-PCB pg WHO-TEQ/g lipids.
- Sum of PCDD/PCDF + DL-PCB pg WHO-TEQ/g lipids.
- Sum of NDL-PCB ng/g lipids.

A volume between 10 to 20 ml of serum for each sample was processed. After the accelerated solvent extraction using a mixture of hexane-acetone, and a clean-up step, the serum was analysed by High Resolution Gas Chromatograph – High Resolution Mass Spectrometer (DFS Magnetic Sector HRGC-HRMS system, Thermo Fisher Scientific) operating at a resolution of at least 10000 at a 10% valley.

Each batch of analysis was monitored for Quality Assurance and Quality Control by testing control samples.

Given the lipophilic nature of the congeners, data obtained from biochemical analysis (through enzymatic method) were considered for the computing of fats percentage: the total lipid (TL) content was estimated using triglycerides (TG), total cholesterol (TC), non-

esterified (free) cholesterol (FC) and phospholipids (PL). For each serum sample, the TL content, expressed in mg dL^{-1}, was calculated by the following equation:

$$TL = 1.677 \times (TC\text{-}FC) + TG + PL$$

where 1.677 is the ratio of the mean molecular weight of esterified cholesterol to unesterified cholesterol (Bernert et al., 2007; Brasseur et al., 2014).

For comparing the estimate of the lipid content in human serum obtained using alternative formulas, the TL contents calculated with the above-mentioned equation was compared with the results calculated using the formula described by Phillips et al. (1989):

$$TL = (2.27 \times TC) + TG + 62.3 \text{ mg } dL^{-1}$$

Two equations provided comparable values for TL.

Analytical results were expressed not only as measured value, but also according to Toxic Equivalent (TEQ) concentrations, by means Toxic Equivalent Factors (TEF) (Van den Berg et al., 2006). Congeners and their sums were also expressed used the approach of Lower, Medium and Upper Bound, that means to replace 0.00 (zero), ½ LOQ (Limit of Quantification) or LOQ, if the detected value was below the LOQ itself (De Felip et al., 2008; Miniero et al., 2017).

3.2.3 Chemical Activated LUciferase gene eXpression (CALUX) Bioassay

These determinations were performed in partnership with Bio-Detection System (BDS) Laboratory (The Netherlands). The analysis was carried out on the 600 subgroup samples. Chemical Activated LUciferase gene eXpression (CALUX) is a ligand-dependent nuclear receptor- bioassay used in the detection of classes of chemicals in food, environmental and biological matrices. It could be considered as a screening method because it estimates the total effect that ligands have on the receptor observed. Engineered cell lines are stably transfected with a luciferase reporter gene. If activated by ligands presence, DNA responsive elements stimulate transcription of the inserted luciferase gene and produce the light-generating enzyme which can be measured by Luciferin addition (Fig.10).

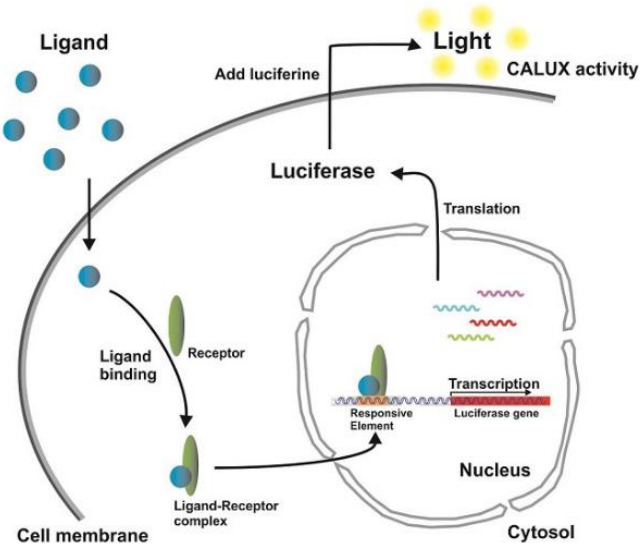


Figure 10. CALUX bioassay mechanism. The presence of the ligand by binding to the nuclear receptor allows its dimerization and transfer to the nucleus, where an intracellular signal transduction cascade activates regulatory regions and responsive elements of the DNA which, associated with the luciferase reporter gene, permit gene expression and luminescence emission as a response to enzymatic activity in the presence of luciferin.

SPES protocol have considered the DR CALUX® (Garrison et al., 1996), PAH- CALUX® (Pieterse et al, 2013) and ERα CALUX® (Sonneveld et al., 2005). They are all based on stably transfected cell lines: DR and PAH CALUX are based on stably transfected rat hepatoma H4IIe cell lines, while the ERα CALUX assay is based on a stably transfected human osteoblastic osteosarcoma U2-OS cell line (American Type Culture Collection). Cells that are exposed to compounds of interest not only express proteins that are under normal circumstances associated to the receptors, but also luciferase. By addition of the appropriate substrate for luciferase, light is emitted. The amount of light produced is proportional to the amount of ligand-specific receptor binding, which is benchmarked against the relevant reference compounds, i.e., 2,3,7,8-TCDD (2,3,7,8-tetrachlorodibenzodioxin) for DR CALUX; B(a)P (benzo(a)pyrene) for PAH CALUX and 17-β-estradiol for ER CALUX.

The results are expressed as bioassay equivalents (BEQ) per gram of lipid (DR, PAH) or per ml of plasma (ERα).

In detail, the human plasma (5 ml obtained from Lithium Heparin Vacutainers) was extracted with hexane/ethyl ether solvent, subsequently purified (clean-up) on a silica column; the extraction product is resuspended in DMSO (dimethyl sulfoxide).

The ERα CALUX cells, U2-OS cells, were routinely sub-cultured every 3–4 days in growth medium consisting of DMEM (Gibco) supplemented with 7.5% dextran-coated charcoal stripped fetal calf serum (DCC-FCS), 1×nonessential amino acids (Gibco) and 10 U/ml

penicillin and 10 µg/ml streptomycin. DR and PAH-CALUX cells were routinely sub-cultured every 3–4 days in growth medium consisting of αMEM (Gibco) supplemented with 10% fetal calf serum. All cell types were always maintained at 37 °C and 5% CO_2.
CALUX® cells were seeded in 96 wells plates in assay medium and fetal calf serum according to conditions indicated in Table 7. Exposure of the cells started when > 95% confluency was reached (DR, PAH) or at 10000 cells per well (ERα).
The cell cultures are then incubated with the DMSO eluted extract for 24 hours. Incubation is in triplicate for each biological sample and takes place in 96-well plates. A solution containing luciferin is added to estimate the luciferase activity by means of a luminometer (Berthold Centro XS3). Each plate contains a calibration curve with reference standards relating to the class of compounds investigated: full dose-response curve of the reference compound, 2,3,7,8-TCDD for DR CALUX, B(a)P for PAH CALUX and 17-β-estradiol for the ERα CALUX.

Table 7. CALUX cell type, culture and exposure information.

Assay	**ERα CALUX**	**DR CALUX**	**PAH CALUX**
Cell type	U2OS	H4IIe	H4IIe
Species	Human	Rat	Rat
%DMSO exposure	0,1%, or 1%*	0,8%	0,8%
%CO_2 incubation	5%	5%	5%
Exposure time	24 hrs	24 hrs	4 hrs
Confluence/cells	10000 cells/well	>95% confluence	>95% confluence
Medium used	DMEM/F12	αMEM	αMEM
% FCS	7,5% DCC-stripped**	10%	10%
Additions to medium	Non-essential amino acids		

* All samples were first tested at 0,1% DMSO; when low values were observed (below Limit of quantification) these samples were re-analyzed using 1% DMSO

** FCS was treated with dextran-coated charcoal (DCC) to strip the endogenously present hormones

Analysis results of plasma extracts expressed as relative light units (RLUs) are interpolated in the calibration curves of each respective bioassay for quantitative determination of the induction potential using the statistical software package GraphPad Prism V5.03. All analysis results are expressed as amount of reference compound equivalents per ml processed sample (BEQ/ml plasma), or per gram of fat (BEQ/g fat).
Fats percentage follows gravimetrically measurement and, in parallel, the same formulas of Dioxins/Furans/PCBs in previous paragraph.

3.3 Biomarkers of Effect and Susceptibility

3.3.1 DNA Methylation Array

The determinations were performed in partnership with Lab. Med. Mol. Ge., University of Salerno. The analysis was carried out on the 600 subgroup samples.

Genomic DNA was extracted from whole blood aliquots (obtained from K_2EDTA vacutainers) stored at -80°C in IZSM Biobank.

The nucleic acid extraction protocol using the QIAsymphony Biorobot (QIAGEN), Custom B_400 protocol, was carried out according to DSP DNA MIDI QIASymphony Kit (REF 937255). A volume of 400 µl of whole blood was the input material. Extracted DNA was finally eluted in 100 µl of nuclease-free H_2O (Invitrogen™ UltraPure™ DNase/RNase-Free Distilled Water).

Genomic DNA was checked for integrity by means of Genomic DNA protocol (REF 5067-5365/5366) on Agilent TapeStation System 4200. Concentration and purity were assessed by spectrophotometer Nanodrop ND-1000: A260/A280 and A260/A230 ratio. Accurate concentration measurement was performed by means of Qubit 2.0 Fluorometer, according to Qubit dsDNA BR – Broad Range kit (REF Q32850).

All samples showed A260/A280 ratio around 1.8, and A260/A230 ratio around 2.00, confirming the good quality of DNA. Tape Station DINs (DNA Integrity Number) were for all samples in acceptable range for integrity 6.00 to 10.00.

The bisulfite conversion step was performed with EZ DNA Methylation Zymo kit (ZYD5001) with 500 ng of genomic DNA as input material, according to manufacturer's instructions.

DNA methylation arrays were performed according to Illumina Infinium HD Arrays procedure. Particularly, Infinium Methylation EPIC 850K (Infinium HD Assay Methylation Protocol Guide Material #20002138 Document #15019519v01), which allows the identification and quantitative analysis of the methylation state of approximately 850000 CpG dinucleotides distributed throughout the entire human genome, was used.

The Infinium HD Assay for Methylation combines bisulfite conversion of genomic DNA and whole-genome amplification (WGA) with direct, array-based capture and scoring of the CpG loci. During the chemical conversion induced by bisulfite, the unmethylated cytosines are deaminated to uracil while the methylated ones are protected and remain so. The enzymatic reaction that occurs on the surface of the array determines the incorporation of complementary fluorescent bases which, thanks to different fluorophores, allows to discriminate one case from another for each dinucleotide analyzed.

This method considers two chemistry technologies, Infinium I and Infinium II. The Infinium I probe design has two probes per site (one for methylated, one for unmethylated) and the 3' end of the probes is positioned directly across from the CpG site (Fig.11A). The

Infinium II probe has one probe per site (for the methylated state) and the 3' end of the probes is positioned immediately adjacent to the site (Fig.11B).

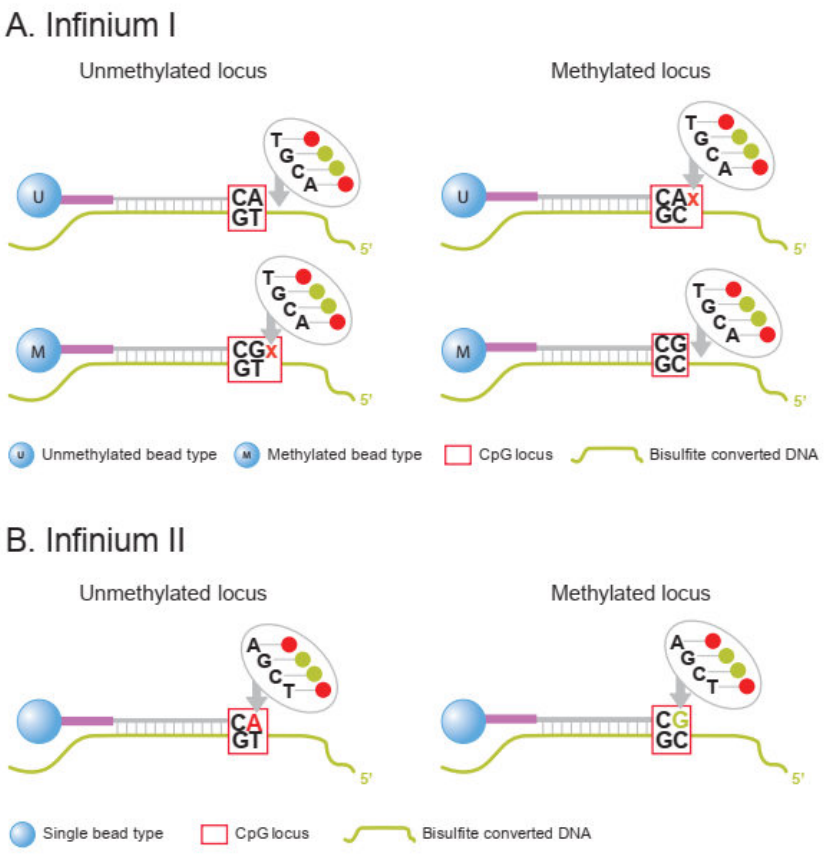


Figure 11. Broader Coverage using Illumina Infinium I and II Assay Design. (A) The Infinium I assay employs 2 probes per CpG locus: 1 "unmethylated" and 1 "methylated" query probe. The 3'terminus of each probe is designed to match either the protected cytosine (methylated design) or the thymine base resulting from bisulfite conversion and whole-genome amplification (unmethylated design). (B) The Infinium II assay design requires only one probe per locus. The 3' terminus of the probe complements the base directly upstream of the query site. A single base extension results in the addition of a labeled G or A base, complementary to either the 'methylated' C or 'unmethylated' T (source www.emea.illumina.com).

This approach allows to evaluate 95% of the cytosines present in the best functionally characterized regions along the whole human genome and, in particular, > 99% of the human gene regulatory sites (RefSeq), including promoters, 5 'UTR, gene body (exons and introns) and 3 'UTR, with excellent coverage also of the enhancer regions (Moran et al, 2016).

3.3.2 DNA Methylation Array data analysis

Signal intensity was measured with the Illumina iScan system. DMAP files (containing the spatial maps of the probes along each bead-chip), were downloaded following the "Decode File Client" instructions, the windows-based software application supplied for this purpose. The input data from the Illumina iScan were analyzed using the Illumina GenomeStudio® Methylation Module v1.9 software, in order to generate Beta Values, the measure of the level of methylation at a locus. Beta values represent the estimate methylation level of the CpG locus as ratio of intensities between methylated and unmethylated alleles; it is computed by means the following expression: $\beta = M/M+U+100$, where M indicates the methylated locus, M+U the sum of methylated and unmethylated loci and 100 a mathematical constant. Beta values are intended to allow sample to sample comparisons for each individual CpG and can be interrogated and compared across samples for powerful large-scale studies.

Genome Studio Software allows to check and validate array sessions through a series of controls: sample independent controls, for quality workflow steps, and sample dependent controls, that evaluate performance across samples. Among them for example, the bisulfite conversion rate check.

First step analysis considers the background subtraction, in order to avoid any non-specific signal, and normalization. Then a filter was applied: only the probes that showed a detection p-value <0.05 (98% for each sample) were selected, considering for the subsequent steps only the CpGs with a high detection quality.

Together with Genome Studio Software, provided by Illumina, thanks to the partnership with The Laboratory of Statistics and computational tools for Bioinformatics (BioinfoLab) – Istituto per le applicazioni del calcolo "Mauro Picone", National Research Council of Naples (CNR), data analysis were also performed by means R software (4.0.3 version) and Bioconductor, with ChAMP package (https://www.bioconductor.org/packages/release/bioc/html/ChAMP.html) as follows:

- Probes normalization;
- Correction for Cell Heterogeneity (step considered when sample matrix is blood);
- Singular Value Decomposition (SVD) for the batch effect correction;
- Batch correction on SVD indication;
- Differential Methylated Probes identification (DMP);
- Differential Methylated Region identification (DMR);
- Gene set enrichment for DMP and DMR.

In addition, both Genome Studio and R Bioconductor-ChAMP display valuable information such as chromosomal coordinates, percent GC, location in CpG Island, and β values useful for perform an in-depth analysis.

3.4 Statistical Analysis

Standard descriptive statistics were used for characterizing the study sample. Numerical variables were described using mean ± standard deviation (std. dev.) or median (25th - 75th percentile) in case of distribution showing a consistent asymmetry. Absolute frequencies and percentages were used to describe categorical factors. Several median regression models were estimated to explore the association between serum exposure biomarkers in the assessed population and impact areas, as well as clusters.

In those models where the impact area was considered as the geographical reference unit, the low impact area was always used as reference category. When analyzing geographical difference at cluster's level, the reference cluster was the one with the median concentration.

Considering Calux bioassay data analysis, statistical significance was set at a level p-value<0,05 and paired datasets were analyzed by the Mann Whitney test, while differences between multiple groups were analyzed using one-way Anova test.

For methylation data analysis, in order to highlight a potential correlation between methylation levels and exposure biomarkers, data analysis proceeded through a robust linear multiple regression model, by M-value transformation from methylation beta values.

All the analyses were performed using the statistical platform R.

3.5 Data management

The study protocol has been supported by web-based software (openspes.campaniatrasparente.it), developed by IZSM, and custom-made for the SPES study. The software interface works in the Linux Apache POSTGRESsql (LAPP) environment to ensure maximum interoperability and scalability. The programming language operated is PHP interfaced with Apache server. The management and permanent storage of data has been carried out using POSTGRESql equipped with PostGis extension in order to ensure maximum interoperability with GIS systems. The graphical interface is based on the twitter Bootstrap framework, so to optimize the web platform for displaying on several kinds of devices of different degrees of resolution. All users are able to access the system using a web browser; they therefore, using a TC/ IP network and communicating via the HTTPs protocol to send data and make requests to the Apache web server, without installing additional software. All restriction and cryptographic policies have been applied to guarantee the protection and privacy of data. The openspes system also communicates with the GIS system to collect environmental data from “Campania Trasparente” and literature references (www.geojizz.it) in order to integrate data into a single software.

4. RESULTS

4.1 Environmental background

A preliminary analysis on the environmental background was performed to compute the environmental pressure index for each municipality.

Data from scientific literature and from "Campania Trasparente" monitoring (www.campaniatrasparente.it; Allocca et al., 2016; Qu et al., 2019;; Sellerino et al., 2019; Nasta et al., 2020) were combined in the "Comparative Risk Assessment Model" (see Paragraph 3.1.2), in order to obtain a synthetic representation of the level of pressure exerted by the pollutant sources, on the municipal territory. This synthetic index is represented by the IRc (see Table 4). IRc values allow to cluster municipalities. Figure 12 shows a graphical representation of the IRcs, according to the geographical distribution.

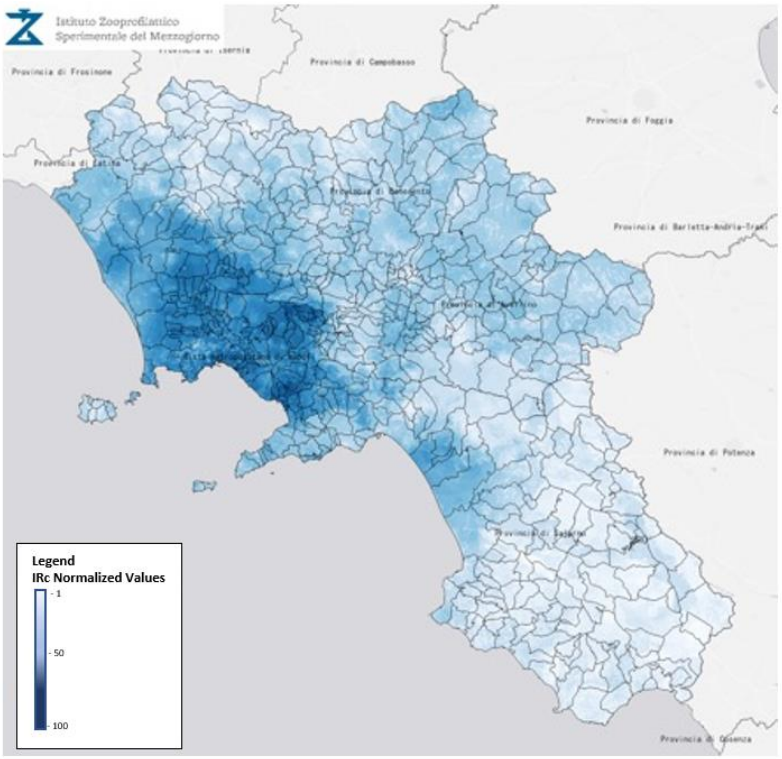


Figure 12. Campania Region and geographical distribution of IRcs. The spatial distribution of the environmental pressures provides an exhaustive representation of the territory, mainly characterized by three increasing levels of impact: high impact, medium impact and low impact.

The evidence of critical issues on the regional territory could be divided, from an environmental point of view, in different macro-areas:

1. Volturno and Regi Lagni areas (Caserta province), where pollutants values above normative referments (Legislative Decree 152/2006; Annex 5, Part IV) are Lead (Pb) and Zinc (Zn) in soil matrix, both attributable to vehicular traffic. Spreading

of Volatil Organic Compounds (VOC) in surface waters is attributable to industrial waste disposal, especially for Villa Literno and Marcianise municipalities. Finally, nitrates peaks are attributable to agricultural and urban contributes.

2. Flegrea-Neapolitan area (Naples and north of Naples), where high levels of toxic metals are recorded both in urban and in agricultural areas. Lead (Pb) and Zinc (Zn) in soils are attributable to vehicular traffic, while polycyclic aromatic hydrocarbons in waters are probably linked to intentional or accidental waste disposal.
3. Sarno-Vesuvio area (south of Naples), where surface pollution for coast line regards heavy metals and organochlorine pesticides, while the most critical issue is registered for Sarno river area and for Terzigno municipality, this last with ground water contamination attributable to landfill leachate (Cava Sari and Cava Ranieri). Finally, the wide presence of copper (Cu) is linked to the indiscriminate use of pesticides in viticulture.
4. Irno and Sabato Valley, where the presence of industrial plants gives an important contribute to the spreading of heavy metals and polycyclic aromatic hydrocarbons.

According to the environmental monitoring and to the IRcs computing, three areas with increasing levels of impact were identified: high-, medium- and low-impact (Paragraph 3.1.2, Fig. 8).

As mentioned in Materials and Methods chapter, the target population in the SPES protocol was identified in a ratio of 4:2:1 in high-, medium- and low-impact areas, respectively: 12 clusters were selected in the high-impact area, 6 in the medium-impact area and 3 in the low-impact area. A subgroup of 600 subjects has been selected among the 4205 enrolled, in ratio of 2:2:1 between high-, medium- and low-impact areas for further investigation.

4.2 People demographic features

Among the 4205 volunteers enrolled for the purposes of the biomonitoring study, a total of 2404 subjects (57.2% of the study sample) were recruited in the high impact area, 1201 (28.6% of the study sample) were recruited in the medium impact area and 600 (14.3% of the study sample) were recruited in the low impact area. Distribution of enrolled subjects according to age, gender, impact area and cluster is shown in table 8.

Mean age of the population was 34.9±8.8 years, with women representing 53.9% of the study sample. Age and gender distributions were consistent across the three study areas.

Median weight of the entire population was 71 kg [IQR: 61 ; 82] , while median height was 170 cm [IQR: 163 ; 175], with a median body mass index of 24.6 [IQR: 22.3 ; 27.7] kg/m^2 (table 9). Median weight in the high-, medium- and low-impact area was 71 [IQR: 60 ; 83], 71 [IQR: 61; 82] and 72 [IQR: 62 ; 82] kg, respectively, while median height was 170 [IQR: 163 ; 176], 168 [IQR: 161.2 ; 175] and 170 [IQR: 163 ; 175] cm, respectively (Table 9).

Blood count, blood biochemistry and endocrinological panel results were consistent with those of a healthy population, with adequate function of bone marrow, liver, kidney and of the endocrine system. No clinically relevant differences were reported among the three areas assessed.

Table 8. Number of participants according to age. gender. impact area and cluster.

Study sample		**Number of subjects (% with respect to the overall sample)**
Total Population		4205 (100.0)
Age		
20-29		1394 (33.2)
30-39		1296 (30.8)
40-50		1515 (36)
Gender		
Male		2268 (53.9)
Female		1937 (46.1)
Area		
High		2404 (57.2)
Low		600 (14.3)
Medium		1201 (28.6)
Cluster	**# Cluster**	
HIGH_1	1	204 (4.9)
HIGH_2	2	199 (4.7)
HIGH_3	3	200 (4.8)
HIGH_4	4	200 (4.8)
HIGH_5	5	200 (4.8)
HIGH_6	6	200 (4.8)
HIGH_7	7	199 (4.7)
HIGH_8	8	201 (4.8)
HIGH_9	9	200 (4.8)
HIGH_10	10	199 (4.7)
HIGH_11	11	200 (4.8)
HIGH_12	12	202 (4.8)
MEDIUM_1	13	179 (4.3)
MEDIUM_2	14	209 (5)
MEDIUM_3	15	212 (5)
IRNO_1	20	188 (4.5)
IRNO_2	21	212 (5)
SABATO	19	201 (4.8)
LOW_1	16	200 (4.8)
LOW_2	17	205 (4.9)
LOW_3	18	195 (4.6)

Table 9. Demographic characteristics of the overall sample population, according to the impact area.

		IMPACT AREAS (HIGH vs MEDIUM vs LOW)		
	Overall (n=4205; 100%)	**High (n=2404; 57.2%)**	**Medium (n=1201; 28.6%)**	**Low (n=600; 14.3%)**
Age; years	34.9+-8.8 (20 to 50)	34.6+-8.8 (20 to 50)	35.5+-8.8 (20 to 50)	35.1+-8.9 (20 to 50)
Gender; female	2268 (53.9)	1306 (54.3)	639 (53.2)	323 (53.8)
Deprivation Index*; %	26+-8.8 (0 to 92.1)	25.5+-8.6 (0 to 74.7)	25.9+-8.9 (5.7 to 92.1)	28+-9.3 (8.6 to 65.3)
Below 1st quartile	1045 (25.1)	652 (27.1)	293 (25.3)	100 (16.7)
Below 2nd quartile	1037 (24.9)	610 (25.4)	293 (25.3)	134 (22.3)
Below 3rd quartile	1071 (25.7)	591 (24.6)	294 (25.4)	186 (31)
Above 3rd quartile	1008 (24.2)	550 (22.9)	278 (24)	180 (30)
Occupational status; %				
Full time	1846 (51.4)	1069 (51.4)	522 (52.8)	255 (48.6)
Partial	957 (26.6)	518 (24.9)	272 (27.5)	167 (31.8)
Unemployed	789 (22)	492 (23.7)	194 (19.6)	103 (19.6)
Current Potential Occupational Exposure; %	605 (14.4)	330 (13.7)	177 (14.7)	98 (16.3)
Weight; (kg)	71 [61; 82] (39 to 166)	71 [60; 83] (42 to 166)	72 [62; 82] (39 to 140)	71 [61; 82] (41 to 135)
Height; (cm)	170 [163; 175] (108 to 198)	170 [163; 176] (108 to 198)	170 [163; 175] (124 to 196)	168 [161.2; 175] (108 to 196)
BMI; (kg/m^2)	24.6 [22.3; 27.7] (15.8 to 63.4)	24.6 [22.2; 27.5] (15.8 to 63.4)	24.6 [22.2; 27.8] (16.9 to 54.6)	25 [22.5; 28.2] (16 to 48.8)

*Deprivation index was computed taking into account both educational characteristics of the study participants (number of years of education, education degree), housing characteristics (number of rooms, number of occupants) and ownership of basic consumption goods (refrigerator and car). These types of data were collected by means EPIC questionnaire (see paragraph 3.1.4).

4.3 Exposure Biomarkers

4.3.1 Heavy Metals

Heavy metals results, expressed as serum concentrations in µg/L, were measured on the overall population. They are reported as geometric mean in the general population and in the three impact areas as well as in the 21 clusters (Tables 10a and 10b). Median and inter-quartile ranges are reported in tables 10c, 10d and 10e.

Geometric mean values were generally comparable among the three impact areas, except for mercury levels that were approximately twice in the medium impact areas vs. the high- and low-impact areas. Considering the 21 clusters defined and the 19 metals assessed, the most relevant differences of a single cluster *vs* average serum levels in the overall study sample, were reported in the Sabato valley cluster for cadmium (Cd111), nickel (Ni60) and antimony (Sb121) and for the two Irno valley clusters for mercury (Hg202), which were approximately five times higher than those found in the entire population.

Percentage fold decreases/increases of interesting 6 of 19 heavy metals, for which percentage fold increase with respect to the general population was ≥ 100% in at least one cluster, are reported in Table 11 and graphically represented in Figures 13 a to f.

In Median regression models, where impact area was the geographical analysis unit (Table 12 a/b), using the low impact area as reference, the high impact area was not associated with higher levels of any of the 19 heavy metals assessed, except for thallium (Tl205) whose median levels were significantly but only slightly higher in the high impact *vs* the low impact area (Difference between medians: 0.009; 95% C.I.: 0.006 to 0.011).

Compared to the low impact area, the medium impact one was instead associated with higher levels of mercury (Hg202) (Difference between medians: 0.87; 95% C.I.: 0.62 to 113), lithium (Li7) (Difference between medians: 0.09; 95% C.I.: 0.03 to 0.16), cobalt (Co59) (Difference between medians: 0.13; 95% C.I.: 0.09 to 0.17), manganese (Mn55) (Difference between medians: 0.15; 95% C.I.: 0.06 to 0.23), antimony (Sb121) (Difference between medians: 0.03; 95% C.I.: 0.01 to 0.04), thallium (Tl205) (Difference between medians: 0.004; 95% C.I.: 0.002 to 0.007) and zinc (Zn66) (Difference between medians: 138; 95% C.I.: 100 to 176). Older age was consistently associated with higher levels of arsenic (As75), cobalt (Co59), copper (Cu63) and strontium (Sr88). Male gender was characterized by higher levels of lithium (Li7), iron (Fe), manganese (Mn55), selenium (Se78) and zinc (Zn66) but lower levels of copper (Cu63).When analyzed by cluster (Table 13 a/b/c), it emerged that the associations found for the medium impact area were mainly due to the contribute of the Irno and Sabato Valley's clusters.

Finally, Tables 13d, 13e, 13f, report a correlation matrix showing how correlated metals are to each other, for impact areas. Noteworthy, there is a significant association between Hg and Cd in high impact areas and between Cd and Mn in medium impact areas.

Table 10a. Geometric mean of the heavy metals assessed in the serum (µg/L) by cluster and area.

Cluster	Li7	Be9	Cd111	As75	Cr52	Co59	Cu63	Fe	Hg202	Mn55
High Impact Area	**1.008**	**0.004**	**0.007**	**0.940**	**0.049**	**0.852**	**981.5**	**2311**	**0.573**	**0.612**
HIGH_1	1.102	0.004	0.006	1.033	0.045	0.948	1106	2363	0.631	0.831
HIGH_2	0.948	0.004	0.007	0.735	0.046	0.806	966.1	2215	0.587	0.518
HIGH_3	0.979	0.004	0.007	1.106	0.058	0.842	978.5	2153	0.686	0.562
HIGH_4	1.002	0.004	0.009	1.082	0.037	0.814	1007	2258	0.55	0.56
HIGH_5	1.151	0.003	0.007	0.744	0.036	0.966	992.4	2378	0.486	0.618
HIGH_6	1.046	0.004	0.008	1.236	0.063	0.856	967.9	2399	0.518	0.685
HIGH_7	0.809	0.005	0.007	0.457	0.048	0.786	942.9	2189	0.473	0.323
HIGH_8	1.021	0.005	0.008	1.17	0.051	0.854	975.9	2433	0.622	0.688
HIGH_9	1.104	0.004	0.007	1.096	0.044	0.905	973.7	2399	0.575	0.422
HIGH_10	0.98	0.003	0.008	1.09	0.078	0.888	916.6	2364	0.628	0.845
HIGH_11	0.96	0.003	0.008	0.911	0.063	0.777	984.2	2256	0.521	0.893
HIGH_12	1.031	0.006	0.009	0.967	0.039	0.799	975.7	2341	0.658	0.666
Medium Impact Area	**1.093**	**0.005**	**0.011**	**1.084**	**0.065**	**0.812**	**966.3**	**2154**	**1.505**	**0.689**
MEDIUM_1	0.958	0.004	0.007	0.765	0.056	0.975	1025	2592	0.648	1.12
MEDIUM_2	0.949	0.004	0.008	0.779	0.063	0.855	1000	2475	0.564	0.592
MEDIUM_3	0.912	0.004	0.005	0.799	0.028	0.833	995	2372	0.463	0.441
IRNO_1	1.287	0.004	0.012	1.566	0.07	0.582	868.1	1656	4.517	0.499
IRNO_2	1.245	0.005	0.011	1.221	0.078	0.697	881	1641	4.971	0.459
SABATO	1.292	0.007	0.053	1.794	0.151	1.021	1044	2425	1.678	1.74
Low Impact Area	**1.037**	**0.004**	**0.009**	**1.182**	**0.061**	**0.904**	**1034.0**	**2379**	**0.756**	**0.922**
LOW_1	1.044	0.005	0.009	1.226	0.036	0.878	1006	2316	0.865	0.658
LOW_2	1.021	0.003	0.008	1.327	0.076	0.941	1049	2487	0.899	1.058
LOW_3	1.047	0.003	0.011	1.007	0.086	0.894	1048	2334	0.561	1.129
Overall	**1.036**	**0.004**	**0.009**	**1.011**	**0.055**	**0.847**	**984.4**	**2274**	**0.804**	**0.671**

Table 10b. Geometric mean of the heavy metals assessed in the serum (µg/L) by cluster and area.

Cluster	Mo98	Ni60	Pb208	Sb121	Se78	Sr88	Tl205	V51	Zn66
High Impact Area	**0.405**	**0.457**	**0.038**	**0.023**	**187**	**28.3**	**0.014**	**0.287**	**1342**
HIGH_1	0.408	0.321	0.036	0.02	191	29.8	0.017	0.374	1434
HIGH_2	0.308	0.315	0.04	0.02	188	27.7	0.015	0.299	1311
HIGH_3	0.466	0.526	0.036	0.02	189	28.5	0.013	0.344	1388
HIGH_4	0.379	0.471	0.043	0.03	183	27.7	0.016	0.275	1270
HIGH_5	0.275	0.586	0.032	0.02	207	30.9	0.014	0.27	1502
HIGH_6	0.611	0.31	0.048	0.03	186	28.4	0.012	0.25	1287
HIGH_7	0.445	0.306	0.03	0.02	165	26.1	0.011	0.213	1246
HIGH_8	0.429	0.648	0.048	0.03	188	28.8	0.014	0.309	1367
HIGH_9	0.53	0.829	0.037	0.04	192	28.5	0.015	0.234	1390
HIGH_10	0.429	0.412	0.039	0.03	184	27.5	0.015	0.37	1287
HIGH_11	0.399	0.36	0.037	0.03	190	28.5	0.019	0.343	1321
HIGH_12	0.299	0.887	0.035	0.02	183	27.4	0.016	0.225	1314
Medium Impact Area	**0.378**	**0.515**	**0.041**	**0.057**	**193**	**29.1**	**0.014**	**0.246**	**1507**
MEDIUM_1	0.391	0.309	0.044	0.04	193	28.7	0.018	0.32	1460
MEDIUM_2	0.374	0.381	0.04	0.02	190	28.0	0.013	0.361	1385
MEDIUM_3	0.423	0.724	0.031	0.02	199	29.2	0.007	0.245	1472
IRNO_1	0.276	0.23	0.043	0.09	184	27.4	0.015	0.171	1580
IRNO_2	0.229	0.226	0.034	0.08	181	27.3	0.009	0.176	1697
SABATO	0.752	3.198	0.06	0.23	209	34.3	0.033	0.26	1451
Low Impact Area	**0.598**	**0.404**	**0.047**	**0.038**	**196**	**29.0**	**0.011**	**0.321**	**1331**
LOW_1	0.464	0.33	0.034	0.03	177	29.0	0.009	0.34	1294
LOW_2	0.663	0.632	0.043	0.05	212	29.6	0.015	0.29	1377
LOW_3	0.698	0.406	0.07	0.03	201	28.4	0.01	0.336	1321
Overall	**0.42**	**0.472**	**0.04**	**0.03**	**190**	**28.6**	**0.014**	**0.279**	**1386**

Table 10c. Median (IQR) of the heavy metals assessed in the serum (µg/L) by cluster and area. Data are reported as median [25th ; 75th percentile].

Cluster	Li7	Be9	Cd111	As75	Cr52	Co59
High Impact Area	**1.01 (0.727 ; 1.444)**	**0.002 (0.002 ; 0.012)**	**0.005 (0.001 ; 0.059)**	**1.446 (0.615 ; 3.237)**	**0.196 (0.002 ; 0.832)**	**0.909 (0.688 ; 1.162)**
HIGH_1	1.048 (0.75 ; 1.427)	0.002 (0.002 ; 0.018)	0.001 (0.001 ; 0.057)	1.565 (0.673 ; 3.182)	0.168 (0.002 ; 0.727)	0.965 (0.744 ; 1.232)
HIGH_2	0.958 (0.655 ; 1.353)	0.002 (0.002 ; 0.011)	0.002 (0.001 ; 0.07)	1.432 (0.483 ; 3.198)	0.149 (0.002 ; 0.727)	0.857 (0.67 ; 1.135)
HIGH_3	0.919 (0.692 ; 1.397)	0.002 (0.002 ; 0.012)	0.001 (0.001 ; 0.062)	1.478 (0.724 ; 2.962)	0.239 (0.002 ; 0.862)	0.914 (0.662 ; 1.167)
HIGH_4	1.02 (0.733 ; 1.385)	0.002 (0.002 ; 0.002)	0.005 (0.001 ; 0.07)	1.749 (0.692 ; 3.405)	0.08 (0.002 ; 0.784)	0.94 (0.706 ; 1.18)
HIGH_5	1.136 (0.789 ; 1.625)	0.002 (0.002 ; 0.002)	0.001 (0.001 ; 0.042)	1.328 (0.589 ; 3.178)	0.088 (0.002 ; 0.633)	1.01 (0.785 ; 1.268)
HIGH_6	1.044 (0.722 ; 1.538)	0.002 (0.002 ; 0.015)	0.001 (0.001 ; 0.073)	1.711 (0.723 ; 3.233)	0.314 (0.002 ; 1.045)	0.914 (0.66 ; 1.207)
HIGH_7	0.961 (0.684 ; 1.429)	0.002 (0.002 ; 0.021)	0.003 (0.001 ; 0.067)	1.151 (0.278 ; 2.904)	0.167 (0.002 ; 1.122)	0.796 (0.6 ; 1.054)
HIGH_8	1.02 (0.736 ; 1.426)	0.002 (0.002 ; 0.028)	0.001 (0.001 ; 0.069)	1.575 (0.724 ; 4.03)	0.189 (0.002 ; 0.882)	0.877 (0.665 ; 1.141)
HIGH_9	1.087 (0.803 ; 1.573)	0.002 (0.002 ; 0.009)	0.009 (0.001 ; 0.029)	1.47 (0.552 ; 3.477)	0.165 (0.002 ; 0.751)	0.925 (0.717 ; 1.192)
HIGH_10	0.972 (0.732 ; 1.358)	0.002 (0.002 ; 0.005)	0.008 (0.001 ; 0.053)	1.591 (0.684 ; 3.23)	0.348 (0.002 ; 0.891)	0.966 (0.778 ; 1.18)
HIGH_11	0.942 (0.654 ; 1.338)	0.002 (0.002 ; 0.006)	0.007 (0.001 ; 0.033)	1.174 (0.577 ; 2.947)	0.253 (0.002 ; 0.777)	0.838 (0.689 ; 1.023)
HIGH_12	1.035 (0.751 ; 1.413)	0.002 (0.002 ; 0.031)	0.006 (0.001 ; 0.102)	1.437 (0.701 ; 3.469)	0.077 (0.002 ; 0.9)	0.89 (0.623 ; 1.129)
Medium Impact Area	**1.099 (0.805 ; 1.581)**	**0.002 (0.002 ; 0.022)**	**0.017 (0.001 ; 0.082)**	**1.69 (0.631 ; 3.89)**	**0.288 (0.002 ; 0.906)**	**1.039 (0.739 ; 1.352)**
MEDIUM_1	0.929 (0.717 ; 1.296)	0.002 (0.002 ; 0.025)	0.005 (0.001 ; 0.04)	1.091 (0.528 ; 2.696)	0.227 (0.002 ; 0.645)	0.996 (0.802 ; 1.242)
MEDIUM_2	0.973 (0.7 ; 1.325)	0.002 (0.002 ; 0.007)	0.007 (0.001 ; 0.069)	1.313 (0.421 ; 3.702)	0.294 (0.002 ; 0.908)	0.965 (0.716 ; 1.25)
MEDIUM_3	0.973 (0.76 ; 1.312)	0.002 (0.002 ; 0.008)	0.001 (0.001 ; 0.031)	1.051 (0.44 ; 2.912)	0.002 (0.002 ; 0.822)	0.895 (0.657 ; 1.19)
IRNO_1	1.244 (0.921 ; 1.938)	0.002 (0.002 ; 0.014)	0.019 (0.001 ; 0.055)	2.843 (0.872 ; 5.978)	0.291 (0.002 ; 0.805)	1.189 (0.772 ; 1.508)
IRNO_2	1.247 (0.854 ; 1.882)	0.002 (0.002 ; 0.032)	0.019 (0.001 ; 0.058)	2.235 (0.926 ; 4.872)	0.345 (0.002 ; 0.81)	1.232 (0.839 ; 1.599)
SABATO	1.197 (0.958 ; 1.643)	0.002 (0.002 ; 0.059)	0.095 (0.038 ; 0.196)	2.319 (1.242 ; 3.69)	0.697 (0.002 ; 1.26)	1.042 (0.812 ; 1.319)
Low Impact Area	**0.99 (0.697 ; 1.537)**	**0.002 (0.002 ; 0.006)**	**0.013 (0.001 ; 0.041)**	**1.45 (0.62 ; 3.168)**	**0.277 (0.002 ; 0.804)**	**0.923 (0.742 ; 1.15)**
LOW_1	1.017 (0.686 ; 1.427)	0.002 (0.002 ; 0.031)	0.013 (0.001 ; 0.059)	1.64 (0.666 ; 3.774)	0.062 (0.002 ; 0.703)	0.906 (0.74 ; 1.162)
LOW_2	0.956 (0.675 ; 1.561)	0.002 (0.002 ; 0.002)	0.012 (0.001 ; 0.029)	1.413 (0.685 ; 3.632)	0.331 (0.002 ; 0.94)	0.935 (0.776 ; 1.155)
LOW_3	0.995 (0.709 ; 1.704)	0.002 (0.002 ; 0.002)	0.017 (0.001 ; 0.044)	1.422 (0.503 ; 2.547)	0.347 (0.002 ; 0.982)	0.925 (0.696 ; 1.101)
Overall	**1.026 (0.736 ; 1.504)**	**0.002 (0.002 ; 0.012)**	**0.009 (0.001 ; 0.062)**	**1.513 (0.624 ; 3.462)**	**0.242 (0.002 ; 0.854)**	**0.941 (0.709 ; 1.213)**

Table 10d. Median (IQR) of the heavy metals assessed in the serum (μg/L) by cluster and area. Data are reported as median [25th ; 75th percentile].

Cluster	Cu63	Fe	Hg202	Mn55	Mo98	Ni60	Pb208
High Impact Area	**985.9 (832.5 ; 1159.9)**	**2389.2 (1794.8 ; 3127.9)**	**0.693 (0.119 ; 1.881)**	**1.031 (0.543 ; 1.506)**	**0.756 (0.305 ; 1.207)**	**1.121 (0.021 ; 2.537)**	**0.018 (0.018 ; 0.044)**
HIGH_1	1093.4 (935.8 ; 1310.8)	2382.7 (1833.3 ; 3216.4)	0.783 (0.119 ; 1.796)	1.18 (0.797 ; 1.538)	0.754 (0.322 ; 1.209)	0.907 (0.021 ; 1.956)	0.018 (0.018 ; 0.02)
HIGH_2	999.7 (814.5 ; 1152.4)	2334.6 (1643 ; 3073.4)	0.792 (0.119 ; 1.798)	0.934 (0.504 ; 1.346)	0.638 (0.128 ; 1.128)	0.765 (0.021 ; 2.092)	0.018 (0.018 ; 0.088)
HIGH_3	989.8 (814.2 ; 1165.7)	2232.8 (1674.4 ; 2867.3)	0.929 (0.119 ; 2.351)	1.015 (0.473 ; 1.449)	0.76 (0.389 ; 1.202)	1.407 (0.021 ; 2.262)	0.018 (0.018 ; 0.036)
HIGH_4	1009.4 (897.2 ; 1159.4)	2327.9 (1770 ; 3069.1)	0.684 (0.119 ; 1.86)	0.981 (0.493 ; 1.47)	0.744 (0.29 ; 1.084)	1.027 (0.021 ; 2.503)	0.018 (0.018 ; 0.073)
HIGH_5	980.6 (819.2 ; 1170.3)	2516.6 (1832.3 ; 3161.3)	0.449 (0.119 ; 1.901)	0.966 (0.506 ; 1.416)	0.628 (0.057 ; 1.151)	1.526 (0.021 ; 3.475)	0.018 (0.018 ; 0.018)
HIGH_6	982 (826.4 ; 1165.9)	2434.7 (1814.8 ; 3321.6)	0.528 (0.119 ; 1.741)	1.148 (0.625 ; 1.575)	0.945 (0.438 ; 1.28)	0.819 (0.021 ; 2.808)	0.018 (0.018 ; 0.142)
HIGH_7	957.6 (815.3 ; 1162.2)	2239.3 (1670.3 ; 3101.3)	0.316 (0.119 ; 1.765)	0.781 (0.156 ; 1.317)	0.774 (0.364 ; 1.223)	0.641 (0.021 ; 2.011)	0.018 (0.018 ; 0.018)
HIGH_8	985.4 (847.2 ; 1151.4)	2472.2 (1924 ; 3111.9)	0.728 (0.142 ; 1.996)	1.019 (0.595 ; 1.52)	0.855 (0.428 ; 1.172)	1.151 (0.021 ; 3.112)	0.018 (0.018 ; 0.098)
HIGH_9	963.2 (851.2 ; 1127.2)	2427.3 (1800 ; 3250.7)	0.697 (0.141 ; 1.534)	0.86 (0.383 ; 1.504)	0.833 (0.387 ; 1.21)	1.753 (0.556 ; 3.056)	0.018 (0.018 ; 0.026)
HIGH_10	921.4 (795.5 ; 1065.5)	2437.4 (1918.2 ; 3123.2)	0.854 (0.119 ; 1.762)	1.135 (0.714 ; 1.611)	0.81 (0.296 ; 1.23)	0.78 (0.021 ; 2.266)	0.018 (0.018 ; 0.048)
HIGH_11	993.7 (832.5 ; 1155.9)	2420.2 (1899.8 ; 3022.3)	0.654 (0.119 ; 1.583)	1.158 (0.737 ; 1.527)	0.742 (0.322 ; 1.168)	0.774 (0.021 ; 1.96)	0.018 (0.018 ; 0.033)
HIGH_12	967.7 (836.5 ; 1120)	2449.7 (1851.1 ; 3261.1)	0.863 (0.146 ; 2.086)	1.081 (0.535 ; 1.595)	0.669 (0.193 ; 1.072)	1.578 (0.692 ; 2.703)	0.018 (0.018 ; 0.021)
Medium Impact Area	**982 (842.7 ; 1148.3)**	**2442.9 (1787 ; 3176.5)**	**1.937 (0.564 ; 4.284)**	**1.309 (0.729 ; 1.78)**	**0.794 (0.27 ; 1.215)**	**1.054 (0.021 ; 3.042)**	**0.018 (0.018 ; 0.06)**
MEDIUM_1	1017.8 (848.6 ; 1169.6)	2597 (2016.9 ; 3408)	0.826 (0.119 ; 2.029)	1.353 (0.847 ; 1.884)	0.789 (0.317 ; 1.209)	0.416 (0.021 ; 2.401)	0.018 (0.018 ; 0.041)
MEDIUM_2	1002.4 (854.6 ; 1200.9)	2534.3 (1820.7 ; 3306.8)	0.568 (0.119 ; 1.792)	1.086 (0.563 ; 1.556)	0.792 (0.176 ; 1.355)	1.068 (0.021 ; 2.47)	0.018 (0.018 ; 0.047)
MEDIUM_3	981.5 (848.3 ; 1172.5)	2441.6 (1876.3 ; 3110)	0.385 (0.119 ; 1.65)	0.952 (0.438 ; 1.368)	0.809 (0.244 ; 1.27)	1.845 (0.021 ; 4.269)	0.018 (0.018 ; 0.018)
IRNO_1	903.8 (812.2 ; 1077.3)	2233.2 (1424.1 ; 3170.7)	5.14 (2.874 ; 8.812)	1.4 (0.614 ; 1.735)	0.728 (0.06 ; 1.112)	0.383 (0.021 ; 1.478)	0.018 (0.018 ; 0.064)
IRNO_2	925.8 (755.3 ; 1083.4)	2326.2 (1286 ; 2998)	5.046 (2.667 ; 10.542)	1.246 (0.643 ; 1.654)	0.757 (0.01 ; 1.165)	0.415 (0.021 ; 1.169)	0.018 (0.018 ; 0.021)
SABATO	1040.3 (900.9 ; 1176.4)	2468.8 (1999.9 ; 3090.4)	1.812 (1.056 ; 2.972)	1.785 (1.421 ; 2.248)	0.895 (0.602 ; 1.168)	3.43 (1.873 ; 11.147)	0.018 (0.018 ; 0.281)
Low Impact Area	**1008.6 (870.8 ; 1210.5)**	**2456 (1847.6 ; 3177.3)**	**0.877 (0.189 ; 2.15)**	**1.157 (0.749 ; 1.622)**	**0.902 (0.541 ; 1.33)**	**0.858 (0.021 ; 3.423)**	**0.018 (0.018 ; 0.144)**
LOW_1	983.4 (839 ; 1202.3)	2381.9 (1810.9 ; 3061.3)	1.059 (0.253 ; 2.635)	0.956 (0.585 ; 1.456)	0.859 (0.405 ; 1.306)	0.653 (0.021 ; 3.29)	0.018 (0.018 ; 0.018)
LOW_2	1034.5 (865.9 ; 1224.6)	2561.7 (1904.3 ; 3301.1)	1.006 (0.333 ; 2.521)	1.187 (0.813 ; 1.594)	0.942 (0.557 ; 1.283)	1.449 (0.021 ; 3.807)	0.018 (0.018 ; 0.085)
LOW_3	1024.3 (899 ; 1191.3)	2378.1 (1840.2 ; 3139.5)	0.734 (0.119 ; 1.398)	1.303 (0.832 ; 1.791)	0.908 (0.624 ; 1.402)	0.81 (0.021 ; 3.243)	0.018 (0.018 ; 0.354)
Overall	**988.1 (841 ; 1163.6)**	**2413.9 (1799.7 ; 3155.7)**	**1.004 (0.119 ; 2.619)**	**1.129 (0.609 ; 1.618)**	**0.795 (0.326 ; 1.224)**	**1.057 (0.021 ; 2.752)**	**0.018 (0.018 ; 0.057)**

Table 10e. Median (IQR) of the heavy metals assessed in the serum (µg/L) by cluster and area. Data are reported as median [25th ; 75th percentile].

Cluster	Pb208	Sb121	Se78	Sr88	Tl205	V51	Zn66
High Impact Area	**0.018 (0.018 ; 0.044)**	**0.063 (0.002 ; 0.168)**	**194.3 (159 ; 228.1)**	**28.3 (22.3 ; 35.3)**	**0.025 (0.001 ; 0.054)**	**0.418 (0.216 ; 0.693)**	**1387 (1178.3 ; 1618.6)**
HIGH_1	0.018 (0.018 ; 0.02)	0.029 (0.002 ; 0.18)	198.3 (162.9 ; 233)	29.2 (24.3 ; 35.9)	0.031 (0.001 ; 0.06)	0.477 (0.292 ; 0.787)	1462.2 (1262.6 ; 1679.6)
HIGH_2	0.018 (0.018 ; 0.088)	0.002 (0.002 ; 0.144)	194.1 (158.5 ; 230.2)	27.7 (22.2 ; 35.2)	0.03 (0.001 ; 0.065)	0.501 (0.232 ; 0.876)	1363.6 (1123.5 ; 1614.8)
HIGH_3	0.018 (0.018 ; 0.036)	0.029 (0.002 ; 0.141)	195.8 (157.4 ; 238.1)	27.6 (20.8 ; 37.2)	0.022 (0.001 ; 0.045)	0.45 (0.251 ; 0.744)	1404.6 (1159.7 ; 1638.7)
HIGH_4	0.018 (0.018 ; 0.073)	0.103 (0.002 ; 0.182)	187 (159 ; 219.3)	28.2 (21.7 ; 35.9)	0.029 (0.001 ; 0.067)	0.451 (0.188 ; 0.774)	1371 (1121.1 ; 1617.6)
HIGH_5	0.018 (0.018 ; 0.018)	0.002 (0.002 ; 0.15)	207.2 (175 ; 243.2)	29.8 (25.4 ; 35.9)	0.022 (0.002 ; 0.046)	0.424 (0.223 ; 0.639)	1529.2 (1308.9 ; 1709.9)
HIGH_6	0.018 (0.018 ; 0.142)	0.071 (0.002 ; 0.155)	195.9 (157.7 ; 228.1)	28.5 (21.8 ; 35.3)	0.022 (0.001 ; 0.049)	0.387 (0.163 ; 0.678)	1380.3 (1114.9 ; 1602.1)
HIGH_7	0.018 (0.018 ; 0.018)	0.002 (0.002 ; 0.132)	175.8 (135.3 ; 206.6)	25.8 (20.3 ; 33.8)	0.017 (0.001 ; 0.058)	0.395 (0.184 ; 0.609)	1304 (1040.8 ; 1568.9)
HIGH_8	0.018 (0.018 ; 0.098)	0.078 (0.002 ; 0.182)	194.7 (154.7 ; 218.9)	27.9 (23 ; 36.5)	0.027 (0.001 ; 0.054)	0.446 (0.218 ; 0.781)	1408.9 (1173.9 ; 1608.5)
HIGH_9	0.018 (0.018 ; 0.026)	0.106 (0.002 ; 0.187)	195.1 (164.2 ; 228.8)	28.5 (22.9 ; 34.6)	0.023 (0.005 ; 0.042)	0.361 (0.184 ; 0.558)	1392.7 (1218.2 ; 1676.3)
HIGH_10	0.018 (0.018 ; 0.048)	0.103 (0.002 ; 0.185)	187.4 (155.5 ; 218.1)	27.9 (22.4 ; 33.6)	0.023 (0.004 ; 0.048)	0.409 (0.266 ; 0.631)	1317.1 (1103.8 ; 1545.1)
HIGH_11	0.018 (0.018 ; 0.033)	0.102 (0.002 ; 0.174)	197.7 (168.4 ; 234.5)	28.9 (22.2 ; 35.4)	0.028 (0.008 ; 0.057)	0.433 (0.299 ; 0.645)	1369.8 (1195.3 ; 1578.4)
HIGH_12	**0.018 (0.018 ; 0.021)**	**0.03 (0.002 ; 0.156)**	**195.2 (157.2 ; 225.9)**	**27.9 (22 ; 33.6)**	**0.026 (0.001 ; 0.061)**	**0.345 (0.14 ; 0.613)**	**1348.2 (1150.6 ; 1533.3)**
Medium Impact Area	**0.018 (0.018 ; 0.06)**	**0.144 (0.002 ; 0.23)**	**196 (163.4 ; 233.9)**	**28.4 (23.5 ; 36.2)**	**0.021 (0.002 ; 0.043)**	**0.446 (0.193 ; 0.731)**	**1539.5 (1320.8 ; 1753.8)**
MEDIUM_1	0.018 (0.018 ; 0.041)	0.143 (0.002 ; 0.228)	195.9 (163.9 ; 233.2)	27.8 (23.6 ; 34.2)	0.026 (0.012 ; 0.045)	0.462 (0.257 ; 0.641)	1493.9 (1322.9 ; 1703.3)
MEDIUM_2	0.018 (0.018 ; 0.047)	0.074 (0.002 ; 0.167)	202.7 (161.2 ; 231.6)	27.4 (22 ; 35.4)	0.02 (0.001 ; 0.046)	0.534 (0.27 ; 0.812)	1428.5 (1224.5 ; 1642.5)
MEDIUM_3	0.018 (0.018 ; 0.018)	0.03 (0.002 ; 0.151)	197.7 (165.8 ; 238.2)	28.6 (24.7 ; 35)	0.008 (0.001 ; 0.028)	0.363 (0.185 ; 0.646)	1457.7 (1265.5 ; 1688.4)
IRNO_1	0.018 (0.018 ; 0.064)	0.138 (0.084 ; 0.215)	184 (155 ; 221.8)	28.2 (22.6 ; 34.3)	0.022 (0.013 ; 0.042)	0.404 (0.009 ; 0.671)	1651.7 (1467.6 ; 1899.3)
IRNO_2	0.018 (0.018 ; 0.021)	0.156 (0.07 ; 0.216)	183.5 (156.3 ; 227.6)	27 (21.8 ; 33.6)	0.017 (0.001 ; 0.03)	0.469 (0.009 ; 0.694)	1708.7 (1522.4 ; 1973.6)
SABATO	0.018 (0.018 ; 0.281)	0.246 (0.183 ; 0.304)	206.6 (184 ; 243.8)	32.8 (27.1 ; 41.3)	0.038 (0.016 ; 0.085)	0.533 (0.186 ; 0.836)	1456.2 (1288.9 ; 1627.7)
Low Impact Area	**0.018 (0.018 ; 0.144)**	**0.118 (0.002 ; 0.192)**	**204 (169.3 ; 240.8)**	**28.1 (22.5 ; 36.2)**	**0.016 (0.003 ; 0.031)**	**0.428 (0.245 ; 0.681)**	**1379.3 (1170.8 ; 1607)**
LOW_1	0.018 (0.018 ; 0.018)	0.094 (0.002 ; 0.176)	188 (152 ; 222.4)	28.3 (22.3 ; 35.7)	0.013 (0.001 ; 0.031)	0.439 (0.248 ; 0.739)	1353.7 (1161.4 ; 1596.6)
LOW_2	0.018 (0.018 ; 0.085)	0.154 (0.002 ; 0.209)	216.2 (181.3 ; 250.8)	28.1 (22.5 ; 37.5)	0.02 (0.009 ; 0.033)	0.416 (0.229 ; 0.584)	1401.6 (1172.1 ; 1598.7)
LOW_3	0.018 (0.018 ; 0.354)	0.105 (0.002 ; 0.196)	207.6 (172.5 ; 236.2)	27.4 (22.6 ; 34.8)	0.015 (0.001 ; 0.027)	0.439 (0.258 ; 0.75)	1387.7 (1187.6 ; 1622.2)
Overall	**0.018 (0.018 ; 0.057)**	**0.103 (0.002 ; 0.194)**	**196.1 (161.9 ; 232.1)**	**28.3 (22.7 ; 35.6)**	**0.022 (0.001 ; 0.046)**	**0.425 (0.215 ; 0.703)**	**1431.4 (1216.8 ; 1661.8)**

Table 11. Percentage fold-increase/decrease of the geometric mean of six of the heavy metals assessed in the population indwelling in each cluster/area vs the entire population by cluster and area. Green and red colors express the fold-decrease and increase, respectively.

Cluster	Cd111	Hg202	Mn55	Ni60	Sb121	Tl205
High Impact Area	-19.1	-28.7	-8.8	-3.2	-28.7	2.8
HIGH_1	-29.4	-21.6	23.9	-32	-36.0	24.3
HIGH_2	-15.9	-27.0	-22.7	-33.3	-50.0	8.6
HIGH_3	-24.3	-14.7	-16.2	11.4	-37.5	-4.9
HIGH_4	-0.4	-31.6	-16.5	-0.2	-14.7	18.3
HIGH_5	-21.7	-39.5	-7.8	24.2	-48.5	-0.2
HIGH_6	-10.6	-35.6	2.1	-34.3	-16.3	-15.1
HIGH_7	-14.4	-41.1	-51.8	-35.2	-50.1	-22.2
HIGH_8	-9.9	-22.6	2.5	37.3	-14.7	3.8
HIGH_9	-14.7	-28.5	-37.0	75.8	23.0	8.1
HIGH_10	-6.8	-21.9	26.0	-12.7	-3.5	8.0
HIGH_11	-13.1	-35.2	33.1	-23.6	-9.6	38.7
HIGH_12	5.5	-18.2	-0.7	88.1	-38.2	16.6
Medium Impact Area	4.0	-5.9	37.5	-14.4	17.7	-19.2
MEDIUM_1	-22.4	-19.3	66.9	-34.5	22.6	34.9
MEDIUM_2	-3.8	-29.8	-11.7	-19.3	-24.3	-7.5
MEDIUM_3	-45.4	-42.4	-34.2	53.4	-32.8	-51.5
IRNO_1	40.7	461.9	-25.7	-51.2	180.9	12.4
IRNO_2	30.9	518.4	-31.5	-52	146.0	-33.8
SABATO	517.5	108.7	159.5	577.6	626.6	142.2
Low Impact Area	27.1	87.2	2.7	9.1	76.6	2.8
LOW_1	9.3	7.6	-1.9	-30	3.8	-34.4
LOW_2	-6.2	11.8	57.8	33.9	48.3	8.6
LOW_3	22.2	-30.3	68.3	-13.9	6.4	-27.9

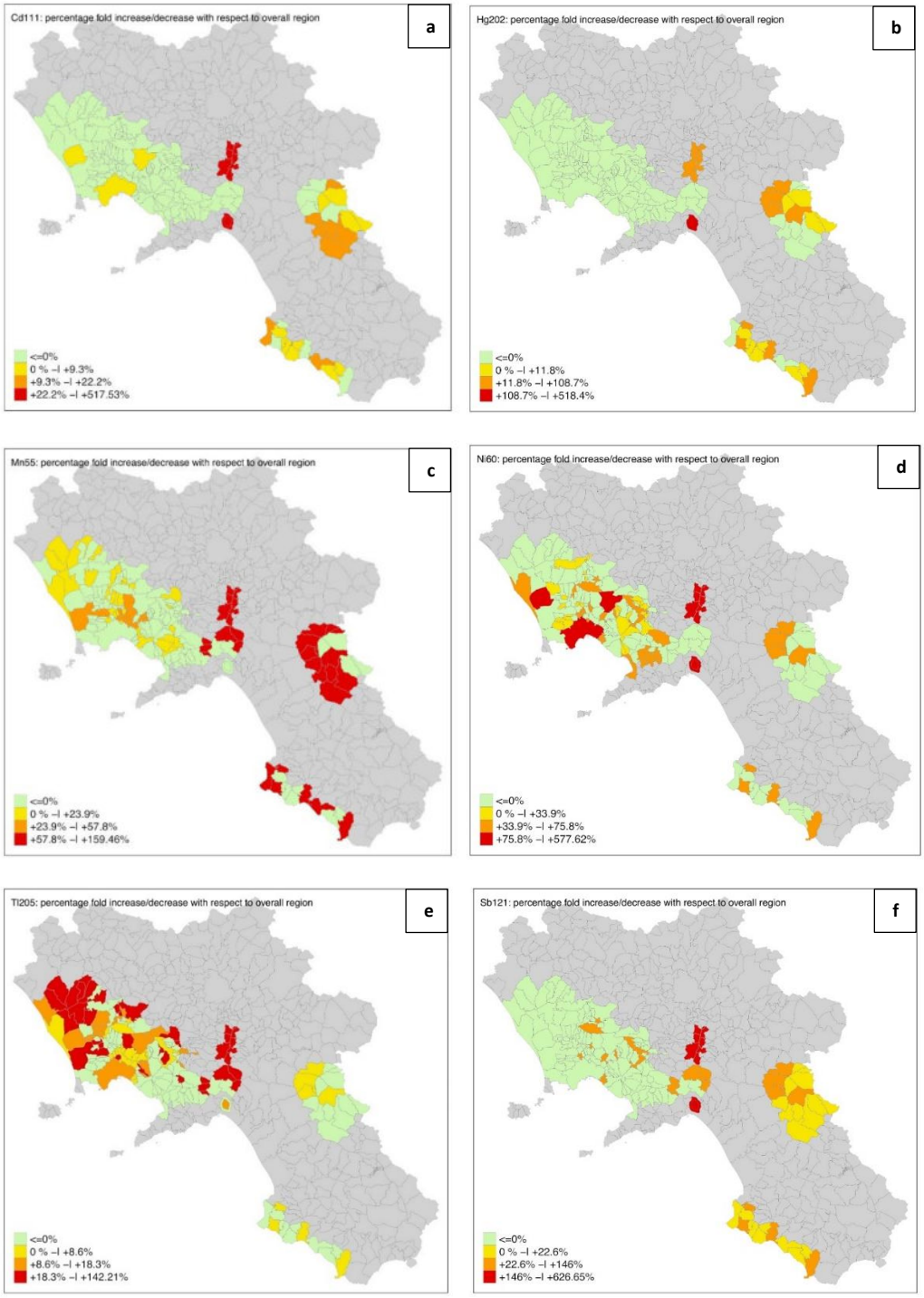


Figure 13. Campania Region representations with percentage fold decreases/increases of heavy metals for which percentage fold increase with respect to the general population was ≥ 100% in at least one cluster: (a) Cadmium – Cd111. (b) Mercury – Hg202. (c) Manganese – Mn55. (d) Nichel – Ni60. (e) Antimony – Sb121. (f) Thallium – Tl205.

Table 12a. Median regression models where low impact area is the geographical analysis unit**.** Likewise, age and gender are considered for the median regression model. Bold results indicate differences with p<0.05. Results are expressed as difference between adjusted medians.

	Se78	**Sr88**	**Tl205**	**V51**	**Zn66**	**Co59**	**Cr52**
Area; Low impact ref.							
Medium impact	**-8.7 (-15 to -2.4) p=0.007**	-0.05 (-1.22 to 1.11) p=0.929	**0.004 (0.002 to 0.007) p=0.002**	0.02 (-0.03 to 0.07) p=0.393	**138.1 (100.1 to 176)** p<0.001	**0.13 (0.09 to 0.17) p<0.001**	0.05 (-0.05 to 0.15) p=0.332
High Impact	**-11.7 (-17.8 to -5.6) p<0.001**	-0.21 (-1.28 to 0.87) p=0.707	**0.009 (0.006 to 0.011) p<0.001**	-0.01 (-0.05 to 0.03) p=0.586	-5.9 (-43.3 to 31.5) p=0.758	-0.01 (-0.04 to 0.03) p=0.677	-0.06 (-0.16 to 0.03) p=0.203
Gender; Male	**21.2 (15.8 to 26.5) p<0.001**	-0.26 (-1.12 to 0.61) p=0.557	-0.002 (-0.005 to 0.001) p=0.153	0 (-0.04 to 0.04) p=0.906	**132.8 (99.3 to 166.2)** p<0.001	0.01 (-0.03 to 0.04) p=0.721	0.03 (-0.03 to 0.1) p=0.321
Age; 20-30 ref							
30-40	-0.1 (-5.2 to 5) p=0.974	**1.23 (0.37 to 2.09) p=0.005**	0.001 (-0.002 to 0.004) p=0.553	-0.01 (-0.04 to 0.02) p=0.428	-18.8 (-52.8 to 15.2) p=0.279	**0.04 (0.01 to 0.08)** p=0.021	0.05 (-0.02 to 0.12) p=0.181
40-50	1.1 (-4.3 to 6.5) p=0.7	**1.78 (0.92 to 2.64) p=0.00**	0.002 (-0.001 to 0.004) p=0.126	-0.02 (-0.06 to 0.01) p=0.192	-0.6 (-32.3 to 31.2) p=0.973	**0.07 (0.03 to 0.1)** p=0.00	0.03 (-0.04 to 0.09) p=0.409
	Hg202	**Li7**	**Cd111**	**As75**	**Cr52**	**Ni60**	**Sb121**
Area; Low impact ref.							
Medium impact	**0.87 (0.62 to 1.13) p<0.001**	**0.09 (0.03 to 0.16) p=0.003**	0.003 (-0.003 to 0.009) p=0.296	0.14 (-0.1 to 0.37) p=0.251	0.05 (-0.05 to 0.15) p=0.332	0.12 (-0.31 to 0.55) p=0.584	**0.03 (0.01 to 0.04) p=0.005**
High Impact	**-0.3 (-0.45 to -0.15) p<0.001**	0.01 (-0.05 to 0.07) p=0.694	**-0.007 (-0.011 to -0.003) p=0.001**	-0.06 (-0.26 to 0.15) p=0.588	-0.06 (-0.16 to 0.03) p=0.203	0.21 (-0.2 to 0.62) p=0.316	**-0.06 (-0.08 to -0.04) p<0.001**
Gender; Male	0.16 (0 to 0.32) p=0.053	**0.12 (0.08 to 0.17) p<0.001**	-0.004 (-0.007 to 0) p=0.058	0.17 (-0.04 to 0.39) p=0.109	0.02 (-0.03 to 0.1) p=0.321	-0.02 (-0.31 to 0.28) p=0.915	0 (-0.02 to 0.02) p=0.779
Age; 20-30 ref							
30-40	0.15 (-0.02 to 0.33) p=0.087	-0.03 (-0.08 to 0.02) p=0.25	0 (-0.002 to 0.002) p=1	**0.25 (0.04 to 0.45) p=0.017**	0.05 (-0.02 to 0.12) p=0.181	-0.02 (-0.31 to 0.28) p=0.905	0.01 (-0.01 to 0.03) p=0.186
40-50	0.15 (-0.03 to 0.33) p=0.107	0.03 (-0.04 to 0.07) p=0.589	(-0.001 to 0.005) p=0.161	**0.26 (0.05 to 0.46) p=0.015**	0.03 (-0.04 to 0.09) p=0.409	0.02 (-0.21 to 0.26) p=0.86	0.01 (-0.01 to 0.03) p=0.272

Table 12b. Median regression models where low impact area is the geographical analysis unit**.** Likewise, age and gender are considered for the median regression model. Bold results indicate differences with p<0.05. Results are expressed as difference between adjusted medians.

	Cu63	**Fe**	**Mn55**	**Mo98**
Area; Low impact ref.				
Medium impact	**-43.8 (-67.4 to -20.2) p<0.001**	6 (-104 to 117) p=0.91	**0.15 (0.06 to 0.23) p=0.001**	**-0.09 (-0.16 to -0.02) p=0.016**
High Impact	**-42.2 (-63.8 to -20.7) p<0.001**	-94 (-200 to 13) p=0.085	**-0.12 (-0.2 to -0.04) p=0.004**	**-0.14 (-0.2 to -0.08) p<0.001**
Gender; Male	**-159 (-183.5 to -135.1) p<0.001**	**544 (440 to 649) p<0.001**	**0.11 (0.02 to 0.19) p=0.013**	0 (-0.06 to 0.06) p=0.991
Age; 20-30 ref				
30-40	**52.4 (33 to 71.7) p=0.00**	34 (-77 to 144) p=0.553	0.01 (-0.06 to 0.08) p=0.77	0.05 (-0.01 to 0.11) p=0.134
40-50	**80 (59.8 to 100.2) p=0.00**	-93 (-198 to 11) p=0.079	0.02 (-0.03 to 0.1) p=0.344	0.03 (-0.02 to 0.11) p=0.169

Table 13a. Median regression models for the heavy metals assessed in the population, using cluster as geographical reference unit. Data are reported as adjusted difference between medians with the corresponding 95% CIs. Bold results indicate differences with p<0.005. It is considered as reference value the cluster with the lower geometric mean for each metal.

Cluster	Hg202	Li7	Cd111	As75	Cr52	Co59
HIGH_1	ref	0 (-0.16 to 0.16)	-0.004 (-0.012 to 0.003)	0.12 (-0.4 to 0.64)	-0.07 (-0.27 to 0.13)	0.04 (-0.06 to 0.13)
HIGH_2	-0.02 (-0.45 to 0.4)	-0.1 (-0.25 to 0.05)	-0.003 (-0.012 to 0.005)	-0.01 (-0.55 to 0.54)	-0.08 (-0.29 to 0.12)	-0.07 (-0.15 to 0.01)
HIGH_3	0.14 (-0.42 to 0.7)	-0.1 (-0.24 to 0.05)	-0.004 (-0.011 to 0.004)	0.22 (-0.31 to 0.75)	-0.01 (-0.21 to 0.18)	-0.01 (-0.1 to 0.09)
HIGH_4	0.02 (-0.46 to 0.5)	-0.03 (-0.18 to 0.12)	-0.002 (-0.013 to 0.009)	0.39 (-0.19 to 0.96)	-0.14 (-0.33 to 0.05)	0.01 (-0.07 to 0.09)
HIGH_5	-0.33 (-0.69 to 0.03)	0.11 (-0.04 to 0.26)	-0.003 (-0.01 to 0.005)	-0.11 (-0.59 to 0.36)	-0.09 (-0.28 to 0.1)	0.08 (-0.02 to 0.18)
HIGH_6	-0.22 (-0.62 to 0.18)	0.06 (-0.1 to 0.22)	-0.003 (-0.011 to 0.005)	0.26 (-0.3 to 0.82)	0.06 (-0.14 to 0.26)	0.01 (-0.1 to 0.11)
HIGH_7	-0.28 (-0.7 to 0.14)	-0.07 (-0.23 to 0.08)	-0.002 (-0.011 to 0.007)	-0.26 (-0.73 to 0.22)	-0.08 (-0.29 to 0.13)	**-0.13 (-0.23 to -0.04)**
HIGH_8	0.03 (-0.37 to 0.43)	0 (-0.13 to 0.13)	-0.004 (-0.012 to 0.005)	0.23 (-0.31 to 0.77)	-0.02 (-0.23 to 0.19)	-0.05 (-0.13 to 0.04)
HIGH_9	-0.1 (-0.48 to 0.29)	0.07 (-0.08 to 0.22)	0.001 (-0.008 to 0.009)	ref	-0.08 (-0.27 to 0.1)	ref
HIGH_10	0.05 (-0.35 to 0.44)	-0.02 (-0.16 to 0.11)	0.001 (-0.009 to 0.01)	0.15 (-0.37 to 0.66)	0.16 (-0.06 to 0.37)	0.05 (-0.04 to 0.14)
HIGH_11	-0.15 (-0.54 to 0.25)	-0.08 (-0.22 to 0.06)	0 (-0.008 to 0.008)	-0.07 (-0.54 to 0.41)	0.01 (-0.21 to 0.23)	**-0.11 (-0.18 to -0.03)**
HIGH_12	0.09 (-0.33 to 0.51)	-0.02 (-0.15 to 0.12)	ref	-0.1 (-0.59 to 0.39)	-0.12 (-0.31 to 0.07)	-0.05 (-0.15 to 0.04)
MEDIUM_1	0.16 (-0.28 to 0.61)	-0.06 (-0.2 to 0.09)	-0.002 (-0.01 to 0.007)	-0.33 (-0.79 to 0.14)	ref	0.07 (-0.02 to 0.17)
MEDIUM_2	-0.07 (-0.58 to 0.44)	-0.03 (-0.16 to 0.1)	-0.002 (-0.011 to 0.008)	-0.14 (-0.65 to 0.36)	0.09 (-0.12 to 0.29)	0.03 (-0.06 to 0.13)
MEDIUM_3	-0.3 (-0.66 to 0.06)	-0.02 (-0.15 to 0.1)	-0.005 (-0.012 to 0.001)	-0.4 (-0.88 to 0.07)	-0.16 (-0.32 to 0.01)	-0.03 (-0.14 to 0.07)
IRNO_1	**4.17 (3.28 to 5.05)**	**0.26 (0.1 to 0.42)**	**0.012 (0.003 to 0.021)**	**1.47 (0.68 to 2.26)**	0.08 (-0.14 to 0.3)	**0.29 (0.16 to 0.42)**
IRNO_2	**4.17 (3.22 to 5.13)**	**0.22 (0.05 to 0.38)**	**0.013 (0.002 to 0.023)**	**0.8 (0.17 to 1.42)**	0.08 (-0.11 to 0.27)	**0.33 (0.22 to 0.45)**
SABATO	**1.03 (0.54 to 1.51)**	**0.17 (0.03 to 0.32)**	**0.092 (0.065 to 0.119)**	**0.82 (0.26 to 1.37)**	**0.45 (0.18 to 0.72)**	0.09 (-0.03 to 0.2)
LOW_1	0.34 (-0.06 to 0.74)	ref	0.004 (-0.005 to 0.012)	0.31 (-0.26 to 0.88)	-0.12 (-0.3 to 0.06)	-0.02 (-0.11 to 0.06)
LOW_2	0.34 (-0.13 to 0.8)	-0.05 (-0.18 to 0.08)	0.003 (-0.005 to 0.012)	0.22 (-0.34 to 0.78)	0.09 (-0.12 to 0.3)	0 (-0.08 to 0.08)
LOW_3	-0.03 (-0.4 to 0.34)	-0.05 (-0.19 to 0.09)	0.009 (0 to 0.018)	-0.02 (-0.5 to 0.45)	0.12 (-0.08 to 0.32)	0 (-0.09 to 0.08)

Table 13b. Median regression models for the heavy metals assessed in the population, using cluster as geographical reference unit. Data are reported as adjusted difference between medians with the corresponding 95% CIs. Bold results indicate differences with p<0.005. It is considered as reference value the cluster with the lower geometric mean for each metal.

Cluster	Cu63	Fe	Mn55	Mo98	Ni60	Sb121
HIGH_1	**79.5 (13.1 to 145.9)**	-4.6 (-287 to 277.8)	0.1 (-0.07 to 0.26)	-0.02 (-0.17 to 0.12)	ref	**-0.05 (-0.1 to -0.01)**
HIGH_2	-36.2 (-97.2 to 24.9)	-58.9 (-343.4 to 225.5)	**-0.18 (-0.36 to 0)**	-0.13 (-0.3 to 0.04)	-0.37 (-1.13 to 0.38)	**-0.07 (-0.1 to -0.05)**
HIGH_3	-46 (-105.3 to 13.4)	-175.1 (-426.5 to 76.3)	-0.08 (-0.24 to 0.08)	0 (-0.15 to 0.16)	0.39 (-0.38 to 1.15)	**-0.06 (-0.1 to -0.02)**
HIGH_4	-2.8 (-57.2 to 51.6)	-99.9 (-351.6 to 151.9)	-0.08 (-0.25 to 0.08)	-0.04 (-0.21 to 0.12)	0.22 (-0.54 to 0.99)	0.01 (-0.03 to 0.05)
HIGH_5	-18.5 (-75 to 38)	-1.4 (-300.6 to 297.7)	-0.12 (-0.3 to 0.05)	-0.17 (-0.37 to 0.02)	0.58 (-0.21 to 1.36)	**-0.08 (-0.1 to -0.05)**
HIGH_6	-0.5 (-59.9 to 59)	67.3 (-210 to 344.5)	0.08 (-0.1 to 0.26)	0.16 (-0.02 to 0.33)	-0.38 (-1.03 to 0.27)	-0.02 (-0.07 to 0.02)
HIGH_7	-38.8 (-94.1 to 16.5)	-121.7 (-413.8 to 170.5)	**-0.29 (-0.46 to -0.12)**	-0.02 (-0.19 to 0.15)	0.04 (-0.74 to 0.83)	**-0.07 (-0.1 to -0.04)**
HIGH_8	ref	53.7 (-221.8 to 329.3)	-0.07 (-0.23 to 0.09)	0.07 (-0.08 to 0.23)	0.17 (-0.54 to 0.89)	-0.01 (-0.06 to 0.04)
HIGH_9	-38.6 (-89.4 to 12.3)	ref	-0.19 (-0.39 to 0.01)	0.03 (-0.13 to 0.19)	0.8 (0.07 to 1.52)	0.01 (-0.02 to 0.04)
HIGH_10	**-62.5 (-115.7 to -9.4)**	-76.7 (-313.8 to 160.5)	0.05 (-0.09 to 0.19)	0.03 (-0.16 to 0.21)	-0.2 (-0.95 to 0.55)	0.01 (-0.03 to 0.04)
HIGH_11	-33.6 (-96.3 to 29.1)	-29.3 (-287.7 to 229.1)	0.08 (-0.07 to 0.22)	-0.06 (-0.22 to 0.11)	-0.17 (-0.85 to 0.51)	0.01 (-0.04 to 0.05)
HIGH_12	-9.5 (-67.5 to 48.4)	-47.4 (-300 to 205.2)	0 (-0.18 to 0.17)	-0.1 (-0.24 to 0.04)	0.54 (-0.18 to 1.25)	-0.06 (-0.09 to -0.02)
MEDIUM_1	-10.3 (-70 to 49.3)	201.5 (-76.4 to 479.4)	**0.26 (0.07 to 0.45)**	ref	-0.35 (-1.09 to 0.4)	0.05 (0.01 to 0.08)
MEDIUM_2	15.8 (-43 to 74.7)	161.6 (-79.9 to 403.2)	ref	0.01 (-0.19 to 0.22)	0 (-0.72 to 0.72)	-0.01 (-0.06 to 0.03)
MEDIUM_3	-36.5 (-89.7 to 16.7)	56.9 (-199.4 to 313.2)	-0.13 (-0.32 to 0.07)	0.04 (-0.14 to 0.21)	0.95 (-0.05 to 1.95)	**-0.06 (-0.1 to -0.03)**
IRNO_1	**-72.5 (-122 to -23)**	11.6 (-294.5 to 317.7)	**0.34 (0.17 to 0.51)**	-0.01 (-0.19 to 0.18)	-0.53 (-1.12 to 0.06)	**0.05 (0.03 to 0.08)**
IRNO_2	**-58.3 (-109.5 to -7.1)**	-33.3 (-296.1 to 229.5)	**0.24 (0.08 to 0.41)**	0.02 (-0.15 to 0.18)	-0.51 (-1.09 to 0.07)	**0.07 (0.04 to 0.09)**
SABATO	23.5 (-34.1 to 81)	17.3 (-206.9 to 241.6)	**0.69 (0.54 to 0.85)**	0.1 (-0.06 to 0.25)	**2.26 (1.58 to 2.95)**	**0.15 (0.12 to 0.18)**
LOW_1	-7.4 (-64.9 to 50.1)	31.1 (-196.7 to 259)	-0.12 (-0.27 to 0.03)	0.07 (-0.09 to 0.23)	-0.33 (-1.19 to 0.54)	ref
LOW_2	37 (-13.7 to 87.6)	132.8 (-111.6 to 377.2)	0.09 (-0.07 to 0.25)	0.14 (0 to 0.28)	0.6 (-0.54 to 1.73)	**0.06 (0.03 to 0.08)**
LOW_3	40.5 (-16.1 to 97.1)	-3.2 (-251.9 to 245.5)	**0.21 (0.04 to 0.38)**	**0.14 (0.01 to 0.28)**	-0.13 (-0.94 to 0.67)	0.01 (-0.02 to 0.05)

Table 13c. Median regression models for the heavy metals assessed in the population, using cluster as geographical reference unit. Data are reported as adjusted difference between medians with the corresponding 95% CIs. Bold results indicate differences with p<0.005. It is considered as reference value the cluster with the lower geometric mean for each metal.

Cluster	Se78	Sr88	Tl205	V51	Zn66
HIGH_1	4.7 (-11.4 to 20.8)	0.45 (-1.59 to 2.49)	0.008 (-0.001 to 0.017)	0.023 (-0.068 to 0.115)	58.8 (-37.5 to 155.2)
HIGH_2	2 (-14.2 to 18.1)	-1.17 (-3.68 to 1.33)	0.007 (-0.004 to 0.017)	0.04 (-0.06 to 0.139)	-37.4 (-129.2 to 54.3)
HIGH_3	ref	-1.32 (-4.17 to 1.52)	0 (-0.01 to 0.011)	-0.016 (-0.106 to 0.073)	-14.3 (-106.8 to 78.1)
HIGH_4	-3.8 (-20.4 to 12.8)	-0.78 (-3.3 to 1.74)	0.007 (-0.005 to 0.018)	-0.002 (-0.105 to 0.101)	-53.6 (-151.8 to 44.6)
HIGH_5	13.7 (-1.8 to 29.2)	1.22 (-0.94 to 3.38)	0 (-0.01 to 0.01)	-0.028 (-0.118 to 0.062)	**103.7 (16.6 to 190.8)**
HIGH_6	-0.4 (-15.5 to 14.8)	0.03 (-2.4 to 2.45)	ref	-0.063 (-0.144 to 0.018)	-23.8 (-107.4 to 59.8)
HIGH_7	**-19.3 (-37 to -1.7)**	-2.42 (-4.87 to 0.03)	-0.004 (-0.017 to 0.008)	-0.056 (-0.131 to 0.02)	-80.6 (-162.9 to 1.7)
HIGH_8	-3.3 (-17.8 to 11.2)	-0.05 (-2.06 to 1.97)	0.005 (-0.004 to 0.015)	-0.003 (-0.109 to 0.103)	-18.1 (-110.6 to 74.4)
HIGH_9	0.4 (-15.7 to 16.6)	-0.14 (-2.2 to 1.92)	-0.001 (-0.009 to 0.008)	**-0.089 (-0.166 to -0.011)**	-22.9 (-114 to 68.1)
HIGH_10	-4.6 (-20.6 to 11.5)	-1.09 (-3.32 to 1.15)	0.001 (-0.008 to 0.01)	-0.04 (-0.116 to 0.036)	**-97.6 (-179.8 to -15.5)**
HIGH_11	4.1 (-12.5 to 20.8)	0.91 (-1.52 to 3.33)	0.007 (-0.001 to 0.016)	-0.023 (-0.099 to 0.054)	-26.3 (-112.7 to 60.1)
HIGH_12	-2.4 (-18.5 to 13.6)	-0.55 (-2.73 to 1.64)	0.004 (-0.005 to 0.013)	**-0.112 (-0.203 to -0.021)**	-54.1 (-144.7 to 36.5)
MEDIUM_1	1.3 (-14.5 to 17.2)	-0.51 (-2.88 to 1.86)	0.002 (-0.006 to 0.011)	0.011 (-0.076 to 0.099)	74.1 (-15 to 163.2)
MEDIUM_2	8.2 (-6.6 to 22.9)	-1.08 (-3.14 to 0.99)	-0.001 (-0.009 to 0.008)	0.084 (-0.02 to 0.188)	-6.2 (-99.4 to 87)
MEDIUM_3	4.5 (-10.4 to 19.5)	0 (-1.94 to 1.93)	**-0.013 (-0.021 to -0.005)**	**-0.096 (-0.178 to -0.014)**	48.2 (-38.8 to 135.2)
IRNO_1	-13.7 (-31.6 to 4.2)	-0.56 (-2.63 to 1.52)	-0.001 (-0.009 to 0.008)	-0.038 (-0.134 to 0.059)	**235.8 (136.4 to 335.3)**
IRNO_2	-11.9 (-27.2 to 3.4)	-1.72 (-3.79 to 0.35)	-0.004 (-0.012 to 0.004)	0.029 (-0.061 to 0.118)	**287.5 (202.1 to 373)**
SABATO	14.7 (-1.9 to 31.2)	4.79 (2.25 to 7.33)	**0.016 (0.006 to 0.026)**	-0.019 (-0.153 to 0.115)	51.2 (-33.7 to 136.2)
LOW_1	-6.5 (-21.9 to 8.9)	0.11 (-2.35 to 2.56)	-0.008 (-0.016 to 0)	ref	-42.3 (-125.7 to 41.2)
LOW_2	**23.4 (8 to 38.8)**	ref	-0.002 (-0.01 to 0.006)	-0.054 (-0.138 to 0.029)	ref
LOW_3	14.4 (-1.4 to 30.1)	-0.31 (-2.71 to 2.1)	-0.008 (-0.016 to 0.001)	-0.019 (-0.107 to 0.07)	-9.4 (-100.7 to 81.9)

Table 13d. Metals correlation matrix for low impact areas and clusters. Data are reported as Spearman Correlation Coefficient (p-value). Red color expresses the significant correlations.

Low	Li7	Be9	Cd111	As75	Cr52	Co59	Cu63	Fe	Hg202	Mn55	Mo98	Ni60	Pb208	Sb121	Se78	Sr88	Tl205	V51	Zn66
Li7	1	0.107 (0.009)	0.178 (<0.001)	0.139 (0.001)	0.061 (0.135)	0.236 (<0.001)	0.104 (0.011)	0.152 (<0.001)	0.23 (<0.001)	0.213 (<0.001)	0.207 (<0.001)	0.141 (0.028)	0.033 (0.418)	0.174 (<0.001)	0.201 (<0.001)	0.526 (<0.001)	0.113 (0.006)	0.173 (<0.001)	0.251 (<0.001)
Be9		1	0.092 (0.025)	0.014 (0.727)	0.08 (0.052)	0.086 (0.036)	0.093 (0.024)	0.03 (0.465)	0.202 (<0.001)	0.011 (0.78)	0.143 (<0.001)	0.087 (0.175)	-0.135 (0.001)	0.018 (0.654)	0.042 (0.306)	0.053 (0.197)	0.074 (0.07)	0.088 (0.031)	0.094 (0.024)
Cd111			1	0.096 (0.019)	0.252 (<0.001)	0.091 (0.027)	0.099 (0.016)	0.068 (0.099)	0.242 (<0.001)	0.265 (<0.001)	0.312 (<0.001)	0.173 (0.007)	0.218 (<0.001)	0.206 (<0.001)	0.056 (0.171)	0.078 (0.056)	0.241 (<0.001)	0.214 (<0.001)	0.105 (0.013)
As75				1	0.045 (0.277)	0.142 (0.001)	0.089 (0.03)	0.096 (0.019)	0.202 (<0.001)	0.005 (0.911)	-0.06 (0.143)	0.149 (0.02)	-0.055 (0.179)	0.032 (0.432)	0.152 (<0.001)	0.116 (0.004)	0.155 (<0.001)	0.004 (0.927)	0.123 (0.003)
Cr52					1	0.034 (0.411)	-0.036 (0.385)	0.051 (0.215)	0.215 (<0.001)	0.332 (<0.001)	0.096 (0.02)	0.283 (<0.001)	0.192 (<0.001)	0.226 (<0.001)	0.033 (0.425)	-0.047 (0.248)	0.178 (<0.001)	0.117 (0.004)	0.067 (0.111)
Co59						1	0.269 (<0.001)	0.212 (<0.001)	0.101 (0.036)	0.227 (<0.001)	0.112 (0.006)	0.111 (0.083)	0.062 (0.13)	0.157 (<0.001)	0.32 (<0.001)	0.25 (<0.001)	0.079 (0.054)	0.116 (0.004)	0.261 (<0.001)
Cu63							1	0.164 (<0.001)	0.055 (0.256)	0.206 (<0.001)	0.17 (<0.001)	0.025 (0.693)	0.026 (0.521)	0.18 (<0.001)	0.415 (<0.001)	0.269 (<0.001)	0.064 (0.121)	0.116 (0.004)	0.293 (<0.001)
Fe								1	0.05 (0.299)	0.322 (<0.001)	0.186 (<0.001)	0.035 (0.586)	0.054 (0.185)	0.172 (<0.001)	0.448 (<0.001)	0.235 (<0.001)	0.073 (0.074)	0.187 (<0.001)	0.424 (<0.001)
Hg202									1	0.068 (0.154)	0.197 (<0.001)	0.314 (<0.001)	0.032 (0.509)	0.271 (<0.001)	-0.018 (0.713)	0.119 (0.013)	0.207 (<0.001)	0.157 (0.001)	0.17 (<0.001)
Mn55										1	0.298 (<0.001)	0.178 (0.005)	0.276 (<0.001)	0.301 (<0.001)	0.258 (<0.001)	0.131 (0.001)	0.142 (<0.001)	0.171 (<0.001)	0.3 (<0.001)
Mo98											1	0.309 (<0.001)	0.151 (<0.001)	0.331 (<0.001)	0.219 (<0.001)	0.158 (<0.001)	0.072 (0.078)	0.087 (0.034)	0.217 (<0.001)
Ni60												1	0.202 (0.001)	0.391 (<0.001)	-0.105 (0.1)	-0.072 (0.265)	0.11 (0.087)	-0.027 (0.671)	-0.031 (0.646)
Pb208													1	0.205 (<0.001)	0.029 (0.472)	-0.025 (0.541)	0.114 (0.005)	0.058 (0.158)	0.046 (0.268)
Sb121														1	0.254 (<0.001)	0.134 (0.001)	0.218 (<0.001)	-0.01 (0.815)	0.186 (<0.001)
Se78															1	0.389 (<0.001)	0.098 (0.017)	0.13 (0.001)	0.542 (<0.001)
Sr88																1	0.05 (0.222)	0.13 (0.001)	0.308 (<0.001)
Tl205																	1	0.061 (0.136)	0.099 (0.018)
V51																		1	0.279 (<0.001)
Zn66																			1

Table 13e. Metals correlation matrix for medium impact areas and clusters. Data are reported as Spearman Correlation Coefficient (p-value). Red color expresses the significant correlations.

Medium	Li7	Be9	Cd111	As75	Cr52	Co59	Cu63	Fe	Hg202	Mn55	Mo98	Ni60	Pb208	Sb121	Se78	Sr88	Tl205	V51	Zn66
Li7	1	0.171 (<0.001)	0.212 (<0.001)	0.055 (0.06)	0.152 (<0.001)	0.04 (0.167)	0.097 (0.001)	0.154 (<0.001)	0.145 (<0.001)	0.061 (0.035)	0.048 (0.154)	0.128 (<0.001)	0.145 (<0.001)	0.138 (<0.001)	0.377 (<0.001)	0.104 (<0.001)	0.029 (0.31)	0.154 (<0.001)	
Be9		1	0.045 (0.118)	0.12 (<0.001)	-0.05 (0.084)	-0.05 (0.085)	-0.014 (0.633)	-0.049 (0.091)	0.074 (0.016)	-0.027 (0.357)	-0.021 (0.459)	-0.072 (0.034)	0.084 (0.004)	0.032 (0.268)	-0.007 (0.82)	-0.06 (0.039)	0.042 (0.147)	-0.071 (0.014)	0.042 (0.151)
Cd111			1	0.084 (0.003)	0.211 (<0.001)	0.153 (<0.001)	0.098 (0.001)	0.048 (0.1)	0.171 (<0.001)	0.322 (<0.001)	0.26 (<0.001)	0.26 (<0.001)	0.164 (<0.001)	0.299 (<0.001)	-0.012 (0.669)	0.14 (<0.001)	0.221 (<0.001)	0.201 (<0.001)	-0.037 (0.208)
As75				1	0.032 (0.275)	-0.016 (0.572)	0.004 (0.883)	0.003 (0.908)	0.276 (<0.001)	0.029 (0.322)	-0.131 (<0.001)	-0.002 (0.953)	0.052 (0.075)	0.023 (0.429)	0.1 (0.001)	0.127 (<0.001)	0.077 (0.008)	-0.002 (0.942)	0.138 (<0.001)
Cr52					1	0.171 (<0.001)	0.037 (0.205)	0.112 (<0.001)	0.142 (<0.001)	0.316 (<0.001)	0.252 (<0.001)	0.279 (<0.001)	0.015 (0.616)	0.286 (<0.001)	-0.102 (<0.001)	0.028 (0.327)	0.074 (0.01)	0.166 (<0.001)	-0.046 (0.12)
Co59						1	0.235 (<0.001)	0.264 (<0.001)	0.165 (<0.001)	0.354 (<0.001)	0.124 (<0.001)	0.204 (<0.001)	0.068 (0.019)	0.235 (<0.001)	0.026 (0.362)	0.203 (<0.001)	0.044 (0.128)	0.233 (<0.001)	0.217 (<0.001)
Cu63							1	0.066 (0.022)	-0.069 (0.024)	0.2 (<0.001)	0.224 (<0.001)	0.178 (<0.001)	0.082 (0.005)	0.089 (0.002)	0.275 (<0.001)	0.235 (<0.001)	0.075 (0.01)	0.19 (<0.001)	0.157 (<0.001)
Fe								1	-0.046 (0.134)	0.392 (<0.001)	0.148 (<0.001)	0.178 (<0.001)	0.116 (<0.001)	0.186 (<0.001)	0.195 (<0.001)	0.147 (<0.001)	0.072 (0.013)	0.182 (<0.001)	0.167 (<0.001)
Hg202									1	0.118 (<0.001)	-0.01 (0.749)	-0.08 (0.02)	0.029 (0.347)	0.234 (<0.001)	-0.188 (<0.001)	0.032 (0.292)	0.165 (<0.001)	0.01 (0.737)	0.205 (<0.001)
Mn55										1	0.379 (<0.001)	0.347 (<0.001)	0.292 (<0.001)	0.451 (<0.001)	0.022 (0.447)	0.204 (<0.001)	0.2 (<0.001)	0.244 (<0.001)	0.079 (0.007)
Mo98											1	0.307 (<0.001)	0.197 (<0.001)	0.307 (<0.001)	0.033 (0.255)	0.091 (0.002)	0.092 (0.001)	0.164 (<0.001)	-0.051 (0.082)
Ni60												1	0.121 (<0.001)	0.306 (<0.001)	-0.012 (0.715)	0.179 (<0.001)	0.188 (<0.001)	0.181 (<0.001)	-0.13 (<0.001)
Pb208													1	0.168 (<0.001)	0.005 (0.87)	0.029 (0.318)	0.078 (0.007)	0.053 (0.065)	-0.029 (0.319)
Sb121														1	-0.017 (0.562)	0.137 (<0.001)	0.242 (<0.001)	0.105 (<0.001)	0.049 (0.093)
Se78															1	0.181 (<0.001)	-0.043 (0.135)	-0.065 (0.024)	0.358 (<0.001)
Sr88																1	0.153 (<0.001)	0.116 (<0.001)	0.068 (0.021)
Tl205																	1	0.098 (0.001)	-0.074 (0.012)
V51																		1	-0.036 (0.215)
Zn66																			1

Table 13f. Metals correlation matrix for high impact areas and clusters. Data are reported as Spearman Correlation Coefficient (p-value). Red color expresses the significant correlations.

High	Li7	Be9	Cd111	As75	Cr52	Co59	Cu63	Fe	Hg202	Mn55	Mo98	Ni60	Pb208	Sb121	Se78	Sr88	Tl205	V51	Zn66
Li7	1	0.073 (<0.001)	0.139 (<0.001)	0.144 (<0.001)	0.077 (<0.001)	0.198 (<0.001)	0.147 (<0.001)	0.179 (<0.001)	0.102 (<0.001)	0.206 (<0.001)	0.116 (<0.001)	0.114 (<0.001)	0.066 (0.001)	0.058 (0.005)	0.196 (<0.001)	0.386 (<0.001)	0.091 (<0.001)	0.16 (<0.001)	0.21 (<0.001)
Be9		1	0.142 (<0.001)	0.087 (<0.001)	0.07 (0.001)	0.076 (<0.001)	0.031 (0.133)	-0.022 (0.286)	0.14 (<0.001)	0.046 (0.026)	0.049 (0.016)	0.037 (0.2)	0.001 (0.951)	0.046 (0.024)	-0.087 (<0.001)	-0.02 (0.336)	0.085 (<0.001)	0.106 (<0.001)	0.033 (0.12)
Cd111			1	0.152 (<0.001)	0.254 (<0.001)	0.114 (<0.001)	0.117 (<0.001)	-0.018 (0.379)	0.344 (<0.001)	0.294 (<0.001)	0.269 (<0.001)	0.273 (<0.001)	0.259 (<0.001)	0.266 (<0.001)	-0.034 (0.1)	0.023 (0.251)	0.23 (<0.001)	0.131 (<0.001)	0.06 (0.004)
As75				1	0.085 (<0.001)	0.079 (<0.001)	0.064 (0.002)	0.087 (<0.001)	0.163 (<0.001)	0.111 (<0.001)	0.055 (0.007)	-0.005 (0.852)	0.034 (0.099)	0.021 (0.31)	0.077 (<0.001)	0.113 (<0.001)	0.103 (<0.001)	0.097 (<0.001)	0.162 (<0.001)
Cr52					1	0.056 (0.006)	0.054 (0.008)	-0.029 (0.158)	0.215 (<0.001)	0.331 (<0.001)	0.185 (<0.001)	0.22 (<0.001)	0.192 (<0.001)	0.33 (<0.001)	-0.062 (0.003)	-0.039 (0.054)	0.074 (<0.001)	0.144 (<0.001)	-0.044 (0.035)
Co59						1	0.288 (<0.001)	0.182 (<0.001)	0.094 (<0.001)	0.207 (<0.001)	0.117 (<0.001)	0.113 (<0.001)	0.021 (0.303)	0.065 (0.001)	0.258 (<0.001)	0.217 (<0.001)	0.036 (0.077)	0.167 (<0.001)	0.31 (<0.001)
Cu63							1	0.126 (<0.001)	0.102 (<0.001)	0.227 (<0.001)	0.211 (<0.001)	0.129 (<0.001)	0.106 (<0.001)	0.106 (<0.001)	0.334 (<0.001)	0.315 (<0.001)	0.109 (<0.001)	0.176 (<0.001)	0.306 (<0.001)
Fe								1	-0.003 (0.907)	0.243 (<0.001)	0.093 (<0.001)	0.001 (0.986)	0.034 (0.1)	0.093 (<0.001)	0.378 (<0.001)	0.2 (<0.001)	0.015 (0.475)	0.087 (<0.001)	0.394 (<0.001)
Hg202									1	0.195 (<0.001)	0.231 (<0.001)	0.34 (<0.001)	0.153 (<0.001)	0.286 (<0.001)	-0.089 (<0.001)	0.059 (0.01)	0.281 (<0.001)	0.135 (<0.001)	0.097 (<0.001)
Mn55										1	0.31 (<0.001)	0.281 (<0.001)	0.264 (<0.001)	0.337 (<0.001)	0.16 (<0.001)	0.197 (<0.001)	0.13 (<0.001)	0.193 (<0.001)	0.224 (<0.001)
Mo98											1	0.209 (<0.001)	0.202 (<0.001)	0.398 (<0.001)	0.097 (<0.001)	0.121 (<0.001)	0.088 (<0.001)	0.063 (0.002)	0.12 (<0.001)
Ni60												1	0.198 (<0.001)	0.24 (<0.001)	-0.017 (0.548)	0.013 (0.658)	0.173 (<0.001)	0.007 (0.805)	0.138 (<0.001)
Pb208													1	0.249 (<0.001)	0.025 (0.224)	0.049 (0.016)	0.073 (<0.001)	0.102 (<0.001)	0.016 (0.443)
Sb121														1	0.01 (0.629)	0.054 (0.008)	0.215 (<0.001)	0.046 (0.024)	0.057 (0.006)
Se78															1	0.353 (<0.001)	0.054 (0.008)	0.03 (0.139)	0.501 (<0.001)
Sr88																1	0.078 (<0.001)	0.141 (<0.001)	0.336 (<0.001)
Tl205																	1	0.032 (0.123)	0.129 (<0.001)
V51																		1	0.136 (<0.001)
Zn66																			1

4.3.2 Dioxins, Furans and PCBs congeners by High Resolution Gas Chromatography – Mass Spectrometry

Results, expressed as serum concentrations, were measured on the subgroup of 600 subjects. All congeners were expressed as estimated value on volume of serum and on gram of fats. It was also considered the TEQ (Toxicity Equivalence) value that is the product of the concentration of the compound and its corresponding TEF (Toxic Equivalency Factor) on gram of fats (Van den Berg et al., 1998; Van den Berg et al., 2006).

To avoid any bias for the substitution of measured values under the Limit of Quantification (LOQ), the "Lower Bound" (LB) approach was applied for statistical analysis. This means that each value registered under the LOQ was substituted with 0.00 (zero).

The figures below show the median of the sum of classes of organic compounds among the three impact areas and among the clusters included in the subgroup.

In order to estimate a cumulative effect on the people potentially exposed, we firstly analyzed the sum of PCDD/PCDF + DL-PCB (pg WHO-TEQ/g lipids) (Fig. 14).

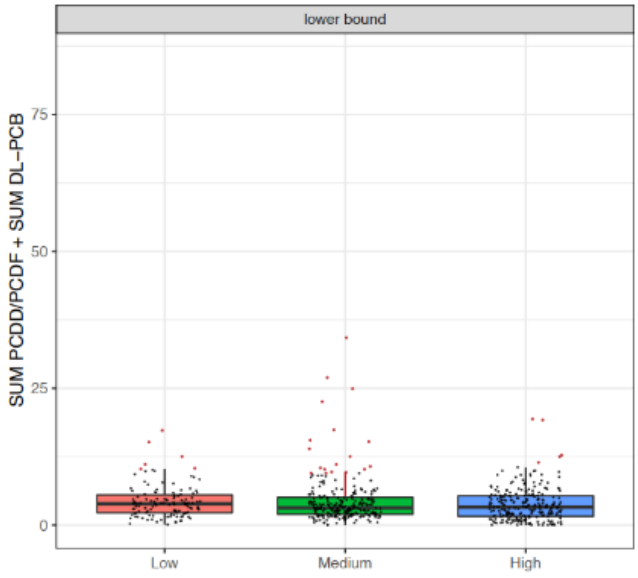


Figure 14. Sum of PCDD/PCDF + sum of DL-PCB expressed as pg WHO-TEQ/g lipids among the three different impact areas: low, medium and high impact area.

No significative evidence was shown among the three impact areas, considering the median values. If the same parameter is observed among clusters (Table 14 and Fig.15), it shows a sensitive increase of median value for Sabato Valley cluster compared to the overall median. Irno 1 and Irno 2 clusters, together with High 9 and High 12, show a slight increase compared to the other clusters.

Table 14. Median of Sum of PCDD/PCDF + sum of DL-PCB assessed in the serum (pg WHO-TEQ/g lipids) by cluster and grouped cluster for areas. LB=Lower Bound; IQR=Interquartile Range.

Cluster	Median [IQR] (Range) LB	CLUSTER GROUPED	Average value
LOW_1	3.919 [2.469; 6.011] (from 0.795 to 12.544)	LOW	3.894
LOW_2	3.869 [1.929; 4.694] (from 0.108 to 17.293)		
MEDIUM_1	3.016 [1.957; 3.826] (from 0.468 to 9.076)	MEDIUM	2.989
MEDIUM_2	2.963 [2.079; 4.13] (from 0.015 to 34.244)		
HIGH_9	3.047 [0.833; 5.079] (from 0.007 to 19.201)	HIGH	3.265
HIGH_10	2.755 [1.342; 6.296] (from 0.013 to 19.386)		
HIGH_11	3.207 [2.212; 4.601] (from 0.219 to 9.899)		
HIGH_12	4.051 [1.797; 5.685] (from 0.052 to 11.466)		
IRNO_1	4.951 [2.371; 8.145] (from 0.355 to 22.572)	IRNO	4.165
IRNO_2	3.38 [1.796; 5.994] (from 0.025 to 15.244)		
SABATO	4.192 [2.027; 6.5] (from 0.374 to 24.923)	SABATO	4.195
OVERALL	3.385 [1.897; 5.317] (from 0.007 to 34.244)	OVERALL	3.385

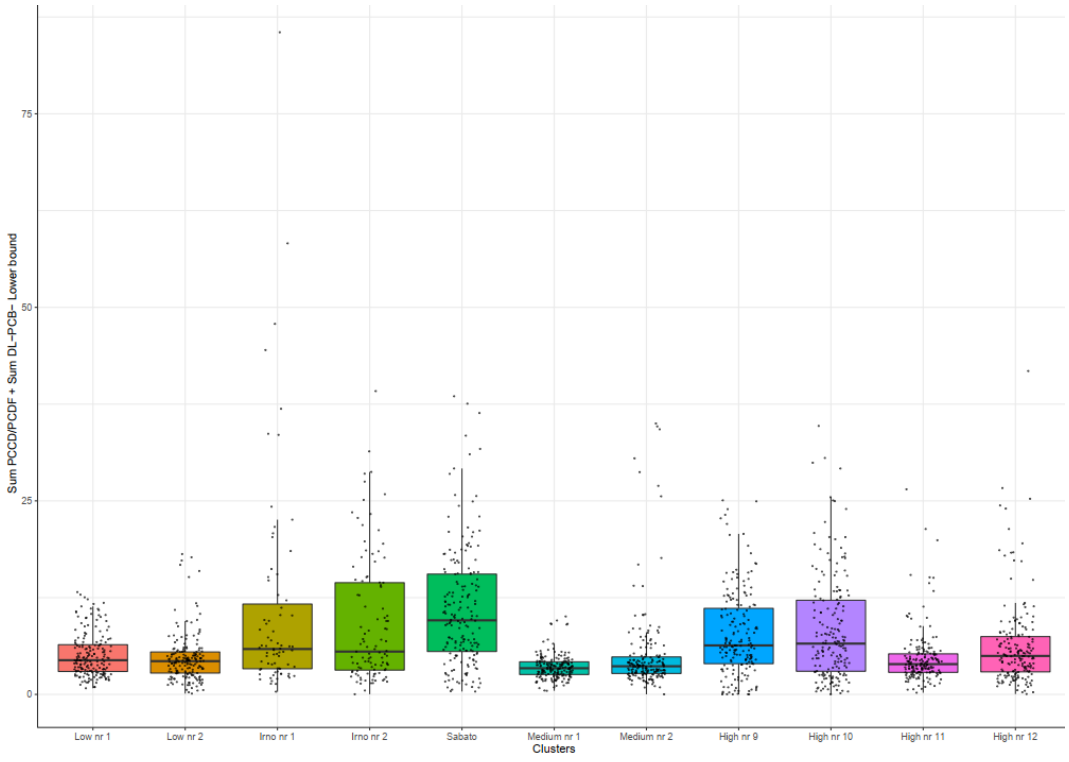


Figure 15. Sum of PCDD/PCDF + sum of DL-PCB expressed as pg WHO-TEQ/g lipids among the clusters.

The same scenario is presented when considering the only contribute of dioxins and furans, without DL-PCB addition (Table 15 and Fig. 16).

Table 15. Median of Sum of PCDD/PCDF assessed in the serum (pg WHO-TEQ/g lipids) by cluster and grouped cluster for areas. LB=Lower Bound; IQR=Interquartile Range.

Cluster	Median [IQR] (Range) LB	CLUSTER GROUPED	Average value
LOW_1	1.041 [0.762; 1.687] (from 0.074 to 3.16)	LOW	1.109
LOW_2	1.177 [0.649; 1.895] (from 0.1 to 6.988)		
MEDIUM_1	1.628 [1.168; 2.413] (from 0.302 to 5.683)	MEDIUM	1.675
MEDIUM_2	1.722 [1.166; 2.664] (from 0 to 26.414)		
HIGH_9	1.069 [0.171; 3.126] (from 0 to 17.691)	HIGH	1.662
HIGH_10	1.727 [0.769; 3.925] (from 0 to 18.996)		
HIGH_11	1.77 [1.229; 2.811] (from 0.158 to 8.81)		
HIGH_12	2.085 [1.012; 3.336] (from 0 to 9.933)		
IRNO_1	1.881 [0.832; 3.572] (from 0.036 to 19.84)	IRNO	1.866
IRNO_2	1.851 [0.995; 3.936] (from 0 to 13.505)		
SABATO	3.127 [0.42; 5.648] (from 0 to 16.347)	SABATO	3.127
OVERALL	1.618 [0.825; 2.891] (from 0 to 26.414)	OVERALL	1.618

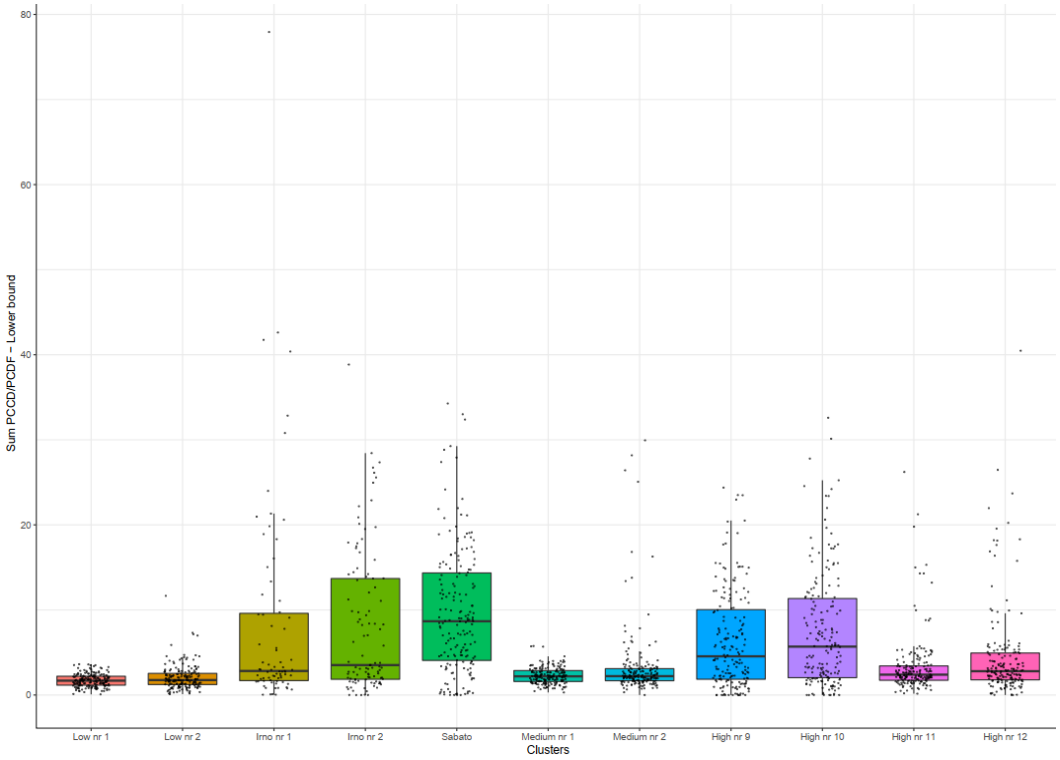


Figure 16. Sum of PCDD/PCDF expressed as pg WHO-TEQ/g lipids among the clusters.

Finally, NDL-PCB are shown in Figure 17. Median of Low 1 cluster is higher compared to those of other clusters (Table 16).

Table 16. Median of Sum of NDL-PCB assessed in the serum (ng/g lipids) by cluster and grouped cluster for areas. LB=Lower Bound; IQR=Interquartile Range.

Cluster	Median [IQR] (Range) LB	CLUSTER GROUPED	Average value
LOW_1	57.802 [36.375; 108.759] (from 14.226 to 342.419)	LOW	1.109
LOW_2	63.646 [29.554; 102.164] (from 4.942 to 355.678)		
MEDIUM_1	38.67 [30.292; 61.9] (from 3.218 to 159.285)	MEDIUM	1.675
MEDIUM_2	40.953 [25.97; 73.585] (from 11.718 to 429.657)		
HIGH_9	23.335 [16.732; 43.549] (from 6.29 to 285.887)	HIGH	1.662
HIGH_10	25.065 [17.488; 47.998] (from 7.91 to 438.94)		
HIGH_11	44.24 [25.263; 65.679] (from 4.571 to 171.898)		
HIGH_12	51.77 [29.158; 101.514] (from 10.2 to 560.58)		
IRNO_1	46.12 [26.369; 82.549] (from 17.95 to 1340.27)	IRNO	1.866
IRNO_2	40.983 [22.835; 67.873] (from 11.53 to 635.98)		
SABATO	47.61 [31.962; 63.532] (from 11.98 to 272.26)	SABATO	3.127
OVERALL	42.611 [24.763; 73.566] (from 3.218 to 1340.27)	OVERALL	1.618

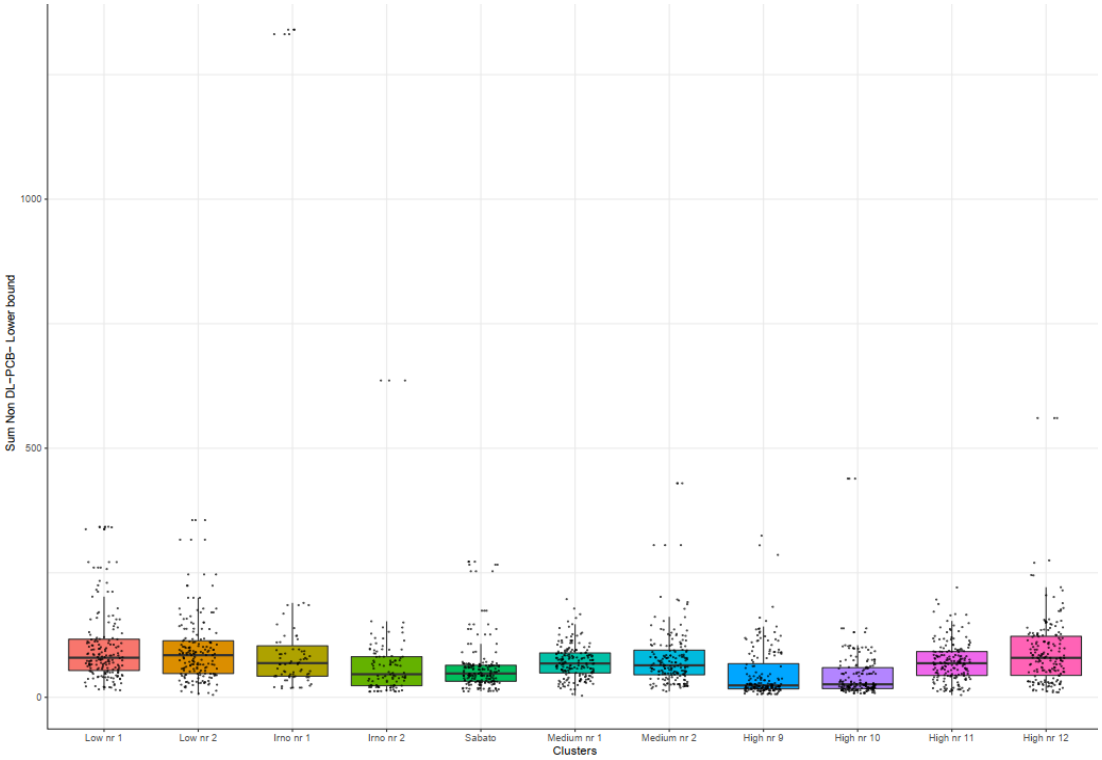


Figure 17. Sum of NDL-PCB expressed as ng/g lipids among the clusters.

Table 17 reports the percentage fold-increase/decrease of the median of compounds assessed per cluster and area. Noteworthy: the Irno and Sabato valley percentage increase for the sum of PCDD/F + DL-PCB, as well as in the high impact cluster 12; the NDL-PCB increase evidence for both clusters low_1 and low_2.

Results are also reported in Table 18 and 19 as median regression models using impact areas or clusters as geographical reference unit, respectively. It emerged that older age was consistently associated with the variable sum NDL-PCB and, to a lesser extent, with sum of PCDD/F, sum of PCDD/F+DL-PCB, sum of NDL-PCB.

Medium and high impact areas revealed association with higher levels of sum PCDD/F, and lower levels of DL-PCB and NDL-PCB.

Concerning median regression model data per clusters, all clusters of medium impact area were positively associated with higher levels of sum of PCDD/F, especially considering Irno_2 and Sabato valley. Also High_10, High_11, High_12 showed the same association. Conversely, low impact clusters were associated with higher levels of DL-PCB and NDL-PCB. Finally, clusters High_11 and High_12, Irno_1, Sabato, Medium_1 and Medium_2 were associated with high levels of NDL-PCB.

Table 17. Percentage fold-increase/decrease of the median of compounds assessed in the population indwelling in each cluster/area vs. the entire population. Green and red colors express the fold-decrease and increase, respectively.

Cluster	SUM PCDD/PCDF pg WHO-TEQ/g lipids	SUM PCDD/PCDF + SUM DL-PCB pg WHO-TEQ/g lipids	SUM DL-PCB pg WHO-TEQ/g lipids	SUM NDL-PCB ng/g lipids
High Impact Area	7,6	-2,3	-24,6	-19
HIGH_9	-33,9	-9,9	-42,8	-45,2
HIGH_10	6,7	-18,6	-59,6	-41,1
HIGH_11	9,3	-5,2	-4,5	3,8
HIGH_12	28,8	19,6	21,3	21,4
Medium Impact Area	16,8	-6,5	-12,8	1,4
MEDIUM_1	0,6	-10,9	-2,9	-9,2
MEDIUM_2	6,4	-12,4	-15,5	-3,8
IRNO_1	16,2	46,2	9,9	8,2
IRNO_2	14,4	-0,1	-37,4	-3,8
SABATO	93,2	23,8	-12,3	11,7
Low Impact Area	-33,1	15,3	93,6	41,9
LOW_1	-35,6	15,7	113,7	35,6
LOW_2	-27,2	14,2	68,9	49,3

Table 18. Results of median regression models for the compounds assessed in the population, using impact area as geographical reference unit. Data are reported as adjusted difference between medians with the corresponding 95% CIs. Bold results indicate differences with p<0.05.

	SUM PCDD/PCDF pg WHO-TEQ/g lipids	SUM PCDD/PCDF + SUM DL-PCB pg WHO-TEQ/g lipids	SUM DL-PCB pg WHO-TEQ/g lipids	SUM NDL-PCB ng/g lipids
Area; Low impact ref.				
Medium impact	**0.68 (0.41 to 1.01); p<0.001**	-0.34 (-0.93 to 0.27); p=0.183	**-1.12 (-1.56 to -0.8); p<0.001**	**-12.28 (-24.62 to -1.48); p=0.032**
High Impact	**0.58 (0.27 to 0.86); p=0.001**	-0.47 (-1.02 to 0.22); p=0.151	**-1.18 (-1.6 to -0.77); p<0.001**	**-16.94 (-29.62 to -6.5); p=0.005**
Gender; Male	**0.32 (0.01 to 0.54); p=0.047**	**0.6 (0.1 to 1.04); p=0.014**	0.08 (-0.17 to 0.46); p=0.636	**11.73 (3.2 to 18.66); p=0.002**
Age; 20-30 ref				
30-40	0.17 (-0.01 to 0.65); p=0.274	**0.63 (0.11 to 1.09); p=0.023**	0.19 (0 to 0.45); p=0.177	**10.17 (4.68 to 16.29); p<0.001**
40-50	**1.06 (0.69 to 1.5); p<0.001**	**2.04 (1.53 to 2.67); p<0.001**	**0.9 (0.59 to 1.15); p= p<0.001**	**43.15 (35.17 to 50.58); p<0.001**

Table 19. Results of median regression models for the compounds assessed in the population, using cluster as geographical reference unit. Data are reported as adjusted difference between medians with the corresponding 95% CIs. Bold results indicate differences with p<0.05.

Cluster	SUM PCDD/PCDF pg WHO-TEQ/g lipids	SUM PCDD/PCDF + SUM DL-PCB pg WHO-TEQ/g lipids	SUM DL-PCB pg WHO-TEQ/g lipids	SUM NDL-PCB ng/g lipids
LOW_1	Reference	0.96 (-0.42 to 2.13); p=0.14	**1.53 (0.75 to 2.34); p<0.001**	**32.33 (11.49 to 48.62); p<0.001**
LOW_2	0.25 (-0.18 to 0.57); p=0.229	0.63 (-0.94 to 1.63); p=0.337	**0.98 (0.32 to 1.78); p=0.012**	**25.54 (12.32 to 47.95); p=0.004**
MEDIUM_1	**0.72 (0.22 to 1.09); p=0.001**	0.12 (-1.49 to 1); p=0.838	0.04 (-0.5 to 0.61); p=0.888	**10.6 (-0.53 to 18); p=0.024**
MEDIUM_2	**0.88 (0.42 to 1.18); p<0.001**	0.57 (-0.99 to 1.41); p=0.37	0.17 (-0.43 to 0.61); p=0.534	**18.13 (6.01 to 29.85); p=0.005**
HIGH_9	0.26 (-0.36 to 1.01); p=0.406	0.33 (-1.64 to 1.46); p=0.688	Reference	Reference
HIGH_10	**0.98 (0.12 to 1.35); p=0.003**	Reference	-0.34 (-0.95 to 0.09); p=0.212	4.61 (-9.32 to 11.33); p=0.394
HIGH_11	**0.70 (0.27 to 1.17); p=0.003**	0.48 (-1.02 to 1.39); p=0.42	0.23 (-0.46 to 0.72); p=0.439	**16.62 (3.97 to 24.67); p=0.002**
HIGH_12	**0.87 (0.33 to 1.37); p=0.002**	0.87 (-0.79 to 2.1); p=0.228	0.49 (-0.19 to 1.02); p=0.107	**16.61 (5.2 to 37.91); p=0.038**
IRNO_1	0.50 (-0.17 to 1.44); p=0.295	0.92 (-0.67 to 3.62); p=0.396	0.68 (-0.53 to 1.68); p=0.238	**17.68 (4.83 to 32.12); p=0.008**
IRNO_2	**1.05 (0.07 to 2.02); p=0.034**	0.59 (-0.96 to 1.87); p=0.419	-0.07 (-0.67 to 0.56); p=0.833	12.09 (-1.03 to 24.5); p=0.079
SABATO	**1.77 (0.37 to 3.13); p=0.016**	1.41 (-0.22 to 2.52); p=0.062	0.09 (-0.56 to 0.59); p=0.753	**19.71 (6.77 to 29.31); p=0.001**

4.3.3 Chemical Activated LUciferase gene eXpression (CALUX) Bioassay

Results of the Calux gene reporter assay, expressed as plasma concentrations depending on the cell line observed, were estimated on the subgroup of 600 subjects. Considering the DR Calux® assay, with Ahrilic (AhR) receptor activation measurement, 597 samples had sufficient amount of plasma for the determinations.

In Figure 18 a Yardstick representation of the results, lined up from lowest to highest value is given, while in Figure 19 a frequency distribution of these data is presented.

A mean value of 44.89 ± 15.74 pg 2,3,67,8-TCDD BEQ/g fat and a median value of 42.19 pg 2,3,7,8-TCDD BEQ/g fat was observed (Fig. 18 and 19, table 20). 90% of the DR CALUX® values were between 24.8 and 73.8 pg 2,3,7,8-TCDD BEQ/g fat, while the 5% highest values ranged from 73.8 to 150.0 pg 2,3,7,8-TCDD BEQ/g fat. The lowest quantifiable result obtained was 6.0 pg 2,3,7,8-TCDD BEQ/g fat, equal to the LOQ of the DR CALUX® method. Two samples had results below the LOQ.

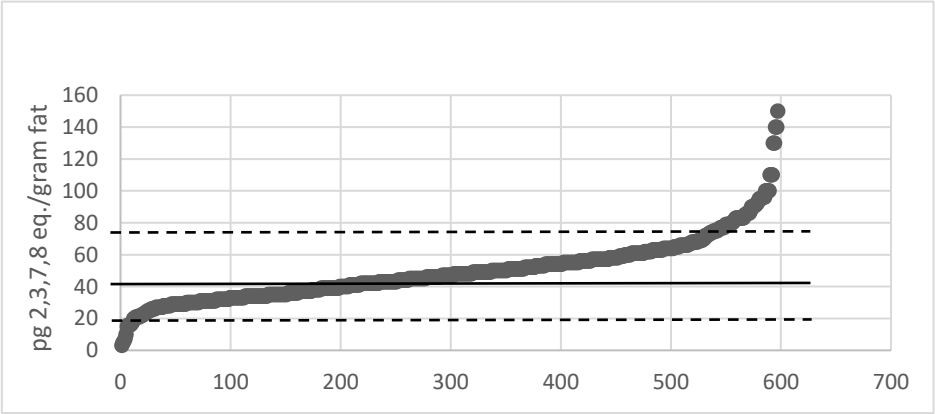


Figure 18. Yardstick-based representation of DR CALUX® results in study sample plasma analyzed. DR CALUX® bioassay results are lined up from lowest to highest obtained result expressed in pg 2,3,7,8-TCDD BEQ/g fat. Back line represents the mean value of 44.89 pg 2,3,7,8-TCDD BEQ/g fat. The top line indicates the 95% percentile and the bottom line the 5% percentile of respectively 73.8 and 24.8 pg 2,3,7,8-TCDD BEQ/g fat.

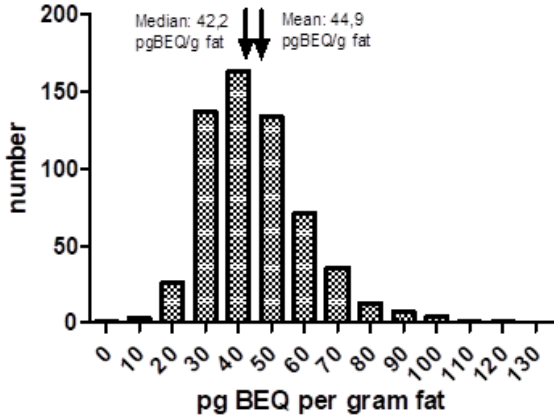


Figure 19. Frequency distribution of DR CALUX® results expressed as pg BEQ/g fat.

In Table 20 a summary of results obtained in plasma, measured by DR CALUX®, is given per impact areas and clusters. Medium impact area is divided in two groups, in order to better highlight differences among clusters.

Table 20. DR CALUX® results (pgBEQ/g fat) compared to areas and cluster of living. In bold are values significantly higher (p-value: >0.001, one-way Anova), as compared to all other clusters; SEM = standard error of mean.

Impact Area	Cluster (grouped)	N. subjects	Mean	Median	SEM	Min.	Max.
Low	16-17	120	41.62	39.66	1.74	12.41	99.51
Medium 1	13-14	118	43.31	41.65	1,58	20.76	88.41
Medium 2	19-21	120	**54.17**	**52.99**	2.82	18.26	100.66
High	9-12	237	42.94	40.07	1.93	36.00	150.0

As can be observed, there is no significant correlation between the area of living and the stable AhR ligands in plasma of volunteers living in those areas. However, there is a significant difference (one-way ANOVA, p=0.0008) between the level of stable AhR ligands in plasma of subjects from clusters 19, 20 and 21 as compared to all other clusters analyzed. These clusters include Irno and Sabato Valleys. Figure 20 reports a box-plot representation of the DR CALUX® results in plasma vs. clusters.

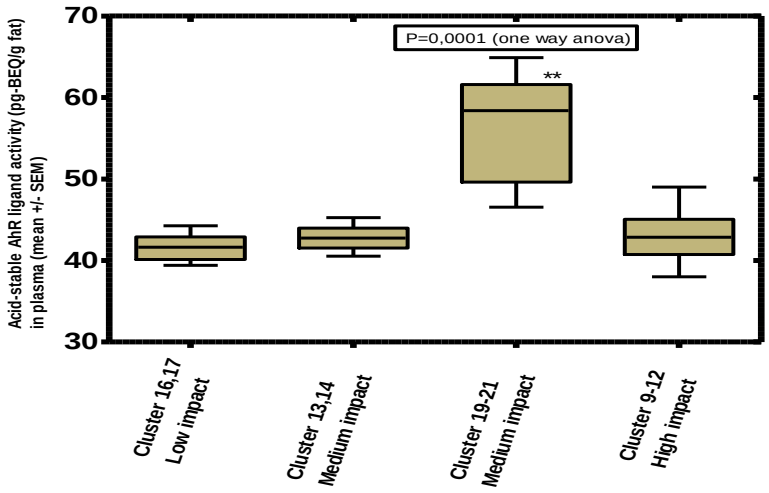


Figure 20. Box-Plot representation of Acid stable AhR ligands measured by DR CALUX® in plasma samples, in relation to clusters and impact areas.

The analysis of plasma samples on the PAH CALUX® bioassay resulted in measurable PAHs and PAH-like responses in all samples analyzed.

A wide range of PAH-like activities were observed in the plasma samples with a mean value of 86.00 ng BaP-EQ/ml, a minimum of 0.3 ng BaP-EQ/ml and a maximum of 7900.00 ng BaP-EQ/ml.

In Table 21 the PAH CALUX® activity in plasma (ngBaP-EQ/ml) is presented per impact areas and clusters. Medium impact area is divided in two groups, in order to better highlight differences among clusters.

As can be observed there is a wide range of PAH activities spanning 3-4 orders of magnitude between the lowest and the highest values per clusters. Overall, there is no relationship between PAH CALUX® levels in plasma and area and cluster of living. Apparently, other factors than the area of living may be more important to explain the internal exposure levels to PAH in ng BaP-EQ/ml as measured with the PAH CALUX® method.

Table 21. PAH CALUX® results (ng BaP-EQ/ml) compared to areas and cluster of living. SEM = standard error of mean.

Impact Area	Cluster (grouped)	N. subjects	Mean	Median	SEM	Min.	Max.
Low	16-17	120	86.00	50.00	13.1	0.4	1400.00
Medium 1	13-14	120	114.00	29.00	65.7	0.3	7900.00
Medium 2	19-21	120	87.00	31.00	27.4	0.1	3200.00
High	9-12	238	58.00	15.00	15.1	0.3	3300.00

The analysis on the ERα CALUX® bioassay resulted in measurable total estrogenic and estrogen-like responses (expressed in pg-EEQ/ml) in all plasma samples analyzed.

In Table 22, the total estrogenic activity data (pg-EEQ/ml) are compared to the main endogenous estrogenic hormone, 17-β-estradiol (17-β) levels measured in plasma of female volunteers from the SPES subgroup selected. Medium impact area is divided in two groups, in order to better highlight differences among clusters.

Table 22. Concentrations of 17-β-estradiol (17-β pg/ml) and total estrogenic activity (pg-EEQ/ml) in female volunteers of the SPES subgroup compared to cluster and area of living. In bold are represented values significantly higher as compared to 17-β value for same cluster (p-value<0.002, Mann-Whitney). SEM = standard error of mean.

Impact Area	Cluster (grouped)	Analyte	N. subjects	Mean	Median	SEM	Min.	Max.
Low	16-17	17-β	60	21.22	21.15	0.77	10.00	53.9
Medium 1	13-14	17-β	60	22.91	21.80	0.93	6.80	51.00
Medium 2	19-21	17-β	57	22.61	22.00	1.51	7.00	70.00
High	9-12	17-β	120	22.28	21.18	0.66	11.00	68.70
Low	16-17	EEQ	60	**45.85**	27.50	6.84	2.10	240.00
Medium 1	13-14	EEQ	60	**94.26** *	59.50	12.30	2.10	510.00
Medium 2	19-21	EEQ	57	**85.44** *	57.00	8.93	8.40	290.00
High	9-12	EEQ	120	**109.39** **	73.00	8.43	2.10	510.00

*Significantly higher as compared to Low Impact Area, Cluster 16-17 (p-value = 0.007, Mann-Whitney).
** significantly higher (p-value= 0.0005, one-way Anova) as compared to all other clusters.

A wide range of total estrogenic activity was observed ranging from 2.10 to 510.00 pg-EEQ/ml. As can be observed there is a significant 3 to 5-fold difference in total estrogenic activity levels measured by the ERα-CALUX® reporter gene assay in comparison to the levels of the natural endogenous hormone 17-β Estradiol. This suggests the presence of endocrine active compounds, that activate the estrogen receptor-alpha in plasma of females, that can to a large extend not be explained by natural endogenous hormone levels and may thus indicate the presence of endocrine active compounds (EACs) in samples.

In addition, the total estrogenic activity levels are significantly higher (p-value = 0.0005, one-way Anova) in clusters of high and medium impact areas (Fig. 21).

Conversely, the plasma levels of 17-β estradiol in females do not differ significantly between volunteers from the various areas of living, suggesting that the additional estrogenic activity may arise from external sources.

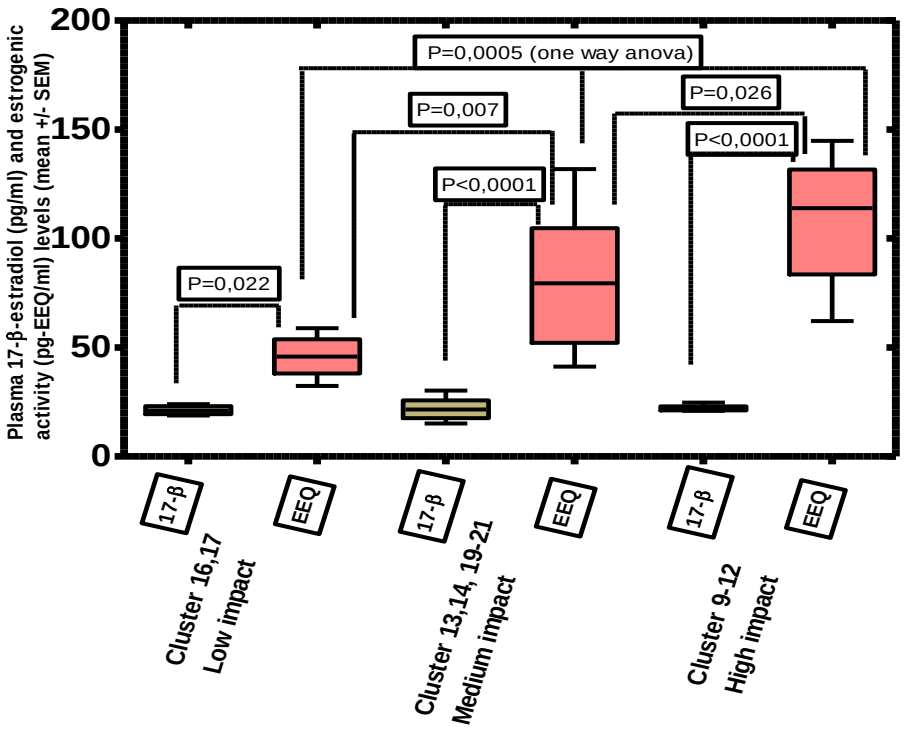


Figure 21. Boxplot of female plasma concentrations of 17-β-estradiol (pg/ml) and total estrogenic activity (pg-EEQ/ml) measured by ERα-CALUX®. Statistical differences were analyzed using one-way Anova for comparisons between clusters, and Mann Whitney test of significance between pairs of boxplots. Level of significance was set at p-value< 0.05 (17-β: 17-β-estradiol; EEQ: estrogen equivalents).

Similarly, it is reported the total estrogenic activity in plasma by the ERα-CALUX® reporter gene assay in male volunteers. These levels have been compared with concentrations of 17-β-estradiol in plasma of the same male individuals (Table 23). The range of total estrogenic activity levels in male plasma is much lower than in female individuals and shows a smaller range of differences i.e., ranging from 1.7 to 150.00 pg-EEQ/ml. Also the differences between the levels of the total estrogenic activity (EEQs) and the natural 17-β estrogen concentrations is much less pronounced, and variable, with a significantly lower EEQ level than 17-β levels in the low impacted area of living and significantly higher EEQ levels than 17-β in the high impacted areas of living (Fig. 22).

Table 23. Concentrations of 17-β-estradiol (17-β pg/ml) and total estrogenic activity (pg-EEQ/ml) in male volunteers of the SPES subgroup compared to cluster and area of living. In bold are represented values significantly lower as compared to 17-β value for same cluster (p-value = 0.002, Mann-Whitney).
SEM = standard error of mean.

Category	Cluster	Analyte	N	Mean	Median	SEM	Min	Max
Low	16-17	17-β	60	19.59	18.35	0.96	10.00	52.00
Medium 1	13,14	17-β	60	19.23	18.20	0.88	10.00	51.00
Medium 2	19-21	17-β	60	20.35	18.15	1.14	7.20	44.00
High	9-12	17-β	127	18.24	18.13	0.36	11.10	40.00
Low	16-17	EEQ	60	**12.72**	15.30	1.51	1.70	37.00
Medium 1	13-14	EEQ	60	24.25*	25.00	1.49	6.80	61.00
Medium 2	19-21	EEQ	60	21.62*	22.50	2.05	6.10	110.00
High	9-12	EEQ	117	27.88**, ***	25.00	2.30	5.00	150.00

* significantly higher as compared to Low Impact Area, Cluster 16-17 (P= 0.0072, Mann-Whitney).
** significantly higher as compared to 17-β value for same cluster (P<0.0001);
*** significantly higher (P-value: 0,0005, one-way Anova) as compared to all other clusters.

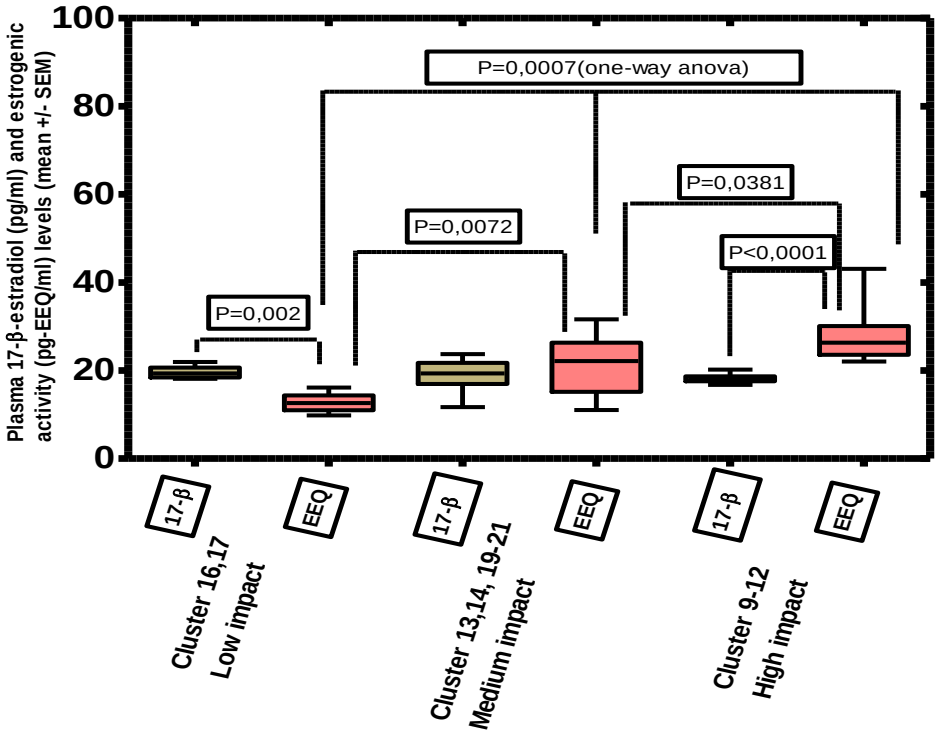


Figure 22. Boxplot of male plasma concentrations of 17-β-estradiol (pg/ml) and total estrogenic activity (pg-EEQ/ml) measured by ERα-CALUX®. Statistical differences were analyzed using one-way Anova for comparisons between clusters, and Mann Whitney test of significance between pairs of boxplots. Level of significance was set at p-value< 0.05 (17-β: 17-β-estradiol; EEQ: estrogen equivalents).

Therefore, also in male subjects the total estrogenic activity levels are significantly higher (p-value = 0.0007, one-way Anova) in clusters of high and medium impact areas (Fig. 22). The plasma levels of the natural hormone, 17-β-estradiol, in males do not differ significantly between volunteers from the diverse areas of living, suggesting that the additional estrogenic activity observed in the high impacted area may arise from external sources. Overall, it is observed a significant higher level of estrogenic activity in plasma from human volunteers as compared to natural endogenous 17-β-estradiol levels, which increased with higher impacted areas of living. This may suggest that the contamination areas give rise to elevated exposure to contaminants with possible endocrine activity, so called Endocrine-disrupting chemicals (EDCs).

Figure 23 shows the geographical distribution per area of living of the EEQ measured by ERα-CALUX® assay on female cohort. Higher values are detected in women from high and medium impact areas. Notably, as previously shown by table22 and figure 21, among medium impact areas, the higher values are geolocated in Irno and Sabato valleys.

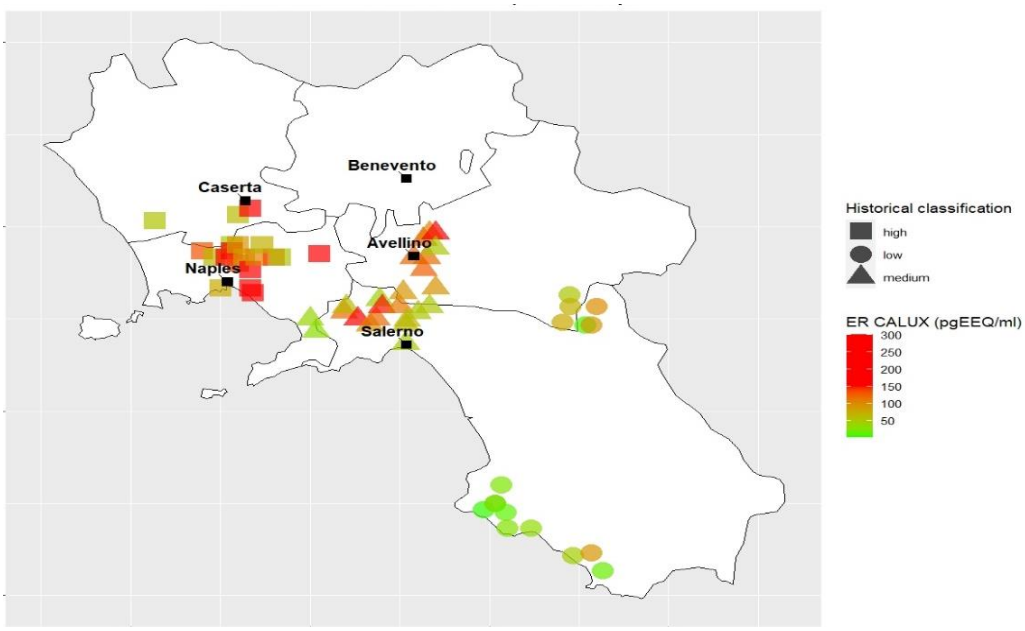


Figure 23. Campania Region maps representing hotspot areas of living for female subjects at municipal level, with their respective total estrogenic activity detected by ERα-CALUX® reporter gene assay (pg EEQ/ml) in plasma.

4.4 Effect Biomarkers

4.4.1 DNA methylation Array Preliminary Analysis

Methylation array data analysis was firstly performed through a quality control step, secondly with a global overview on methylation differences among the three impact areas and, finally, by means of a regression model method to explore the correlation between the exposure biomarkers assessed and the epigenetic profile.

As mentioned in paragraph 3.3.2, DNA methylation array analysis consists of a preliminary control quality check, with following filtering and normalization steps.

Samples with a lot of missing probes and probes with poor quality were removed from the analysis. Finally, 592 of 600 samples and 677886 of 865918 probes were included in the analysis.

In order to have an overview of the dataset, the most variable 1000 probes were graphically represented in a multidimensional plot (MDS-plot) before and after normalization step (Fig. 24 a/b), as well as all probes raw data were represented in a density plot (Fig. 25 a/b).

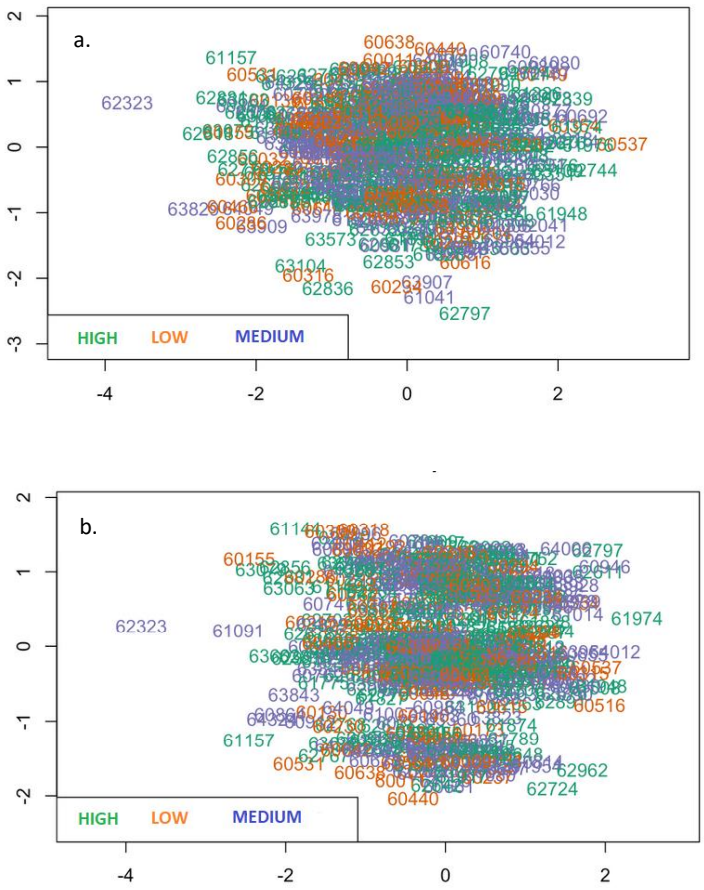


Figure 24. 1000 most variable probes represented in a multidimensional plot (MDS), before (a) and after (b) the normalization step. The 1000 probes were selected computing variance in the beta value matrix and sorting them in descending order. The figures were obtained by means a multidimensional scaling approach.

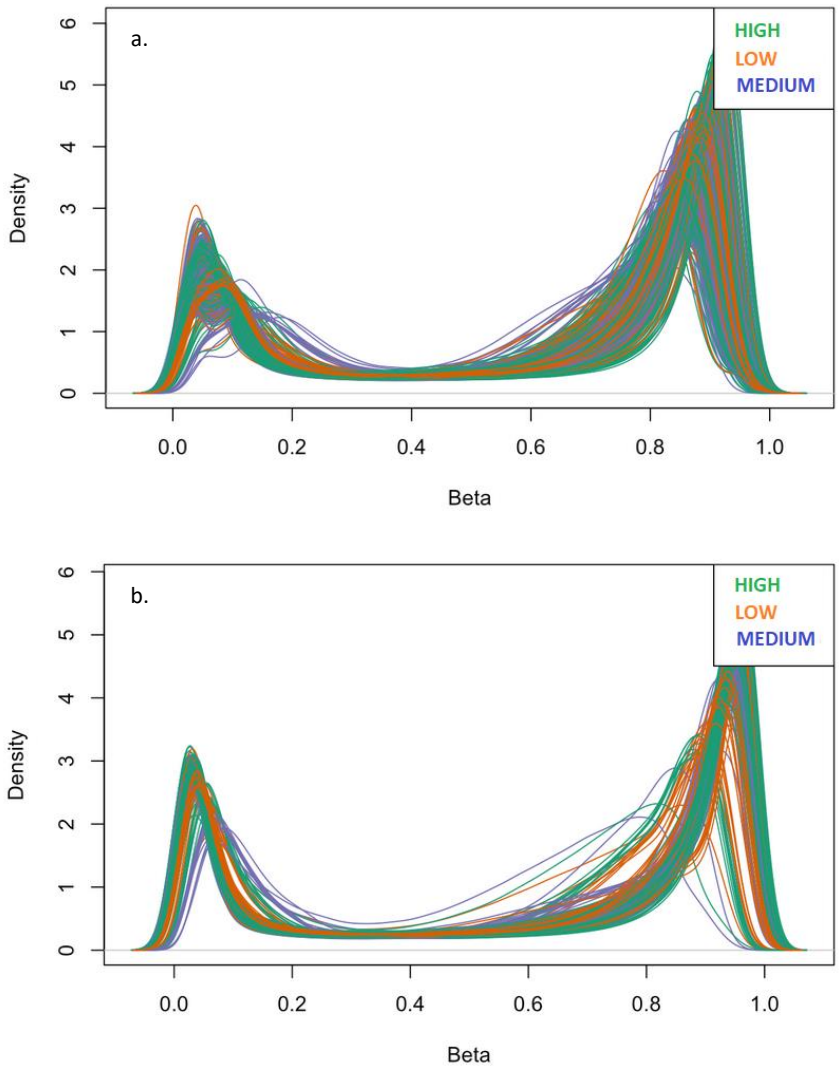


Figure 25. Density plot of raw data (677886 probes) before (a) and after (b) the normalization step.

It can be observed a better probe distribution after the normalization step, even though it is not completely satisfying. A total number of 188032 probes were removed through the normalization step. Profiles are consistent with reference methylation dataset distribution of beta values between 0.0 and 1.0. The following "cell heterogeneity" step improves the probes distribution from a qualitative point of view.

Then a Singular Value Decomposition Analysis (SVD) was performed in order to allow the batch correction (Fig. 26 a/b). Among the most important confounding variables it can be highlighted Run_ID, Slide, Cluster, Age and Gender. Run_ID and Slide are linked to the analytical process, while Cluster, Age and Gender are linked to the population features.

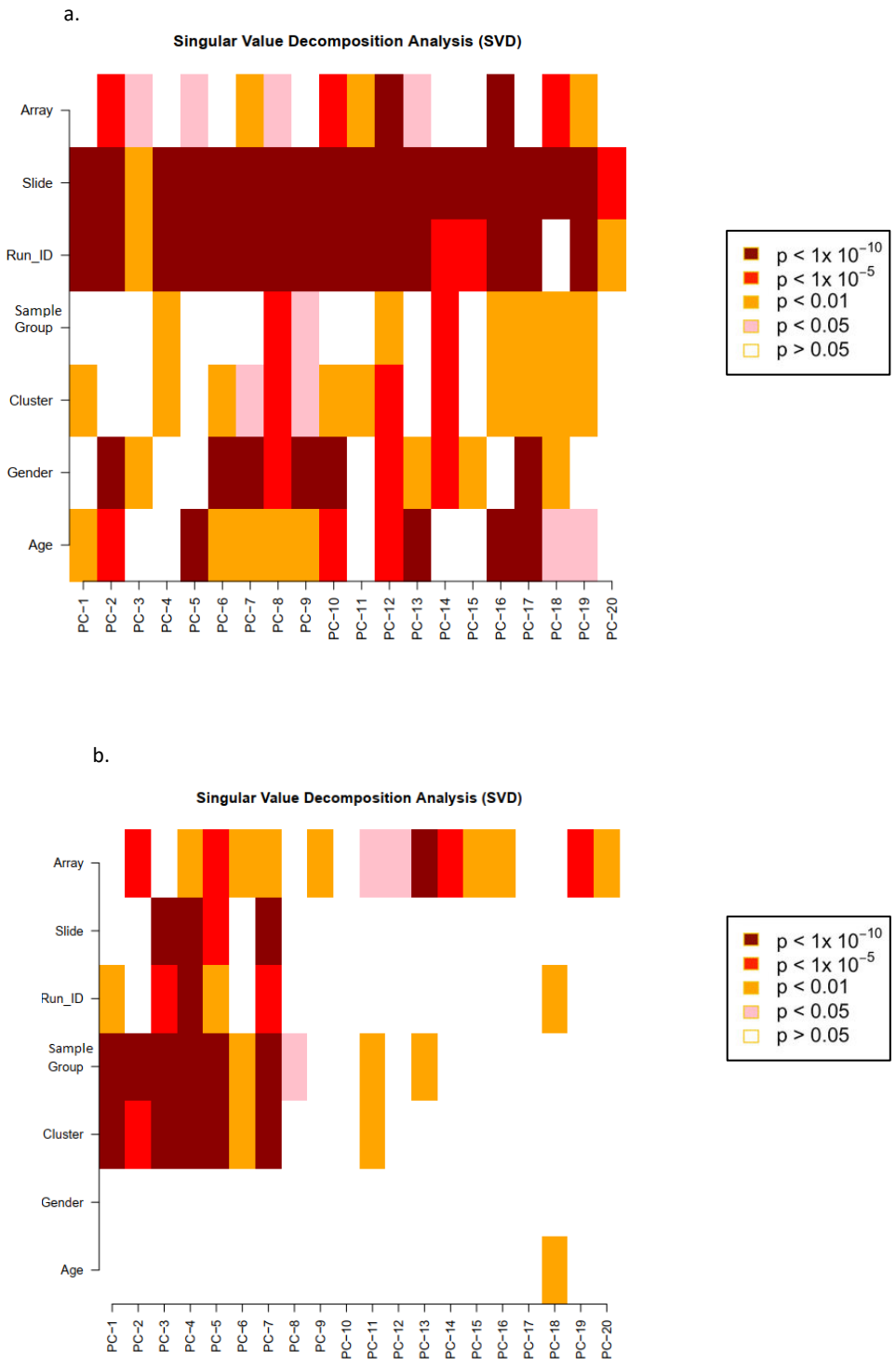


Figure 26. Singular Value Decomposition Analysis (SVD) before (a) and after (b) the batch correction considering Slide, Age and Gender as factors to correct.

Considering the nature of samples (genomic DNA from whole blood), it becomes fundamental a step of cell heterogeneity correction (Houseman et al., 2012).

In the Figure 27 a/b, the Multidimensional Plot on the most variables 1000 probes and the Density Plot on the whole dataset after batch and cell heterogeneity corrections, show a notably improvement of distribution from a qualitative point of view.

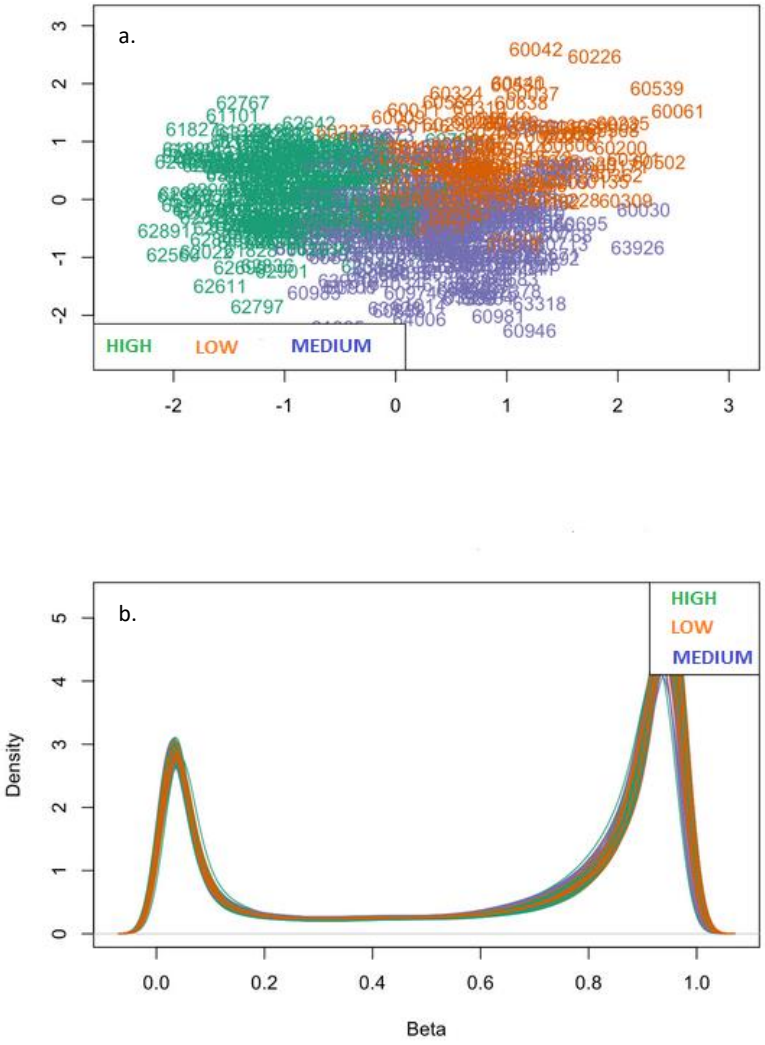


Figure 27. (a) 1000 most variable probes represented in a multidimensional plot (MDS) after normalization, batch and cell heterogeneity correction steps. (b) Density plot of raw data (677886 probes) after normalization, batch and cell heterogeneity correction steps.

4.4.2 Differentially Methylated Probes and Differentially Methylated Regions

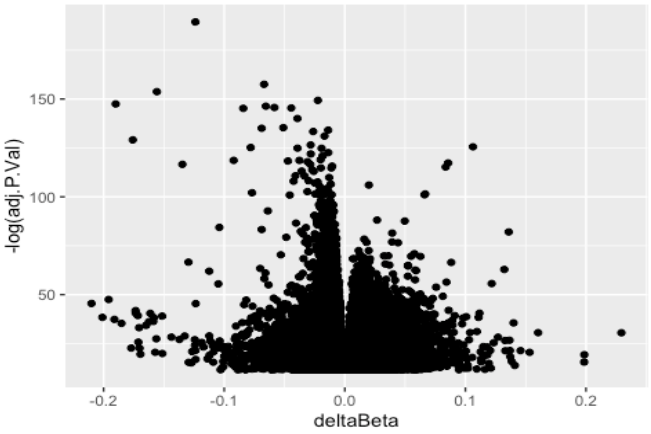
Differentially Methylated Probes (DMPs) were identified in the following comparisons, among impact areas:

1. High impact area vs Low impact area;
2. Medium impact area vs High impact area;
3. Medium impact area vs Low impact area.

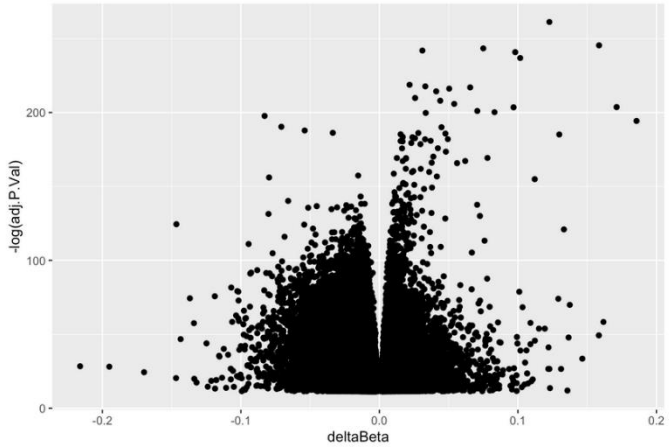
The following Volcano plots (Fig. 28 a/b/c) show an overview of the global methylation, representing on the x the delta beta values (Δβ) between the two impact areas observed in each

comparison, and on the y the statistical significance. A strict cut-off was selected for adjusted p-value (0.00001) for the selection of the probes.

a. High vs Low



b. Medium vs High



c. Medium vs Low

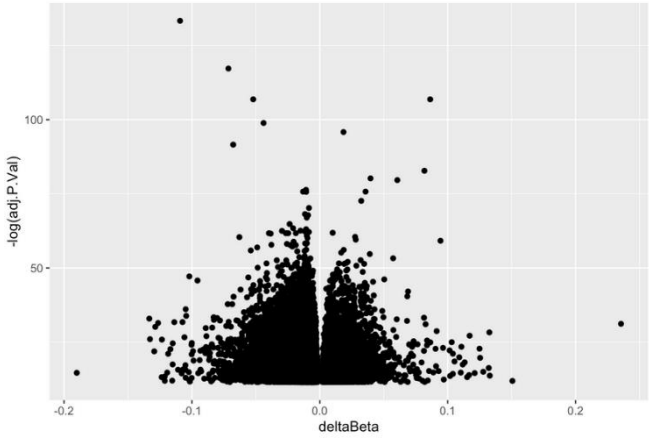


Figure 28. Volcano plot for DMP comparison among impact areas (a) high vs low; Δβ>0 means higher value for high impact area; (b) medium vs high; Δβ>0 means higher value for medium impact area; (c) medium vs low; Δβ>0 means higher value for medium impact area.

Volcano plot 28a shows relevant statistical significance of global hypomethylation of high impact area compared to the low one.

Less visible difference on global methylation is shown in Fig. 28b and 28c: slight hypomethylation for high impact area vs medium impact area, slight global hypomethylation for medium impact area vs low impact area.

Frequently the delta beta values are not very high, and this means a difficulty in biological understanding of the phenomena. For the purposes of this dissertation the focus is particularly on DMPs mapping on promoters' regions.

All analysis were performed considering the M-values (the logit transformation of the beta-values), even if the results are shown in tables as β-values, in order to have an easier interpretation.

Tables below report DMPs filtered from impact areas comparisons, considering a $\Delta\beta \geq 0.10$.

Table 24a shows DMPs positioned on proximal promoter region (5'UTR, TSS200, TSS1500, 1stExon), comparing low vs high impact areas.

Table 24b report the same object referring to comparison between medium vs low impact areas. For both tables, hypermethylated CpGs in areas of increasing pollution risk were selected. Interestingly, among DMPs in table 24a, mapping on promotor regions in low vs high comparison, it is possible to list the following genes involved:

- RHOJ, a GTP binding protein, regulating angiogenesis and involved in insulin signaling pathway;
- MYCBPAP, encoding for a MYCBP associated protein, which is concern for spermatogenesis;
- WBSCR26, encoding for a lncRNA ABHD11-AS1, which play a role in gastric and bladder cancer;
- LOXL3, encoding for a Lysil Oxidase Like 3, which has a role in copper ion binding, oxidoreductase activity and regulation of inflammatory response;
- CLDN14, playing a role in cell-to-cell adhesion in epithelial or endothelial cell sheets;
- CLEC4C, implied in cell adhesion, cell-to-cell signaling, inflammation and immune responses;
- COQ5, with its methyltransferase activity;
- COX7A1, encoding for Cytochrome C Oxidase subunit 7A1;
- FAT1, encoding for adhesion molecule and signaling receptor, with a role in development processes and cell communication;
- PTPRE, encoding for a tyrosine phosphatase protein family member, which has a role in signaling regulating cell growth, differentiation, mitotic cycle and oncogenic transformation.

Considering table 24b, reporting the medium vs low impact area comparison, beyond RHOJ and LOXL3 genes, in common with the results of high vs low comparison, it is possible to count:

- NRIP1 gene, encoding a nuclear protein that modulate the transcriptional activity of the estrogen receptor, and has also a role in inflammation pathway lead by NF-kB;
- S100A13 gene, which encodes for a S100 family protein, involved in cell cycle progression and differentiation. Its expression was documented by Azimi et al., in 2014, in chemotherapy resistance in melanoma.

Table 24a. High impact areas vs low impact area. Differential Methylated Probes with delta beta value ≥ 0.10, positioned on promoter region. Statistical significance is expressed as FDR (False Discovery Rate) adjusted p-value (adj.P.Val). $\Delta\beta > 0$ means the methylation is higher in high impact areas.

CpG	adj.P.Val	deltaBeta	gene	feature	cgi	UCSC_Islands_Name
cg11775521	1.13E-51	0.0857	KCNE3	TSS200	island	chr11:74178175-74178801
cg18297203	1.34E-44	0.0661	LPCAT2	TSS200	shore	chr16:55542920-55543274
cg20094837	6.36E-31	0.0627	NBEA	TSS1500	opensea	
cg08370546	3.20E-25	0.0843	PHKG1	TSS200	opensea	
cg01216565	4.55E-22	0.0813	WDR45B	TSS1500	shore	chr17:80605301-80606324
cg23980942	1.05E-21	0.0606	SEMA6D	5'UTR	shore	chr15:47476369-47477499
cg20833838	2.25E-21	0.0649	GPRIN1	5'UTR	shore	chr5:176036458-176037557
cg25432232	5.89E-19	0.0861	AURKC	5'UTR	island	chr19:57741955-57742457
cg01668281	7.77E-19	0.0605	CLDN14	5'UTR	opensea	
cg09340198	3.73E-18	0.0677	ANKRD28	TSS1500	shore	chr3:15900497-15901990
cg21695771	3.85E-18	0.0735	COX7A1	TSS200	shore	chr19:36642327-36643585
cg12569497	7.14E-18	0.06	COL21A1	TSS1500	shore	chr6:56112188-56112444
cg11659652	1.51E-17	0.0708	PTPRE	5'UTR	opensea	
cg16443366	5.89E-17	0.0617	SOGA3	5'UTR	shore	chr6:127836730-127837706
cg09548084	3.20E-16	0.06	SLC35B3	TSS1500	shore	chr6:8435314-8436141
cg02176678	8.71E-16	0.0814	TTLL4	5'UTR	shore	chr2:219575448-219576080
cg19603903	1.12E-15	0.0893	AURKC	TSS200	island	chr19:57741955-57742457
cg15176213	1.20E-15	0.0673	COX7A1	5'UTR	island	chr19:36642327-36643585
cg05524354	1.94E-15	0.071	PTPRE	5'UTR	opensea	
cg06643849	3.26E-15	0.085	AURKC	5'UTR	island	chr19:57741955-57742457
cg05275307	1.27E-14	0.064	CLDN14	TSS200	opensea	
cg19052272	3.00E-14	0.1095	ALLC	TSS1500	opensea	
cg24824686	2.23E-13	0.0615	PTPRE	5'UTR	opensea	
cg00161556	2.38E-13	0.0711	ZDHHC13	TSS1500	shore	chr11:19138497-19139216
cg25932599	3.48E-13	0.0669	FGFRL1	TSS1500	island	chr4:1003106-1005455
cg08161306	5.39E-13	0.127	TBC1D22A	TSS1500	opensea	
cg00255882	8.73E-13	0.0729	WDR33	TSS1500	shore	chr2:128568321-128569064
cg19825600	1.04E-12	0.0681	ALLC	TSS1500	opensea	
cg00999904	2.67E-12	0.0768	ALLC	TSS1500	opensea	
cg21183455	2.71E-12	0.0626	C22orf34	TSS1500	opensea	
cg24335895	3.86E-12	0.0693	COX7A1	TSS200	shore	chr19:36642327-36643585
cg14409029	4.17E-12	0.0837	ZDHHC13	TSS1500	shore	chr11:19138497-19139216
cg06888460	7.22E-12	0.0693	LDHAL6A	TSS200	island	chr11:18477298-18477756
cg23786580	8.73E-12	0.0723	HECW1	5'UTR	opensea	
cg26740318	9.84E-12	0.0608	LDHAL6A	TSS1500	shore	chr11:18477298-18477756
cg16168363	1.44E-11	0.0609	FAT1	5'UTR	shore	chr4:187644319-187648253
cg07889765	1.47E-11	0.0864	NOC3L	TSS1500	shore	chr10:96122540-96122904
cg21517946	2.67E-11	0.0985	FKBP5	TSS1500	shore	chr6:35699166-35699953
cg24527560	3.28E-11	0.0915	LOC441666	TSS200	shore	chr10:42862953-42863216
cg25605888	4.06E-11	0.0618	COQ5	TSS1500	shore	chr12:120966788-120967038
cg21042336	7.07E-11	0.0794	OSBPL1A	TSS1500	shore	chr18:21977275-21978110
cg17240725	7.94E-11	0.118	WBSCR26	TSS200	shelf	chr7:73152564-73153716

cg01485177	3.12E-10	0.0611	LOC441666	TSS200	shore	chr10:42862953-42863216
cg04413036	3.37E-10	0.0626	PLA2G6	5'UTR	opensea	
cg07157030	5.19E-10	0.1457	RHOJ	5'UTR	opensea	
cg19148731	9.85E-10	0.1126	LOXL3	5'UTR	shore	chr2:74781494-74782685
cg14259717	9.95E-10	0.0841	LDHC	TSS200	opensea	
cg12759523	1.14E-09	0.0707	CYFIP1	5'UTR	opensea	
cg07189587	1.24E-09	0.1532	RHOJ	5'UTR	opensea	
cg02511570	1.30E-09	0.0788	THCAT158	TSS1500	shore	chr17:45400874-45401440
cg10645314	2.26E-09	0.0701	ALLC	TSS1500	opensea	
cg20111217	2.27E-09	0.0992	MYCBPAP	TSS1500	shore	chr17:48585385-48586167
cg09408571	3.36E-09	0.0916	GPR88	TSS200	shore	chr1:101004471-101005885
cg00901687	4.44E-09	0.1182	MYCBPAP	TSS1500	shore	chr17:48585385-48586167
cg19272349	4.71E-09	0.0793	CSNK1A1L	TSS1500	opensea	
cg21471707	1.26E-08	0.068	LDHC	TSS200	opensea	
cg11440486	1.62E-08	0.0902	MYCBPAP	TSS1500	shore	chr17:48585385-48586167
cg25251562	2.44E-08	0.076	ALLC	TSS1500	opensea	
cg03343083	6.77E-08	0.0888	LEPREL1	5'UTR	opensea	
cg11474464	7.46E-08	0.06	GABRB3	TSS200	island	chr15:26874097-26874528
cg21367232	2.05E-07	0.064	SCAMP2	TSS1500	shore	chr15:75165373-75165903
cg19149522	2.09E-07	0.0609	ZDHHC4	TSS1500	shore	chr7:6616918-6617277
cg00614959	3.08E-07	0.0681	PCGF3	5'UTR	shelf	chr4:724335-726070
cg11079896	3.43E-07	0.0912	RHOJ	5'UTR	opensea	
cg24480555	8.18E-07	0.0656	C8orf33	TSS1500	shore	chr8:146277801-146278396
cg07195891	2.48E-06	0.0981	CLEC4C	5'UTR	opensea	
cg24521141	2.91E-06	0.0791	LOC101929596	TSS1500	opensea	

Table 24b. Medium impact areas vs low impact area. Differential Methylated Probes with delta beta value ≥ 0.10, positioned on promoter region. Statistical significance is expressed as FDR (False Discovery Rate) adjusted p-value (adj.P.Val). $\Delta\beta > 0$ means the methylation is higher in medium impact areas.

CpG	adj.P.Val	deltaBeta	gene	feature	cgi	UCSC_Islands_Name
cg08370546	2.02E-26	0.0944	PHKG1	TSS200	opensea	
cg01708150	3.46E-13	0.0914	RHOJ	5'UTR	opensea	
cg11079896	5.28E-13	0.1326	RHOJ	5'UTR	opensea	
cg16102102	1.43E-11	0.1032	HAPLN1	TSS1500	shore	chr5:83017894-83018358
cg04526399	2.93E-11	0.0765	HAPLN1	TSS1500	shore	chr5:83017894-83018358
cg16287373	6.85E-11	0.0723	NRIP1	5'UTR	opensea	
cg00318109	6.89E-10	0.1041	MYH7	5'UTR	opensea	
cg16406611	8.01E-09	0.0688	IFITM5	TSS1500	shore	chr11:299389-299635
cg19148731	9.59E-09	0.1054	LOXL3	5'UTR	shore	chr2:74781494-74782685
cg01347250	1.20E-08	0.0624	S100A13	TSS200	opensea	
cg24443494	4.74E-08	0.0909	CHST15	5'UTR	opensea	
cg07189587	3.11E-07	0.1279	RHOJ	5'UTR	opensea	
cg07157030	9.40E-07	0.1167	RHOJ	5'UTR	opensea	
cg18008345	1.89E-06	0.1153	POLR2E	TSS1500	shore	chr19:1094797-1095649
cg08296601	3.01E-06	0.0777	MRPL28	5'UTR	island	chr16:419860-420256
cg22110428	4.01E-06	0.0655	CEACAM18	TSS1500	opensea	
cg17422692	6.31E-06	0.0804	MRPL28	5'UTR	island	chr16:419860-420256

The analysis was carried out through the identification of the Differentially Methylated Regions (DMRs), with the same comparisons among impact areas used for DMPs:

1. High impact area vs Low impact area;
2. Medium impact area vs High impact area;
3. Medium impact area vs Low impact area.

The DMRs identification was performed by means of two independent methods (Bumpunther and DMRcate, implemented on R software), then only the "consensus DMRs", reached out with both two methods, were selected: 392 consensus DMRs were identified for high impact vs low impact comparison, 1082 for medium vs high, 537 for medium vs low.

In order to understand the biological implication given by the differential methylation levels for each impact areas, a gene set enrichment analysis was performed for both DMPs and DMRs. This analysis consists of a Gene Ontology (GO) by "gometh" method and a Pathway Analysis by "KEGG" method (Adjusted P-value for GO was set at 0.001, for KEGG at 0.05).

Table 25 is an extract of the GO annotated terms, showing the most significative molecular functions come out from each impact area comparison for DMPs. Among the most enriched: "metal ion binding", "heterocyclic compound binding", and "organic cyclic compound binding" were counted.

Otherwise, in Table 26, among biological processes enriched in the GO, it is possible to find "nitrogen compound metabolic process", "organonitrogen compound metabolic process", "cell development" and " nervous system development".

Figure 29 a/b/c and 30 a/b/c graphically represents GO most enriched terms for each impact area comparison for molecular function and biological process, respectively.

Fig. 29d and Fig. 30d show the comparing bar plots of common terms and number of DMPs genes involved.

Table 25. Gene Ontology for DMPs and most significant Molecular Function (MF) reported for each impact area comparison (High vs Low; Medium vs Low; Medium vs High). DE=number of genes from DMPs annotated; FDR=statistical significance of the term.

High vs Low				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0005515	MF	protein binding	11033	2.64E-115
GO:0005488	MF	binding	12817	5.66E-94
GO:0043167	MF	ion binding	5154	2.26E-46
GO:0003824	MF	catalytic activity	4641	5.20E-46
GO:0019899	MF	enzyme binding	1985	8.51E-39
GO:1901363	MF	heterocyclic compound binding	4941	2.23E-36
GO:0097159	MF	organic cyclic compound binding	5007	4.81E-36
GO:0043168	MF	anion binding	2393	6.76E-29
GO:1901265	MF	nucleoside phosphate binding	1826	3.75E-28
GO:0000166	MF	nucleotide binding	1825	4.11E-28
GO:0016740	MF	transferase activity	1886	1.20E-25
GO:0043169	MF	cation binding	3503	2.06E-25
GO:0046872	MF	metal ion binding	3439	2.69E-25
GO:0017076	MF	purine nucleotide binding	1621	1.64E-24
GO:0032555	MF	purine ribonucleotide binding	1610	1.64E-24
GO:0032553	MF	ribonucleotide binding	1624	2.22E-24
GO:0035639	MF	purine ribonucleoside triphosphate binding	1554	2.98E-24
GO:0140096	MF	catalytic activity, acting on a protein	1828	6.86E-23
Medium vs Low				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0005515	MF	protein binding	10657	1.15E-46
GO:0005488	MF	binding	12442	3.31E-38
GO:0043167	MF	ion binding	5000	1.84E-22
GO:0003824	MF	catalytic activity	4482	3.14E-20
GO:0043168	MF	anion binding	2331	2.32E-17
GO:0008092	MF	cytoskeletal protein binding	852	2.34E-14
GO:0019899	MF	enzyme binding	1891	2.36E-14
GO:0036094	MF	small molecule binding	2073	2.63E-14
GO:1901265	MF	nucleoside phosphate binding	1762	6.03E-14
GO:0000166	MF	nucleotide binding	1761	6.42E-14
GO:0005524	MF	ATP binding	1248	7.49E-14
GO:0140096	MF	catalytic activity, acting on a protein	1780	7.49E-14
GO:0097367	MF	carbohydrate derivative binding	1846	1.83E-13
GO:0043169	MF	cation binding	3408	2.85E-13
GO:0046872	MF	metal ion binding	3344	5.00E-13
GO:0030554	MF	adenyl nucleotide binding	1302	5.50E-13
GO:0032559	MF	adenyl ribonucleotide binding	1292	5.71E-13
GO:0035639	MF	purine ribonucleoside triphosphate binding	1503	8.16E-13
GO:0032555	MF	purine ribonucleotide binding	1555	2.01E-12

Medium vs High				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0005488	MF	binding	14711	6.19E-213
GO:0005515	MF	protein binding	12497	8.44E-215
GO:0043167	MF	ion binding	5791	2.32E-69
GO:0097159	MF	organic cyclic compound binding	5676	1.53E-42
GO:1901363	MF	heterocyclic compound binding	5593	8.61E-41
GO:0003824	MF	catalytic activity	5215	8.32E-73
GO:0043169	MF	cation binding	3964	1.05E-41
GO:0046872	MF	metal ion binding	3889	1.52E-40
GO:0003676	MF	nucleic acid binding	3811	6.71E-13
GO:0098772	MF	molecular function regulator	3444	8.37E-29
GO:0043168	MF	anion binding	2640	1.09E-34
GO:0036094	MF	small molecule binding	2361	5.05E-32
GO:0003677	MF	DNA binding	2269	1.70E-13
GO:0019899	MF	enzyme binding	2139	1.84E-29
GO:0097367	MF	carbohydrate derivative binding	2091	3.39E-18
GO:0016740	MF	transferase activity	2078	1.07E-28
GO:0140096	MF	catalytic activity, acting on a protein	2021	3.26E-30
GO:1901265	MF	nucleoside phosphate binding	1992	4.77E-28
GO:0140110	MF	transcription regulator activity	1824	7.05E-13

Table 26. Gene Ontology for DMPs and most significant Biologic Process (BP) reported for each impact area comparison (High vs Low; Medium vs Low; Medium vs High). DE=number of genes from DMPs annotated; FDR=statistical significance of the term.

High vs Low				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0006996	BP	organelle organization	3445	9.06E-61
GO:0044237	BP	cellular metabolic process	8880	9.07E-57
GO:0048522	BP	positive regulation of cellular process	4708	1.25E-52
GO:0050794	BP	regulation of cellular process	8839	1.19E-50
GO:1901576	BP	organic substance biosynthetic process	5051	2.44E-50
GO:0044267	BP	cellular protein metabolic process	4400	3.95E-50
GO:0043412	BP	macromolecule modification	3696	4.56E-50
GO:0006464	BP	cellular protein modification process	3520	7.09E-50
GO:0036211	BP	protein modification process	3520	7.09E-50
GO:0031323	BP	regulation of cellular metabolic process	5195	7.92E-50
GO:1901564	BP	organonitrogen compound metabolic process	5566	1.44E-49
GO:0006807	BP	nitrogen compound metabolic process	8244	2.65E-48
GO:0048518	BP	positive regulation of biological process	5133	1.60E-46
GO:0019538	BP	protein metabolic process	4808	1.88E-46
GO:0009059	BP	macromolecule biosynthetic process	4158	6.74E-46
GO:0051171	BP	regulation of nitrogen compound metabolic process	4852	1.49E-45
GO:0080090	BP	regulation of primary metabolic process	5006	3.42E-45
GO:0034645	BP	cellular macromolecule biosynthetic process	4057	2.20E-44

GO:0044085	BP	cellular component biogenesis	2738	1.06E-43
GO:0007399	BP	nervous system development	2141	7.97E-43
Medium vs Low				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0016043	BP	cellular component organization	5214	4.28E-33
GO:0071840	BP	cellular component organization or biogenesis	5356	7.17E-32
GO:0032502	BP	developmental process	5277	8.76E-24
GO:0051179	BP	localization	5399	1.62E-23
GO:1901564	BP	organonitrogen compound metabolic process	5389	1.84E-22
GO:0048856	BP	anatomical structure development	4891	3.74E-22
GO:0048518	BP	positive regulation of biological process	4970	5.64E-21
GO:0051128	BP	regulation of cellular component organization	2045	7.35E-21
GO:0048468	BP	cell development	1877	2.06E-20
GO:0048523	BP	negative regulation of cellular process	4078	5.42E-20
GO:0120036	BP	plasma membrane bounded cell projection organization	1334	1.09E-19
GO:0048522	BP	positive regulation of cellular process	4524	1.09E-19
GO:0044260	BP	cellular macromolecule metabolic process	6518	1.35E-19
GO:0030030	BP	cell projection organization	1368	2.13E-19
GO:0007275	BP	multicellular organism development	4495	2.79E-19
GO:0019538	BP	protein metabolic process	4638	3.34E-19
GO:0051234	BP	establishment of localization	4206	6.23E-19
GO:0009653	BP	anatomical structure morphogenesis	2349	1.08E-18
GO:0000902	BP	cell morphogenesis	916	2.16E-18
GO:0007399	BP	nervous system development	2053	2.16E-18
Medium vs High				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0071840	BP	cellular component organization or biogenesis	6170	7.01E-100
GO:0016043	BP	cellular component organization	5989	9.49E-97
GO:0009987	BP	cellular process	15580	5.87E-96
GO:0044260	BP	cellular macromolecule metabolic process	7660	1.26E-88
GO:0050794	BP	regulation of cellular process	10131	4.10E-81
GO:1901564	BP	organonitrogen compound metabolic process	6287	2.87E-75
GO:0065007	BP	biological regulation	11442	2.47E-69
GO:0044237	BP	cellular metabolic process	10151	3.15E-69
GO:0051179	BP	localization	6254	6.77E-68
GO:0032502	BP	developmental process	6111	2.24E-65
GO:0044238	BP	primary metabolic process	9898	3.69E-64
GO:0048518	BP	positive regulation of biological process	5756	7.13E-61
GO:0006807	BP	nitrogen compound metabolic process	9441	1.22E-60
GO:0050789	BP	regulation of biological process	10843	6.15E-60
GO:0048522	BP	positive regulation of cellular process	5232	3.44E-59
GO:0019538	BP	protein metabolic process	5402	8.02E-59
GO:0048856	BP	anatomical structure development	5657	1.97E-58
GO:1901576	BP	organic substance biosynthetic process	5657	1.06E-55
GO:0009058	BP	biosynthetic process	5734	1.87E-55
GO:0006996	BP	organelle organization	3758	2.36E-54

Figure 29. GO most enriched terms for Molecular Function in (a) high vs low area, (b) medium vs low area, (c) medium vs high area. On the x the number of genes from DMPs annotated. (d) Common GO terms and number of DMPs annotated genes for each area comparison.

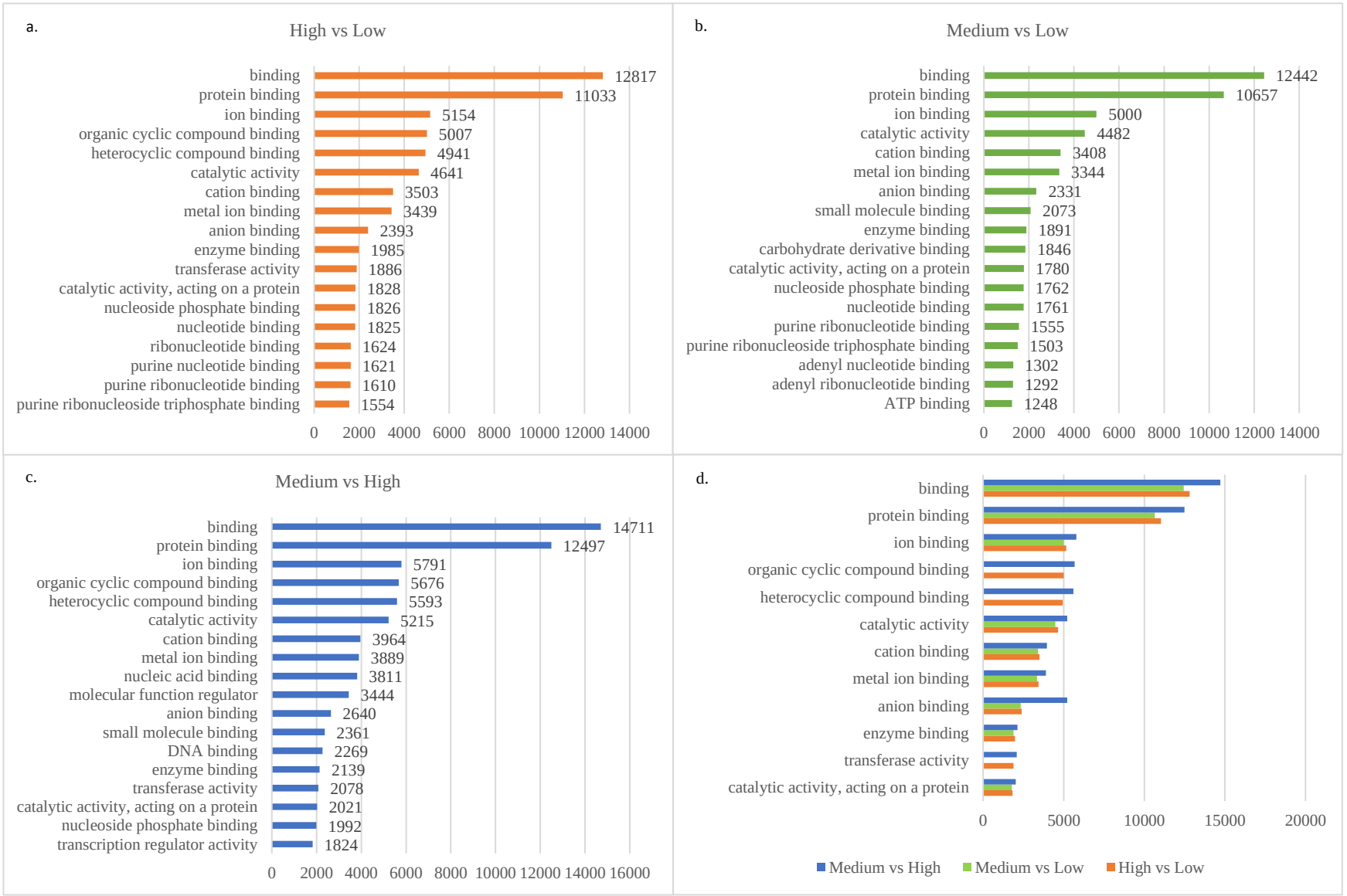


Figure 30. GO most enriched terms for Biological Process in (a) low vs high areas, (b) medium vs low area, (c) medium vs high area. On the x the number of genes from DMPs annotated. (d) Common GO terms and number of DMPs annotated genes for each area comparison.

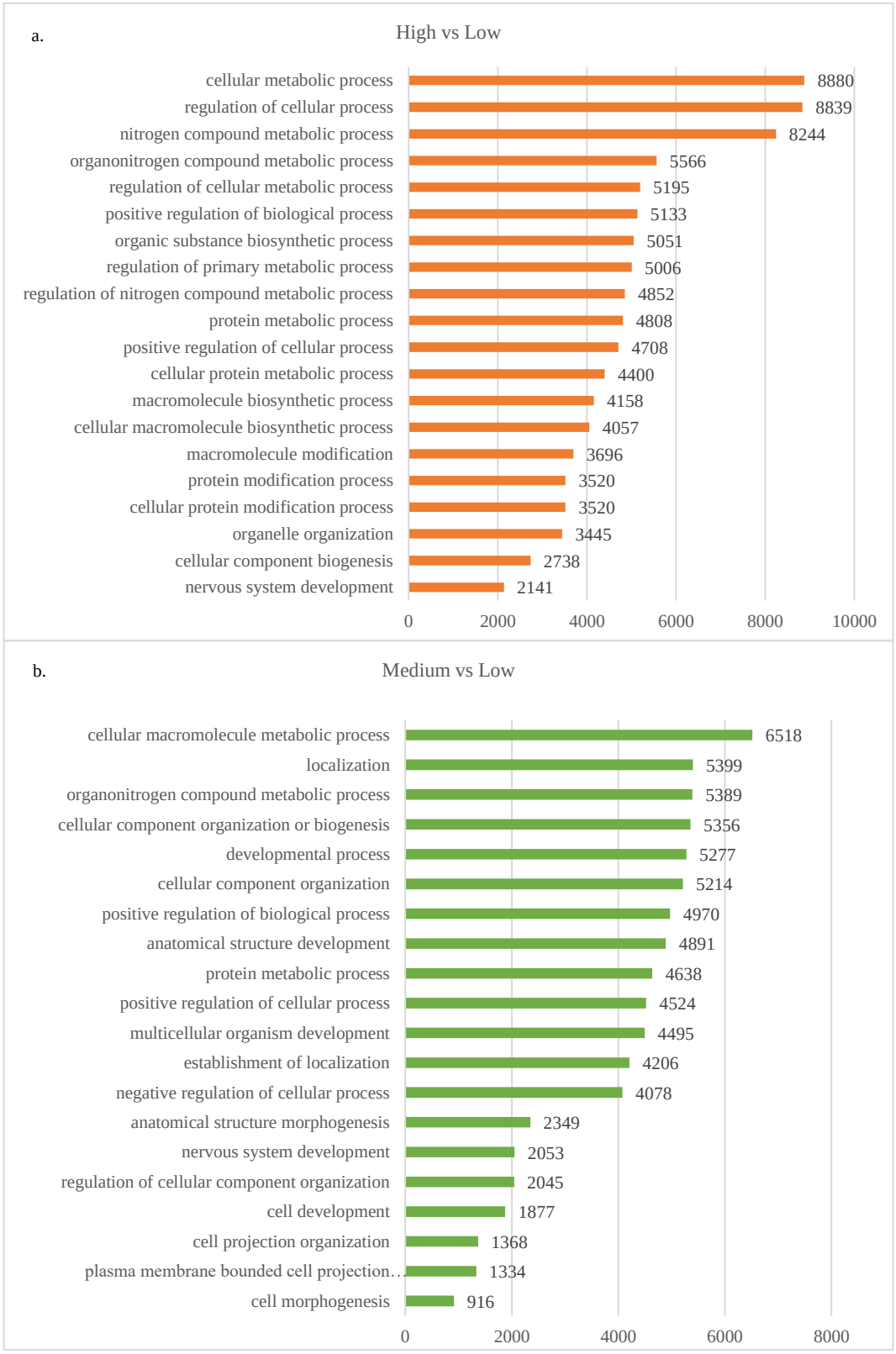
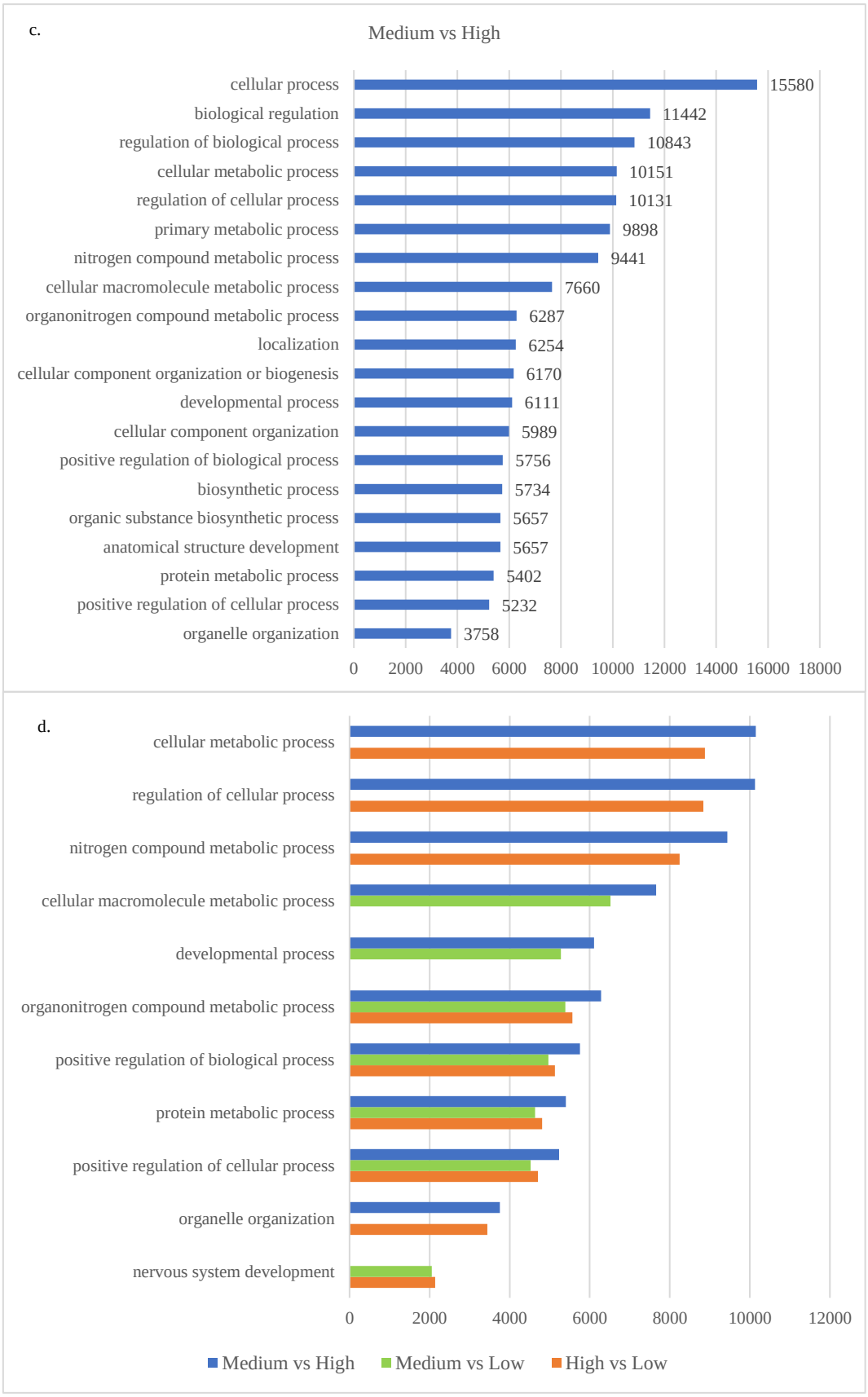


Figure 30. (Continuing from previous page).



Tables 27 a/b/c report the KEGG pathway analysis on DMPs, performed for each impact area comparison.

Noteworthy, it is possible to underline for high and medium impact area vs low one the enrichment of gene-set involved in metabolic and cancer pathways, transduction signaling and inflammation pathways, as well as small and non-small cell lung cancer, pancreatic and prostate cancer, cell cycle and p53 signaling pathways. For all conditions breast cancer and gastric cancer pathways are enriched. Interestingly, for high vs low and for medium vs high are counted some neurodegenerative disorders pathways (e.g. Alzheimer and Parkinson diseases), as well as a series of infectious disease pathways. Medium vs low comparison shows, in addition, enrichment for estrogenic signaling and thyroid hormone pathways, as well as endocrine resistance one.

Considering the DMRs consensus, the GO and the KEGG analysis show significant gene-set enrichment evidence only for medium vs high and medium vs low impact areas comparison (Table 28). All terms reported involve gene silencing processes and post-transcriptional regulation.

The “neuroactive ligand receptor interaction” is the only pathway that resulted enriched in medium vs high comparison by KEGG analysis on DMRs consensus (Table 29).

Table 27a. KEGG for DMPs reported for High vs Low impact area comparison. DE=number of genes from DMPs annotated; FDR=statistical significance of the pathway.

High vs Low							
TermID	**Description**	**DE**	**FDR**	**TermID**	**Description**	**DE**	**FDR**
path:hsa01100	Metabolic pathways	1206	1.56E-06	path:hsa03040	Spliceosome	116	4.46E-03
path:hsa05200	Pathways in cancer	468	1.29E-10	path:hsa04110	Cell cycle	115	2.11E-04
path:hsa05022	Pathways of neurodegeneration - multiple diseases	408	2.19E-07	path:hsa04926	Relaxin signaling pathway	112	3.13E-02
path:hsa05010	Alzheimer disease	313	6.13E-06	path:hsa04071	Sphingolipid signaling pathway	111	1.17E-03
path:hsa05014	Amyotrophic lateral sclerosis	308	4.46E-06	path:hsa04611	Platelet activation	111	3.13E-02
path:hsa04151	PI3K-Akt signaling pathway	302	2.78E-05	path:hsa04142	Lysosome	110	3.71E-02
path:hsa05165	Human papillomavirus infection	293	2.99E-06	path:hsa04722	Neurotrophin signaling pathway	108	8.57E-04
path:hsa04010	MAPK signaling pathway	262	6.11E-06	path:hsa04919	Thyroid hormone signaling pathway	108	1.19E-02
path:hsa05016	Huntington disease	255	2.89E-04	path:hsa03010	Ribosome	108	4.50E-02
path:hsa04144	Endocytosis	232	8.93E-09	path:hsa04935	Growth hormone synthesis, secretion and action	105	1.33E-02
path:hsa05020	Prion disease	222	2.53E-03	path:hsa04725	Cholinergic synapse	103	1.19E-02
path:hsa05132	Salmonella infection	217	2.13E-04	path:hsa04114	Oocyte meiosis	103	4.96E-02
path:hsa05131	Shigellosis	215	8.19E-05	path:hsa04668	TNF signaling pathway	98	2.91E-02
path:hsa04014	Ras signaling pathway	203	8.57E-04	path:hsa04066	HIF-1 signaling pathway	97	9.79E-03
path:hsa05012	Parkinson disease	202	2.78E-03	path:hsa04916	Melanogenesis	92	2.82E-02
path:hsa04810	Regulation of actin cytoskeleton	197	1.94E-06	path:hsa01522	Endocrine resistance	90	2.78E-03
path:hsa04015	Rap1 signaling pathway	196	6.13E-06	path:hsa04713	Circadian entrainment	90	5.77E-03
path:hsa05205	Proteoglycans in cancer	188	7.34E-06	path:hsa04750	Inflammatory mediator regulation of TRP channels	90	1.38E-02
path:hsa05163	Human cytomegalovirus infection	187	1.41E-02	path:hsa04660	T cell receptor signaling pathway	90	3.08E-02
path:hsa05166	Human T-cell leukemia virus 1 infection	184	3.29E-02	path:hsa04070	Phosphatidylinositol signaling system	89	3.28E-04
path:hsa04510	Focal adhesion	183	1.50E-05	path:hsa05215	Prostate cancer	89	2.78E-03
path:hsa04024	cAMP signaling pathway	183	4.34E-02	path:hsa04933	AGE-RAGE signaling pathway in diabetic complications	88	1.38E-02
path:hsa04020	Calcium signaling pathway	176	8.57E-04	path:hsa05231	Choline metabolism in cancer	88	1.75E-02
path:hsa05130	Pathogenic Escherichia coli infection	167	1.22E-02	path:hsa05142	Chagas disease	88	3.75E-02

path:hsa05169	Epstein-Barr virus infection	166	1.91E-02	path:hsa04350	TGF-beta signaling pathway	87	1.11E-02
path:hsa05203	Viral carcinogenesis	164	4.51E-03	path:hsa04925	Aldosterone synthesis and secretion	86	3.08E-02
path:hsa04360	Axon guidance	163	3.38E-04	path:hsa05222	Small cell lung cancer	84	2.06E-03
path:hsa05167	Kaposi sarcoma-associated herpesvirus infection	161	3.66E-02	path:hsa04666	Fc gamma R-mediated phagocytosis	84	2.16E-02
path:hsa04141	Protein processing in endoplasmic reticulum	155	1.50E-05	path:hsa05032	Morphine addiction	82	1.24E-02
path:hsa05225	Hepatocellular carcinoma	154	2.99E-06	path:hsa05210	Colorectal cancer	81	8.57E-04
path:hsa04022	cGMP-PKG signaling pathway	151	2.11E-04	path:hsa04914	Progesterone-mediated oocyte maturation	81	1.09E-02
path:hsa04390	Hippo signaling pathway	148	1.10E-05	path:hsa04512	ECM-receptor interaction	80	1.81E-02
path:hsa04310	Wnt signaling pathway	145	4.43E-04	path:hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	80	3.76E-02
path:hsa04934	Cushing syndrome	144	2.13E-04	path:hsa04540	Gap junction	79	3.50E-02
path:hsa04218	Cellular senescence	143	3.33E-04	path:hsa03015	mRNA surveillance pathway	78	1.09E-02
path:hsa05224	Breast cancer	139	1.80E-05	path:hsa04727	GABAergic synapse	78	1.63E-02
path:hsa05226	Gastric cancer	139	1.46E-04	path:hsa04012	ErbB signaling pathway	75	1.75E-02
path:hsa04921	Oxytocin signaling pathway	139	7.40E-03	path:hsa05100	Bacterial invasion of epithelial cells	74	7.33E-04
path:hsa04150	mTOR signaling pathway	138	3.34E-03	path:hsa01521	EGFR tyrosine kinase inhibitor resistance	73	1.12E-02
path:hsa04072	Phospholipase D signaling pathway	136	3.06E-04	path:hsa04721	Synaptic vesicle cycle	73	1.41E-02
path:hsa03013	RNA transport	135	6.39E-04	path:hsa05214	Glioma	72	8.57E-04
path:hsa05161	Hepatitis B	134	1.63E-02	path:hsa05220	Chronic myeloid leukemia	72	8.57E-04
path:hsa05017	Spinocerebellar ataxia	133	3.95E-05	path:hsa04115	p53 signaling pathway	71	7.33E-04
path:hsa04550	Signaling pathways regulating pluripotency of stem cells	132	9.99E-05	path:hsa05223	Non-small cell lung cancer	70	3.28E-04
path:hsa04261	Adrenergic signaling in cardiomyocytes	131	1.25E-02	path:hsa05212	Pancreatic cancer	70	4.41E-03
path:hsa04120	Ubiquitin mediated proteolysis	125	1.73E-04	path:hsa05412	Arrhythmogenic right ventricular cardiomyopathy	70	1.38E-02
path:hsa04140	Autophagy - animal	124	1.30E-04	path:hsa00562	Inositol phosphate metabolism	68	3.45E-04
path:hsa04723	Retrograde endocannabinoid signaling	122	8.71E-03	path:hsa05218	Melanoma	67	4.51E-03
path:hsa05418	Fluid shear stress and atherosclerosis	121	9.64E-04	path:hsa04520	Adherens junction	67	7.00E-03
path:hsa04728	Dopaminergic synapse	120	8.57E-04	path:hsa05211	Renal cell carcinoma	62	1.19E-02
path:hsa05135	Yersinia infection	120	4.41E-03	path:hsa05217	Basal cell carcinoma	61	3.44E-03
path:hsa04210	Apoptosis	119	1.41E-02	path:hsa01524	Platinum drug resistance	61	1.28E-02

path:hsa04270	Vascular smooth muscle contraction	118	2.32E-02	path:hsa05031	Amphetamine addiction	61	4.34E-02
path:hsa04910	Insulin signaling pathway	118	2.72E-02	path:hsa05213	Endometrial cancer	55	2.50E-03
path:hsa04068	FoxO signaling pathway	117	4.51E-03				

Table 27b. KEGG for DMPs reported for Medium vs Low impact area comparison. DE=number of genes from DMPs annotated; FDR=statistical significance of the pathway.

Medium vs Low							
TermID	**Description**	**DE**	**FDR**	**TermID**	**Description**	**DE**	**FDR**
path:hsa01100	Metabolic pathways	1169	4.18E-03	path:hsa04916	Melanogenesis	93	1.45E-02
path:hsa05200	Pathways in cancer	443	7.15E-04	path:hsa04625	C-type lectin receptor signaling pathway	93	2.40E-02
path:hsa04151	PI3K-Akt signaling pathway	292	7.31E-03	path:hsa05231	Choline metabolism in cancer	92	6.00E-04
path:hsa04010	MAPK signaling pathway	259	1.21E-04	path:hsa01522	Endocrine resistance	91	1.35E-03
path:hsa04014	Ras signaling pathway	202	1.81E-03	path:hsa04660	T cell receptor signaling pathway	90	3.25E-02
path:hsa04015	Rap1 signaling pathway	194	1.21E-04	path:hsa05146	Amoebiasis	88	2.19E-02
path:hsa04810	Regulation of actin cytoskeleton	191	5.71E-04	path:hsa04713	Circadian entrainment	88	3.14E-02
path:hsa05166	Human T-cell leukemia virus 1 infection	184	2.82E-02	path:hsa04925	Aldosterone synthesis and secretion	87	1.69E-02
path:hsa04714	Thermogenesis	183	3.42E-02	path:hsa05414	Dilated cardiomyopathy	86	1.12E-02
path:hsa05205	Proteoglycans in cancer	180	5.59E-03	path:hsa04070	Phosphatidylinositol signaling system	85	2.19E-02
path:hsa04510	Focal adhesion	179	7.77E-04	path:hsa05410	Hypertrophic cardiomyopathy	83	3.13E-03
path:hsa04360	Axon guidance	164	4.65E-04	path:hsa04512	ECM-receptor interaction	82	4.18E-03
path:hsa05225	Hepatocellular carcinoma	148	1.35E-03	path:hsa04914	Progesterone-mediated oocyte maturation	79	4.61E-02
path:hsa04310	Wnt signaling pathway	146	4.65E-04	path:hsa04012	ErbB signaling pathway	77	3.27E-03
path:hsa04921	Oxytocin signaling pathway	139	7.51E-03	path:hsa04911	Insulin secretion	76	4.61E-02
path:hsa04934	Cushing syndrome	139	1.03E-02	path:hsa05412	Arrhythmogenic right ventricular cardiomyopathy	73	6.00E-04
path:hsa04072	Phospholipase D signaling pathway	135	1.19E-03	path:hsa01521	EGFR tyrosine kinase inhibitor resistance	73	1.17E-02
path:hsa05224	Breast cancer	134	3.13E-03	path:hsa04971	Gastric acid secretion	71	1.12E-02

path:hsa05226	Gastric cancer	134	8.78E-03	path:hsa05212	Pancreatic cancer	70	5.75E-03
path:hsa04261	Adrenergic signaling in cardiomyocytes	129	3.42E-02	path:hsa05214	Glioma	70	1.17E-02
path:hsa04915	Estrogen signaling pathway	122	1.17E-02	path:hsa05220	Chronic myeloid leukemia	70	1.17E-02
path:hsa04371	Apelin signaling pathway	121	4.09E-02	path:hsa05100	Bacterial invasion of epithelial cells	70	4.84E-02
path:hsa05135	Yersinia infection	118	1.58E-02	path:hsa04520	Adherens junction	69	7.15E-04
path:hsa04140	Autophagy - animal	117	3.25E-02	path:hsa05223	Non-small cell lung cancer	67	1.92E-02
path:hsa04270	Vascular smooth muscle contraction	117	3.69E-02	path:hsa00562	Inositol phosphate metabolism	65	2.04E-02
path:hsa04068	FoxO signaling pathway	116	9.90E-03	path:hsa05218	Melanoma	65	3.32E-02
path:hsa05418	Fluid shear stress and atherosclerosis	116	3.25E-02	path:hsa04664	Fc epsilon RI signaling pathway	64	2.72E-03
path:hsa04926	Relaxin signaling pathway	113	1.87E-02	path:hsa04929	GnRH secretion	61	1.35E-03
path:hsa04071	Sphingolipid signaling pathway	112	7.15E-04	path:hsa04137	Mitophagy - animal	61	1.92E-02
path:hsa04919	Thyroid hormone signaling pathway	109	6.67E-03	path:hsa05211	Renal cell carcinoma	61	3.25E-02
path:hsa04152	AMPK signaling pathway	108	1.00E-02	path:hsa05221	Acute myeloid leukemia	59	3.37E-02
path:hsa04724	Glutamatergic synapse	107	1.51E-03	path:hsa04370	VEGF signaling pathway	57	7.48E-03
path:hsa04935	Growth hormone synthesis, secretion and action	106	7.48E-03	path:hsa04730	Long-term depression	56	1.00E-02
path:hsa04722	Neurotrophin signaling pathway	105	1.17E-02	path:hsa05213	Endometrial cancer	54	1.17E-02
path:hsa04725	Cholinergic synapse	104	7.31E-03	path:hsa04340	Hedgehog signaling pathway	48	2.79E-02
path:hsa04066	HIF-1 signaling pathway	97	1.00E-02	path:hsa04960	Aldosterone-regulated sodium reabsorption	36	8.78E-03

Table 27c. KEGG for DMPs reported for Medium vs High impact area comparison. DE=number of genes from DMPs annotated; FDR=statistical significance of the pathway.

Medium vs High							
TermID	**Description**	**DE**	**FDR**	**TermID**	**Description**	**DE**	**FDR**
path:hsa01100	Metabolic pathways	1386	2.13E-13	path:hsa05225	Hepatocellular carcinoma	159	1.95E-02
path:hsa05200	Pathways in cancer	497	1.58E-04	path:hsa04310	Wnt signaling pathway	155	5.65E-03
path:hsa05022	Pathways of neurodegeneration - multiple diseases	445	6.62E-06	path:hsa04390	Hippo signaling pathway	154	6.15E-03
path:hsa05010	Alzheimer disease	343	9.29E-05	path:hsa04934	Cushing syndrome	152	2.42E-02
path:hsa05014	Amyotrophic lateral sclerosis	337	4.18E-04	path:hsa04921	Oxytocin signaling pathway	151	2.42E-02
path:hsa05165	Human papillomavirus infection	312	1.73E-02	path:hsa04150	mTOR signaling pathway	149	2.42E-02
path:hsa04080	Neuroactive ligand-receptor interaction	310	2.42E-02	path:hsa05226	Gastric cancer	147	6.87E-03
path:hsa05016	Huntington disease	281	6.87E-03	path:hsa05224	Breast cancer	145	6.87E-03
path:hsa04010	MAPK signaling pathway	280	9.75E-04	path:hsa04072	Phospholipase D signaling pathway	143	3.06E-02
path:hsa05020	Prion disease	248	1.95E-02	path:hsa05017	Spinocerebellar ataxia	139	9.06E-03
path:hsa04144	Endocytosis	240	4.85E-03	path:hsa04550	Signaling pathways regulating pluripotency of stem cells	138	3.53E-02
path:hsa05132	Salmonella infection	235	2.42E-02	path:hsa04723	Retrograde endocannabinoid signaling	133	4.26E-02
path:hsa05131	Shigellosis	231	3.06E-02	path:hsa04210	Apoptosis	132	1.09E-02
path:hsa05012	Parkinson disease	225	1.57E-02	path:hsa04270	Vascular smooth muscle contraction	131	1.19E-02
path:hsa04014	Ras signaling pathway	221	6.87E-03	path:hsa04120	Ubiquitin mediated proteolysis	131	4.53E-02
path:hsa04714	Thermogenesis	209	2.42E-02	path:hsa04140	Autophagy - animal	129	4.89E-02
path:hsa04015	Rap1 signaling pathway	207	6.19E-04	path:hsa04071	Sphingolipid signaling pathway	118	1.95E-02
path:hsa04810	Regulation of actin cytoskeleton	207	6.19E-04	path:hsa01200	Carbon metabolism	108	2.42E-02
path:hsa04024	cAMP signaling pathway	206	2.42E-02	path:hsa04066	HIF-1 signaling pathway	105	2.85E-02
path:hsa05205	Proteoglycans in cancer	198	5.65E-03	path:hsa01522	Endocrine resistance	95	4.34E-02
path:hsa04510	Focal adhesion	192	6.87E-03	path:hsa05231	Choline metabolism in cancer	95	4.45E-02
path:hsa04020	Calcium signaling pathway	189	2.34E-02	path:hsa04350	TGF-beta signaling pathway	94	4.45E-02
path:hsa04360	Axon guidance	174	2.94E-03	path:hsa00564	Glycerophospholipid metabolism	93	4.53E-02
path:hsa04141	Protein processing in endoplasmic reticulum	163	1.73E-02				

Table 28. Gene Ontology for DMRs consensus, reported for the enriched comparison (Medium vs High and Medium vs Low). DE=number of genes from DMRs annotated; FDR=statistical significance of the term.

Medium vs High				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0016458	BP	gene silencing	56	1.73E-05
GO:0031047	BP	gene silencing by RNA	54	1.14E-06
GO:0016441	BP	posttranscriptional gene silencing	50	1.86E-06
GO:0035194	BP	post-transcriptional gene silencing by RNA	50	1.86E-06
GO:0035195	BP	gene silencing by miRNA	49	1.86E-06
GO:0150100	MF	RNA binding involved in posttranscriptional gene silencing	21	4.24E-04
GO:1903231	MF	mRNA binding involved in posttranscriptional gene silencing	21	4.24E-04
GO:0008191	MF	metalloendopeptidase inhibitor activity	7	2.89E-02
Medium vs Low				
TermID	**ONTOLOGY**	**TERM**	**DE**	**FDR**
GO:0007156	BP	homophilic cell adhesion via plasma membrane adhesion molecules	21	5.64E-04
GO:0098742	BP	cell-cell adhesion via plasma-membrane adhesion molecules	25	1.19E-02

Table 29. KEGG for DMRs consensus reported for Medium vs High impact area comparison. DE=number of genes from DMRs annotated; FDR=statistical significance of the pathway.

Medium vs High			
TermID	**Description**	**DE**	**FDR**
path:hsa04080	Neuroactive ligand-receptor interaction	38	9.54E-04

4.4.3 Regression analysis for methylation and exposure biomarkers correlation

In order to highlight a potential correlation between methylation levels and exposure biomarkers evaluated in the subgroup of 600 volunteers, data analysis proceeded through a "Robust Linear Multiple Regression Model", by M-value transformation from beta values, considering the following predictive parameters:

- exposure biomarkers;
- clusters;
- gender;
- age;
- BMI;
- Lymphocytes;
- Neutrophils.

Among the 19 heavy metals assessed as biomarkers of exposure, those showing a statistical significance in correlation to one of the cluster or to one of the impact areas in their independent data analysis were selected (Mercury – Hg202, Arsenic – As75, Cadmium – Cd111, Thallium – Tl205 (µg/L)).

For dioxins and dioxin-like compounds the sum of PCDD/F+DL-PCB (pg WHO-TEQ/g lipids) was considered, as an explorative example of cumulative effect of more congeners.

For Calux gene reporter, all the three assays were included in the regression analysis.

The robust linear multiple regression model, for computing power reasons, was performed on the most variable 5000 differentially methylated probes.

Heatmaps below show the results of the regression model as estimate association between each probe and each exposure biomarker selected, represented by position on gene (Fig. 31a) and by relation to CpG features (Fig. 32a). The same is represented also with the indication of the statistical significance level (Adjusted P-value) (Fig. 31b and 32b).

The model explains how the methylation is affected by exposure biomarkers, given the contribute of the other factors selected as predictive parameters.

Estimate > 0 means that, with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Conversely, if estimate < 0, the increase of the biomarker causes a decrease in the methylation level for the probe observed.

The Adjusted P-value gives support in understanding if the contribute of the exposure biomarker on the methylation level for each probe has a statistical significance.

No particular patterns of hyper- or hypo-methylation are shown in the heatmaps, nor considering the link with gene position, neither the relation to the CpG features.

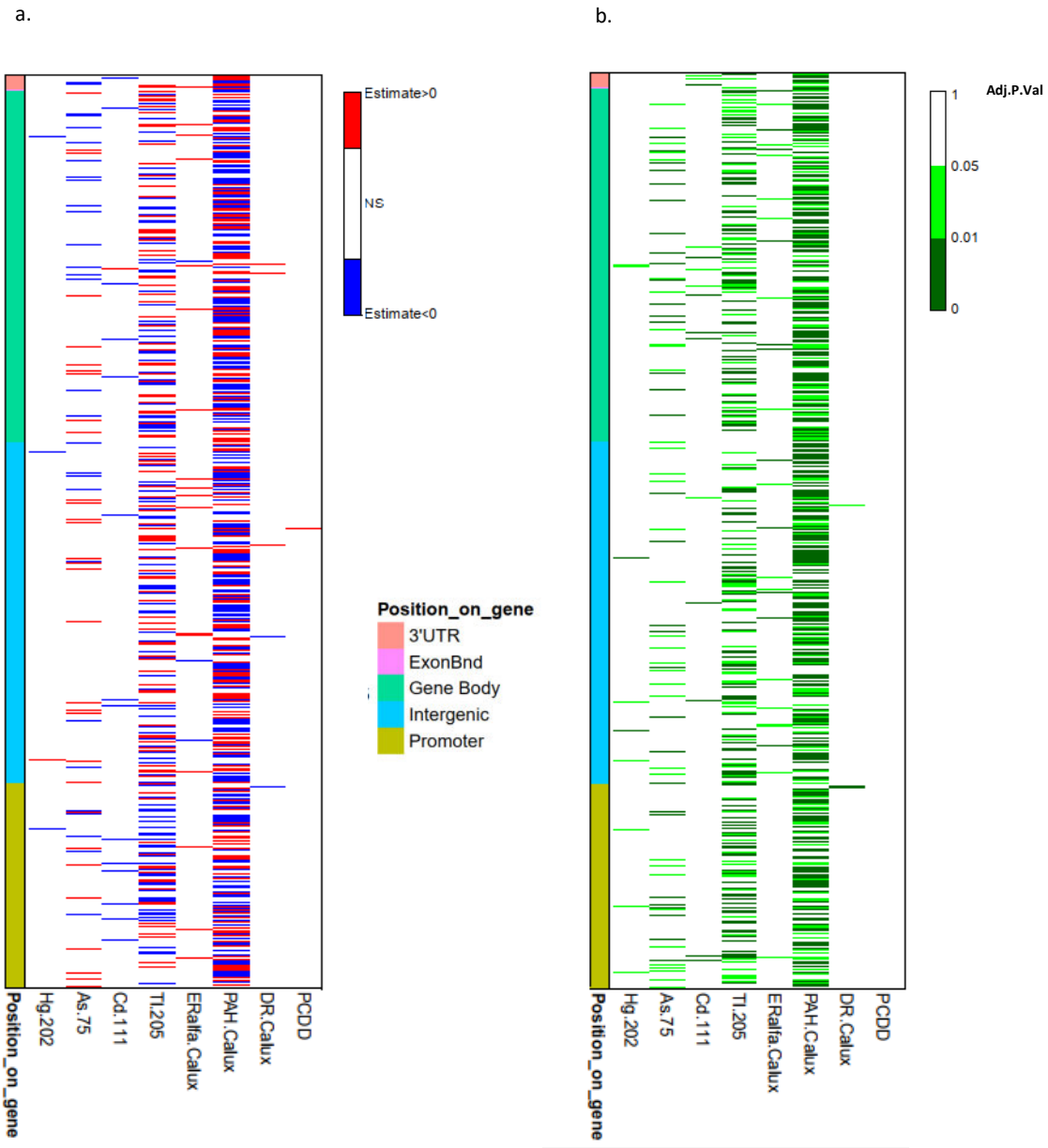


Figure 31. Heatmaps of robust linear multiple regression model linked to probe position on gene. (a) Estimate association between each probe and biomarkers of exposure. (b) Adjusted p-value as indicator of statistical significance for the correlation between the probe and the biomarkers of exposure.

Hg.202 = Mercury (µg/L); As.75 = Arsenic (µg/L); Cd.111 = Cadmium (µg/L); Tl.205 = Thallium (µg/L); ERalfa.Calux = Erα Calux assay (pg-EEQ/ml) ; PAH.Calux = PAH Calux assay (ngBaP-EQ/ml); DR.Calux = DR Calux assay (pg BEQ/g fats); PCDD = sum of PCDD/F+DL-PCB (pg WHO-TEQ/g fats).

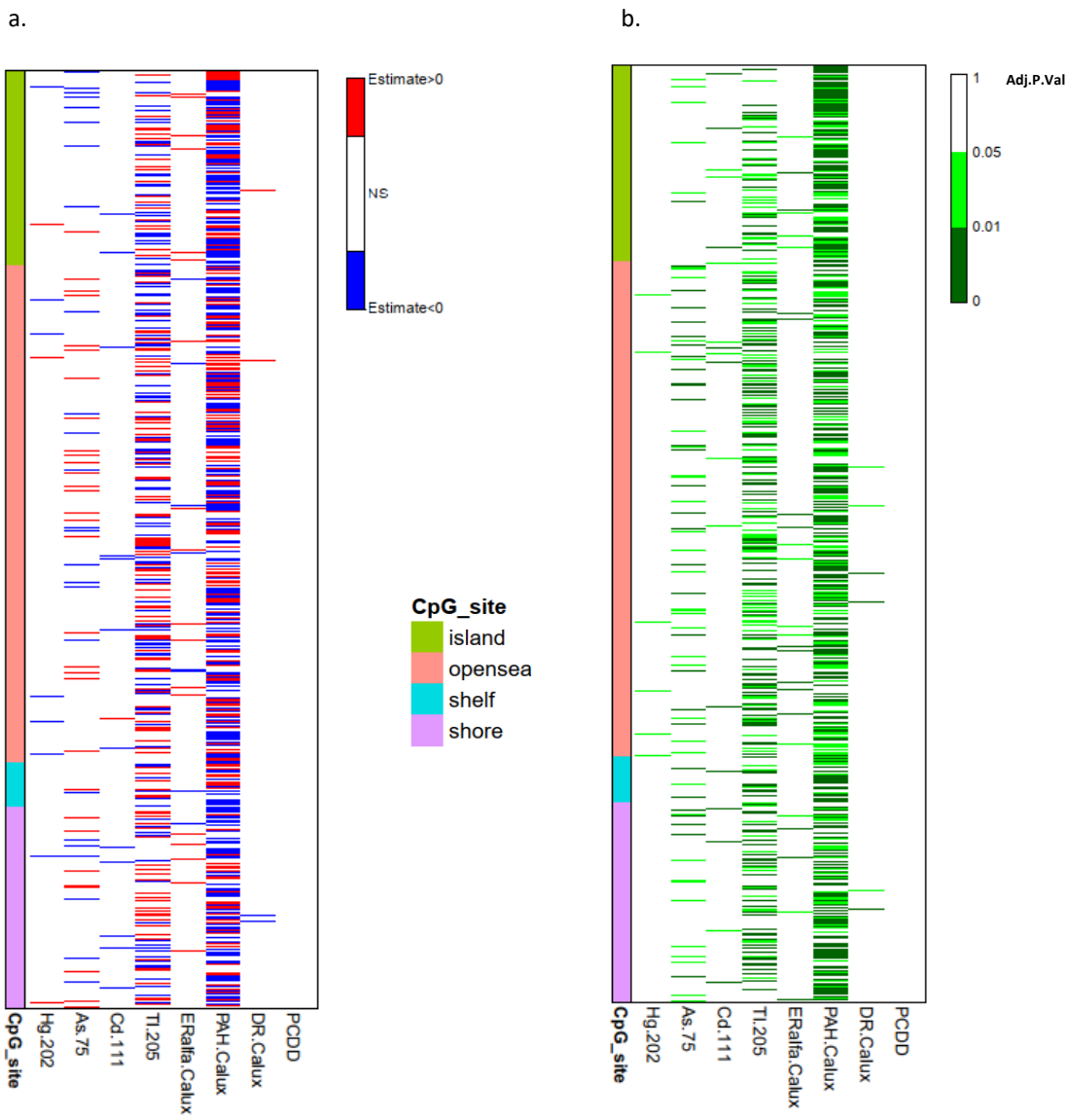


Figure 32. Heatmaps of robust linear multiple regression model linked to CpG features. (a) Estimate association between each probe and biomarkers of exposure. (b) Adjusted p-value as indicator of statistical significance for the correlation between the probe and the biomarker of exposure.

Hg.202 = Mercury (µg/L); As.75 = Arsenic (µg/L); Cd.111 = Cadmium (µg/L); Tl.205 = Thallium (µg/L); ERalfa.Calux = Erα Calux assay (pg-EEQ/ml) ; PAH.Calux = PAH Calux assay (ngBaP-EQ/ml); DR.Calux = DR Calux assay (pg BEQ/g fats); PCDD = sum of PCDD/F+DL-PCB (pg WHO-TEQ/g fats).

In the regression model for the mercury (Hg202), the contribute of the metal to explain the change in methylation is significant for the probes showed in table 30. The focus is on those located on promoter regions, linked to CpG islands and shores.

As reported in 2016 by Martin and Fry, the ATHL1 gene shows a mercury-associated hypomethylation for probes located on CpG shore-promoter region. Likewise, probes mapping on RUNX1 gene (leukemia linked gene) and RASGEF1C show mercury-associated hypomethylation on promoter proximal region. RASGEF1C was linked to differentially methylated site in Alzheimer's disease by Li et al. in 2020. SLC43A2 gene, associated to the transport of aminoacids, shows mercury-associated hypermethylation in promoter region, that means transcriptional repression.

Table 30. Top significant probes in robust linear regression model for Mercury (Hg202). Metal.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Metal.Pr expresses the statistical significance ≤ 0.05.

CpG	gene	feature	cgi	UCSC_Islands_0.00me	metal.Estimate	metal.Pr
cg26423139	SLC43A2	TSS200	shore	chr17:1509882-1510101	2.05E+14	0.0370
cg15747436	ATHL1	TSS1500	shore	chr11:289771-290010	-1.39E+14	0.0190
cg08829299	ATHL1	TSS1500	shore	chr11:288847-289239	-2.31E+14	0.0510
cg18727742	ATHL1	TSS1500	shore	chr11:288847-289239	-2.67E+14	0.0330
cg12955789	RASGEF1C	5'UTR	island	chr5:179588400-179588773	-1.13E+14	0.0500
cg06758350	RUNX1	5'UTR	island	chr21:36258952-36259472	-1.20E+14	0.0500
cg12477880	RUNX1	1stExon	island	chr21:36258952-36259472	-1.69E+14	0.0300

In the regression model for the arsenic (As75), the contribute of the metal to explain the change in methylation is significant for probes showed in table 31. The focus is on those located on promoter regions, linked to CpG islands and shores. Interestingly, among the genes involved, those showing arsenic-related hypermethylation for multiple probes are:

- C1orf109, which encode for a CK2 (casein-kinase 2) substrate that is involved in cell cycle control (G1 to S phase transition) and DNA repair;
- MRPS18A, with its mitochondrial function;
- PIWIL1, which plays a role in gene silencing by RNA; Sellitto et al. in 2019 reported that the PIWIL1-piRNA pathway actively contribute to the clinic-pathological features of the colorectal cancers. Also Amaar and Reeves in 2020 documented the RASSF1C-PIWIL1-piRNA pathway involvement in modulation of key oncogenes and tumor suppressors;
- NAV1 involved in neuronal development;
- CYP1A1, encoding for a member of the Cytochrome P450 family; its expression is induced by some polycyclic aromatic hydrocarbons (PAHs) and has been associated with lung cancer risk.

Conversely, PON1 and SDHAP3 genes show arsenic-related hypomethylation. PON1 encodes for an enzyme able to bind “paraoxon”, a metabolite of Parathion, a toxic organophosphorus pesticide. SDHAP3 is linked to pancreatic cancer (Wolpin et al., 2014). Noteworthy, these results report a total arsenic estimation that need to be careful interpreted because of the presence of non-toxic organic arsenic for the food intake, which could hide the association between the toxic inorganic arsenic species. Indeed, over the purposes of the present dissertation, the SPES protocol have considered the speciation of Arsenic in the biological samples (urine) collected in the study.

Table 31. Top significant probes in robust linear regression model for Arsenic (As75). Metal.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Metal.Pr expresses the statistical significance ≤0.05.

CpG	gene	feature	cgi	UCSC_Islands_Name	metal.Estimate	metal.Pr
cg26052728	ADAP1	TSS1500	shore	chr7:960048-960351	5.88E+14	0.03236
cg14644031	APOB	TSS1500	shore	chr2:21266669-21266961	3.02E+14	0.01887
cg24088508	C1orf109	TSS1500	shore	chr1:38157295-38158586	1.21E+14	0.00065
cg06917450	C1orf109	TSS1500	shore	chr1:38157295-38158586	4.02E+13	0.00445
cg04944931	COMTD1	TSS1500	shore	chr10:76993892-76995953	2.49E+14	0.03363
cg22549041	CYP1A1	TSS1500	island	chr15:75018186-75019336	2.45E+14	0.00374
cg14644378	DNAJB5	TSS1500	shore	chr9:34989555-34989875	1.97E+14	0.01310
cg26038582	HIVEP3	1stExon	shore	chr1:42383640-42384003	2.75E+14	0.00008
cg20290983	MRPS18A	1stExon	island	chr6:43655271-43655618	5.62E+14	0.01002
cg04842962	MRPS18A	1stExon	island	chr6:43655271-43655618	4.22E+14	0.01002
cg04124281	MTHFS	TSS1500	shore	chr15:80189006-80189695	2.09E+14	0.01887
cg14920846	NAV1	1stExon	island	chr1:201617041-201619788	1.54E+14	0.04120
cg08010094	NXPH2	TSS1500	shore	chr2:139537692-139538650	4.11E+14	0.01599
cg18782210	PAMR1	TSS1500	shore	chr11:35547139-35547396	1.82E+14	0.00158
cg24838063	PIWIL1	TSS200	island	chr12:130822360-130822696	5.04E+14	0.01111
cg19424457	PIWIL1	TSS200	shore	chr12:130823502-130824083	4.21E+14	0.00313
cg26677194	PIWIL1	TSS200	island	chr12:130822360-130822696	3.97E+14	0.00290
cg27630820	PIWIL1	TSS200	shore	chr12:130823502-130824083	3.53E+14	0.01635
cg11931211	PIWIL1	TSS200	island	chr12:130822360-130822696	3.33E+14	0.03954
cg11649415	PTPRS	5'UTR	shore	chr19:5335044-5335261	2.98E+14	0.04571
cg03760072	SRPRB	TSS200	island	chr3:133502561-133502963	2.46E+14	0.04261
cg10426581	SRRT	TSS1500	island	chr7:100472375-100473393	4.18E+14	0.02306
cg25093797	UBASH3B	TSS1500	shore	chr11:122526278-122527479	1.25E+14	0.05157
cg03012280	ZFYVE19	TSS1500	shore	chr15:41099243-41099878	1.45E+14	0.00008
cg15721243	ZNF468	TSS1500	shore	chr19:53360201-53360460	3.56E+14	0.03373
cg01147067	RNF213	TSS1500	shore	chr17:78233888-78235283	-2.74E+13	0.02352
cg22355889	ELMOD1	TSS1500	shore	chr11:107461743-107462648	-3.07E+14	0.03847
cg25605443	METTL24	TSS1500	shore	chr6:110679417-110679833	-1.89E+14	0.01846
cg00785941	OR2L13	1stExon	island	chr1:248100325-248100726	-5.78E+13	0.00516
cg09506600	OR2L13	TSS1500	shore	chr1:248100325-248100726	-2.13E+14	0.02507
cg08944170	OR2L13	1stExon	island	chr1:248100325-248100726	-7.47E+14	0.02352
cg22190023	PCDHB7	1stExon	shore	chr5:140553681-140554637	-2.03E+14	0.00430
cg18008345	POLR2E	TSS1500	shore	chr19:1094797-1095649	-4.80E+14	0.00537
cg19678392	PON1	1stExon	island	chr7:94953769-94953971	-3.38E+14	0.00179
cg20119798	PON1	TSS1500	shore	chr7:94953769-94953971	-3.61E+14	0.00013
cg17330251	PON1	TSS200	island	chr7:94953769-94953971	-3.72E+14	0.02789
cg01874867	PON1	TSS200	shore	chr7:94953769-94953971	-3.91E+14	0.00818
cg15571646	RAMP1	TSS1500	shore	chr2:238768073-238768831	-3.97E+14	0.00088
cg06665890	RNF213	TSS1500	shore	chr17:78233888-78235283	-3.28E+14	0.01640
cg24960960	SDHAP3	TSS200	island	chr5:1594238-1595027	-4.50E+14	0.02820
cg21931717	SDHAP3	TSS200	island	chr5:1594238-1595027	-4.51E+14	0.04767
cg19397885	VWDE	TSS1500	shore	chr7:12443163-12443705	-3.07E+14	0.00528
cg03579179	VWDE	TSS1500	shore	chr7:12443163-12443705	-3.12E+14	0.00820

In the regression model for the cadmium (Cd111), the contribute of the metal to explain the change in methylation is significant for probes showed in table 32. The focus is on those located on promoter regions, linked to CpG islands and shores. Among genes showing cadmium-related hypermethylation C1orf109, the same listed for arsenic regression model, which is involved in cell proliferation by control on G1-S phase transition.
Instead, the HNF1B gene, involved in type II diabetes mellitus pathway and with altered expression in some cancer types, shows cadmium-related hypomethylation.

Table 32. Top significant probes in robust linear regression model for Cadmium (Cd111). Metal.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Metal.Pr expresses the statistical significance ≤ 0.05.

CpG	gene	feature	cgi	UCSC_Islands_Name	metal.Estimate	metal.Pr
cg22355889	ELMOD1	TSS1500	shore	chr11:107461743-107462648	5.31583E+14	0.0048
cg24088508	C1orf109	TSS1500	shore	chr1:38157295-38158586	5.08453E+14	0.0085
cg15655868	DZIP1L	TSS1500	shore	chr3:137833848-137834592	-2.18796E+13	0.0495
cg04198914	HNF1B	TSS1500	shore	chr17:36102033-36104766	-1.70354E+14	0.0100
cg17729891	SNAR-F	TSS1500	island	chr19:51107425-51107742	-1.90419E+14	1.69E-05
cg07107916	PPP2R5A	TSS1500	shore	chr1:212457967-212459694	-1.96994E+14	0.0184
cg21140145	ZC2HC1A	TSS1500	shore	chr8:79578112-79578667	-3.67212E+14	0.0002
cg03983883	ZC2HC1A	TSS1500	shore	chr8:79578112-79578667	-4.28353E+14	0.0002
cg04842962	MRPS18A	1stExon	island	chr6:43655271-43655618	-6.09795E+14	1.69E-05

In the regression model for the Thallium (Tl205), the contribute of the metal to explain the change in methylation is significant for probes showed in table 33. The focus is on those located on promoter regions, linked to CpG islands and shores.
Among thallium related CpGs, those showing multiple hypermethylated probes within promoter regions, involve genes as:

- COLEC11, playing a role in innate immunity, apoptosis and embryogenesis;
- PRDM8, encoding a protein belonging to histone methyltransferase family, which acts as negative regulators of transcription. It is involved in steroidogenesis and neural development, a particular interesting function considering that thallium exposure could affect nervous system.

Considering the hypermethylation related genes:

- DUSP22, activating the JNK signaling pathway;
- NOC3L, involved in adipogenesis, gastric and colorectal cancers;
- PON1, encoding the enzyme able to bind paraoxon, a metabolite of Parathion, a toxic organophosphorus pesticide, then with a role in toxic compounds response;
- RRAS2, with its important role in activating signal transduction pathways controlling cell proliferation.

Table 33. Top significant probes in robust linear regression model for Thallium (Tl205). Metal.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Metal.Pr expresses the statistical significance ≤ 0.05.

CpG	gene	feature	cgi	UCSC_Islands_Name	metal.Estimate	Metal.Pr
cg13286116	ARNTL	5'UTR	shore	chr11:13298796-13300735	4.63E+14	1.18E-03
cg08837037	C22orf27	TSS200	shore	chr22:31318228-31318489	1.54E+14	8.57E-05
cg19867917	COLEC11	TSS200	island	chr2:3642547-3642751	8.93E+14	5.31E-05
cg12135269	COLEC11	1stExon	island	chr2:3642547-3642751	6.76E+14	5.56E-05
cg18032014	DYNC1I1	5'UTR	island	chr7:95401691-95402432	1.08E+14	8.37E-04
cg16631618	GPR150	TSS1500	shore	chr5:94955630-94957244	2.16E+14	3.83E-12
cg02104434	LOC148824	TSS1500	island	chr1:247694035-247694501	2.14E+14	2.74E-04
cg12758973	LOC148824	TSS200	island	chr1:247694035-247694501	1.72E+14	5.06E-06
cg12754571	LOC148824	TSS200	island	chr1:247694035-247694501	1.86E+13	8.22E-06
cg00112952	OR2B11	TSS1500	shore	chr1:247614695-247614944	8.33E+14	7.54E-04
cg14830002	OR2B11	TSS1500	shore	chr1:247614695-247614944	5.61E+14	4.32E-04
cg20717585	PCDHGA2	1stExon	island	chr5:140720138-140720829	1.75E+14	1.90E-04
cg05059566	PRDM8	TSS200	shore	chr4:81119095-81119391	1.88E+14	1.58E-04
cg05452645	PRDM8	TSS1500	shore	chr4:81118137-81118603	1.77E+14	3.09E-04
cg26299084	PRDM8	5'UTR	island	chr4:81118137-81118603	1.66E+14	1.18E-03
cg10694470	PRDM8	TSS200	island	chr4:81118137-81118603	1.59E+14	3.40E-06
cg06373870	PRDM8	TSS1500	shore	chr4:81118137-81118603	1.44E+14	1.61E-04
cg06307913	PRDM8	5'UTR	shore	chr4:81119095-81119391	1.33E+14	5.70E-04
cg09419670	PSMD5	TSS1500	shore	chr9:123605043-123605585	3.69E+14	1.21E-05
cg10411850	TDRD10	TSS1500	shore	chr1:154474107-154475699	2.45E+14	6.22E-10
cg07732735	TTC23	TSS200	shore	chr15:99791328-99792042	2.01E+14	1.90E-04
cg22618164	WDR66	TSS200	island	chr12:122356315-122356853	6.03E+14	4.32E-04
cg04074835	ACCS	5'UTR	shore	chr11:44087518-44088165	-5.44E+14	1.93E-05
cg08742575	C21orf56	5'UTR	shore	chr21:47602431-47602740	-9.14E+14	1.23E-04
cg19427746	CAMTA2	5'UTR	shore	chr17:4890282-4890991	-2.00E+14	5.37E-08
cg24324837	CCDC155	1stExon	island	chr19:49891269-49891604	-1.51E+14	2.50E-04
cg02656474	CMAS	TSS1500	shore	chr12:22199062-22199589	-2.09E+13	2.76E-09
cg03680517	CSGALNACT2	TSS1500	island	chr10:43632966-43634222	-1.55E+14	1.12E-03
cg05185784	DEF8	5'UTR	island	chr16:90015551-90016120	-1.84E+14	1.17E-03
cg03395511	DUSP22	TSS200	shore	chr6:291948-292839	-2.05E+13	1.00E-04
cg05064044	DUSP22	1stExon	island	chr6:291948-292839	-1.18E+14	7.68E-04
cg18110333	DUSP22	1stExon	island	chr6:291948-292839	-2.00E+14	3.17E-04
cg21548813	DUSP22	TSS1500	shore	chr6:291948-292839	-2.15E+14	3.62E-06
cg11245928	EIF3L	TSS1500	shore	chr22:38245341-38245681	-1.41E+14	5.30E-05
cg24578493	EIF3L	TSS1500	shore	chr22:38245341-38245681	-1.87E+14	1.69E-07
cg01359822	ETS2	TSS1500	shore	chr21:40177002-40178667	-1.77E+13	5.23E-06
cg14748380	FAM163A	5'UTR	island	chr1:179711947-179713951	-1.36E+14	2.43E-05
cg13607226	GGTA1	TSS1500	shore	chr9:124261806-124262587	-1.33E+14	9.20E-04
cg04945312	GLB1L	5'UTR	shore	chr2:220107844-220108348	-1.88E+14	6.18E-04
cg10602248	GLB1L	1stExon	island	chr2:220107844-220108348	-2.46E+14	1.86E-08
cg24061197	GLB1L	5'UTR	shore	chr2:220107844-220108348	-2.46E+14	7.38E-05
cg24076236	ICE2	TSS1500	shore	chr15:60771068-60771558	-2.10E+14	2.33E-04
cg07520074	KCTD11	TSS1500	shore	chr17:7253298-7253764	-2.76E+14	1.84E-04
cg03185704	LRRC61	TSS200	island	chr7:150019950-150020752	-2.51E+14	2.99E-05
cg06099315	MIR8078	TSS200	island	chr18:112309-112529	-1.29E+14	3.20E-04

cg21537187	MTPN	TSS1500	shore	chr7:135661841-135662272	-3.16E+14	3.55E-07
cg19968001	NFKBIB	5'UTR	shore	chr19:39390033-39390895	-2.83E+14	1.11E-06
cg27514038	NIPA2	TSS1500	shore	chr15:23033933-23034710	-1.69E+14	3.58E-04
cg07889765	NOC3L	TSS1500	shore	chr10:96122540-96122904	-2.14E+14	4.17E-05
cg08923894	NOC3L	TSS1500	shore	chr10:96122540-96122904	-2.68E+14	5.70E-04
cg15157241	PKNOX1	5'UTR	shore	chr21:44394084-44395850	-1.23E+14	3.66E-04
cg02802029	PLOD2	TSS1500	shore	chr3:145878430-145879287	-1.85E+14	1.75E-04
cg17330251	PON1	TSS200	island	chr7:94953769-94953971	-2.79E+14	1.76E-04
cg07107916	PPP2R5A	TSS1500	shore	chr1:212457967-212459694	-1.96E+14	1.28E-07
cg13824270	PRPF4B	TSS1500	shore	chr6:4021444-4022004	-4.37E+14	7.00E-04
cg09293560	REPIN1	TSS200	shore	chr7:150068756-150070051	-1.59E+14	4.25E-09
cg19336497	RRAS2	TSS1500	shore	chr11:14379808-14380737	-1.59E+14	5.96E-04
cg15356281	SEPT2	5'UTR	shore	chr2:242254299-242256026	-1.24E+13	6.58E-04
cg08602346	TUBGCP3	TSS1500	shore	chr13:113241711-113241925	-1.32E+14	1.51E-04
cg14853367	ZNF329	TSS1500	shore	chr19:58661737-58662287	-1.71E+14	1.56E-04
cg21464363	ZSWIM4	TSS1500	shore	chr19:13905810-13906649	-2.48E+14	5.31E-05

In the regression model for the DRCalux, the contribute of the biomarker to explain the change in methylation is significant for probes showed in table 34. The focus is on those located on promoter regions, linked to CpG islands and shores.

The only two probes with significant hypomethylation DR Calux-linked, are cg00321850 and cg00523161, with no particular documented links with dioxin compounds exposure. However, because its role in fatty acids metabolism, LP1N1 gene could potentially be related to the biological processes involved, given the lipophilic nature of the dioxins.

Table 34. Top significant probes in robust linear regression model for DRCalux. Calux.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Calux.Pr expresses the statistical significance ≤ 0.05.

CpG	**gene**	**feature**	**cgi**	**UCSC_Islands_Name**	**calux.Estimate**	**calux.Pr**
cg00321850	KIAA0040	TSS200	shore	chr1:175161900-175162364	-2.88E+14	0.0012
cg00523161	LPIN1	5'UTR	shore	chr2:11886249-11887207	-2.84E+14	0.0046

In the regression model for the ERαCalux, the contribute of the biomarker to explain the change in methylation is significant for probes showed in table 35. The focus is on those located on promoter regions, linked to CpG islands and shores. Genes with ErαCalux related hypermethylation are:

- C21orf56 (SPATC1L), associated with Protein Kinase A regulatory subunit and male infertility (Kim et al., 2018);
- PLSCR2, with its role in blood coagulation and apoptosis;
- RUNX1, leukemia linked gene.

Among hypomethylated:

- DDX4 which encodes for a DEAD box protein, involved in embryogenesis, spermatogenesis, cellular growth and division;
- ZNF714 which plays a role in transcriptional regulation of zinc finger proteins.

Table 35. Top significant probes in robust linear regression model for ERαCalux. Calux.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Calux.Pr expresses the statistical significance ≤ 0.05.

CpG	gene	feature	cgi	UCSC_Islands_Name	calux.Estimate	calux.Pr
cg07747299	C21orf56	5'UTR	shore	chr21:47602431-47602740	7.06E+14	0.018
cg12016809	C21orf56	5'UTR	shore	chr21:47602431-47602740	6.60E+14	0.024
cg22355889	ELMOD1	TSS1500	shore	chr11:107461743-107462648	1.65E+14	0.041
cg24005645	PLSCR2	TSS200	island	chr3:146187108-146187710	5.89E+14	0.038
cg14937633	RUNX1	1stExon	island	chr21:36258952-36259472	5.17E+14	0.022
cg20418711	RUNX1	1stExon	island	chr21:36258952-36259472	4.85E+14	0.027
cg17729891	SNAR-F	TSS1500	island	chr19:51107425-51107742	2.40E+14	0.041
cg09352518	ZNF714	5'UTR	island	chr19:21265164-21265433	-1.05E+14	0.011
cg05202629	C17orf97	TSS1500	shore	chr17:259754-260306	-5.42E+14	0.016
cg27560391	DDX24	5'UTR	shore	chr14:94547001-94547665	-8.89E+14	4.09E-06
cg06238316	ZNF714	5'UTR	island	chr19:21265164-21265433	-7.17E+14	0.034

In the regression model for the PAHCalux, the contribute of the biomarker to explain the change in methylation is significant for probes showed in table 36. The focus is on those located on promoter regions, linked to CpG islands and shores. Genes with PAHCalux related hypermethylation are:

- AURKC, encoding for a member of Aurora subfamily of serine/threonine protein kinase. It is overexpressed in several cancer lines, suggesting a role in oncogenic signal transduction;
- GPR88 with its role in cognitive disorders;
- PF4 with its role in platelet aggregation;

- RNF213 involved in wnt signaling;
- UHMK1 encoding a serine/threonine protein kinase that promotes cell cycle progression through G1, by phosphorylation of the cyclin-dependent kinase inhibitor 1B. The encoded protein has also a role in the adult nervous system.

Among hypomethylated:

- CDKN1A, regulator of cell cycle progression at G1 (expression of gene controlled by p53) and DNA damage;
- HOX4, with its role in gene expression, morphogenesis and differentiation, for instance during embryonic development;
- MIR886, pre-miRNA, direct inhibitor of protein kinase R, plays an important role in regulation of cell growth. It is repressed in some cancers.
- TACSTD2, encoding a carcinoma-associated antigen, a cell receptor that transduces calcium signals.
- TEKT4, which contributes to sperm motility, being a structural component of the sperm flagellum.

Table 36. Top significant probes in robust linear regression model for PAHCalux. Calux.Estimate > 0 means that with the increase of the measured biomarker, there is an increase in the methylation for the probe observed. Calux.Pr expresses the statistical significance ≤ 1.0^{-5}.

CpG	gene	feature	cgi	UCSC_Islands_Name	calux.Estimate	calux.Pr.
cg04450003	AQP5	1stExon	island	chr12:50354839-50356163	3.23E+08	2.00E-10
cg18315870	ATG101	TSS1500	shore	chr12:52463348-52464202	4.74E+08	5.88E-06
cg07314988	ATHL1	5'UTR	island	chr11:289771-290010	7.57E+08	2.58E-07
cg14370847	AURKC	TSS200	island	chr19:57741955-57742457	4.25E+08	3.17E-06
cg02349264	CARHSP1	5'UTR	shore	chr16:8962189-8963190	2.83E+08	4.03E-07
cg03651525	CCDC25	TSS1500	shore	chr8:27629753-27630234	4.18E+08	6.08E-06
cg07187855	DDR1	5'UTR	shore	chr6:30852102-30852676	3.25E+08	1.65E-05
cg05185784	DEF8	5'UTR	island	chr16:90015551-90016120	2.67E+08	4.33E-05
cg27638615	FZD7	1stExon	shore	chr2:202897386-202901046	5.92E+08	5.27E-11
cg06223162	GPR88	TSS200	shore	chr1:101004471-101005885	1.04E+09	1.13E-09
cg07412545	GPR88	1stExon	shore	chr1:101004471-101005885	1.03E+09	1.23E-12
cg09408571	GPR88	TSS200	shore	chr1:101004471-101005885	8.07E+08	1.29E-08
cg19978674	HOPX	5'UTR	shore	chr4:57521621-57522703	6.65E+08	8.84E-05
cg16406611	IFITM5	TSS1500	shore	chr11:299389-299635	4.70E+08	8.30E-09
cg00944873	KIAA0319	TSS1500	shore	chr6:24645931-24646591	5.13E+08	3.50E-05
cg18049933	LOC100996579	TSS1500	shore	chr2:162101439-162101650	3.54E+08	3.71E-07
cg25755428	MRI1	TSS1500	island	chr19:13875044-13875951	1.39E+09	1.06E-06
cg08989084	P2RX2	TSS1500	shore	chr12:133195922-133196556	3.75E+08	2.26E-05
cg05509609	PF4	TSS200	island	chr4:74847528-74847830	6.10E+08	1.65E-09
cg06834998	PF4	TSS200	island	chr4:74847528-74847830	5.51E+08	4.95E-10
cg02530824	PF4	TSS200	island	chr4:74847528-74847830	5.12E+08	2.65E-08
cg16072462	PF4	TSS200	island	chr4:74847528-74847830	4.67E+08	5.90E-09
cg15398841	PF4	TSS200	island	chr4:74847528-74847830	4.57E+08	4.96E-09
cg21043213	PF4	1stExon	island	chr4:74847528-74847830	4.11E+08	1.92E-08
cg11649415	PTPRS	5'UTR	shore	chr19:5335044-5335261	4.53E+08	2.03E-06
cg08742290	RNASET2	TSS1500	shore	chr6:167369439-167370588	4.42E+08	2.31E-07
cg06665890	RNF213	TSS1500	shore	chr17:78233888-78235283	4.69E+08	5.92E-06
cg01147067	RNF213	TSS1500	shore	chr17:78233888-78235283	4.48E+08	4.16E-07
cg22198044	SH3BP2	TSS1500	island	chr4:2819499-2820429	2.48E+07	7.25E-07
cg26390081	TMEM204	5'UTR	shore	chr16:1583809-1584641	3.97E+08	1.49E-07
cg27594616	TMEM204	5'UTR	shore	chr16:1583809-1584641	2.90E+08	8.56E-05
cg09777637	TMEM229A	TSS1500	shore	chr7:123672063-123673691	3.28E+07	2.03E-09
cg06179011	TSPAN4	TSS1500	shore	chr11:842293-843396	3.11E+08	1.61E-05
cg19814518	UHMK1	TSS1500	shore	chr1:162467599-162468027	6.61E+08	1.44E-14
cg00859178	UHMK1	TSS1500	shore	chr1:162467599-162468027	5.80E+08	6.98E-14
cg03012280	ZFYVE19	TSS1500	shore	chr15:41099243-41099878	5.98E+08	3.33E-09
cg26751094	ABCC4	TSS1500	shore	chr13:95953337-95954211	-2.60E+08	4.44E-05
cg01324343	ABCC5	5'UTR	shore	chr3:183735183-183736103	-5.58E+08	7.36E-08
cg18349420	AQP11	TSS1500	shore	chr11:77300360-77301391	-4.24E+07	3.44E-12
cg06061081	AQP11	TSS1500	shore	chr11:77300360-77301391	-3.70E+08	2.02E-08
cg04518186	AQP11	TSS1500	shore	chr11:77300360-77301391	-4.83E+08	1.01E-08
cg15747436	ATHL1	TSS1500	shore	chr11:289771-290010	-3.74E+08	1.67E-10
cg08829299	ATHL1	TSS1500	shore	chr11:288847-289239	-5.64E+08	5.71E-09
cg18727742	ATHL1	TSS1500	shore	chr11:288847-289239	-6.52E+08	3.98E-08
cg04567952	C2orf69	TSS1500	shore	chr2:200775746-200776548	-6.12E+08	1.41E-06
cg15855900	CA7	TSS1500	shore	chr16:66878172-66879072	-2.42E+08	6.91E-05
cg24425727	CDKN1A	TSS1500	shore	chr6:36646244-36648465	-2.34E+08	8.19E-05
cg02452435	COG1	TSS1500	shore	chr17:71188407-71189458	-4.84E+08	1.55E-07
cg02271943	COMTD1	TSS1500	shore	chr10:76993892-76995953	-1.19E+09	8.54E-70
cg27560391	DDX24	5'UTR	shore	chr14:94547001-94547665	-8.17E+08	9.01E-17

cg00481259	DECR2	TSS1500	island	chr16:450834-451140	-3.88E+08	2.63E-09
cg07973095	DECR2	TSS1500	island	chr16:450834-451140	-5.33E+08	1.66E-05
cg24578493	EIF3L	TSS1500	shore	chr22:38245341-38245681	-3.80E+08	1.37E-10
cg11245928	EIF3L	TSS1500	shore	chr22:38245341-38245681	-4.35E+08	2.00E-11
cg08291335	FANCG	TSS1500	shore	chr9:35079628-35080213	-4.81E+08	6.40E-11
cg14785464	HLCS	TSS200	island	chr21:38362015-38362868	-3.25E+08	4.91E-06
cg17457637	HOXA4	TSS1500	shore	chr7:27169572-27170638	-2.99E+08	6.91E-05
cg07317062	HOXA4	5'UTR	island	chr7:27169572-27170638	-3.47E+08	3.28E-05
cg24426391	HTATIP2	TSS1500	shore	chr11:20385161-20385673	-8.09E+08	1.18E-06
cg19521832	KCNE1	TSS200	island	chr21:35831697-35832365	-3.71E+08	6.49E-07
cg04177826	KLHDC7B	TSS1500	island	chr22:50984972-50988141	-2.57E+08	2.20E-06
cg15509177	LOC100101938	TSS200	island	chr13:19918585-19919221	-3.05E+07	2.52E-05
cg15469871	LOC387646	TSS200	island	chr10:27541297-27541883	-3.74E+08	1.44E-05
cg18806716	MAP3K8	TSS1500	shore	chr10:30722378-30723707	-4.23E+08	5.04E-05
cg26896946	MIR886	TSS200	island	chr5:135416204-135416475	-1.07E+07	9.48E-38
cg18797653	MIR886	TSS1500	shore	chr5:135415069-135415307	-9.63E+08	7.58E-30
cg00124993	MIR886	TSS200	island	chr5:135416204-135416475	-1.15E+09	5.26E-22
cg08745965	MIR886	TSS1500	shore	chr5:135415069-135415307	-1.20E+09	3.22E-48
cg18678645	MIR886	TSS200	island	chr5:135416204-135416475	-1.30E+09	1.99E-33
cg23605644	MOBP	TSS200	shore	chr3:39543771-39544223	-3.91E+08	3.06E-06
cg03054684	MOBP	5'UTR	shore	chr3:39543771-39544223	-5.16E+08	1.09E-06
cg07878407	MOBP	1stExon	shore	chr3:39543771-39544223	-6.31E+08	2.02E-09
cg05633900	MOBP	1stExon	shore	chr3:39543771-39544223	-6.58E+08	2.13E-10
cg14666699	MRGPRG-AS1	TSS1500	shore	chr11:3239217-3239988	-2.36E+08	1.42E-05
cg18705301	NDUFAF1	TSS1500	shore	chr15:41694493-41694756	-5.54E+08	3.07E-05
cg13348653	PCDHB3	TSS200	shore	chr5:140481546-140482460	-3.68E+08	2.24E-08
cg03780733	PCDHB7	1stExon	shore	chr5:140553681-140554637	-1.69E+09	7.31E-05
cg15157241	PKNOX1	5'UTR	shore	chr21:44394084-44395850	-5.91E+08	6.34E-12
cg24530147	PRR23C	TSS200	island	chr3:138762689-138764113	-3.42E+08	9.50E-09
cg05895034	PRSS22	TSS1500	shore	chr16:2907574-2907836	-4.19E+08	2.66E-11
cg09745688	PRSS22	TSS1500	shore	chr16:2907574-2907836	-4.67E+08	1.49E-05
cg09309269	PSMD11	TSS1500	island	chr17:30770960-30772137	-3.24E+08	4.49E-06
cg00871487	RANBP17	TSS200	shore	chr5:170288879-170289737	-5.18E+08	1.72E-07
cg06091474	SEMA3C	TSS1500	shore	chr7:80548306-80548726	-3.09E+08	1.95E-07
cg10589385	SETDB1	TSS1500	shore	chr1:150898619-150898839	-2.73E+08	5.05E-05
cg26748578	SLC1A4	1stExon	island	chr2:65215598-65217212	-3.65E+07	3.69E-06
cg05845376	SLC25A2	TSS200	island	chr5:140683195-140683773	-2.93E+07	3.46E-05
cg17729891	SNAR-F	TSS1500	island	chr19:51107425-51107742	-8.53E+08	3.55E-06
cg22023531	SORD	TSS1500	shore	chr15:45315201-45315543	-4.42E+08	1.14E-09
cg16699148	TACSTD2	TSS200	island	chr1:59042013-59043295	-3.67E+08	3.23E-06
cg16080552	TACSTD2	TSS200	island	chr1:59042013-59043295	-6.19E+08	6.32E-05
cg04863005	TACSTD2	TSS200	island	chr1:59042013-59043295	-6.40E+08	3.06E-06
cg06282964	TEKT4	1stExon	island	chr2:95537196-95537810	-3.68E+07	8.82E-10
cg05852568	TEKT4	1stExon	island	chr2:95537196-95537810	-4.24E+08	6.40E-09
cg24354818	TNFRSF6B	5'UTR	island	chr20:62328013-62328558	-3.82E+08	6.60E-05
cg00152799	TXNL1	TSS1500	shore	chr18:54305419-54306087	-3.80E+08	2.17E-06
cg17751872	ZNF714	5'UTR	shore	chr19:21265164-21265433	-4.13E+08	6.71E-05
cg21464363	ZSWIM4	TSS1500	shore	chr19:13905810-13906649	-4.33E+08	2.78E-06

In the regression model for the sum of PCDD/F + DL-PCB, no significant probes on promoter regions and linked to CpG islands and shores were reported.

5. DISCUSSION

Human Biomonitoring can be considered as a form of public health surveillance (Salines, 2012). It represents an important and useful tool to investigate human exposure to environmental chemicals, providing a reliable measurement of the internal dose of pollutants, through different exposure pathways (Calafat et al., 2006).

Its methodological approach, aimed to estimate exposure importance and extension, is widely used in the risk assessment for occupational health and for monitoring of general population. According to EU in 2013, HB can serve the chemicals surveillance by providing "authorities with a more comprehensive view of actual exposure of the population to pollutants, […] and can provide better evidence from guiding appropriate responses" (Ganzleben et al., 2017).

Assessing chemicals in body tissues and fluids, HB can help to understand the toxic-kinetic profile of compounds and their interactions, the correlation with effects on biological systems and the potential health outcomes (CDC, 2005). Moreover, from an epidemiological point of view, it can help to define what is the total exposure to the substances from different sources, how does the chemical exposure evolve over time and among regions or population groups, and incidence and mortality rates of putative linked diseases. Thus, HB evidence is able to support both the advance of scientific knowledge and the effectiveness of policy and regulatory actions (Louro et al., 2019).

Despite of the great potential, the conventional HB protocols often suffer the lack of data integration and methodological sharing between different science fields, as well as the lack of a holistic approach oriented to a "one health" perspective (Frazzoli et al., 2015; Falzon et al., 2018; Sleeman et al., 2019).

To support a complete understanding of the phenomena, HB protocols need to integrate a single multilevel analysis. The environmental background characterization, the anamnestic and epidemiological data collection, together with the omics approach, become fundamental for the evaluation of the exposure risk and for depicting a framework of the real degree of contamination (Pierri et al., 2020). This means to set up a systemic overview, in which the multidimensionality and the complexity of the observations are solved by dividing the phenomenon in several levels of analysis that, in the final intersection, could provide the global reading and interpretation.

Especially considering the Campania Region and the Land of Fire case study, the controversy emerging from scientific literature (Cantoni, 2016; Mazza et al., 2018) does not allow to obtain a clear definition of the level of pollution in the regional territory.

For these reasons, the main purpose of the present PhD dissertation is to introduce and describe an innovative HB integrated approach, the basis of the SPES protocol, the

observational study on Campania Region healthy population promoted by the IZSM. A particular focus in the dissertation comes from exploring the role of epigenetic biomarkers, as the DNA methylation, and their integration with several levels of analysis by means statistical model of data elaboration.

Starting from the paradigm of the method (which considers three main elements: the identification of pollution sources, of pollution migration pathways and of final target organisms), the Campania environmental context has been characterized, highlighting sites of particular interest. By means of a "Pressure Index" (IRc), the environmental background provides a synthetic representation of the level of pressure exerted by the contamination sources on the municipalities. The three increasing levels of impact derived were consistent with areas with the already known reputation of "triangle of death" (Triassi et al., 2015), expression firstly coined by Senior and Mazza on The Lancet Oncology in 2004 and referred to area with documented illegal disposal of urban, toxic and industrial wastes. Interestingly, the IRc methods showed a medium impact area, where the presence of industrial plants gives an important contribute to the spreading of heavy metals and polycyclic aromatic hydrocarbons in the ecosystem. It is well known that workers surveillance and study monitoring on people indwelling areas close to industrial plants document adverse health effects linked to the spreading of pollutants. Scientific literature describes a series of renowned italian sites, as well as Taranto city, with its steel foundry, that is one of the most famous of the Southern Italy (Vimercati et al., 2017; Lovreglio et al., 2018; Lucchini et al., 2019). The medium impact areas, identified with the computed IRc in SPES protocol, consist of some sites of particular interest, given the presence of industrial plants: tanneries in Solofra area, a steel foundry in the Irno Valley, metallurgic, iron and steel industries, asbestos disposal plant in Sabato Valley and, finally, the area known as "agro-nocerino-sarnese", mainly characterized by groundwater contamination caused by landfill leachate and wide presence of compounds linked to the use of pesticides. This suggests other areas, over those included in the LOF, where industrial sources of contamination could represent a risk for residing people. Data acquired by HB support this hypothesis. The assessment of heavy metals in serum of healthy volunteers, resulting from SPES human biomonitoring, revealed that the average levels (µg/L) of cadmium (Cd111) and mercury (Hg202) were approximately 5 times higher in Irno and Sabato Valley clusters (Medium Impact Area), compared to the entire population assessed.

The high impact areas, where are present municipalities belonging to the so-called "Land of Fire", were only characterized by thallium (Tl205) higher values, probably due to the natural background levels of this compound in the soil, because of the volcanic origin of the region (Karbowska, 2016).

In line with these results, also organic compounds measurement showed a consistent statistical significance in medium impact area, especially for Irno and Sabato Valley clusters. Considering high resolution gas chromatography results, the median value of the sum of PCDD/PCDF + DL-PCB (pg WHO-TEQ/g lipids) showed a sensitive increase for Sabato Valley cluster, compared to the overall median. A slighter increase is also evident for Irno Valley cluster and High 12 cluster. This last includes the municipalities of Acerra, Villa Literno and Naples, where the presence of a waste incinerator for Acerra, a busy commercial road (Nola-Villa Literno provincial road - SP333), and the complex urbanization of Naples metropolis, could explain the higher values of the sum of dioxins congeners (Marques and Domingo, 2019).

Comparable data were obtained considering the DR-Calux® bioassay, expression of stable ligands of Ahrilic receptor (e.g. dioxins congeners), which found a significant correlation between DR-Calux® median values (pgBEQ/g lipids) and Irno and Sabato Valley clusters. In addition, the bioassay showed an interesting evidence of total estrogenic and estrogenic-like responses by the ER-Calux® assessment (pgEEQ/ml), with higher values detected in females and males from high and medium impact areas. The total estrogenic activity observed in females is 3 to 5-fold increase compared to the levels of the natural endogenous hormone 17-β-estradiol, and this could suggest the presence of compounds with possible endocrine activity (Endocrine disrupting chemicals EDCs).

Morgan et al. in 2017 highlighted a link between environmental exposure to PCB138 and breast cancers, as well as a correlation with the sum of non-dioxin like PBCs and breast cancer. Calaf et al. in 2020 reviewed xenoestrogens compounds in relation to breast cancer incidence, showing that Bisphenol A, Dichlorodiphenyltrichloroethane (DDT) and PCBs are involved in alteration of cell proliferation and of cell death, are immunosuppressive agents and induce oxidative stress and epigenetic alterations.

The Report AIOM-AIRTUM of 2019 (www.salute.gov.it) estimated approximately 4050 new diagnosis of breast cancers among women in Campania Region, with an incidence rate of 140.5 per 100.000 residents (Friuli Venezia Giulia Region has the higher incidence rate at 203.9 per 100.000 people, while Calabria Region has the lower rate at 124 per 100.000 people). Fusco et al., in 2013, already showed an increasing trend in Campania for this type of cancer, as well as for lung cancer and skin melanoma in women, for colorectal cancer and skin melanoma for men.

Some studies reported a link between exposure to PCBs and risk of cutaneous melanoma (Zani et al., 2017; Boffetta et al., 2018; Cao et al., 2019).

Lung and colorectal cancers were also reported in literature as possible adverse outcome linked to pollution (de Groot et al., 2012; Juloski et al., 2020).

Certainly, the incidence trends from literature need to be explained also considering life habits, prevention culture and familiar history. Indeed, these anamnestic considerations are possible confounding factors.

This is the reason why in the SPES protocol the subgroup of 600 volunteers, selected for deepened investigation, were chosen according to inclusion criteria that allowed to avoid bias given by the confounding factors (see paragraph 3.1.5).

Data obtained with exposure biomarkers investigation are supported by the deep analysis of effect biomarkers. Considering the epigenetic profile resulted from methylation array data analysis, it is important to underline that the KEGG pathway analysis on the DMPs revealed, in medium impact areas vs low impact areas, an interesting enrichment for estrogenic signaling and thyroid hormone pathways, as well as for the endocrine resistance one. These results are consistent with the higher levels of organic compounds detected in the high and medium clusters.

The same KEGG pathway analysis has shown for high and medium impact areas vs low one, the enrichment of gene-set involved in metabolic and cancer pathways, transduction signaling and inflammation, as well as breast cancer, gastric cancer, small and non-small lung cancer, melanoma, cell cycle and p53 signaling.

The KEGG pathway analysis performed on the DMRs also revealed the enrichment of "neuroactive ligand receptor interaction", supported by biological processes terms enriched in the gene ontology analysis, "nervous system development", and by the KEGG neurodegenerative disorders pathways reported for DMPs in high and medium vs low impact area comparison.

As a matter of fact, there are a lot of scientific evidences that pre- and post-natal exposures to environmental factors could predispose to the onset of neurodegenerative diseases in later life (e.g. Alzheimer and Parkinson diseases). According to Chin-Chan et al., neurotoxic metals such as lead, mercury, aluminum, cadmium and arsenic, as well as some pesticides have involved in Alzheimer etiology, due to their ability to increase the formation of beta amyloids plaques. The exposure to lead, manganese, solvents and some pesticides has related to Parkinson disease. Epigenetic alteration is one of the involved mechanisms (Chin-Chan et al., 2015).

The use of epigenetics as a key to understand the molecular mechanism subtended to toxicity, gives the chance of characterize the biological response to environmental stressors and identify possible molecular biomarkers or patterns of exposure (Vineis et al., 2013).

From the epigenetic analysis, carried out for the purposes of the present dissertation, emerges a global hypomethylation in the high impact areas and in the medium one compared to the low impact areas. This is consistent with the exposure to several toxic compounds as benzene (Bollati et al., 2007), POPs persistent Organic Pollutants (Rusiecki

et al., 2008), Particular Matters (Baccarelli et al., 2009; Plusquin et al., 2017), Endocrine Disruptors (La Rocca et al., 2014) and, at the same time, with genomic instability and a series of chronic and degenerative conditions (Van Togelen et al., 2017; Zang et al., 2020). Additionally, the regression analysis performed to highlight possible correlations between DNA methylation levels and exposure biomarkers, revealed some promoter regions, linked to CpG island and shore, where hypermethylated sites could be involved in a transcriptional repression mechanisms, and could contribute to generate the final response of the biological system to the external stressors. Similarly, hypomethylated sites could contribute to generate a reaction to the external input.

As reported in the Results chapter, heavy metals included in the robust linear regression models (Hg, As, Cd and Tl) showed a significant link with genes involved in cell cycle regulation, carcinogenesis, neuronal development and this is consistent with the known possible effect of some of them (WHO, 2015).

Among them, an interesting correlation emerged between arsenic and PIWIL1 gene, which plays a role in gene silencing by RNA; as previously described the PIWIL1-piRNA pathway is involved in modulation of key oncogenes and tumor suppressors (Amaar and Reeves, 2020), and, as documented by Sellitto et al. in 2019, in colorectal cancer cells contributes to the clinic-pathological features of the disease. Noteworthy, as previously described (see paragraph 4.4.3), these results report a total arsenic estimation that need to be careful interpreted because of the presence of non-toxic organic arsenic for the food intake, which could hide the association between the toxic inorganic arsenic species. Indeed, over the purposes of the present dissertation, the SPES protocol have considered the speciation of Arsenic in the biological samples (urine) collected in the study.

Regarding arsenic and thallium, both the two metals show a correlation with the gene PON1: arsenic-related hypomethylation and thallium-related hypermethylation of its promoter region. This gene encodes for the enzyme able to bind the paraoxon, a metabolite of Parathion, a toxic organophosphorus pesticide. Marhooz et al. in 2019 described PON1 as a fundamental antiatherogenic and antioxidant enzyme in the circulation that has been associated with adverse health outcomes, particularly cardiovascular disease. It is also been analyzed by Cardenas et al., in 2017, in order to assess an association with the maternal prenatal mercury exposure in cord blood samples. According to their findings, DNA methylation at the PON1 locus could regulate the association between prenatal mercury exposure, cognitive development and other health outcomes in children. The researchers expressed that the persistent epigenetic disruption of the PON1 gene might serve as a biomarker of mercury exposure and disease susceptibility. Moreover, the PON1 activity shows a polymorphic distribution in patients with Alzheimer's disease (Saeidi et al., 2017), with different frequencies in group of people exposed or not to organophosphate pesticides

(Javeres et al., 2020). This could suggest the methylation and the expression of PON1 gene as a biomarker of effect in response to heavy metals exposure, and as an early sentinel of health consequences, given its involvement in multiple biological processes.

Finally, the organic compounds included in the robust linear regression models (DR Calux, ER Calux, PAH Calux) showed a significant link with genes involved in fertility, embryogenesis, sperm motility, but also oncogenic signal transduction (AURCK), cell cycle progression (UHMK1 and CDKN1A) and cognitive disorders (GPR88).

These results are consistent with the possible action of the compounds as endocrine disruptors (Schug et al., 2011).

In view of these, the multilevel analysis carried out by means of a holistic approach, from the environmental background characterization, through human biomonitoring assessment, to epidemiological considerations, has been able to depict a frame of the regional state of contamination. The spatial and geographical reference given to pollutants allows to identify a new area of interest where compounds associated to industrial plants and the presence of a particular orography of the valleys, could represents a risk for residing people.

The hypothesis is supported by HB data and by an early epidemiological evaluation of incidence rates of diseases potentially related with.

Notwithstanding the preliminary results encourage the theory, the integration of the observational SPES dataset with *in vitro* validation of the effects of compounds, as well as a comparative meta-analysis with other HB or omics datasets, could be useful to consolidate and support the evidence, avoiding false positive correlations, potentially generating from multiple comparisons.

Furthermore, a potential contribute could come from integrative bio-statistics models of data elaboration, considering mixture of elements and their cumulative effects, together with other biomarkers of effects assessed and not object of the present dissertation. Indeed, as previously described (see paragraph 3.1.5), a set of other effect biomarkers are object of further investigations in the SPES protocol: inflammatory cytokines and redox status aimed to explore early evidence of phlogosis and immunity involvement, telomeric length and polymorphisms in order to observe genetic consequences of exposure and individual susceptibility, miRNA to deepen epigenetic alterations, gene expression profiling to identify transcriptomic variations in response to environmental exposure.

Model of big data analysis could be the right choice to implement the multiparametric evaluation, and the integrative approach could provide a fully detailed insight into the biological meaning of exposure to several compounds. Novel biomarkers of diagnostic, prognostic, predictive and therapeutic interest could be found out from this multi-level analysis.

5.1 Applications and future perspectives

Considering the long-term purpose of the study, all data collected with the SPES protocol and its integrated approach represent an important contribute to define a full and dynamic map of the environmental pollution of the Campania Region. This could constitute a scientific support for public health policy and for decision making, giving to the model the chance to be translated in different contexts and for various applications.

Indeed, susceptible cohorts of people, or people residing in critical areas, involved in environmental emergencies for spreading of pollutants, or for the presence of a known source of contamination, could be investigated through this innovative approach.

The systemic overview, from environmental matrices, through biomonitoring and molecular insight, to epidemiology, could be used to characterize the context and acquire the real level of pressure exerted by pollutants in that area or on that population.

From a mere scientific and research point of view, the “one health” perspective, intrinsic in the novel human biomonitoring model, reveals the ability of the methodological approach in integrate different science fields, in order to gain an inter-disciplinary reading and interpretation of phenomena observed.

6. REFERENCES

1. Albertini R.J., Anderson D., Douglas G.R., Hagmar L., Hemminki K., Merlo F., Natarajan A.T., Norppa H., Shuker D.E., Tice R., Waters M.D., Aitio A. (2000). IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. International Programme on Chemical Safety. *Mutat Res.* 463(2):111-72.
2. Allocca, V., Corniello, A., De Vita, P., Ducci, D., Pizzolante, A., Romano, N. (2016). "Hydrogeological and vadose zone studies in the framework of the "Campania Trasparente" project (southern Italy)." *88th Congress of the Italian Geological Society on "Geosciences on a changing planet: Learning from the past, exploring the future",* Naples (Italy), Sept. 7th-9th, doi: 10.13140/RG.2.2.25861.81122.
3. Almeida D.L., Pavanello A., Saavedra L.P., Pereira T.S., de Castro-Prado M.A.A., de Freitas Mathias P.C. (2019). Environmental monitoring and the developmental origins of health and disease. *J Dev Orig Health Dis.* 27:1-8.
4. Ameer, S. S., Engström, K., Hossain, M. B., Concha, G., Vahter, M., & Broberg, K. (2017). Arsenic exposure from drinking water is associated with decreased gene expression and increased DNA methylation in peripheral blood. *Toxicology and Applied Pharmacology*, *321*, 57–66. https://doi.org/10.1016/j.taap.2017.02.019.
5. Angerer, J., Ewers, U., Wilhelm, M., (2007). Human biomonitoring: state of the art. *Int. J. Hyg. Environ. Health* 210, 201-228.
6. Anway, M.D., Cupp, A.S., Uzumcu, M., Skinner, M.K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science*. 308:1466-1469.
7. Anway, M.D., Skinner, M.K. (2006). Epigenetic transgenerational actions of endocrine disruptors. *Endocrinology*. 147: S43-S49.
8. Baccarelli, A., Wright, R.O., Bollati, V. (2009). Rapid DNA methylation changes after exposure to traffic particles. *Am J Respir Crit Care Med*. 179(7):572-578.
9. Balassone, G., Aiello, G., Barra, D., Cappelletti, P., De Bonis, A., Donadio, C., Guida, M., Melluso, L., Morra, V., Parisi, R., Pennetta, M., Siciliano, A. (2016). Effects of anthropogenic activities in a Mediterranean coastland: the case study of the Falerno-Domitio littoral in Campania, Tyrrhenian Sea (southern Italy). *Mar Pollut Bull*.112(1-2):271-290. doi: 10.1016/j.marpolbul.2016.08.004.
10. Balmain, A., Gray, J., Ponder, B. (2003). The genetics and genomics of cancer. *Nat Genet.* 33(Suppl):238-244.
11. Bailey, KA, Wu, MC, Ward, WO, Smeester, L, Rager JE, García-Vargas, G, Del Razo, LM, Drobná, Z, Stýblo, M, Fry, RC. (2013). Arsenic and the epigenome: interindividual differences in arsenic metabolism related to distinct patterns of DNA methylation. *J Biochem Mol Toxicol*. 27(2):106-15. doi:10.1002/jbt.21462.
12. Barba, M., Mazza, A., Guerriero, C., Di Maio, M., Romeo, F., Maranta, P…Giordano, A. (2011). Wasting lives: The effects of toxic waste exposure on health - The case of

Campania, Southern Italy. *Cancer Biology and Therapy*, *12*(2), 106–111. https://doi.org/10.4161/cbt.12.2.16910.

13. Basile, A., Sorbo, S., Aprile, G., Conte, B., Castaldo Cobianchi, R., Pisani, T., & Loppi, S. (2009). Heavy metal deposition in the Italian "triangle of death" determined with the moss Scorpiurum circinatum. *Environmental Pollution*.*157*(8–9),2255–2260. https://doi.org/10.1016/j.envpol.2009.04.001.
14. Belpomme, D., Irigaray, P., Hardell, L., Clapp, R., Montagnier, L., Epstein, S., & Sasco, A. J. (2007a). The multitude and diversity of environmental carcinogens. *Environmental Research*. *105*(3), 414–429. https://doi.org/10.1016/j.envres.2007.07.002.
15. Belpomme, D., Irigaray, P., Sasco, A. J., Newby, J. A., Howard, V., Clapp, R., & Hardell, L. (2007b). The growing incidence of cancer: Role of lifestyle and screening detection (Review). *International Journal of Oncology*. *30*(5), 1037–1049. https://doi.org/10.3892/ijo.30.5.1037.
16. Bernert, J.T., Turner, W.E., Patterson Jr., D.G., Needham, L.L. (2007). Calculation of serum ''total lipid'' concentrations for the adjustment of persistent organohalogen toxicant measurements in human samples. *Chemosphere*. 68, 824–831. https://doi.org/10.1016/j.chemosphere.2007.02.043.
17. Bianchi, F., Comba, P., Martuzzi, M., Palombino, R., & Pizzuti, R. (2004). Italian "Triangle of death." *Lancet Oncology*, *5*(12), 710. https://doi.org/10.1016/S1470-2045(04)01645-6.
18. Boffetta, P., Catalani, S., Tomasi, C., Pira, E., Apostoli, P. (2018). Occupational exposure to polychlorinated biphenyls and risk of cutaneous melanoma: a meta-analysis. *Eur J Cancer Prev*. 27(1):62-69. doi: 10.1097/CEJ.0000000000000316.
19. Bollati, V, Baccarelli, A, Hou, L, Bonzini, M, Fustinoni, S, Cavallo, D, Byun, HM, Jiang, J, Marinelli, B, Pesatori, AC, Bertazzi, PA, Yang, AS. (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer Res*. 67(3):876-80. doi:10.1158/0008-5472.
20. Bollati, V, Marinelli, B, Apostoli, P, Bonzini, M, Nordio, F, Hoxha, M, Pegoraro, V, Motta, V, Tarantini, L, Cantone, L, Schwartz, J, Bertazzi, PA, Baccarelli, A. (2010) Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. *Environ Health Perspect*. 118(6):763-8.doi: 10.1289/ehp.0901300.
21. Borghini, A., Faita, F., Mercuri, A., Minichilli, F., Bustaffa, E., Bianchi, F., & Andreassi, M. G. (2016). Arsenic exposure, genetic susceptibility and leukocyte telomere length in an Italian young adult population. *Mutagenesis*, *31*(5), 539–546. https://doi.org/10.1093/mutage/gew017.
22. Brambilla P, Gerthoux P, Beretta C, Cazzaniga M, Ferrari E, Mocarelli P. (1998). Seveso: biological monitoring of dioxin pollution. An example of environmental medicine. Multicentric and interdisciplinary applied research]. *G Ital Med Lav Ergon*. 20(3):197-9.

23. Brasseur, C., Pirard, C., Scholl, G., De Pauw, E., Viel, J.F., Shen, L., Reiner, E.J., Focant, J.F., (2014). Levels of dechloranes and polybrominated diphenyl ethers (PBDEs) in human serum from France. *Environ Int*. 65, 33-40. https://doi.org/10.1016/j.envint.2013.12.014.
24. Brucker, N., Moro, A. M., Charão, M. F., Durgante, J., Freitas, F., Baierle, M., Garcia, S. C. (2013). Biomarkers of occupational exposure to air pollution, inflammation and oxidative damage in taxi drivers. *Science of the Total Environment, 463–464*, 884–893. https://doi.org/10.1016/j.scitotenv.2013.06.098.
25. Buononato E.V., De Luca D., Galeandro I.C., Congedo M.L., Cavone D., Intranuovo G., Guastadisegno C.M., Corrado V., Ferri G.M. (2016). Assessment of environmental and occupational exposure to heavy metals in Taranto and other provinces of Southern Italy by means of scalp hair analysis. *Environ Monit Assess*. 188(6):337.
26. Calaf, G.M., Ponce-Cusi, R., Aguayo, F., Muñoz, J.P., Bleak, T.C. (2020). Endocrine disruptors from the environment affecting breast cancer. *Oncol Lett*. 20(1):19-32. doi: 10.3892/ol.2020.11566.
27. Calafat, A.M., Ye, X., Silva, M.J., Kuklenyik, Z., Needham, L.L. (2006). Human exposure assessment to environmental chemicals using biomonitoring. *Int J Androl.* (1):166-71. doi: 10.1111/j.1365-2605.2005.00570.
28. Cantoni, R. (2016). The waste crisis in Campania, South Italy: A historical perspective on an epidemiological controversy. *Endeavour*. *40*(2), 102–113. https://doi.org/10.1016/j.endeavour.2016.03.003.
29. Cao, J., Fan, T., Li, W., Xiao, S. (2019). Association study between plasma levels of polychlorinated biphenyls and risk of cutaneous malignant melanoma. *Environ Int.* 126:298-301. doi: 10.1016/j.envint.2019.02.014.
30. Cao, Y. (2015). Environmental pollution and DNA methylation: carcinogenesis, clinical significance, and practical applications. *Front Med.* 9(3):261-74. doi:10.1007/s11684-015-0406-y.
31. Cardenas, A., Rifas-Shiman, S.L., Agha, G., Hivert, M.-F., Litonjua, A.A., DeMeo, D.L., Lin, X., Amarasiriwardena, C.J., Oken, E., Gillman, M.W. Baccarelli, A.A. (2017). Persistent DNA methylation changes associated with prenatal mercury exposure and cognitive performance during childhood. *Sci. Rep*. 7(1) 288.
32. Center for Disease Control and Prevention (CDC). (2005). Third National Report on Human Exposure to Environmental Chemicals.
33. Chen, M., Li, J., Zhuang, C., Cai, Z. (2017). Increased lncRNA ABHD11-AS1 represses the malignant phenotypes of bladder cancer. *Oncotarget*. 8(17):28176-28186. doi: 10.18632/oncotarget.14945.
34. Chighizola C, Meroni PL. (2012). The role of environmental estrogens and autoimmunity. *Autoimmun Rev*. 11(6-7): A493-501.

35. Chin-Chan, M., Navarro-Yepes, J., Quintanilla-Vega, B. (2015). Environmental pollutants as risk factors for neurodegenerative disorders: Alzheimer and Parkinson diseases. *Front Cell Neurosci*. 9:124. doi: 10.3389/fncel.2015.00124.
36. Clapp, R. W., Jacobs, M. M., & Loechler, E. L. (2008). Environmental and occupational causes of cancer: New evidence 2005-2007. *Reviews on Environmental Health*. 23(1), 1–37. https://doi.org/10.1515/reveh.2008.23.1.1.
37. Clark, S. J., Harrison, J. & Frommer, M. (1995). Cpnpg methylation in mammalian cells. *Nat Genet*.10, 20-7.
38. Claus, S. P., Guillou, H., & Ellero-Simatos, S. (2016). The gut microbiota: A major player in the toxicity of environmental pollutants? *Npj Biofilms and Microbiomes*, *2* (March), 1–12. https://doi.org/10.1038/npjbiofilms.2016.3..
39. Comba, P.; Ricci, P.; Iavarone, I.; Pirastu, R.; Buzzoni, C.; Fusco, M.; Ferretti, S.; Fazzo, L.; Pasetto, R.; Zona, A.; et al. (2014). Cancer incidence in Italian contaminated sites. *Ann. Ist. Super. Sanità*, 50, 186–191.
40. Crispo, A., Barba, M., Malvezzi, M., Arpino, G., Grimaldi, M., Rosso, T., Montella, M. (2013). Cancer mortality trends between 1988 and 2009 in the metropolitan area of Naples and Caserta, Southern Italy: Results from a joinpoint regression analysis. *Cancer Biology and Therapy*, *14*(12), 1113–1122. https://doi.org/10.4161/cbt.26425
41. Decreto-Legge convertito con modificazioni dalla L. 6 febbraio 2014, n. 6 (in G.U. 8/2/2014 n. 32). Disposizioni urgenti dirette a fronteggiare emergenze ambientali e industriali ed a favorire lo sviluppo delle aree interessate. Italian Parliament, Available from: https://www.gazzettaufficiale.it/eli/id/2014/02/08/14A00744/sg
42. Decreto Legislativo 3 aprile 2006, n. 152. Norme in materia ambientale. (GU n.88 del 14-4-2006 - Suppl. Ordinario n. 96). Italian Parliament, Available from: https://www.minambiente.it/sites/default/files/archivio/normativa/dlgs_03_04_2006_152.pdf
43. De Bord, D. G., Burgoon, L., Edwards, S. W., Haber, L. T., Kanitz, M. H., Kuempel, E., …Yucesoy, B. (2015). Systems Biology and Biomarkers of Early Effects for Occupational Exposure Limit Setting. *Journal of Occupational and Environmental Hygiene*, *12*, S41–S54. https://doi.org/10.1080/15459624.2015.1060324.
44. De Felice, B., Nappi, C., Zizolfi, B., Guida, M., Sardo, A. D. S., Bifulco, G., & Guida, M. (2012). Telomere shortening in women resident close to waste landfill sites. *Gene*. 500(1), 101–106. https://doi.org/10.1016/j.gene.2012.03.040.
45. De Felip, E., Abballe, A., Casalino, F., di Domenico, A., Domenici, P., Iacovella, N., Ingelido, A.M., Pretolani, E., Spagnesi, M. (2008). Serum levels of PCDDs, PCDFs and PCBs in non-occupationally exposed population groups living near two incineration plants in Tuscany, Italy. *Chemosphere*.72,25-33. https://doi.org/10.1016/j.chemosphere.2008.02.046.
46. De Felip, E., Bianchi, F., Bove, C., Cori, L., D'Argenzio, A., D'Orsi, G., di Domenico, A. (2014). Priority persistent contaminants in people dwelling in critical areas of

Campania Region, Italy (SEBIOREC biomonitoring study). *Science of the Total Environment*, *487*(1), 420–435. https://doi.org/10.1016/j.scitotenv.2014.04.016.

47. de Groot, P., Munden, R.F. (2012). Lung cancer epidemiology, risk factors, and prevention. *Radiol Clin North Am*. 50(5):863-76. doi:10.1016/j.rcl.2012.06.006.
48. De Roma, A., Neola, B., Serpe, F. P., Sansone, D., Picazio, G., Cerino, P., & Esposito, M. (2017). Land Snails (Helix aspersa) as Bioindicators of Trace Element Contamination in Campania (Italy). *OA Lib*. 04(02), 1–12. https://doi.org/10.4236/oalib.1103339.
49. Demanelis, K., Argos, M., Tong, L., Shinkle, J., Sabarinathan, M., Rakibuz-Zaman, M., … Pierce, B. L. (2019). Association of arsenic exposure with whole blood DNA methylation: An epigenome-wide study of Bangladeshi adults. *Environmental Health Perspectives*, *127*(5), 1–15. https://doi.org/10.1289/EHP3849.
50. Dennis, K. K., Marder, E., Balshaw, D. M., Cui, Y., Lynes, M. A., Patti, G. J., Barr, D. B. (2017). Biomonitoring in the era of the exposome. *Environmental Health Perspectives*, *125*(4), 502–510. https://doi.org/10.1289/EHP474.
51. Dent BD. (1999). Cartography: Thematic Map Design. *McGraw-Hill*.
52. Dinis, M. D. L., & Fiúza, A. (2009). Methodology for Exposure and Risk Assessment in Complex Environmental Pollution Situations. *Chapter in NATO Security through Science Series C: Environmental Security*.111–132. https://doi.org/10.1007/978-90-481-2335-3_7.
53. Ducci, D., Albanese, S., Boccia, L., Celentano, E., Cervelli, E., Corniello, A., Fagnano, M. (2017). An integrated approach for the environmental characterization of a wide potentially contaminated area in southern Italy. *International Journal of Environmental Research and Public Health*. 14(7). https://doi.org/10.3390/ijerph14070693.
54. Ducci D, Della Morte R, Mottola A, Onorati G, Pugliano G. (2019). Nitrate trends in groundwater of the Campania region (southern Italy). *Environ Sci Pollut Res Int.*26(3):2120-2131. doi: 10.1007/s11356-017-0978-y. Epub 2017 Dec 14.
55. EEA. (n.d.). *EEA Risk assessement part1-section1.pdf*.
56. Espín-Pérez, A., Krauskopf, J., Chadeau-Hyam, M., van Veldhoven, K., Chung, F., Cullinan, P., de Kok, T. M. C. M. (2018). Short-term transcriptome and microRNAs responses to exposure to different air pollutants in two population studies. *Environmental Pollution*, *242*, 182–190. https://doi.org/10.1016/j.envpol.2018.06.051..
57. Esposito, M., Serpe, F. P., Diletti, G., Messina, G., Scortichini, G., La Rocca, C., … Monda, M. (2014). Serum levels of polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans and polychlorinated biphenyls in a population living in the Naples area, southern Italy. *Chemosphere*, *94*, 62–69. https://doi.org/10.1016/j.chemosphere.2013.09.013.
58. Esteban M, Castaño A. (2009) Non-invasive matrices in human biomonitoring: a review. *Environ Int*. 35(2):438-49. doi: 10.1016/j.envint.2008.09.003.
59. Ehrlich, M., Gama-Sosa, MA., Huang, LH., Midgett, RM., Kuo, KC., McCune, RA., Gehrke, C. (1982) Amount and distribution of 5-methylcytosine in human DNA from

different types of tissues of cells. *Nucleic Acids Res*. 10(8):2709-21. doi:10.1093/nar/10.8.2709 .

60. Esteller, M., Herman, JG. (2002) Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J Pathol*. 2002 Jan;196(1):1-7. doi:10.1002/path.1024.
61. EU. (2013). Decision No 1386/2013/EU of the European Parliament and of the Council of 20 November 2013 on a General Union Environment Action Programme to 2020 'Living well, within the limits of our planet'. *Official Journal of the European Union*. L. 354/171.
62. EU 679/2016 del Parlamento Europeo e del Consiglio. Relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati). https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex%3A32016R0679.
63. Fairman R., Mead C.D. and Williams W.P. (1998), Environmental Risk Assessment – Approaches, Experiences and Information Sources. Published by *European Environment Agency – EEA*. Online: http://reports.eea.europa.eu/GH-07-97-595-EN-C2/en/riskindex.html.
64. Falzon, L.C., Lechner, I., Chantziaras, I., Collineau, L., Courcoul, A., Filippitzi, M.E., Laukkanen-Ninios, R., Peroz, C., Pinto Ferreira, J., Postma, M., Prestmo, P.G., Phythian, C.J., Sarno, E., Frazzoli, C., Bocca, B., Mantovani, A. (2015). The One Health Perspective in Trace Elements Biomonitoring. *J Toxicol Environ Health B Crit Rev*. 18(7-8):344-70. doi: 10.1080/10937404.2015.1085473.
65. Fang, J., Nyberg, E., Bignert, A., & Bergman, Å. (2013). Temporal trends of polychlorinated dibenzo-p-dioxins and dibenzofurans and dioxin-like polychlorinated biphenyls in mothers' milk from Sweden, 1972-2011. *Environment International*. 60, 224–231. https://doi.org/10.1016/j.envint.2013.08.019.
66. Fatima P, Debnath BC, Hossain MM, Rahman D, Banu J, Begum SA, Rahman MW. (2010) Relationship of blood and semen lead level with semen parameter. *Mymensingh Med J*. 19(3):405-14.
67. Fatima P, Hossain MM, Rahman D, Rahman MW, Mugni CR, Sumon GM, Hossain HB, Hossain HN. (2015) Association of Blood and Semen Lead and Zinc Level with Semen Parameter in the Male Partner of Infertile Couple. *Mymensingh Med J*. 24(3):537-41.
68. Fazzo L, de Santis M, Mitis F, Benedetti M, Martuzzi M, Comba P, Fusco M. (2011). Ecological studies of cancer incidence in an area interested by dumping waste sites in Campania (Italy). *Ann. Ist. Super. Sanità*. 47, 181–191.
69. Feil, R., Fraga, M.F. (2012). Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet*. 13(2):97-109. doi: 10.1038/nrg3142.
70. Forouzanfar, M. H., Afshin, A., Alexander, L. T., Biryukov, S., Brauer, M., Cercy, K…Zhu, J. (2016). Global, regional, and national comparative risk assessment of 79

behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1659–1724. https://doi.org/10.1016/S0140-6736(16)31679-8.

71. Forum, R. A. (2019). Guidelines for Human Exposure Assessment.
72. Forrest, M.S., Lan, Q., Hubbard, A.E., Zhang, L., Vermeulen, R., Zhao, X., Smith M.T., (2005) Discovery of novel biomarkers by microarray analysis of peripheral blood mononuclear cell gene expression in benzene-exposed workers. *Environ Health Perspect*. 113:801-807.
73. Frank, A. L., & Joshi, T. K. (2014). The global spread of asbestos. *Annals of Global Health*. 80(4), 257–262. https://doi.org/10.1016/j.aogh.
74. Fraumeni, J. F. (2011). Molecular epidemiology: principles and practices. Foreword. In *IARC scientific publications*.
75. Fusco, M., De Angelis, R., Senatore, G., Zigon, G., Rossi, S. (2013). Estimates of cancer burden in Campania. *Tumori*. 99(3):374-81. doi: 10.1700/1334.14802.
76. Ganzleben, C., Antignac, J.P., Barouki, R., Castaño, A., Fiddicke, U., Klánová, J., Lebret, E., Olea, N., Sarigiannis, D., Schoeters, G.R., Sepai, O., Tolonen, H., Kolossa-Gehring, M. (2017). Human biomonitoring as a tool to support chemicals regulation in the European Union. *Int J Hyg Environ Health*. 220(2 Pt A):94-97. doi: 10.1016/j.ijheh.2017.01.007.
77. Garrison, P.M., Tullis, K., Aarts, J.M., Brouwer, A., Giesy, J.P., Denison M.S. (1996) Species-specific recombinant cell lines as bioassay systems for the detection of 2,3,7,8-tetrachlorodibenzo-p-dioxin-like che-micals. *Fundam Appl Toxicol*. 30: 194-203.
78. Gilbert KM, Blossom SJ, Erickson SW, Reisfeld B, Zurlinden TJ, Broadfoot B, West K, Bai S, Cooney CA. (2016) Chronic exposure to water pollutant trichloroethylene increased epigenetic drift in CD4(+) T cells. *Epigenomics*. 8(5):633-49.
79. Gruzieva, O., Merid, SK., Gref, A., Gajulapuri, A., Lemonnier, N., Ballereau, S., Gigante, B., Kere, J., Auffray, C., Melén, E., Pershagen, G. (2017). Exposure to Traffic-Related Air Pollution and Serum Inflammatory Cytokines in Children. *Environ Health Perspect*. 125(6): 067007.
80. Hay A. (1980). Health monitoring endangered at Séveso. *Nature*. 283(5744):237.
81. Hassan L, Pecht T, Goldstein N, Haim Y, Kloog I, Yarza S, Sarov B, Novack V. (2019) The effects of ambient particulate matter on human adipose tissue. *J Toxicol Environ Health A*.26:1-13.
82. Heibati, B, Godri Pollitt, KJ, Charati, JY, Ducatman, A, Shokrzadeh, M, Karimi, A, Mohammadyan, M. (2018). Biomonitoring-based exposure assessment of benzene, toluene,ethylbenzene and xylene among workers at petroleum distribution facilities. *Ecotoxicol Environ Saf.* 149:19-25. doi: 10.1016/j.ecoenv.2017.10.070.
83. Hou, L., Zhang, X., Wang, D., Baccarelli, A. (2012). Environmental chemical exposures and human epigenetics. *Int J Epidemiol*. 41:79-105.

84. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., et al. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 13: 86. doi: 10.1186/1471-2105-13-86.
85. Hunter, DJ. (2005). Gene-environment interactions in human diseases. *Nat Rev Genet*. 6:287-298.
86. ICOH. (2002). International Code of Ethics for Occupational Health Professionals, Internationl Commission on Occupational Health.
87. Ingelido, AM, Abate, V, Abballe, A, Albano, FL, Battista, T, Carraro, V, Conversano, M, Corvetti, R, De Luca, S, Franchini, S, Fulgenzi, AR, Giambanco, L, Iacovella, N, Iamiceli, AL, Maiorana, A, Maneschi, F, Marra, V, Pirola, F, Porpora, MG, Procopio, E, Suma, N, Valentini, S, Valsenti, L, Vecchiè, V, De Felip, E. (2017). Concentrations of polychlorinated dibenzodioxins, polychlorodibenzofurans, and polychlorobiphenyls in women of reproductive age in Italy: A human biomonitoring study. *Int J Hyg Environ Health*. 220(2 Pt B):378-386.
88. Irigaray, P., Newby, J. A., Clapp, R., Hardell, L., Howard, V., Montagnier, L…Belpomme, D. (2007). Lifestyle-related factors and environmental agents causing cancer: An overview. *Biomedicine and Pharmacotherapy*, *61*(10), 640–658. https://doi.org/10.1016/j.biopha.2007.10.006.
89. Javierre, B.M., Esteller, M. & Ballestar, E. (2008). Epigenetic connections between autoimmune disorders and haematological malignancies. *Trends Immunol*. 29, 616–623 .
90. Jirtle, R., Skinner, M.K. (2007). Environmental epigenomics and disease susceptibility. *Nat Rev Genet.* 8(4):253-262.
91. Joas, A., Schwedler. G., Choi, J., Kolossa-Gehring, M. (2017). Human biomonitoring: Science and policy for a healthy future. *Int J Hyg Environ Health*. 220(2 Pt A):299-304. doi: 10.1016/j.ijheh.2017.01.013.
92. Joseph, P. G., Pare, G. & Anand, S. S. (2013). Exploring gene-environment relationships in cardiovascular disease. *Can J Cardiol*, 29, 37-45.
93. Juloski, J.T., Rakic, A., Ćuk, V.V., Ćuk, V.M., Stefanović, S., Nikolić, D., Janković, S., Trbovich, A.M., De Luka, S.R. (2020). Colorectal cancer and trace elements alteration. *J Trace Elem Med Biol*. 59:126451. doi: 10.1016/j.jtemb.2020.126451.
94. Karbowska, B. (2016). Presence of thallium in the environment: sources of contaminations, distribution and monitoring methods. *Environ Monit Assess*. 188(11): 640. doi: 10.1007/s10661-016-5647-y.
95. Kreiss, K. (2007). Emerging opportunities to prevent occupational lung disease. *Occupational and Environmental Medicine*, *64*(8), 499–500. https://doi.org/10.1136/oem.2006.029918.
96. Krishanu (2011). Reproductive and Developmental Toxicology.
97. Kumarathasan, P., Vincent, R., Blais, E., Bielecki, A., Guénette, J., Filiatreault, A., Dales, R. (2018). Cardiovascular and inflammatory mechanisms in healthy humans exposed to

air pollution in the vicinity of a steel mill. *Particle and Fibre Toxicology*, *15*(1), 1–17. https://doi.org/10.1186/s12989-018-0270-4.

98. La Rocca, J., Binder A.M., McElarath, T.F., Michels, K.B. (2014). The impact of first trimenster phthalate and phenol exposure on IGF2/H19 genomic imprinting and birth outcomes. *Environ Res*. 133:396-406.
99. La Rocca J, Binder AM, McElrath TF, Michels KB. (2016). First-Trimester Urine Concentrations of Phthalate Metabolites and Phenols and Placenta miRNA Expression in a Cohort of U.S. Women. *Environ Health Perspect*. 124(3):380-7. doi: 10.1289/ehp.1408409.
100. Ladd-Acosta, C., & Fallin, M. D. (2016). The role of epigenetics in genetic and environmental epidemiology. *Epigenomics*, *8*(2), 271–283. https://doi.org/10.2217/epi.15.102.
101. Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N. (Nil), … Zhong, M. (2018). The Lancet Commission on pollution and health. *The Lancet*, *391*(10119), 462–512. https://doi.org/10.1016/S0140-6736(17)32345-0
102. Lane, N. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis.* 35:88-93.
103. Legambiente. Report Land of Fires. (2013). Available online: http://www.legambiente.it/sites/default/files/docs/ dossier_terradeifuochi_2013.pdf (accessed on 12 March 2016).
104. Legambiente. Rapporto Ecomafia. (2015). Available online: https://www.legambiente.it/contenuti/dossier/ ecomafia-2015 (accessed on 12 March 2016).
105. Leonel Javeres, M.N., Habib, R., Judith, N., Iqbal, M., Nepovimova, E., Kuca, K., Batool, S., Nurulain, S.M. (2020). Analysis of PON1 gene polymorphisms (rs662 and rs854560) and inflammatory markers in organophosphate pesticides exposed cohorts from two distinct populations. *Environ Res*. 191:110210. doi: 10.1016/j.envres.2020.110210.
106. Lewalter J, Korallus U, Harzdorf C, Weidemann H. (1985) Chromium bond detection in isolated erythrocytes: a new principle of biological monitoring of exposure to hexavalent chromium. *Int Arch Occup Environ Health*. 55(4):305-18. doi:10.1007/BF00377689.
107. Li, M. (2009) Sensitive digital quantification of DNA methylation in clinical samples. *Nat. Biotechnol*. 27, 858–863.
108. Li, P., Guo, S., Zhao, J., Gao, Y., & Li, Y. F. (2019a). Human Biological Monitoring of Mercury Through Hair Samples in China. *Bulletin of Environmental Contamination and Toxicology*, *102*(5), 701–707. https://doi.org/10.1007/s00128-019-02563-8.
109. Li, Q.S., Sun, Y., Wang, T. (2020). Epigenome-wide association study of Alzheimer's disease replicates 22 differentially methylated positions and 30 differentially methylated regions. *Clin Epigenetics.* 12(1):149. doi:10.1186/s13148-020-00944-z.
110. Li, X., Brejnrod, A. D., Ernst, M., Rykær, M., Herschend, J., Olsen, N. M. C., Sørensen, S. J. (2019b). Heavy metal exposure causes changes in the metabolic health-associated

gut microbiome and metabolites. *Environment International*. 454–467. https://doi.org/10.1016/j.envint.2019.02.048.

111. Lionetto, M. G., Caricato, R., & Giordano, M. E. (2019). Pollution Biomarkers in Environmental and Human Biomonitoring. *The Open Biomarkers Journal*. 9(1), 1–9. https://doi.org/10.2174/1875318301909010001.
112. Liu S.S., Zheng H.X., Jiang H.D., He J., Yu Y., Qu Y.P., Yue L., Zhang Y., Li Y. (2012) Identification and characterization of a novel gene, c1orf109, encoding a CK2 substrate that is involved in cancer cell proliferation. *J Biomed Sci*. 19(1):49. doi: 10.1186/1423-0127-19-49.
113. Lovreglio, P., Barbierato, M., Crociata, F., Tomao, E., Diomede, L., Gallo, E., Scaramuzzo, P., Drago. I., Paganelli, M., Apostoli, P., Soleo, L. (2018). Biological monitoring of exposure to polycyclic aromatic hydrocarbons and to metallic elements in Italian Navy workers operating near the industrial area in Taranto (South Italy). *Med Lav*. 110(5):339-362. doi: 10.23749/mdl.v110i5.7123.
114. Louro, H., Heinälä, M., Bessems, J., Buekers, J., Vermeire, T., Woutersen, M., van Engelen, J., Borges, T., Rousselle, C., Ougier, E., Alvito, P., Martins, C., Assunção, R., Silva, M.J., Pronk, A., Schaddelee-Scholten, B., Del Carmen Gonzalez, M., de Alba, M.,Castaño, A., Viegas, S., Humar-Juric, T., Kononenko, L., Lampen, A., Vinggaard, A.M., Schoeters, G., Kolossa-Gehring, M., Santonen, T. (2019). Human biomonitoring in health risk assessment in Europe: Current practices and recommendations for the future. *Int J Hyg Environ Health*. 222(5):727-737. doi: 10.1016/j.ijheh.2019.05.009.
115. Lucchini, R.G., Guazzetti, S., Renzetti, S., Conversano, M., Cagna, G., Fedrighi, C., Giorgino, A., Peli, M., Placidi, D., Zoni, S., Forte, G., Majorani, C., Pino, A., Senofonte, O., Petrucci, F., Alimonti, A. (2019). Neurocognitive impact of metal exposure and social stressors among schoolchildren in Taranto, Italy. *Environ Health*. 18(1):67. doi: 10.1186/s12940-019-0505-3.
116. Macchi C, Ferri N, Favero C, Cantone L, Vigna L, Pesatori AC, Lupo MG, SirtoriCR, Corsini A, Bollati V, Ruscica M. (2019) Long-term exposure to air pollution raises circulating levels of proprotein convertase subtilisin/kexin type 9 in obese individuals. *Eur J Prev Cardiol*. 26(6):578-588.
117. Manno, M., Viau, C., Cocker, J., Colosio, C., Lowry, L., Mutti, A., … Wang, S. (2010). Biomonitoring for occupational health risk assessment (BOHRA). *Toxicology Letters*. 192(1), 3–16. https://doi.org/10.1016/j.toxlet.2009.05.001..
118. Manno, M., Sito, F., Licciardi, L., Ethics in biomonitoring for occupational health. (2014). *Toxicology Letters*. 231.111-121.
119. Margetts, B. M., & Pietinen, P. (1997). European Prospective Investigation into Cancer and Nutrition: Validity studies on dietary assessment methods. *International Journal of Epidemiology*. 26 (SUPPL. 1), 1S – 5. https://doi.org/10.1093/ije/26.suppl_1.s1.

120. Marquès, M., Domingo, J.L. (2019). Concentrations of PCDD/Fs in Human Blood: A Review of Data from the Current Decade. *Int J Environ Res Public Health*. 16(19):3566. doi: 10.3390/ijerph16193566.
121. Martin, E.M., Fry, R.C. (2016). A cross-study analysis of prenatal exposures to environmental contaminants and the epigenome: support for stress-responsive transcription factor occupancy as a mediator of gene-specific CpG methylation patterning. *Environ Epigenet*. 2(1):dvv011. doi: 10.1093/eep/dvv011.
122. Martino, D., & Saffery, R. (2015). Characteristics of DNA methylation and gene expression in regulatory features on the Infinium 450k Beadchip. *BioRxiv*, 032862. https://doi.org/10.1101/032862.
123. Martuzzi, M., Mitis, F., Bianchi, F., Minichilli, F., Comba, P., & Fazzo, L. (2009). Cancer mortality and congenital anomalies in a region of Italy with intense environmental pressure due to waste. *Occupational and Environmental Medicine*. 66(11), 725–732. https://doi.org/10.1136/oem.2008.044115.
124. Mauriello, M.C., Sbordone, C., Montuori, P., Alfano, R., Triassi, M., Iavicoli, I., Manno, M. (2017). Biomonitoring of toxic metals in incinerator workers: A systematic review. *Toxicol Lett*. 272:8-28. doi: 10.1016/j.toxlet.2017.02.021.
125. Maximilian, J., Brusseau, M.L., Glenn, E.P. and Matthias, A.D. (2019) Chapter 25 - Pollution and Environmental Perturbations in the Global System. *Environmental and Pollution Science (Third Edition).* 457-476. https://doi.org/10.1016/B978-0-12-814719-1.00025-2.
126. Mazza, A., Piscitelli, P., Falco, A., Santoro, M. L., Colangelo, M., Imbriani, G., Colao, A. (2018). Heavy environmental pressure in campania and other italian regions: A short review of available evidence. *International Journal of Environmental Research and Public Health*, *15*(1), 1–12. https://doi.org/10.3390/ijerph15010105.
127. McHale, C.M., Zhang, L., Lan, Q., Li, G., Hubbard, A.E., Forrest, M.S., Vermeulen, R., Chen, J., Shen, M., Rappaport, S.M., Rothman, N. (2009). Changes in peripheral blood transcriptome associated with occupational benzene exposure identified by cross-comparison on two microarray platforms. *Genomics*. 93:343-349.
128. McHale, C.M., Zhang, L., Hubbard, A.E.., Smith, MT. (2010). Toxicogenomic profiling of chemically exposed humans in risk assessment. *Mutat Res.* 705(3):172-83. doi: 10.1016/j.mrrev.2010.04.001.
129. Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Lander, E. S. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, *454*(7205), 766–770. https://doi.org/10.1038/nature07107.
130. Migliore L. 2018. Genomica e Mutagenesi Ambientale. *Zanichelli*.
131. Miller, K. Michael Pollard, Christine G. Parks, Dori R. Germolec, Patrick S.C. Leung, Carlo Selmi, Michael C. Humble, Noel R. Rose. (2012). Criteria for Environmentally Associated Autoimmune Diseases. *J Autoimmun*; 39(4): 253–258. doi: 10.1016/j.jaut.2012.05.001.

132. Miniero, R., Ingelido, A.M., Abballe, A., di Domenico, A., Valentini, S., Marra, V., Barbieri, P.G., Garattini, S., Speziani, F., De Felip, E. (2017). Occupational exposure to PCDDs, PCDFs, and DL-PCBs in metallurgical plants of the Brescia (Lombardy Region, northern Italy) area. *Chemosphere.* 166, 418 - 421. https://doi.org/10.1016/j.chemosphere.2016.09.134.
133. Minolfi, G., Albanese, S., Lima, A., Tarvainen, T., Fortelli, A., & De Vivo, B. (2018). A regional approach to the environmental risk assessment - Human health risk assessment case study in the Campania region. *Journal of Geochemical Exploration*, *184*(October 2017), 400–416. https://doi.org/10.1016/j.gexplo.2016.12.010.
134. Mitchell, C., Schneper, L.M., Notterman, D.A. (2016). DNA methylation, early life environment, and health outcomes. *Pediatr Res*. 79(1-2):212-9. doi:10.1038/pr.2015.193.
135. Moran, S., Arribas, C., Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics*. 8(3):389-99. doi: 10.2217/epi.15.114.
136. Morgan, HD., Sutherland, HGE., Martin, DIK., Whitelaw, E. (1999). Epigenetic inheritance at the Agouti locus in the mouse. *Nature Genet*. 23:314-318.
137. Morgan, M., Deoraj, A., Felty, Q., Roy, D. (2017). Environmental estrogen-like endocrine disrupting chemicals and breast cancer. *Mol Cell Endocrinol*. 457:89-102. doi: 10.1016/j.mce.2016.10.003.
138. Moura, M. De, & Houten, B. Van. (2010). Review Article. *Environmental and Molecular Mutagenesis*. 405(April), 391–405. https://doi.org/10.1002/em.
139. Muenyi, C. S., Carrion, S. L., Jones, L. A., Kennedy, L. H., Slominski, A. T., Sutter, C.H., Sutter, T.R. (2014). Effects of in utero exposure of C57BL/6J mice to 2,3,7,8-tetrachlorodibenzo-p-dioxin on epidermal permeability barrier development and function. *Environ Health Perspect*.122(10):1052-8. doi: 10.1289/ehp.1308045.
140. Musmeci L, Bellino M, Cicero MR, Falleni F, Piccardi A, Trinca S. (2010). The impactmeasure of solid waste management on health: the hazard index. *Ann Ist Super Sanità*. 46(3):293-8.
141. Naha, N., Bhar, R. B., Mukherjee, A., & Chowdhury, A. R. (2005). Structural alteration of spermatozoa in the persons employed in lead acid battery factory. *Indian Journal of Physiology and Pharmacology*, *49*(2), 153–162.
142. Nasta, P., Palladino, M., Sica, B., Pizzolante, A., Trifuoggi, M., Toscanesi, M…Romano, N. (2020). Evaluating pedotransfer functions for predicting soil bulk density using hierarchical mapping information in Campania, Italy. *Geoderma Regional, 21,* e00267. https://doi.org/10.1016/j.geodrs.2020.e00267.
143. Nawrot, T. S., Saenen, N. D., Schenk, J., Janssen, B. G., Motta, V., Tarantini, L… Bollati, V. (2018). Placental circadian pathway methylation and in utero exposure to fine particle air pollution. *Environment International*. 114: 231–241. https://doi.org/10.1016/j.envint.2018.02.034.

144. Ning, B., Zhenqiang, S., Mei, N., Hong, H., Deng, H., Shi, L., Fuscoe, J. C., Tolleson, W. (201) Toxicogenomics and Cancer Susceptibility: Advanced with Next-Generation Sequencing. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev*. 32(2):121-158.
145. National Research Council (US) Committee on Applications of Toxicogenomic Technologies to Predictive Toxicology. (2007). Applications of Toxicogenomic Technologies to Predictive Toxicology and Risk Assessment. *Washington (DC): National Academies Press (US).*
146. NRC. (2008). Science and Decisions: Advancing risk assessment.
147. Onul NM, Biletska EM, Stus VP, Polion MY. (2018). The role of lead in the etiopathogenesis of male fertility reduction. *Wiad Lek*. 71(6):1155-1160.
148. Pacini, N., Abate, V., Brambilla, G., De Felip, E., De Filippis, S. P., De Luca, S…Iamiceli, A. L. (2013). Polychlorinated dibenzodioxins, dibenzofurans, and biphenyls in freshwater fish from Campania Region, southern Italy. *Chemosphere*, *90*(1), 80–88. https://doi.org/10.1016/j.chemosphere.2012.07.039.
149. Pastor, W. A., Aravind, L. & Rao, A. (2013). Tetonic shift: Biological roles of tet proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol*. 14, 341-56.
150. Pembrey, ME. (2006). Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet.* 14:159-166.
151. Perera, F.P. (1997). Environment and Cancer: who are susceptible? *Science*. 278:1068-1073.
152. Perera F, Tang WY, Herbstman J, Tang D, Levin L, Miller R, Ho SM. (2009) Relation of DNA methylation of 5'-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. *PLoS One*. 4(2):e4488. doi: 10.1371/journal.pone.0004488.
153. Phillips, D.L., Pirkle, J.L., Burse, V.W., Bernert, J.T. Jr, Henderson, L.O., Needham, L.L., (1989). Chlorinated hydrocarbon levels in human serum: effects of fasting and feeding. *Arch. Environ. Contam. Toxicol.* 18, 495–500. https://doi.org/10.1007/bf01055015.
154. Pino, A., Bocca, B., Forte, G., Majorani, C., Petrucci, F., Senofonte, O., Alimonti, A. (2018). Determination of mercury in hair of children. *Toxicol Lett*. 298:25-32. doi: 10.1016/j.toxlet.2018.06.1215.
155. Pierri, B., Buonerba, C., Coppola, A., Pizzolante, A., Di Stasio, A., Cerino, P. (2020). Population-based human biomonitoring in the ‘ Land of Fires ’ area: innovations in study design and procedures. *Future Sci OA*. doi:10.2144/fsoa-2020-0164.
156. Pieterse, B, Felzel, E, Winter, R, van der Burg, B, Brouwer, A. (2013). PAH-CALUX, an optimized bioassay for AhR-mediated hazard identification of polycyclic aromatic hydrocarbons (PAHs) as individual compounds and in complex mixtures. *Environ Sci Technol.* 47(20):11651-9. doi: 10.1021/es403810w.
157. Pirastu, R., Pasetto, R., Zona, A., Ancona, C., Iavarone, I., Martuzzi, M., Comba, P. (2013). The health profile of populations living in contaminated sites: Sentieri approach.

Journal of Environmental and Public Health, *2013*. https://doi.org/10.1155/2013/939267.

158. Pirastu, R., Ricci, P., Comba, P., Bianchi, F., Biggeri, A., Conti, S… Crocetti, E. (2014). SENTIERI Project: discussion and conclusions | Discussione e conclusioni del Progetto SENTIERI. *Epidemiologia e Prevenzione*, *38*(2), 125–133.
159. Plusquin, M., Guida, F., Polidoro, S., Vermeulen, R., Raaschou-Nielsen, O., Campanella, G., Chadeau-Hyam, M. (2017). DNA methylation and exposure to ambient air pollution in two prospective cohorts. *Environment International*.108, 127–136. https://doi.org/10.1016/j.envint.2017.08.006.
160. Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. *Nature Biotechnology*. 28(10), 1057–1068. https://doi.org/10.1038/nbt.1685.
161. Prüss-Üstün, A., Bonjour, S., & Corvalán, C. (2008). The impact of the environment on health by country: A meta-synthesis. *Environmental Health: A Global Access Science Source*. 7, 1–10. https://doi.org/10.1186/1476-069X-7-7.
162. Prüss-Ustün, A., Neira, M. (2016). Preventing disease through healthy environments: A global assessment of the environmental burden of disease. *Toxicology Letters*, *259*, S1. https://doi.org/10.1016/j.toxlet.2016.07.028.
163. Pruss-Ustun, A., Wolf, J., Corvalan, C., Bos, R., Neira, M. (2016). Preventing disease through health environments:a global assessment of the burden of disease from environmental risks. *World Health Organization, Geneva, Switzerland*. https://www.who.int/quantifying ehimpacts/publications/preventing-disease/en/.
164. Qiao, X., Lv, S.X., Qiao, Y., Li, Q.P., Ye, B., Wang, C.C., Miao, L.(2018). Long noncoding RNA ABHD11-AS1 predicts the prognosis of pancreatic cancer patients and serves as a promoter by activating the PI3K-AKT pathway. *Eur Rev Med Pharmacol Sci*. 22(24):8630-8639. doi: 10.26355/eurrev_201812_16627.
165. Qu, C., Albanese, S., Li, J., Cicchella, D., Zuzolo, D., Hope, D., … De Vivo, B. (2019). Organochlorine pesticides in the soils from Benevento provincial territory, southern Italy: Spatial distribution, air-soil exchange, and implications for environmental health. *Science of the Total Environment*.674,159–170. doi:10.1016/j.scitotenv.2019.04.029.
166. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
167. Riboli, E. (1992). Nutrition and cancer: Background and rationale of the European prospective investigation into cancer and nutrition (EPIC). *Annals of Oncology*. 3(10), 783–791. https://doi.org/10.1093/oxfordjournals.annonc.a058097.
168. Rivezzi, G., Piscitelli, P., Scortichini, G., Giovannini, A., Diletti, G., Migliorati, G…Giani, U. (2013). A general model of dioxin contamination in breast milk: Results from a study on 94 women from the caserta and Naples Areas in Italy. *International Journal of Environmental Research and Public Health*. 10(11), 5953–5970. https://doi.org/10.3390/ijerph10115953.

169. Rossner, P., Rossnerova, A., Spatova, M., Beskid, O., Uhlirova, K., Libalova, H., … Sram, R. J. (2013). Analysis of biomarkers in a Czech population exposed to heavy air pollution. Part II: Chromosomal aberrations and oxidative stress. *Mutagenesis*. 28(1), 97–106. https://doi.org/10.1093/mutage/ges058.
170. Ruiz-Hernandez, A., Kuo, C. C., Rentero-Garrido, P., Tang, W. Y., Redon, J., Ordovas, J. M., … Tellez-Plaza, M. (2015). Environmental chemicals and DNA methylation in adults: A systematic review of the epidemiologic evidence. *Clinical Epigenetics*. 7(1). https://doi.org/10.1186/s13148-015-0055-7.
171. Rusiecki, JA., Baccarelli, A., Bollati, V. (2008). Global DNA hypomethylation is associated with high serum-persistent organic pollutants in Greenlandic Inuit. *Environ Health Perspect.* 116(11): 1547-1552.
172. Saberi HF, Pechlivanis A, Keun HC, Portengen L, Bueno-de-Mesquita HB, Heederik D, Vermeulen R. (2013). Serum metabolomic pertubations among workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). *Environ Mol Mutagen.* 54(7):558-65. doi: 10.1002/em.21802.
173. Saaty Thomas L. (2008). Decision making with the Analytic Hierarchy Process. *Int. J. Services Sciences.* Vol. 1, No. 1.
174. Saeidi, M., Shakeri, R., Marjani, A., Khajeniazi, S. (2017). Alzheimer's Disease and Paraoxonase 1 (PON1) Gene Polymorphisms. *Open Biochem J*. 11:47-55. doi: 10.2174/1874091X01711010047.
175. Salines, G. (2012). Biosurveillance humaine, biomarqueurs et biosurveillance environnementale [Human biomonitoring, biomarkers, environmental biomonitoring]. *Ann Pharm Fr*. 70(4):199-203. doi: 10.1016/j.pharma.2012.03.004.
176. Sakthivel S, Balasubramanian P, Nakamura M, Ko S, Chakraborty P. (2016). CALUX bioassay: a cost-effective rapid screening technique for screening dioxins like compounds. *Rev Environ Health*. 31(1):149-52. doi: 10.1515/reveh-2015-0078. PMID: 26943601.
177. Schuhmacher, M., Mari, M., Nadal, M., & Domingo, J. L. (2019). Concentrations of dioxins and furans in breast milk of women living near a hazardous waste incinerator in Catalonia, Spain. *Environment International*. 334–341. https://doi.org/10.1016/j.envint.2019.01.074.
178. Schug, T., Janesick, A., Blumberg, B., Heindel, J.J. (2011). Endocrine Disrupting Chemicals and Disease Susceptibility Thaddeus. *J Steroid Biochem Mol Biol*. 127(35): 204–215. doi: 10.1016/j.jsbmb.2011.08.007.
179. Schwedler G, Seiwert M, Fiddicke U, Ißleb S, Hölzer J, Nendza J, Wilhelm M, Wittsiepe J, Koch HM, Schindler BK, Göen T, Hildebrand J, Joas R, Joas A, Casteleyn L, Angerer J, Castano A, Esteban M, Schoeters G, Den Hond E, Sepai O, Exley K, Bloemen L, Knudsen LE, Kolossa-Gehring M. (2017). Human biomonitoring pilot study DEMOCOPHES in Germany: Contribution to a harmonized European approach. *Int J Hyg Environ Health*. Jun;220(4):686-696. doi: 10.1016/j.ijheh.2017.01.01.

180. Sellerino, M., Forte, G., & Ducci, D. (2019). Identification of the natural background levels in the Phlaegrean fields groundwater body (Southern Italy). *Journal of Geochemical Exploration*. 200, 181–192. https://doi.org/10.1016/j.gexplo.2019.02.007.
181. Senior, K., & Mazza, A. (2004). Italian "Triangle of death" linked to waste crisis. *Lancet Oncology*. 5(9), 525–527. https://doi.org/10.1016/S1470-2045(04)01561-X.
182. Serrano Cardona, L., & Muñoz Mata, E. (2013). Paraninfo Digital. *Early Human Development*. 83(1), 1–11. https://doi.org/10.1016/j.earlhumdev.2006.05.022.
183. Sharma, S., Kelly, T. K. & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis.* 31**,** 27-36.
184. Shiba, Y., Hauch, K., & Laflamme, M. (2009). Cardiac Applications for Human Pluripotent Stem Cells. *Current Pharmaceutical Design*. 15(24), 2791–2806. https://doi.org/10.2174/138161209788923804.
185. Shock, L. S., Thakkar, P. V., Peterson, E. J., Moran, R. G. & Taylor, S. M. (2011). DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proceedings of the National Academy of Sciences of the United States of America*. 108, 3630-3635.
186. Siemiatycki, J., (2007). Investigating cancer risks related to asbestos and other occupational carcinogens. *Occup Environ Med*. 2007 Aug; 64(8): 500–501. doi: 10.1136/oem.2006.030759.
187. Signorini S., Gerthoux P.M., Dassi C., Cazzaniga M., Brambilla P., Vincoli N., Mocarelli P. (2000). Environmental exposure to dioxin: the Seveso experience. *Andrologia*. 32(4-5):263-70.
188. Siroux, V., Agier, L., & Slama, R. (2016). The exposome concept: A challenge and a potential driver for environmental health research. *European Respiratory Review*. *25*(140), 124–129. https://doi.org/10.1183/16000617.0034-2016.
189. Sleeman, J.M., Richgels, K.L.D., White, C.L., Stephen, C. (2019). Integration of wildlife and environmental health into a One Health approach. *Rev Sci Tech*. 38(1):91-102. doi: 10.20506/rst.38.1.2944.
190. Sonneveld, E., Jansen, H.J., Riteco, J.A., Brouwer, A., van der Burg, B. (2005). Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. *Toxicol Sci*. 83(1):136-48. doi: 10.1093/toxsci/kfi005.
191. Skvortsova, K., Iovino, N., & Bogdanović, O. (2018). Functions and mechanisms of epigenetic inheritance in animals. *Nature Reviews Molecular Cell Biology*. 19(12), 774–790. https://doi.org/10.1038/s41580-018-0074-2.
192. Slocum TA. (1999). Thematic Cartography and Visualization. *Prentice-Hall*.
193. Smolders, R., Schramm, K. W., Nickmilder, M., & Schoeters, G. (2009). Applicability of non-invasively collected matrices for human biomonitoring. *Environmental Health: A Global Access Science Source*. 8(1), 1–10. https://doi.org/10.1186/1476-069X-8-8.

194. Tarantini, L., Bonzini, M., Apostoli, P., Pegoraro, V., Bollati, V., Marinelli, B., Cantone, L., Rizzo, G., Hou, L., Schwartz, J., Bertazzi, PA., Baccarelli, A. (2009). Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. *Environ Health Perspect.* 117(2):217-22. doi:10.1289/ehp.11898. Erratum in: Environ Health Perspect.117(4): A143.
195. Thiombane M, Petrik A, Di Bonito M, Albanese S, Zuzolo D, Cicchella D et al. (2018). Status, sources and contamination levels of organochlorine pesticide residues in urban and agricultural areas: a preliminary review in central-southern Italian soils. *Environ Sci Pollut Res Int*. 25(26):26361-26382. doi:10.1007/s11356-018-2688-5.
196. Tiffon, C. (2018). The impact of nutrition and environmental epigenetics on human health and disease. *International Journal of Molecular Sciences*. 19(11). https://doi.org/10.3390/ijms19113425.
197. Trevors, J. T. (2010). What is a global environmental pollution problem? *Water, Air, and Soil Pollution*. 210(1–4), 1–2. https://doi.org/10.1007/s11270-010-0337-9.
198. Triassi, M., Alfano, R., Illario, M., Nardone, A., Caporale, O., & Montuori, P. (2015). Environmental pollution from illegal waste disposal and health effects: A review on the "triangle of death." *International Journal of Environmental Research and Public Health*. 12(2), 1216–1236. https://doi.org/10.3390/ijerph120201216.
199. Ulaszewska, M. M., Zuccato, E., Capri, E., Iovine, R., Colombo, A., Rotella, G…Fanelli, R. (2011). The effect of waste combustion on the occurrence of polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and polychlorinated biphenyls (PCBs) in breast milk in Italy. *Chemosphere*, *82*(1), 1–8. https://doi.org/10.1016/j.chemosphere.2010.10.044.
200. Urdinguio, R.G., Sanchez-Mut, J.v. & Esteller, M. (2009). Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. *Lancet Neurol*. 8, 1056–1072.
201. U.S. EPA (U.S. Environmental Protection Agency). (2019). Guidelines for Human Exposure Assessment. (EPA/100/B-19/001). Washington, D.C.: Risk Assessment Forum.
202. US EPA Methods 1613. Available: https://www.epa.gov/sites/production/files/201702/documents/sop_hwss_25_revision_3.pdf
203. US EPA Methods 1668. Available: https://19january2017snapshot.epa.gov/sites/production/files/201509/documents/method_1668c_2010.pdf
204. Vanantwerpen, G., Vergne, T., Grindlay, D.J.C., Brennan, M.L. (2018). Quantitative Outcomes of a One Health approach to Study Global Health Challenges. *Ecohealth.* 15(1):209-227. doi: 10.1007/s10393-017-1310-5.
205. Vaissière, T, Hung, RJ, Zaridze, D, Moukeria, A, Cuenin, C, Fasolo, V, Ferro, G, Paliwal, A, Hainaut, P, Brennan, P, Tost, J, Boffetta, P, Herceg, Z. (2009) Quantitative analysis of DNA methylation profiles in lung cancer identifies aberrant DNA methylation of specific

genes and its association with gender and cancer risk factors. *Cancer Res.* 69(1):243-52. doi: 10.1158/0008-5472.CAN-08-2489.

206. Van den Berg, M., Birnbaum, L., Bosveld, A.T., Brunström, B., Cook, P., Feeley, M.,et al., (1998). Toxic equivalency factors (TEFs) for humans and wildlife. *Environ. Health Perspect*. 106 (12), 775–792.
207. Van den Berg. M., Birnbaum, L.S, Denison, M., De Vito, M., Farland, W., Feeley, M., Fiedler, H., Hakansson, H., Hanberg, A., Haws, L., Rose, M., Safe, S., Schrenk, D., Tohyama, C., Tritscher, A., Tuomisto, J., Tysklind, M., Walker, N., Peterson, R.E. (2006). The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Toxicol Sci*. 93, 223-241. https://doi.org/10.1093/toxsci/kfl055.
208. Van Leeuwen, D.M., Pedersen, M.., Hendriksen, P.J., Boorsma, A., Van Herwijnen, M.H., Gottschalk, R.W., Kleiinjans, J.C. (2008). Genomic analysis suggests higher susceptibility of children to air pollution. *Carcinogenesis*. 29:977-983.
209. Van Tongelen. A., Loriot, A., De Smet, C. (2017). Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. *Cancer Lett*. 396:130-137. doi:10.1016/j.canlet.2017.03.029.
210. Vimercati L., Baldassarre A., Gatti M.F., Gagliardi T., Serinelli M., De Maria L., Caputi A., Dirodi A.A., Galise I., Cuccaro F., Assennato G. (2016). Non-occupational exposure to heavy metals of the residents of an industrial area and biomonitoring. *Environ Monit Assess*.188(12):673.
211. Vimercati, L., Gatti, M.F., Gagliardi, T., Cuccaro, F., De Maria, L., Caputi, A., Quarato, M., Baldassarre, A. (2017). Environmental exposure to arsenic and chromium in an industrial area. *Environ Sci Pollut Res Int*. 24(12):11528-11535. doi:10.1007/s11356-017-8827-6.
212. Vineis P., Van Veldhoven K., Chadeau-Hyam M., Athersuch T.J. (2013). Advancing the application of omics based biomarkers in environmental epidemiology. *Environ Mol Mutagen*. 54(7): 461-467.
213. Vrijens, K., Bollati, V., & Nawrot, T. S. (2015). MicroRNAs as potential signatures of environmental exposure or effect: A systematic review. *Environmental Health Perspectives*. 123(5), 399–411. https://doi.org/10.1289/ehp.1408459.
214. Wan, E.S., Qiu, W., Baccarelli, A., Carey, V.J., Bacherman, H., Rennard, S.I., Agusti, A., Anderson, W., Lomas, D.A., Demeo, D.L. (2012) Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet*. 1(13):3073-82. doi: 10.1093/hmg/dds135.
215. Warner, M., Eskenazi, B., Patterson, D. G., Clark, G., Turner, W. E., Bonsignore, L.…Gerthoux, P. M. (2005). Dioxin-like TEQ of women from the Seveso, Italy area by ID-HRGC/HRMS and CALUX. *Journal of Exposure Analysis and Environmental Epidemiology*. 15(4), 310–318. https://doi.org/10.1038/sj.jea.7500407.

216. Warner M., Mocarelli P., Brambilla P., Wesselink A., Samuels S., Signorini S., Eskenazi B. (2013) Diabetes, metabolic syndrome, and obesity in relation to serum dioxin concentrations: the Seveso women's health study. *Environ Health Perspect*. 121(8):906-11.
217. Waters, M.D., Fostel, J.M. (2004). Toxicogenomics and system toxicology: aims and prospects. *Nat Rev Genet*.5:936-948.
218. WHO. (2015). Human biomonitoring: facts and figures. *World Health Organization*, 1–88. Retrieved from http://www.euro.who.int/__data/assets/pdf_file/0020/276311/Human-biomonitoring-facts-figures-en.pdf.
219. Wild, C. P. (2005). Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiology Biomarkers and Prevention*.14(8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456.
220. Wild, C. P. (2011). Future research perspectives on environment and health: The requirement for a more expansive concept of translational cancer research. *Environmental Health: A Global Access Science Source*. 10(SUPPL. 1), S15. https://doi.org/10.1186/1476-069X-10-S1-S15.
221. Wild, C. P. (2012). The exposome: From concept to utility. *International Journal of Epidemiology*. 41(1), 24–32. https://doi.org/10.1093/ije/dyr236.
222. Wild CP, Scalbert A, Herceg Z. (2013). Measuring the exposome: a powerful basis for evaluating environmental exposures and cancer risk. *Environ Mol Mutagen*. 54(7):480-99. doi: 10.1002/em.21777.
223. Wolpin, B.M., Rizzato, C., Kraft, P… Amundadottir, L.T. (2014) Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet.* 46(9):994-1000. doi: 10.1038/ng.3052.
224. Zani, C., Ceretti, E., Covolo, L., Donato, F. (2017). Do polychlorinated biphenyls cause cancer? A systematic review and meta-analysis of epidemiological studies on risk of cutaneous melanoma and non-Hodgkin lymphoma. *Chemosphere*. 183:97-106. doi:10.1016/j.chemosphere.2017.05.053.
225. Zeng, X.Y., Jiang, X.Y., Yong, J.H., Xie, H., Yuan, J., Zeng, D., Dou, Y.Y., Xiao, S.S. (2019). lncRNA ABHD11-AS1, regulated by the EGFR pathway, contributes to the ovarian cancer tumorigenesis by epigenetically suppressing TIMP2. *Cancer Med.* 8(16):7074-7085. doi: 10.1002/cam4.2586.
226. Zhang, W., Klinkebiel, D., Barger, C.J., Pandey, S., Guda, C., Miller, A., Akers, S.N., Odunsi, K., Karpf, A.R. (2020). Global DNA Hypomethylation in Epithelial Ovarian Cancer: Passive Demethylation and Association with Genomic Instability. *Cancers (Basel).* 12(3):764. doi: 10.3390/cancers12030764.

227. Zhau, R, Su, S, Lu, X, Liao, R, Ge, X, Huang, Y, Mai, S, Lu, X, Christiani, D. (2005). Proteomic profiling in the sera workers occupationally exposed to arsenic and lead: identification of potential biomarkers. *Biometals.* 18:603-613.
228. Zhao CN, Xu Z, Wu GC, Mao YM, Liu LN, Qian-Wu, Dan YL, Tao SS, Zhang Q, Sam NB, Fan YG, Zou YF, Ye DQ, Pan HF. (2019) Emerging role of air pollution in autoimmune diseases. *Autoimmun Rev.* 18(6):607-614.
229. Zidek, A., Macey, K., MacKinnon, L., Patel, M., Poddalgoda, D., Zhang, Y., (2017). A review of human biomonitoring data used in regulatory risk assessment under Canada's Chemicals Management Program. *Int. J. Hyg. Environ. Health 220.*167-178.
230. Ziech, D., Franco, R., Pappa, A., Malamou-Mitsi, V., Georgakila, S., Georgakilas, A.G. & Panayiotidis, M. I. (2010). The role of epigenetics in environmental and occupational carcinogenesis. *Chem Biol Interact*. 188, 340-349.
231. Zona A, Iavarone I, Buzzoni C, Conti S, Santoro M, Fazzo L, et al. (2019). Gruppo di lavoro SENTIERI; Gruppo di lavoro AIRTUM-SENTIERI; Gruppo di lavoro Malformazioni congenite-SENTIERI. SENTIERI: Studio epidemiologico nazionale dei territori e degli insediamenti esposti a rischio da inquinamento. Quinto Rapporto [SENTIERI: Epidemiological Study of Residents in National Priority Contaminated Sites. Fifth Report]. *Epidemiol Prev*. 43(2-3 Suppl 1):1-208. Italian. doi: 10.19191/EP19.2-3.S1.032. Erratum in: Epidemiol Prev. 43(4):219.
232. https://monographs.iarc.fr/agents-classified-by-the-iarc/
233. www.bioconductor.org/packages/release/bioc/html/ChAMP.html
234. www.biobancaizsm.it
235. www.campaniatrasparente.it
236. www.emea.illumina.it
237. www.geojizz.it
238. www.hbm4eu.eu
239. www.salute.gov.it

7. ACKNOWLEDGMENTS

I would like to acknowledge everyone who had contributed to achieve this moment.
My gratitude goes to the PhD course coordinator, Prof. Palmiero Monteleone and to my supervisor Prof. Roberta Tarallo for her helpfulness and guidance during the time of the PhD course. I would like to extend my thanks also to the Director of the Lab. Med. Mol. Ge., Prof. Alessandro Weisz, for the opportunity to carry out part of the research activity in the laboratory he manages.
Thanks to Dr. Claudia Angelini and Dr. Annamaria Carissimo for their collaboration on the bioinformatic and statistical analysis.
I owe a deep and sincere sense of gratitude to the Istituto Zooprofilattico Sperimentale del Mezzogiorno, with its "National Reference Center for Analysis and Correlation Study on Environment, Animals and Humans". I would like to express my very great appreciation to the General Manager Dr. Antonio Limone, for giving me the chance to contribute to the research activity and for having faith in my work. I am extremely grateful to Dr. Rino Cerino, for his professional and personal support, for believing in my abilities and for the continuous work challenges he sets me.
Above all, a heartfelt thanks to all my family, for their devoted support.
Thanks to my mother, my father and my brother Andrea for their unconditional love.
Thanks to my brother Mattia and my grandmother Bianca for filling my heart with their constant memory. Thanks to my little love Oliver.
A special thanks to Lucia and to Antonio's family for their encouragement and affection.
Finally, thanks to Antonio, my confidant, my life partner…thank you for always being by my side.

7. RINGRAZIAMENTI

Vorrei ringraziare tutti coloro che hanno contribuito alla realizzazione di questo momento. La mia gratitudine va al coordinatore del corso di dottorato, Prof. Palmiero Monteleone e al mio supervisore, Prof.ssa Roberta Tarallo, per la sua disponibilità e guida lungo il percorso del dottorato. Vorrei estendere il ringraziamento anche al Direttore del Lab. Med. Mol. Ge., Prof. Alessandro Weisz, per avermi dato l'opportunità di condurre parte dell'attività di ricerca nel laboratorio da lui diretto.

Un ringraziamento alla Dr.ssa Claudia Angelini e alla Dr.ssa Annamaria Carissimo, per la collaborazione sull'analisi statistica e bioinformatica.

Devo una profonda e sincera gratitudine all' Istituto Zooprofilattico Sperimentale del Mezzogiorno, con il suo "Centro di Referenza Nazionale per l'analisi e gli studi di correlazione su ambiente, animale ed uomo". Un ringraziamento particolare al Direttore Generale Dr. Antonio Limone, per avermi dato la possibilità di contribuire alle attività di ricerca e per aver riposto fiducia nel mio lavoro. Sono estremamente grata al Dr. Rino Cerino per il suo supporto professionale e personale, per aver creduto nelle mie capacità e per le continue sfide lavorative che pone sul mio cammino.

Oltre tutto e tutti, un grazie di cuore alla mia famiglia per il loro devoto supporto.

Grazie a mia madre, mio padre e mio fratello Andrea per il loro amore incondizionato.

Grazie a mio fratello Mattia e alla mia nonnina Bianca, per riempire il mio cuore con il loro ricordo costante. Grazie al mio piccolo cuoricino Oliver.

Un ringraziamento particolare a Lucia e alla famiglia di Antonio per il loro sostegno ed affetto.

Infine, grazie ad Antonio, mio confidente e compagno di vita…grazie per essere sempre al mio fianco.