HEALS

Health and Environment-wide Associations based on Large population Surveys

FP7-ENV-2013- 603946
http://www.heals-eu.eu/

13.1 Report on exposure biomarker profiles/biological responses/health outcomes

WP 13 Exposure and health association studies

Version 0.1

Lead beneficiary: UPMC

Date: 31/03/2017

Nature: Report

Dissemination level: Public

HEALS FP7-ENV-2013-603946	D13.1 - Report on exposure biomarker profiles/ biological responses/ health outcomes		
	WP13: Exposure and health association studies	**Security**:	
	Author(s): Denis A. Sarigiannis	**Version**:	0.1

Document Information

Grant Agreement Number	ENV-603946	**Acronym**	HEALS
Full title	Health and Environment-wide Associations based on Large population Surveys		
Project URL	http://www.heals-eu.eu/		
EU Project Officer	Tuomo Karjalainen,- Tuomo.KARJALAINEN@ec.europa.eu		

Deliverable	**Number**	13.1	**Title**	Report on exposure biomarker profiles/biological responses/health outcomes
Work Package	**Number**	13	**Title**	Exposure and health association studies

Delivery date	**Contractual**	M24	**Actual**	09/03/2017
Status	Draft ☐		Final X	
Nature	Demonstrator ☐	Report X	Prototype ☐	Other ☐
Dissemination level	Confidential ☐	Public X		

Author (Partners)	Denis A. Sarigiannis, S. Karakitsios, I. Furxhi, E. Handakas, N. Papaioannou, A. Gotti (AUTH)			
Responsible Author	Denis A. Sarigiannis		**Email**	denis@eng.auth.gr
	Partner	AUTH	**Phone**	+30-2310-994562

Document History

Name	Date	Version	Desciption

TABLE OF CONTENTS

Summary

A key novelty in HEALS is the integrated use of advanced statistic and computational tools supporting environmental and biological data analyses for comprehensive data interpretation, aiming at investigating causality between environment and health. The aim of this report is to provide a review of biomarkers of early biological effects and to present the methodology on the environment-wide association between exposure data and biological effect data/health outcomes. Overall, what is actually needed in the environmental-health domain is a synthesis of multiple type of information (environmental exposure, biomonitoring, toxicokinetics and omics), towards a comprehensive understanding of the multifactorial causality of disease (Chapter 1). Towards this aim, different omics responses (mainly metabolomics and transcriptomics) for identifying early biological effects, as well as susceptibility parameters, where comprehensively reviewed (Chapter 2), and these results are tabulated per health endpoint of interest (Chapter 3). The statistical methods for data clustering and more importantly for associating the different type of environmental, exposure, dietary, sociodemographic and omics data are established and described in Chapter 4. Among the various statistical techniques, the Environment Wide Associations Study (EWAS) was found to serve in the best way the aims of exposome analysis, that requires (a) an unbiased stance towards initial hypothesis between exposure and disease and (b) the use of multiple type of data, that all of them are considered as potential contributors to overall health or eventually pre-clinical or disease states. In practice, HEALS introduces a novel approach towards defining causal associations between health status and environmental stressors through the integrated use of EWAS. Environmental factors that are correlated are not considered confounders; rather they are co-variates, which are in “linkage disequilibrium” with each other. EWAS findings could then be used to identify further factors that may be in “disequilibrium”, for further detailed measurement and causal identification. This was clearly illustrated in the application of the methodology developed in HEALS in two different pre-existing cohorts (presented in Chapter 5), carried out in Poland (REPRO_PL study) and Greece (HERACLES study). Although these studies aimed at the associations among environmental pressures and neurodevelopmental effects, different type of environmental, exposure sociodemographic data and biospeciments were available, while evaluation of neurodevelopment was carried out using different test batteries. However, despite the differences in the availability of the data, the strength of the methodology allowed us to better identify associations between different exposure factors and neurodevelopment. In the case of Poland, it was clearly illustrated that several gestation factors have a beneficial (e.g. the concentration of selenium in maternal blood) or a negative influence (e.g. maternal bodyweight) in child neurodevelopment. On the other hand, child exposure to phthalates itself has a stronger negative influence in child neurodevelopment than maternal exposure. From the HERACLES study, it was found that children exposure to heavy metals and proximity to waste management sites have a negative influence in child neurodevelopment, however these effects are significantly modified by sociodemographic parameters (such as children SES and parents educational level, as well as diet. Further steps of analysis in both cohorts will include the results of the toxicity pathways identified to be perturbed from the omics analysis. Finally, a short discussion is given in Chapter 6 on the common nodes of toxicity among the various health endpoints of interest in HEALS how this can be facilitated using biomarkers of early effects.

1 Introduction

1.1 General considerations

The exposome (Wild, 2005) represents the totality of exposures from conception onwards, simultaneously identifying, characterizing and quantifying the exogenous and endogenous exposures and modifiable risk factors that predispose to and predict diseases throughout a person's life span. Exposome came as a complement to the human genome; although decoding of human genome (Schmutz et al., 2004) increased our understanding of the underlying causes of disease, genome explains only a percentage of population burden. Thus, it is evident that environmental factors are equally or eventually more important and what is actually critical is the interaction of environmental factors with the biological systems. Towards a better understanding of the causal links among genome, environment and disease, unraveling the exposome implies that both environmental exposures and genetic variation are reliably measured simultaneously.

HEALS (Health and Environment-wide Associations based on Large population Surveys) brings together a comprehensive array of novel technologies, data analysis and modeling tools that support the efficient design and execution of large-scale exposome studies. The HEALS approach brings together and organizes environmental, socio-economic, exposure, biomarker and health effect data; in addition, it includes all the procedures and computational sequences necessary for applying advanced bioinformatics coupling advanced data mining, biological and exposure modeling so as to ensure that environmental exposure-health associations are studied comprehensively. The overall approach will be verified in a series of population studies across Europe, tackling various levels of environmental exposure, age windows and gender differentiation of exposure, and socio-economic and genetic variability. The main objective of HEALS is the refinement of an integrated methodology and the application of the corresponding analytical and computational tools for performing environment-wide association studies in support of EU-wide environment and health assessments. For the first time, HEALS will try to reverse the paradigm of "nature versus nurture" and adopt one defined by complex and dynamic interactions between DNA sequence, epigenetic DNA modifications, gene expression and environmental factors that all combine to influence disease phenotypes. HEALS will start from analysis of data collected in on-going epidemiological EU studies involving mother/infant pairs, children, or adults including the elderly to evidence relevant environmental exposure/health outcome associations. These associations will aid in designing pilot surveys using an integrated approach, where the selection of biomarkers of exposure, effects and individual susceptibility results in integrated risk assessment.

The overall methodological concept of HEALS and the different arrays involved is graphically illustrated in Figure 1. This includes a wide array of state of the art technologies across all major disciplines of the environmental exposure, biochemistry, molecular biology, toxicology, bioinformatics and epidemiology arena.

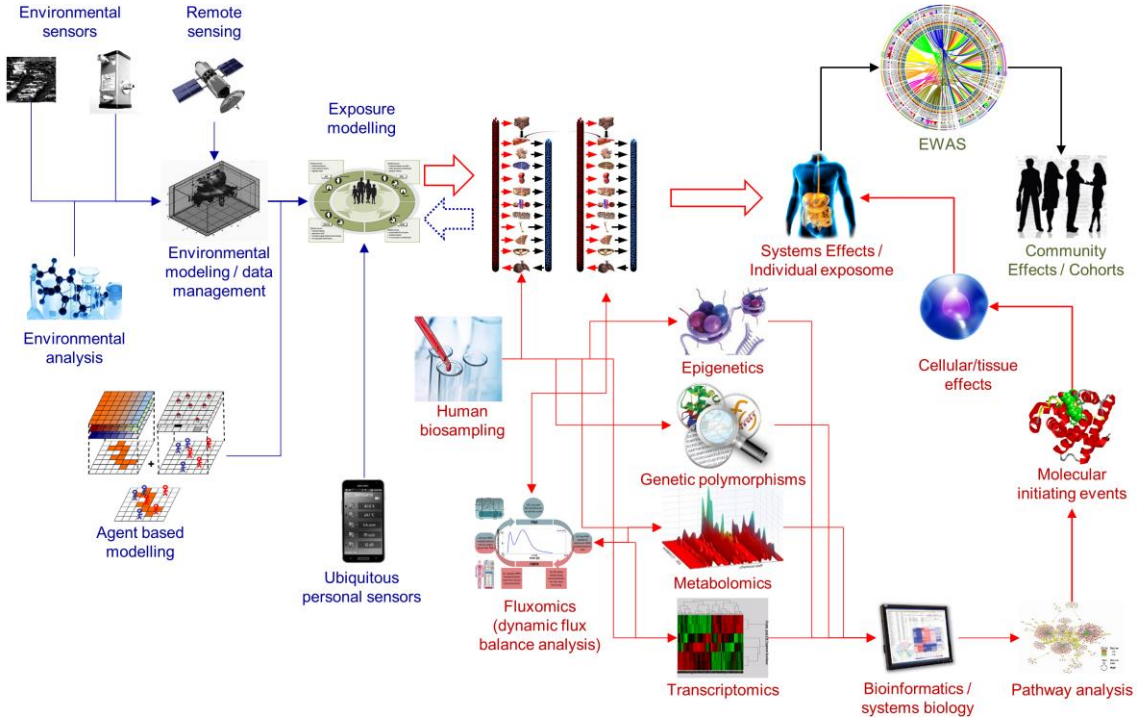

Figure 1: Overall HEALS conceptual methodological framework

1.2 Methodological overview

Towards the aim of identifying the predictive biomarkers for EWAS regarding the health endpoints of interest in HEALS and based on the approach outlined above, a comprehensive approach requires the following steps:

- Starting from Agnostic metabolomics / transcriptomics
- Joint pathway analysis
- Identification of plausible pathways pf toxicity
- Corroboration / checking of POT hypothesis using targeted multi-omics (transcriptomics, proteomics, metabolomics)
- Phenotypic anchoring with observed adverse health outcomes of human biochemical markers of effect
- Selection of metabolic networks / gene regulatory networks linked with validated PoT
- Use these networks (in the metabolome or/and transcriptome) as categorical variables (activation or not / to what context) in EWAS algorithm to derive associations with:
    - Exposure determinants
    - Exposure markers (e.g. biochemical markers in biospecimens)
    - Exposure modifiers (e.g. SES, gender, age)
    - Adverse health outcomes (e.g. neurological / cognitive development test scores)

2 Review of associations between health perturbations and early biological events

2.1 Review of associations between environmental exposures and asthma and allergies

2.1.1 Review of early biological responses / biomarkers identification

Asthma is highly induced by environmental factors such as ambient air particles. Xia et al. (2015b) identified components of the Notch pathway, most notably Jagged 1 (Jag1), as targets of PM induction in human monocytes and murine dendritic cells. PM, especially ultrafine particles, upregulated TH cytokine levels, IgE production, and allergic airway inflammation in mice in a Jag1- and Notch-dependent manner, especially in the context of the proasthmatic IL-4 receptor allele Il4raR576. PM-induced Jag1 expression was mediated by the aryl hydrocarbon receptor (AhR), which bound to and activated AhR response elements in the Jag1 promoter. Pharmacologic antagonism of AhR or its lineage-specific deletion in CD11c+ cells abrogated the augmentation of airway inflammation by PM. In conclusion PM activates an AhR-Jag1-Notch cascade to promote allergic airway inflammation in concert with proasthmatic alleles. In another study, the effect of Diesel exhaust particulates (DEPs) on house dust mite (HDM)–specific memory responses was determined by using an asthma model (Brandt et al., 2015). Data from children enrolled in the Cincinnati Childhood Allergy and Air Pollution Study birth cohort were analyzed to determine the effect of DEP exposure on asthma outcomes. DEP coexposure with HDM resulted in persistent TH2/TH17 CD1271 effector/memory cells in the lungs, spleen, and lymph nodes of adult and neonatal mice. After 7 weeks of rest, a single exposure to HDM resulted in airway hyperresponsiveness and increased TH2 cytokine levels in mice that had been previously exposed to both HDM and DEPs versus those exposed to HDM alone. On the basis of these data, Brant et al. (2015) examined whether DEP exposure was similarly associated with increased asthma prevalence in children in the presence or absence of allergen exposure/sensitization in the Cincinnati Childhood Allergy and Air Pollution Study birth cohort. Early-life exposure to high DEP levels was associated with significantly increased asthma prevalence among allergic children but not among non-allergic children. These findings suggested that DEP exposure results in accumulation of allergen-specific TH2/TH17 cells in the lungs, potentiating secondary allergen recall responses and promoting the development of allergic asthma. One major proposed mechanism of air pollutants toxicity involves the activation of t-helper type 2 (Th2) immune responses which are linked to asthma pathogenesis. Secretion of Th2 inflammatory cytokines (IL-4, IL-5, IL-9, IL-13) may lead to mucus hypersecretion and thickening and contraction of the airway smooth muscle in atopic asthmatics (Larché et al., 2003). The main direct targets of inhaled PM are airway epithelial cells and dendritic cells (DC). Recently, Xia et. al.(2015a) identified components of the Notch pathway, most notably Jagged 1 (Jag1), as targets of PM induction in human monocytes and murine dendritic cells. PM, especially ultrafine particles, upregulated TH cytokine levels, IgE production, and allergic airway inflammation in mice in a Jag1- and Notch-dependent manner, especially in the context of the proasthmatic IL-4 receptor allele Il4raR576. PM-induced Jag1 expression was mediated by the aryl hydrocarbon receptor (AhR), which bound to and activated AhR response elements in the Jag1 promoter. Many of the PM components (i.e., polycyclic aromatic hydrocarbons and metals) are redox-active and capable of inducing cellular

oxidative stress and injuries including inflammation and cell death. Airway epithelial cells and antigen-presenting dendritic cells (DC) are the major and direct targets of inhaled PM. The epithelial cells can further enhance the DC response to allergen and PM through several immune regulatory cytokines including thymic stromal lymphopoietin (TSLP), IL-33, and IL-25. Among these cytokines TSLP is particularly relevant to the mechanisms by which particulate air pollutants contribute to asthma pathogenesis. Studies have found that TSLP released by PM-exposed human airway epithelial cells could polarize the DC towards a T-helper 2 immune response, which is one of the key immunological mechanisms in asthma pathogenesis. The convergence of regulatory signals generated by PM-induced oxidative stress in DC and the interactions among them may be one of the major mechanisms that are specifically related to the contribution of PM towards asthma pathogenesis (Li and Buglak, 2015). Hinks et al. (2015) performed a comprehensive assessment of TH17 cells, regulatory T cells, mucosal-associated invariant T (MAIT) cells, other T-cell subsets, and granulocyte mediators in asthmatic patients. Sixty patients with mild-to-severe asthma and 24 control subjects underwent detailed clinical assessment and provided induced sputum, endobronchial biopsy, bronchoalveolar lavage, and blood samples. Adaptive and invariant T-cell subsets, cytokines, mast cells, and basophil mediators were analyzed. Significant heterogeneity of T-cell phenotypes was observed, with levels of IL-13–secreting T cells and type 2 cytokines increased at some, but not all, asthma severities. TH17 cells and γδ-17 cells, proposed drivers of neutrophilic inflammation, were not strongly associated with asthma, even in severe neutrophilic forms. MAIT cell frequencies were strikingly reduced in both blood and lung tissue in relation to corticosteroid therapy and vitamin D levels, especially in patients with severe asthma in whom bronchoalveolar lavage regulatory T-cell numbers were also reduced. Bayesian network analysis identified complex relationships between pathobiologic and clinical parameters. Topological data analysis identified 6 novel clusters that are associated with diverse underlying disease mechanisms, with increased mast cell mediator levels in patients with severe asthma both in its atopic (type 2 cytokine–high) and nonatopic forms. The evidence for a role for TH17 cells in patients with severe asthma is limited. Severe asthma is associated with a striking deficiency of MAIT cells and high mast cell mediator levels. This study provides proof of concept for disease mechanistic networks in asthmatic patients with clusters that could inform the development of new therapies.

The systemic cysteine oxidation and its association with inflammatory and clinical features in healthy children and children with difficult-to-treat asthma has been also investigated (Stephenson et al., 2015). It was hypothesized that cysteine oxidation would be associated with increased markers of oxidative stress and inflammation, increased features of asthma severity, decreased clinically defined glucocorticoid responsiveness, and impaired GR function. PBMCs were collected from healthy children (n = 16) and children with asthma (n = 118) aged 6 to 17 years. Children with difficult-to-treat asthma underwent glucocorticoid responsiveness testing with intramuscular triamcinolone. Cysteine, cystine, and inflammatory chemokines and reactive oxygen species generation were quantified, and expression and activity of the GR were assessed. Cysteine oxidation was present in children with difficult-to-treat asthma and accompanied by increased reactive oxygen species generation and increased CCL3 and CXCL1 mRNA expression. Children with the greatest extent of cysteine oxidation had more features of asthma severity, including poorer symptom control, greater medication use, and less glucocorticoid responsiveness despite inhaled glucocorticoid therapy. Cysteine oxidation also modified the GR protein by decreasing available sulfhydryl groups and decreasing nuclear GR

expression and activity. A highly oxidized cysteine redox state promotes a posttranslational modification of the GR that might inhibit its function. Given that cysteine oxidation is prevalent in children with difficult-to-treat asthma, the cysteine redox state might represent a potential therapeutic target for restoration of glucocorticoid responsiveness in this population.

To better understand the inter- and intra-individual variability and seasonal variation of IgE, and high (FcεRI)- and low-affinity (CD23) IgE receptor expression in blood of seasonal allergic rhinitis (SAR) subjects, thirty-two otherwise healthy subjects with a history of SAR to birch pollen and a positive skin prick test to birch pollen were sampled three times out of the pollen season and three times during the pollen season (Carlsson et al., 2015). FcεRI and CD23 expressions were analysed using flow cytometry. Total IgE was analysed using ImmunoCAP(®) and free IgE was analysed with a novel customised research assay using an IgG-FcεRI-chimera protein coupled to ImmunoCAP as capture reagent, ImmunoCAP-specific IgE conjugate and ImmunoCAP IgE calibrators. The performance of the free IgE assay was compared well with the reference ImmunoCAP total IgE assay. The working range of the assay was 0.35-200 kU/l IgE. FcεRI expression on basophils and CD23 expression on B cells showed low intrasubject variability both in and out of the pollen season (<10% CV). There was a small seasonal difference with lower total IgE levels (120 versus 128 kU/l; P = 0.004) and FcεRI expression (283 versus 325 mean fluorescence intensity (MFI); P < 0.001) during the pollen season. IgE, FcεRI expression and CD23 expression fulfilled biomarker and assay requirements of variability, and allergen exposure affected the biomarkers only to a minor degree. The free IgE assay may be used for measurement of free IgE levels in patients after anti-IgE antibody treatment

Specific metabolic signatures have also been used for identifying pathologic conditions. To better understand the metabolic phenotypes of asthma, a plasma metabolic signature associated with allergic asthma in ovalbumin (OVA)-sensitized mice by using UPLC-Q-TOF/MS was investigated using sixteen metabolites, characterized as potential pathological biomarkers related to asthma (Yu et al., 2016). The identified potential biomarkers were involved in 6 metabolic pathways and achieved the most entire metabolome contributing to the formation of allergic asthma. Purine metabolism was the most prominently influenced in OVA-induced asthma mice according to the metabolic pathway analysis (MetPA), suggesting that significantly changes in inflammatory responses in the pathophysiologic process of asthma. The metabolites of purine metabolism, especially uric acid (P12) and inosine (P13), may denote their potential as targeted biomarkers related to experimental asthma. The decreased plasma uric acid (P12) suggested that inflammation responses of allergic asthma inhibited the activity of xanthine oxidase in purine metabolism, and manifested the severity of asthma exacerbation. The increased level of inosine (P13) suggests that inflammatory cells induce adenosine triphosphate (ATP) breakdown, resulting in excessive expression of adenosine deaminase (ADA) in the formation of allergic asthma. These findings provided a novel perspective on the metabolites signatures related to allergic asthma, which provided us with new insights into the pathogenesis of asthma, and the discovery of targets for clinical diagnosis and treatment. Chinese patients with mild persistent asthma using GC-MS coupled with a series of multivariate statistical analyses were investigated and clear intergroup separations existed between the asthmatic patients and control subjects (Chang et al., 2015). A list of differential metabolites and several top altered metabolic pathways were identified. The levels of succinate (an intermediate in tricarboxylic acid cycle) and

inosine were highly upregulated in the asthmatic patients, suggesting a greater effort to breathe during exacerbation and hypoxic stress due to asthma. Other differential metabolites, such as 3,4-dihydroxybenzoic acid and phenylalanine, were also identified. Furthermore, the differential metabolites possessed higher values of area under the ROC curve (AUC), suggesting an excellent clinical ability for the prediction of asthma. An integrated approach of combination of LC-GC metabolomics was used to analyze BALF from experimental asthma (Ho et al., 2013). The author investigated the effects of dexamethasone on metabolic profiles of BALF in the murine model of asthma. The findings revealed substantial alterations in energy, lipid, and sterol metabolism in BALF from an experimental murine model of asthma, including potentially important metabolites for phenotyping asthma. Urine samples of 41 atopic asthmatic children (further subdivided in sub-groups according to the symptoms) and 12 age-matched controls were analyzed. Untargeted metabolic profiles were collected by LC-MS, and studied by multivariate analysis (Mattarucchi et al., 2012). The group of the asthmatics was differentiated by a model that proved to be uncorrelated with the chronic assumption of controller drugs on the part of the patients. The distinct sub-groups were also appropriately modeled. NMR-based metabolomics indicates that obese asthmatic (OA) patients are characterized by a respiratory metabolic fingerprint fully different from patients independently affected by asthma or obesity. Such phenotypic difference strongly suggests unique pathophysiological pathways involved in the pathogenesis of asthma in adult obese subjects (Maniscalco et al., 2016). Furthermore, the OA metabotype could define a strategy for patient stratification based on unbiased biomarkers, with important diagnostic and therapeutic implications.

The relationships between infections in early life and asthma are not completely understood. Likewise, the clinical relevance of microbial communities present in the respiratory tract is only partially known. A number of microbiome studies analyzing respiratory tract samples have found increased proportions of gamma-Proteobacteria including Haemophilus influenzae, Moraxella catarrhalis, and Firmicutes such as Streptococcus pneumoniae (Castro-Nallar et al., 2015). A new approach that combines RNA microbial identification with host gene expression was presented, for characterization and validation of metagenomic taxonomic profiling in individuals with asthma. Using whole metagenomic shotgun RNA sequencing, the microbial communities of individuals, children and adolescents, with asthma and controls were characterized and compared. The resulting data were analyzed by partitioning human and microbial reads. Microbial reads were then used to characterize the microbial diversity of each patient, and potential differences between asthmatic and healthy groups. Human reads were used to assess the expression of known genes involved in the host immune response to specific pathogens and detect potential differences between those with asthma and controls. For the first time, the power of combining RNA taxonomic profiling and host gene expression signatures for microbial identification was showed. This approach not only identifies microbes from metagenomic data, but also adds support to these inferences by determining if the host is mounting a response against specific infectious agents. In particular, M. catarrhalis is abundant in asthma patients but not in controls, and that its presence is associated with a specific host gene expression signature.

Whole-transcriptome RNA sequencing was performed on nasal airway brushings from 10 control subjects and 10 asthmatic subjects, which were compared with established bronchial and small-airway transcriptomes (Poole et al., 2014). Targeted RNA sequencing nasal expression analysis was used to

profile 105 genes in 50 asthmatic subjects and 50 control subjects for differential expression and clustering analyses. High-depth whole-transcriptome sequencing was used to comprehensively determine the degree to which the nasal airway serves as a biologic proxy for the bronchial airway. A novel targeted RNA sequencing (RNA-seq) technology was used, to profile gene expression of candidate airway biomarkers in a larger group of well-characterized children with asthma and healthy control subjects. These data were used to determine the relationship between the nasal transcriptome and subphenotypes of asthma. 90.2% overlap in expressed genes and strong correlation in gene expression (ρ = .87) between the nasal and bronchial transcriptomes was found. Previously observed asthmatic bronchial differential expression was strongly correlated with asthmatic nasal differential expression.

2.1.2 Susceptibility parameters

Genes involved in conferring susceptibility to the development of asthma can be grouped under four headings:

• Genes controlling factors involved in airway development and repair, including remodelling.12 For example, polymorphism of ADAM33, which has been linked with airway wall remodelling, is strongly associated with asthma in diverse populations.

• Genes involved in controlling the responses of the immune system. For example, a polymorphism in the TNF promoter region has been associated with asthma (and its severity).

• Genes controlling bronchial hyperresponsiveness. For example, airway smooth muscle cells producing inadequate levels of C/EBPa, due to a genetic variant, are more susceptible to contractile stimuli than normal cells.

• Genes controlling the production of endogenous anti-oxidants by the airways.18 For example, polymorphic variation of glutathione S-transferase M1, glutathione S-transferase P1 and NAD(P)H: quinine reductase, has been linked with differences in baseline lung function, with airway responsiveness to ozone and with the influence of maternal smoking on asthma.

In addition, a model was based on four SNPs (rs9522789, rs7147228, rs2701423, rs759582) and two metabolites—monoHETE_0863 and sphingosine-1-phosphate (S1P) which could predict asthma control with an AUC of 95% (McGeachie et al., 2015). Integrative ORA identified 17 significantly enriched pathways related to cellular immune response, interferon signaling, and cytokine-related signaling, for which arachidonic acid, PGE2 and S1P, in addition to six genes (CHN1, PRKCE, GNA12, OASL, OAS1, and IFIT3) appeared to drive the pathway results. Of these predictors, S1P, GNA12, and PRKCE were enriched in the results from integrative and metabolic ORAs.

2.2 Review of associations between environmental exposures and neurodevelopmental and neurodegenerative disorders

2.2.1 Review of early biological responses / biomarkers identification

The spectrum of neurodevelopmental disorders includes several clinical observations such as autism, Fetal alcohol spectrum disorders, as well as poor performance in several tests (e.g. Bayley test, Wechsler Intelligence Scale for children). In the latest years, the use of omics has resulted in the identification of early biological events (mainly transcriptomics and metabolomics) related to these outcomes.

With regard to metabolomics, different metabolite profiles have been identified in children with fetal alcohol spectrum disorders (FASD). The mechanisms underlying FASD are incompletely understood, and biomarkers to identify those at risk are lacking. From a metabolomic analysis of embryoid bodies and neural lineages derived from human embryonic stem (hES) to identify the neural secretome produced in response to ethanol (EtOH) exposure (Palmer et al., 2012). It was found that EtOH treatment induced statistically significant changes to metabolite abundance in human embryoid bodies (180 features), neural progenitors (76 features), and neurons (42 features). There were no shared significant features between different cell types. Fifteen features showed a dose-response to EtOH. Four chemical identities were confirmed: L-thyroxine, 5'-methylthioadenosine, and the tryptophan metabolites, L-kynurenine and indoleacetaldehyde. One feature with a putative annotation of succinyladenosine was significantly increased in both EtOH treatments. As a result, it was found that EtOH exposure induces statistically significant changes to the metabolome profile of human embryoid bodies, neural progenitors, and neurons. Several of these metabolites are normally present in human serum, suggesting their usefulness as potential serum FASD biomarkers. These findings suggest the biochemical pathways that are affected by EtOH in the developing nervous system and delineate mechanisms of alcohol injury during human development.

A variety of possible mechanisms by which neurotoxicants can lead to neurodevelopmental abnormalities is supported by the scientific literature. These mechanisms involve induction of oxidative stress, interfering calcium signaling, effects on neurotransmitter pathways, neuroendocrine effects and epigenetic control (Chen et al., 2011). Especially in the critical stages of nervous system development, the abovementioned effects can impact neuronal growth, differentiation, migration, synaptogenesis, and myelination, leading to an array of neurodevelopmental deficits. Many environmental toxicants, (heavy metals, PBDEs, PCBs, some pesticides etc) posses the ability to generate ROS and deplete antioxidant capacity. Neuronal cells are especially vulnerable to oxidative stress because of the high amount of ROS generated during normal metabolism and neuronal activity (Milou et al. 2011). In vitro studies have demonstrated that environmental toxicant such as (Pb, Hg, PCBs, PBDEs, …) cause oxidative stress in neuronal cells leading to apoptotic cell death. Evidence is also present for the effect of many subclasses of pollutants to neurotransmitter pathways. Recently, an animal study discovered that repetitive postnatal PCB exposure resulted in increased levels of homovanillic and 5-hydroxyindoleacetic acid, metabolites of DA and 5-HT, in the neostriatum of young adult animals, without changing levels of the transmitters themselves. Moreover, analysis of DA-synapses demonstrated specific effects on a restricted number of specific synaptic proteins, including

the presynaptic DAT, the postsynaptic D5 receptor and the PSD-95 scaffolding synapse protein (Dervola et al., 2015). Exposure to PCBs have been found to affect the levels of urinary homovanillic acid, a DA metabolite, in humans (Putschögl et al., 2015).

Autism Spectrum Disorders (ASD) are a group of developmental disorders caused by environmental and genetic factors. Diagnosis is based on behavioral and developmental signs detected before 3 years of age with no reliable biological marker. The potential use of a 2D NMR-based approach to express the global biochemical signature of autistic individuals compared to normal controls was investigated by Mavel et al. (2013). This technique has greater spectral resolution than to 1D H NMR spectroscopy, which is limited by overlapping signals. The urinary metabolic profiles of 30 autistic and 28 matched healthy children were obtained using a 1H–13C NMR-based approach. The data acquired were processed by multivariate orthogonal partial least-squares discriminant analysis (OPLS-DA). Some discriminating metabolites were identified: β-alanine, glycine, taurine and succinate concentrations were significatively higher, and creatine and 3-methylhistidine concentrations were lower in autistic children than in controls. Also, differences in several other metabolites that were unidentified but characterized by a cross peak correlation in 1H–13C HSQC were noted. Statistical models of 1H and 1H–13C analyses were compared and only 2D spectra allowed the characterization of statistically relevant changes [R2Y(cum)¼ 0.78 and Q2(cum)¼ 0.60] in the low abundance metabolites. This method has the potential to contribute to the diagnosis of neurodevelopment disorders but needs to be validated on larger cohorts and on other developmental disorders to define its specificity. Similarly, Wang et al. (2016) performed a metabolomics analysis of serum to identify potential biomarkers for the early diagnosis and clinical evaluation of autism. They analyzed a discovery cohort of patients with autism and participants without autism in the Chinese Han population using ultra-performance liquid chromatography quadrupole time-of-flight tandem mass spectrometry (UPLC/Q-TOF MS/MS) to detect metabolic changes in serum associated with autism. The potential metabolite candidates for biomarkers were individually validated in an additional independent cohort of cases and controls. They built a multiple logistic regression model to evaluate the validated biomarkers, including 73 patients and 63 controls in the discovery cohort and 100 cases and 100 controls in the validation cohort. Metabolomic analysis of serum in the discovery stage identified 17 metabolites, 11 of which were validated in an independent cohort. A multiple logistic regression model built on the 11 validated metabolites fit well in both cohorts. The model consistently showed that autism was associated with 2 particular metabolites: sphingosine 1-phosphate and docosahexaenoic acid. In another study (Emond et al., 2013)., GC-MS urinary metabolic profiles of 26 autistic and 24 healthy children were obtained by liq/liq extraction, and were or were not subjected to an oximation step, and then were subjected to a persilylation step. These metabolic profiles were then processed by multivariate analysis, in particular orthogonal partial least-squares discriminant analysis (OPLS-DA, R2Y(cum) = 0.97, Q2(cum) = 0.88). Discriminating metabolites were identified. The relative concentrations of the succinate and glycolate were higher for autistic than healthy children, whereas those of hippurate, 3-hydroxyphenylacetate, vanillylhydracrylate, 3-hydroxyhippurate, 4-hydroxyphenyl-2-hydroxyacetate, 1H-indole-3-acetate, phosphate, palmitate, stearate, and 3-methyladipate were lower. Eight other metabolites, which were not identified but characterized by a retention time plus a quantifier and its qualifier ion masses, were found to differ between the two groups. Comparison of statistical models

leaded to the conclusion that the combination of data obtained from both derivatization techniques leads to the model best discriminating between autistic and healthy groups of children.

In terms of transcriptomics signatures, recent studies of genomic variation associated with autism have suggested the existence of extreme heterogeneity. Large-scale transcriptomics should complement these results to identify core molecular pathways underlying autism. Gupta et al. (2014) reported results from a large-scale RNA sequencing effort, utilizing region-matched autism and control brains to identify neuronal and microglial genes robustly dysregulated in autism cortical brain. A gene expression module corresponding to M2-activation states in microglia is negatively correlated with a differentially expressed neuronal module, implicating dysregulated microglial responses in concert with altered neuronal activity-dependent genes in autism brains. These observations provided pathways and candidate genes that highlight the interplay between innate immunity and neuronal activity in the etiology of autism.

In addition to metabolomics and transcriptomics signatures, Heberling and Dhurjati (2015) paper proposed a computational approach whereby metagenomes characteristic of "healthy" and autistic individuals are artificially constructed via genomic information, analyzed for the enzymes coded within, and then these enzymes are compared in detail. This is a text mining application. A custom-designed online application was built and used for the comparative metabolomics study and made publically available. Several of the enzyme-catalyzing reactions involved with the amino acid glutamate were curiously missing from the "autism" microbiome and were coded within almost every organism included in the "control" microbiome. Interestingly, there exists a leading hypothesis regarding autism and glutamate involving a neurological excitation/inhibition imbalance; but the association with this study is unclear. The results included data on the transsulfuration and transmethylation pathways, involved with oxidative stress, also of importance to autism. The results from this study are in alignment with leading hypotheses in the field, which is impressive, considering the purely in silico nature of this study. The present study provides new insight into the complex metabolic interactions underlying autism, and this novel methodology has potential to be useful for developing new hypotheses. However, limitations include sparse genome data availability and conflicting literature experimental data. We believe our software tool and methodology has potential for having great utility as data become more available, comprehensive and reliable. In terms of computation methods, starting from plasma metabolome analysis from children aged 4 to 6, 52 with ASD and 30 age-matched TD children, a broad range of metabolites were identified (West et al., 2014). Univariate, multivariate and machine learning methods were used to develop models to rank the importance of features that could distinguish ASD from TD. A set of 179 statistically significant features resulting from univariate analysis were used for multivariate modeling. Subsets of these features properly classified the ASD and TD samples in the 61-sample training set with average accuracies of 84% and 86%, and with a maximum accuracy of 81% in an independent 21-sample validation set. In another study, a supervised multivariate model to classify the metabolome alterations between autistic spectrum disorders (ASD) patients and controls, siblings of autistic patients, had been realized and used to realize a network model of the ASD patients' metabolome (Noto et al., 2014). In this experiment, propose of a quantification of urinary metabolites with the Mass Spectroscopy technique couple to Gas Chromatography was made. A multivariate model has been used to extrapolate the variables of

importance for a network model of interaction between metabolites. In this way we are able to propose a network-based approach to ASD description. Plasma metabolomic phenotyping was obtained for 23 premutation carriers and 16 age- and sex-matched controls. Three biomarkers, phenylethylamine normalized by either aconitate or isocitrate and oleamide normalized by isocitrate, exhibited excellent model performance (Giulivi et al., 2016). The lower phenylethylamine and oleamide plasma levels in carriers may indicate, respectively, incipient nigrostriatal degeneration and higher incidence of substance abuse, anxiety and sleep disturbances. Higher levels of citrate, isocitrate, aconitate, and lactate may reflect deficits in both bioenergetics and neurotransmitter metabolism (Glu, GABA).

Beyond neurodevelopmental disorders, omics biomarkers have been used for identifying metabolic signatures of neurodegenerative diseases such as Parkinson Disease (PD). PD is a multifactorial disease that is characterized by the progressive loss of dopaminergic neurons of the substantia nigra pars compacta (SN). This progressive loss of dopamine input from the SN to the striatum results in degenerative loss of motor function that manifests in bradykinesia, postural instability, tremor and rigidity (Roede et al., 2014). Epidemiological studies have identified that factors such as living in a rural area, consuming well water, farming, and pesticide exposure may be risk factors for developing PD. These observations have led to the development of an “environmental hypothesis” of PD. This hypothesis states that there are chemicals in the environment that are capable of selectively damaging the dopaminergic neurons of the SN, thus contributing to the development of PD. Two common pesticides, paraquat (PQ) and maneb (MB), have been demonstrated in vivo to preferentially alter the nigrostriatal dopamine system. A combination of the herbicide paraquat (PQ) and fungicide maneb (MB) has been linked to Parkinson's disease. Previous studies show that this involves an additive toxicity with at least two different mechanisms (Roede et al., 2014). However, detailed understanding of mixtures is often difficult to elucidate because of the multiple ways by which toxic agents can interact. In the present study, we used a combination of transcriptomics and metabolomics to investigate mechanisms of toxicity of PQ and MB in a neuroblastoma cell line. Conditions were studied with concentrations of PQ and MB that each individually caused 20% cell death and together caused 50% cell death. Transcriptomic and metabolomic samples were collected at time points prior to significant cell death. Statistical and bioinformatic methods were applied to the resulting 30,869 transcripts and 1358 metabolites. Results showed that MB significantly changed more transcripts and metabolites than PQ, and combined PQ + MB impacted more than MB alone. Transcriptome–metabolome-wide association study (TMWAS) showed that significantly changed transcripts and metabolites mapped to two network substructures, one associating with significant effects of MB and the other included features significantly associated with PQ + MB. The latter contained 4 clusters of genes and associated metabolites, with one containing genes for two cation transporters and a cation transporter regulatory protein also recognized as a pro-apoptotic protein. Other clusters included stress response genes and transporters linked to cytoprotective mechanisms. MB also had a significant network structure linked to cell proliferation. Together, the results show that the toxicologic mechanism of the combined neurotoxicity of PQ and MB involves network level interactions and that TMWAS provides an effective approach to investigate such complex mechanisms. For further elucidation of the mechanisms related to PD, cerebrospinal fluid biomarker studies focused on different disease pathways: oxidative stress, neuroinflammation, lysosomal dysfunction and proteins

involved in PD and other neurodegenerative disorders, focusing on four clinical domains: their ability to (1) distinguish PD from healthy subjects and other neurodegenerative disorders as well as their relation to (2) disease duration after initial diagnosis, (3) severity of disease (motor symptoms) and (4) cognitive dysfunction. Oligomeric alpha-synuclein might be helpful in the separation of PD from controls (Andersen et al., 2016). Through metabolomics, changes in purine and tryptophan metabolism have been discovered in patients with PD. Neurofilament light chain (NfL) has a significant role in distinguishing PD from other neurodegenerative diseases. Several oxidative stress markers are related to disease severity, with the antioxidant urate also having a prognostic value in terms of disease severity. Increased levels of amyloid and tau-proteins correlate with cognitive decline and may have prognostic value for cognitive deficits in PD. Urine metabolic phenotyping has also been associated with the development of PD. A metabolomic study with was performed using gas chromatography - mass spectrometry (GC - MS) and liquid chromatography - mass spectrometry (LC - MS) to characterize the urinary metabolic phenotypes of idiopathic PD patients at three stages (early, middle and advanced) and normal control subjects, with the aim of discovering potential urinary metabolite markers for the diagnosis of idiopathic PD by Luan et al. (2015). Both GC-MS and LC-MS metabolic profiles of idiopathic PD patients differed significantly from those of normal control subjects. 18 differentially expressed metabolites were identified as constituting a unique metabolic marker associated with the progression of idiopathic PD. Related metabolic pathway variations were observed in branched chain amino acid metabolism, glycine derivation, steroid hormone biosynthesis, tryptophan metabolism, and phenylalanine metabolism.

With regard to amyotrophic lateral sclerosis (ALS), plasma biomarkers can aid in distinguishing patients with ALS from those with disease mimics were identified (Lawton et al., 2014). In a multi-center study, plasma samples were collected from 172 patients recently diagnosed with ALS, 50 healthy controls, and 73 neurological disease mimics. Samples were analyzed using metabolomics. Using all identified biochemicals detected in > 50% of all samples in the metabolomics analysis, samples were classified as ALS or mimic with 65% sensitivity and 81% specificity by LASSO analysis (AUC of 0.76). A subset panel of 32 candidate biomarkers classified these diagnosis groups with a specificity of 90%/sensitivity 58% (AUC of 0.81). Creatinine was lower in subjects with lower revised ALS Functional Rating Scale (ALSFRS-R) scores. In conclusion, ALS can be distinguished from neurological disease mimics by global biochemical profiling of plasma samples. Our analysis identified ALS versus mimics with relatively high sensitivity. From the study, a subset of 32 metabolites were identify, discriminating patients with ALS with a high specificity. Interestingly, lower creatinine was found to correlate significantly with a lower ALSFRS-R score. Finally, molecules previously reported to be important in disease pathophysiology, such as urate, were also included in our metabolite panel. The evolution of metabolism alteration and its link with disease progression has also been described by Patin et al. (2016). They ran a study focused on (1) the evolution of metabolism disturbance during disease progression through omics approaches and (2) the relation between metabolome profile and clinical evolution. SOD1-G93A (mSOD1) transgenic mice (n = 11) and wild-type (WT) littermates (n = 17) were studied during 20 weeks. Metabolomic profile of muscle and cerebral cortex was analysed at week 20, and plasma samples were assessed at four time points over 20 weeks. The relevant metabolic pathways highlighted by metabolomic analysis were explored by a targeted transcriptomic approach in mice. Plasma metabolomics were also performed in 24 ALS patients and 24 gender- and age-matched controls.

Metabolomic analysis of muscle and cerebral cortex enabled an excellent discrimination between mSOD1 and WT mice ($p < 0.001$). In another study (Gray et al., 2015), metabolomic analysis of cerebrospinal fluid (CSF) using proton nuclear magnetic resonance (1H-NMR) spectroscopy for revealing nervous system cellular pathology was used. The 1H-NMR CSF metabolomic signature of ALS was sought in a longitudinal cohort. Six-monthly serial collection was performed in ALS patients across a range of clinical sub-types ($n = 41$) for up to two years, and in healthy controls at a single time-point ($n = 14$). A multivariate statistical approach, partial least squares discriminant analysis, was used to determine differences between the NMR spectra from patients and controls. Significantly predictive models were found using those patients with at least one year's interval between recruitment and the second sample. Glucose, lactate, citric acid and, unexpectedly, ethanol were the discriminating metabolites elevated in ALS. It is concluded that 1H-NMR captured the CSF metabolomic signature associated with derangements in cellular energy utilization connected with ALS, and was most prominent in comparisons using patients with longer disease duration. The specific metabolites identified support the concept of a hypercatabolic state, possibly involving mitochondrial dysfunction specifically. Endogenous ethanol in the CSF may be an unrecognized novel marker of neuronal tissue injury in ALS.

With regard to Alzheimer's disease (AD), an unbiased analysis of steroid-related compounds to identify novel plasma biomarkers using liquid chromatography-atmospheric pressure chemical ionization-mass spectroscopy was carried out by Sato et al. (2012). The analysis revealed that desmosterol was found to be decreased in AD plasma versus controls. To precisely quantify variations in desmosterol, we established an analytical method to measure desmosterol and cholesterol. Using this LC-based method, we discovered that desmosterol and the desmosterol/cholesterol ratio are significantly decreased in AD. Finally, the validation of this assay using 109 clinical samples confirmed the decrease of desmosterol in AD as well as a change in the desmosterol/cholesterol ratio in AD. Difference between mild cognitive impairment and control was also observed. In addition, the decrease of desmosterol was somewhat more significant in females. Receiver operating characteristic (ROC) analysis between controls and AD, using plasma desmosterol shows a score of 0.80, indicating a good discrimination power for this marker in the two reference populations and confirms the potential usefulness of measuring plasma desmosterol levels for diagnosing AD. Further analysis showed a significant correlation of plasma desmosterol with Mini-Mental State Examination scores. Plasma and serum biochemical markers proposed for Alzheimer disease (AD) are based on pathophysiologic processes such as amyloid plaque formation [amyloid β-protein (Aβ), Aβ autoantibodies, platelet amyloid precursor protein (APP) isoforms], inflammation (cytokines), oxidative stress (vitamin E, isoprostanes), lipid metabolism (apolipoprotein E, 24S-hydroxycholesterol), and vascular disease [homocysteine, lipoprotein (a)]. Most proteins or metabolites evaluated in plasma or serum thus far are, at best, biological correlates of AD: levels are statistically different in AD versus controls in some cohorts, but they lack sensitivity or specificity for diagnosis or for tracking response to therapy (Irizarry, 2004). Approaches combining panels of existing biomarkers or surveying the range of proteins in plasma (proteomics) show promise for discovering biomarker profiles that are characteristic of AD, yet distinct from nondemented patients or patients with other forms of dementia. In another study, plasma from 26 AD patients (mean MMSE 21) and 26 cognitively normal controls in a non-targeted approach using multi-dimensional mass spectrometry-based shotgun lipidomics was analysed by Han et al. (2011b) to

determine the levels of over 800 molecular species of lipids. These data were then correlated with diagnosis, apolipoprotein E4 genotype and cognitive performance. Plasma levels of species of sphingolipids were significantly altered in AD. Of the 33 sphingomyelin species tested, 8 molecular species, particularly those containing long aliphatic chains such as 22 and 24 carbon atoms, were significantly lower ($p<0.05$) in AD compared to controls. Levels of 2 ceramide species (N16:0 and N21:0) were significantly higher in AD ($p<0.05$) with a similar, but weaker, trend for 5 other species. Ratios of ceramide to sphingomyelin species containing identical fatty acyl chains differed significantly between AD patients and controls. MMSE scores were correlated with altered mass levels of both N20:2 SM and OH-N25:0 ceramides ($p<0.004$) though lipid abnormalities were observed in mild and moderate AD. Within AD subjects, there were also genotype specific differences.

2.2.2 Susceptibility parameters

Neurodevelopmental disorders with periventricular nodular heterotopia (PNH) are etiologically heterogeneous, and their genetic causes remain in many cases unknown. Missense mutations in NEDD4L mapping to the HECT domain of the encoded E3 ubiquitin ligase lead to PNH associated with toe syndactyly, cleft palate and neurodevelopmental delay (Broix et al., 2016). Cellular and expression data showed sensitivity of PNH-associated mutants to proteasome degradation. Moreover, an in utero electroporation approach showed that PNH-related mutants and excess wild-type NEDD4L affect neurogenesis, neuronal positioning and terminal translocation. Further investigations, including rapamycin-based experiments, found differential deregulation of pathways involved. Excess wild-type NEDD4L leads to disruption of Dab1 and mTORC1 pathways, while PNH-related mutations are associated with deregulation of mTORC1 and AKT activities. Altogether, these data provide insights into the critical role of NEDD4L in the regulation of mTOR pathways and their contributions in cortical development. De novo mutations in CHD8 are strongly associated with autism spectrum disorder, but the basic biology of CHD8 remains poorly understood. Chd8 knockdown during cortical development results in defective neural progenitor proliferation and differentiation that ultimately manifests in abnormal neuronal morphology and behaviors in adult mice (Durak et al., 2016) . Transcriptome analysis revealed that while Chd8 stimulates the transcription of cell cycle genes, it also precludes the induction of neural-specific genes by regulating the expression of PRC2 complex components. Furthermore, knockdown of Chd8 disrupts the expression of key transducers of Wnt signaling, and enhancing Wnt signaling rescues the transcriptional and behavioral deficits caused by Chd8 knockdown. These roles of Chd8 and the dynamics of Chd8 expression during development help negotiate the fine balance between neural progenitor proliferation and differentiation. Together, these observations provide new insights into the neurodevelopmental role of Chd8. Caubit et al. (2016) identified TSHZ3 as the critical region for a syndrome associated with heterozygous deletions at 19q12-q13.11, which includes autism spectrum disorder (ASD). In Tshz3-null mice, differentially expressed genes include layer-specific markers of cerebral cortical projection neurons (CPNs), and the human orthologs of these genes are strongly associated with ASD. Furthermore, mice heterozygous for Tshz3 show functional changes at synapses established by CPNs and exhibit core ASD-like behavioral abnormalities. These findings highlight essential roles for Tshz3 in CPN development and function, whose alterations can account for ASD in the newly defined TSHZ3 deletion syndrome. A genome-wide microRNA (miRNA) expression profiling in post-mortem brains from individuals with ASD and controls and identified miRNAs and co-regulated modules that were perturbed in ASD was performed by Wu

et al. (2016). Putative targets of these ASD-affected miRNAs were enriched for genes that have been implicated in ASD risk. A regulatory relationship between several miRNAs and their putative target mRNAs in primary human neural progenitors was confirmed. These include hsa-miR-21-3p, a miRNA of unknown CNS function that is upregulated in ASD and that targets neuronal genes downregulated in ASD, and hsa_can_1002-m, a previously unknown, primate-specific miRNA that is downregulated in ASD and that regulates the epidermal growth factor receptor and fibroblast growth factor receptor signaling pathways involved in neural development and immune function. In a recent application of genome-wide association studies (GWAS) to ASD, Inoue and Inoue (2016) indicated significant associations with the single nucleotide polymorphisms (SNPs) on chromosome 5p14.1, located in a non-coding region between cadherin10 (CDH10) and cadherin9 (CDH9). An in vivo bacterial artificial chromosome (BAC) based enhancer-trapping strategy in mice to scan the gene desert for spatiotemporal cis-regulatory activities was applied. The results showed, that the ASD-associated interval harbors the cortical area, striatum, and cerebellum specific enhancers for a long non-coding RNA, moesin pseudogene1 antisense (MSNP1AS) during the brain developing stages. Mouse moesin protein levels are not affected by exogenously expressed human antisense RNAs in transgenic brains, demonstrating the difficulty in modeling rather smaller effects of common variants. This in vivo evidence for the spatiotemporal transcription of MSNP1AS however provides a further support to connect this intergenic variant with the ASD susceptibility. With regard to human studies, to identify candidate genes for intellectual disability a meta-analysis on 2,637 de novo mutations, identified from the exomes of 2,104 patient–parent trios was performed (Lelieveld et al., 2016). Statistical analyses identified 10 new candidate ID genes: DLG4, PPM1D, RAC1, SMAD6, SON, SOX5, SYNCRIP, TCF20, TLK2 and TRIP12. In addition, it was showed that these genes are intolerant to nonsynonymous variation and that mutations in these genes are associated with specific clinical ID phenotypes. In addition, mutations in the aspartate/glutamate mitochondrial transporter, SLC25A12, have been associated with ASD (West et al., 2014).

Attention-deficit hyperactivity disorder (ADHD) is a prevalent and highly heritable disorder of childhood with negative lifetime outcomes. Although candidate gene and genome-wide association studies have identified promising common variant signals, these explain only a fraction of the heritability of ADHD. The observation that rare structural variants confer substantial risk to psychiatric disorders suggests that rare variants might explain a portion of the missing heritability for ADHD. A large-scale next-generation targeted sequencing study of ADHD in 152 child and adolescent cases and 188 controls across an a priori set of 117 genes was performed by Hawi et al. (2016). A multi-marker gene-level analysis of rare (<1% frequency) single-nucleotide variants (SNVs) revealed that the gene encoding brain-derived neurotrophic factor (BDNF) was associated with ADHD at Bonferroni corrected levels. Sanger sequencing confirmed the existence of all novel rare BDNF variants. BDNF is a genetic risk factor for ADHD, potentially by virtue of its critical role in neurodevelopment and synaptic plasticity.

For PD, mutations in the PINK1 gene are associated with early onset autosomal recessive parkinsonism (EOP), which is characterized by a phenotypic presentation that, although variable, generally overlaps with that of idiopathic Parkinson Disease (PD) (Rango et al., 2013). Rango et al. detailed clinical and brain metabolomic analysis of a sporadic Italian patient carrying the novel association of compound heterozygous A168P/W437X mutations in PINK1, including brain magnetic resonance spectroscopy

(MRS) with assessment of brain mitochondrial function, PET, and SPECT. They also studied the brain metabolomics of a control group of healthy subjects and a group of PD patients; both groups were age-matched and sex-matched., Also, several proteins encoded by PD-related genes are associated with mitochondria including PTEN-induced putative kinase 1 (PINK1), which was first identified as a gene that is upregulated by PTEN. Loss-of-function PINK1 mutations induce mitochondrial dysfunction and, ultimately, neuronal cell death. Mutations in the DJ-1 gene PARK7 are associated with hereditary recessive early onset PD (Andersen et al., 2016). Mutations in the PARK2 gene coding for Parkinson, an enzyme in the ubiquitin-proteasome system, are responsible for autosomal recessive juvenile parkinsonism. Genetic mutations of transthyretin (TTR), a protein transporting thyroxin (T4) and retinol, are related to familial amyloid polyneuropathy, and TTR seems to be related to diseases such as AD, PD and psychiatric disorders. In patients with early PD, higher urate levels conferred a milder clinical and radiographic progression of the disease (Mehta and Adler, 2015). An interesting genetic study went a step further to assess causality. DNA from 808 PD patients was genotyped for 3 SLC2A9 single-nucleotide polymorphisms (SNPs) that identify an allele associated with lower urate concentrations. They found that SNPs in SLC2A9 predicted differences in serum urate and the rate of progression to a level of disability requiring dopaminergic treatment was faster among those patients carrying the SLC2A9 genotypes associated with lower serum urate.

Regarding ALS, patients with D90A mutation on superoxide dismutase (SOD1) gene conformed the group of subjects more differentiated compared to other mutations and to sporadic ALS cases. In part that different fingerprint was due to decreased levels of primarily amino acids in CSF from carriers of D90A mutation (Ibanez et al., 2015)

2.3 Review of associations between environmental exposures and obesity and childhood diabetes

2.3.1 Review of early biological responses / biomarkers identification

With regard to diabetes, it has been found that diabetic men showed higher circulating levels of glucose, triglyceride, oxidized low-density lipoprotein (LDL), high-sensitivity C-reactive protein, interleukin (IL)-6, tumor necrosis factor-alpha (TNF-α), homeostasis model assessment-insulin resistance, urinary 8-epi-prostaglandin F_{2a} (8-epi-$PGF_{2\alpha}$) and ba-PWV than nondiabetic man (Ha et al., 2012). In plasma, 19 metabolites including three amino acids, eight acylcarnitines, six lysophosphatidylcholines (lysoPCs), and two lysophosphatidylethanolamines (lysoPEs; C18:2 and C22:6) significantly increased in diabetes men, whereas serine and lysoPE (C18:1) decreased. Decanoyl carnitine, lysoPCs (C14:0, C16:1, C18:1 and C22:6) and lysoPE (C18:1) with variable importance in the projection values >1·0 were major plasma metabolites that distinguished nondiabetic and diabetic men. Decanoyl carnitine positively correlated with oxidized LDL, 8-epi-PGF2α, IL-6, TNF-α and ba-PWV. ba-PWV correlated positively with lysoPCs C14:0 and C16:1, and negatively with lysoPE C18:1. 8-epi-PGF2α correlated positively with lipoprotein-associated phospholipase A2, ba-PWV and lysoPCs (C14:0 and C16:1). The receiver operating characteristic curve estimation suggested that decanoyl carnitine and lysoPC (C14:0) are the best metabolites for predicting the risk of developing diabetes. Circulating lipid-related intermediate metabolites can be closely associated with inflammation, oxidative stress

and arterial stiffness in early diabetes. The human plasma phospholipids in T2DM and DN have been also characterized by Zhu et al. (2011) to identify potential biomarkers of T2DM and DN. Normal phase liquid chromatography coupled with time of flight mass spectrometry (NPLC–TOF/MS) was applied to the plasma phospholipids metabolic profiling of T2DM and DN. As a result, 18 compounds in 7 PL classes with significant regulation in patients compared with healthy controls were regarded as potential biomarkers for T2DM or DN. Among them, 3 DM-specific biomarkers, 8 DN-specific biomarkers and 7 common biomarkers to DM and DN were identified. Ultimately, 2 novel biomarkers, i.e., PI C18:0/22:6 and SM dC18:0/20:2, can be used to discriminate healthy individuals, T2DM cases and DN cases from each other group. NMR-based metabolomic analysis in conjunction with multivariate statistics was applied to examine the urinary metabolic changes in two rodent models of type 2 diabetes mellitus as well as unmedicated human sufferers. The db/db mouse and obese Zucker (fa/fa) rat have autosomal recessive defects in the leptin receptor gene, causing type 2 diabetes (Salek et al., 2007). 1H-NMR spectra of urine were used in conjunction with uni- and multivariate statistics to identify disease-related metabolic changes in these two animal models and human sufferers. Metabolic similarities between the three species examined, including metabolic responses associated with general systemic stress, changes in the TCA cycle, and perturbations in nucleotide metabolism and in methylamine metabolism. All three species demonstrated profound changes in nucleotide metabolism, including that of N-methylnicotinamide and N-methyl-2-pyridone-5-carboxamide, which may provide unique biomarkers for following type 2 diabetes mellitus progression. A total of 19 serum amino acids in T2DM patients and non-diabetics were measured by Drabkova et al. (2015) and ther were 9 amino acids, which were significantly different in these groups ($p<0.05$). Significantly decreased levels of arginine, asparagine, glycine, serine, threonine and significantly increased levels of alanine, isoleucine, leucine, valine in diabetics were found. Significant difference in metabolism of amino acids between diabetics and non-diabetics were observed. The altered levels of amino acids in diabetic patients could be a suitable predictor of diabetes. Using comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry (GC × GC–TOFMS) coupled with pattern recognition methods, diabetic patients and healthy controls could be correctly distinguished based on the metabolic abnormity in plasma (Li et al., 2009). Five potential biomarkers including glucose, 2-hydroxyisobutyric acid, linoleic acid, palmitic acid and phosphate were identified. It was found that elevated free fatty acids were essential pathophysiological factors in diabetes mellitus which reflected either the hyperglycemia or the deregulation of fatty acids metabolism. These potential biomarkers in plasma, e.g. palmitic acid, linoleic acid and 2-hydroxybutyric acid might be helpful in the diagnosis or further study of diabetes mellitus. Similar results have been also obtained from the metabolomics study of Zhao et al. (2010), where normal glucose tolerant (NGT) and IGT subjects were clustered in two distinct groups independent of the investigated metabonome clearly. Pre-diabetes associated alterations in fatty acid-, tryptophan-, uric acid-, bile acid-, and lysophosphatidylcholine-metabolism, as well as the TCA cycle were identified, while Han et al. (2011a) showed that several species of arachidonic acids especially the class of C20 fatty acids might be useful indicators for distinguishing pathological abnormalities among populations in the developments of T2D and obesity. Using 2 highly sensitive metabolomic techniques, distinct serum profile change of a wide range of metabolites from healthy persons to type 2 diabetes mellitus have been reported (Xu et al., 2013). Apart from glucose, IFG and diabetes mellitus are characterized by abnormalities in amino acid, fatty acids, glycerophospholipids, and sphingomyelin metabolism. These early broad-spectrum metabolic changes

emphasize the complex abnormalities present in a disease defined mainly by elevated blood glucose levels. Similarly, Kim et al. (2012) demonstrated that oxidized -LDLM lysoPC (14:0), lycoPC (16:1), high sensitive-C reactive protein (hs-CRP) and interleukin-6 (IL-6) were strongly linked to the events of diabetic complications and higher arterial stiffness in T2D. Also, significant differences on global metabolism were also identified in the a metabonomic study involving 74 patients who were newly diagnosed with type 2 diabetes mellitus (T2DM) and a 48-week single drug treatment of repaglinide, metformin or rosiglitazone (Bao et al., 2009). As compared with the biochemical indices (FPG, 2hPG, HbA1c), T-predicted score showed different impacts on global metabolism by three treatments, respectively. Metabonomic analysis can reveal different treatment effects and provide a novel nonglucose based evaluation strategy for T2DM. A study investigated the differences in plasma metabolomic profiling between overweight/obese and normal-weight men using UPLC−Q-TOF MS (Kim et al., 2010). Overweight/obese (n = 30) and age-matched, normal-weight men (n = 30) were included. Three lyso-phosphatidylcholine (lysoPC) (lysoPC C14:0, lysoPC C18:0, and lysoPC C18:1) were identified as potential plasma markers and confirmed eight known metabolites (valine, leucine, propionyl carnitine, butyryl carnitine, tryptophan, hexanoyl carnitine, and l-carnintine) for overweight/obesity. Differences in plasma concentrations of >350 metabolites in fasted obese T2DM vs. obese non-diabetic African-American women were investigated principal components analysis to identify 158 metabolite components that strongly correlated with fasting HbA1c over a broad range of the latter ($r = -0.631$; $p<0.0001$) were utilized by Fiehn et al. (2010). In addition to many unidentified small molecules, specific metabolites that were increased significantly in T2DM subjects included certain amino acids and their derivatives (i.e., leucine, 2-ketoisocaproate, valine, cystine, histidine), 2-hydroxybutanoate, long-chain fatty acids, and carbohydrate derivatives. Leucine and valine concentrations rose with increasing HbA1c, and significantly correlated with plasma acetylcarnitine concentrations. It is hypothesized that this reflects a close link between abnormalities in glucose homeostasis, amino acid catabolism, and efficiency of fuel combustion in the tricarboxylic acid (TCA) cycle. Metabolomic profiling of obese versus lean humans reveals a branched-chain amino acid (BCAA)-related metabolite signature that is suggestive of increased catabolism of BCAA and correlated with insulin resistance. To test its impact on metabolic homeostasis, we fed rats on high-fat (HF), HF with supplemented BCAA (HF/BCAA) or standard chow (SC) diets (Newgard et al., 2009). Despite having reduced food intake and weight gain equivalent to the SC group, HF/BCAA rats were equally insulin resistant as HF rats. Pair-feeding of HF diet to match the HF/BCAA animals or BCAA addition to SC diet did not cause insulin resistance. Insulin resistance induced by HF/BCAA feeding was accompanied by chronic phosphorylation of mTOR, JNK, and IRS1(ser307), accumulation of multiple acylcarnitines in muscle, and was reversed by the mTOR inhibitor, rapamycin. the Findings showed that in the context of a poor dietary pattern that includes high fat consumption, BCAA contributes to development of obesity-associated insulin resistance.

Additional findings have been obtained using transcriptomics. It is commonly used to refer to the expression of genome-wide protein coding genes and we follow that convention here, although our thesis applies equally to the expression of non-protein coding genes (Jenkinson et al., 2016). T2D is strongly correlated with other complex disorders such as obesity, insulin resistance, the metabolic syndrome, dyslipidemia, cardiovascular disease, and a variety of diabetic complications. Transcriptomics data from different tissues including beta-cells, pancreatic islets, arterial tissue,

peripheral blood mononuclear cells, liver, and skeletal muscle of 228 samples were integrated with protein-protein interaction data and genome scale metabolic models to unravel the molecular and tissue-specific biomarker signatures of type 2 diabetes mellitus (Calimlioglu et al., 2015). Classifying differentially expressed genes, reconstruction and topological analysis of active protein-protein interaction subnetworks indicated that genomic reprogramming depends on the type of tissue, whereas there are common signatures at different levels. Basal fasting RNA was extracted from adipose tissue biopsies from a subset of 75 unrelated individuals, and gene expression data generated on the Illumina BeadArray platform(Winnier et al., 2015). The number of gene probes with significant expression above baseline was approximately 31,000. A multiple regression analysis of all probes with 15 metabolic traits was performed. Adipose tissue had 3,012 genes significantly associated with the traits of interest (false discovery rate, FDR ≤ 0.05). The significance of gene expression changes was used to select 52 genes with significant (FDR ≤ 10-4) gene expression changes across multiple traits. Gene sets/Pathways analysis identified one gene, alcohol dehydrogenase 1B (ADH1B) that was significantly enriched (P < 10-60) as a prime candidate for involvement in multiple relevant metabolic pathways.

2.3.2 Susceptibility parameters

Metabolic changes and the differences in inflammatory markers, oxidative markers and arterial stiffness between early diabetic and nondiabetic subjects remains relatively unstudied (Ha et al., 2012). Compared to non-diabetic men, patients with newly diagnosed type 2 diabetes also showed higher concentrations of TNF-a, IL-6 and hs-CRP. Higher arterial stiffness assessed by ba-PWV was also identified in patients with diabetes and represents a composite risk factor to identify patients with early atherosclerotic change. Along with the expected glycosuria, changes in the excretion of TCA cycle intermediates, polyols, amines, and amino acids were detected. A profound perturbation in nucleotide metabolism, previously linked with peroxisome proliferation, was also observed and may indicate a metabolic consequence of substrate excess in many tissues, especially the liver (Salek et al., 2007).

In terms of genetic polyrphisms, large-scale sample populations have revealed that there are more than 56 SNPs (Abu Bakar et al., 2015). The SNPs associated with diabetes are mostly found in non-coding regions of the genome that do not encode protein sequences. Regulation of the genomes inherently affects the gene expression and transcription. The SNPs in these metabolic genes are metabolically linked to particular enzymes that expressively lead to the production of specific metabolites of associated genes of interest. Transcription factor 7-like 2 (TCF7L2) polymorphism rs7903146 is identified to be the strongest genetic marker in type 2 diabetes, especially among the Caucasians (Abu Bakar et al., 2015). A SNP in the minor T-allele of rs1260326 in glucokinase (hexokinase 4) regulator (GCKR) might play its role in reducing the risk of type 2 diabetes susceptibility by lowering triglyceride accumulation and dyslipidemia and improving the fasting insulin and glucose level of the subjects. In addition, GCKR is a major pleiotropic risk locus associated with diabetes (Suhre et al., 2011). The melatonin-receptor gene, MTNR1B is linked to the changes in fasting glucose concentrations in T2D (Abu Bakar et al., 2015). Alos, the dysregulation of the phenylalanine hydroxylase gene by the hepatocyte nuclear factor 1a (Hnf1a) in phenylalanine metabolism is observed in T2D. Rare mutations in both KCNJ11 and PPARG are also known to be causal for certain rare monogenic syndromes (neonatal diabetes and lipodystrophies) characterized by severe metabolic

disturbance of β-cell function and insulin resistance (Prokopenko et al., 2008). Specific defects in glucagon-like peptide 1-stimulated insulin secretion, glucose-stimulated insulin secretion, insulin exocytosis, insulin granule docking or post-transcriptional processing of insulin have been demonstrated to be associated with different variants, supporting the notion that a range of biological processes are involved in the pathogenesis of type 2 diabetes (Grarup et al., 2014). Novel loci (Figure 1) such as KCNQ1 and C2CD4A, have been also associated with type 2 diabetes in Japanese individuals (Grarup et al., 2014).

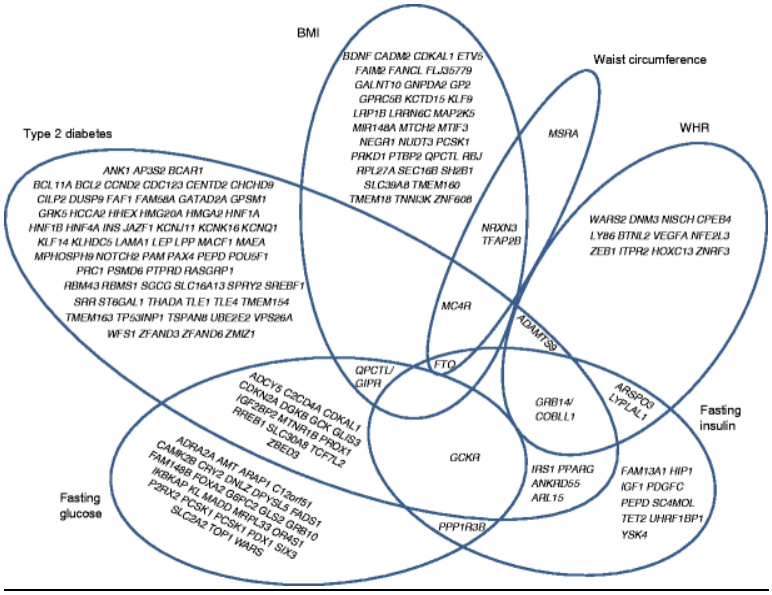

Figure. 1. Venn diagram of intersection between loci associated at genome-wide significance with type 2 diabetes, measures of adiposity and glucose homeostasis. Genome-wide significant associations for six metabolic traits are shown. Gene symbols shown in the plot are by convention the closest gene and not necessarily the functional gene (Grarup et al., 2014).

FTO and MC4R are two identified obesity locus. PCSK1, GP2 and GALNT10 loci in Asian or African populations have been identified for risk variants for obesity (Grarup et al., 2014). Other loci are: MC4R, POMC, LEPR, BDNF, SH2B1, PCSK1 and NTRK2. Five loci are shared both in diabetes and obesity: FTO, MC4R, ADAMTS9, GRB14/COBLL1 and QPCTL/GIPR. GRB14/COBLL1 is an example of an obesity-associated locus with pleiotropic effects on a range of phenotypes related to type 2 diabetes where all associations follow the expected metabolically unhealthy profile. Among genes identified through adipose tissue expression profiling to be regulated by obesity, 16 genes were selected that were encoded in four QTLs for human obesity susceptibility. In a comprehensive allelic association analysis (Jiao et al., 2008), SNPs in PFKFB3 and one haplotype in IRF5 were associated with obesity in cohort 1 with a nominal P-value <0.01. Though no statistically significant association with obesity was observed in a second cohort, PFKFB3 rs1064891 displayed a similar trend of association with obesity in both cohorts and for this SNP 95% confidence intervals of odds ratios from the two cohorts overlapped.

Also, a total of 118 polymorphisms in 16 genes were analyzed for association with obesity (Jiao et al., 2008). Single nucleotide polymorphism rs1064891, located in the 3' UTR of the 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3 (PFKFB3) gene, was nominally associated with obesity in combined analysis of cohorts 1 and 2 (P=0.007) and, in men that were lean or had severe obesity, with BMI (P=or<0.005).

Known obesity quantitative trait loci (QTL) information for the model allowed to further filter genes for increased likelihood of being causal or secondary for obesity. This successfully identified several genes previously linked to obesity (C1qr1, and Np3r) as positional QTL candidate genes elevated specifically in F line adipose tissue. A number of novel obesity candidate genes were also identified (Thbs1, Ppp1r3d, Tmepai, Trp53inp2, Ttc7b, Tuba1a, Fgf13, Fmr) that have inferred roles in fat cell function (Morton et al., 2011). Quantitative microarray analysis was then applied to the most phenotypically divergent adipose depot after exaggerating F and L strain differences with chronic high fat feeding which revealed a distinct gene expression profile of line, fat depot and diet-responsive inflammatory, angiogenic and metabolic pathways. Selected candidate genes Npr3 and Thbs1, as well as Gys2, a non-QTL gene that otherwise passed our enrichment criteria were characterized, revealing novel functional effects consistent with a contribution to obesity. Known obesity quantitative trait loci (QTL) information for the model allowed to further filter genes for increased likelihood of being causal or secondary for obesity. This successfully identified several genes previously linked to obesity (C1qr1, and Np3r) as positional QTL candidate genes elevated specifically in F line adipose tissue. A number of novel obesity candidate genes were also identified (Thbs1, Ppp1r3d, Tmepai, Trp53inp2, Ttc7b, Tuba1a, Fgf13, Fmr) that have inferred roles in fat cell function (Morton et al., 2011). Quantitative microarray analysis was then applied to the most phenotypically divergent adipose depot after exaggerating F and L strain differences with chronic high fat feeding which revealed a distinct gene expression profile of line, fat depot and diet-responsive inflammatory, angiogenic and metabolic pathways. Selected candidate genes Npr3 and Thbs1, as well as Gys2, a non-QTL gene that otherwise passed our enrichment criteria were characterized, revealing novel functional effects consistent with a contribution to obesity. The linkage of type 2 diabetes mellitus to chromosome 10q1 has been suggestive. 228 microsatellite markers were genotyped in Icelandic individuals with type 2 diabetes and controls throughout a 10.5-Mb interval on 10q. A microsatellite, DG10S478, within intron 3 of the transcription factor 7–like 2 genes (TCF7L2; formerly TCF4) was associated with type 2 diabetes (P = 2.1 times 10-9). This was replicated in a Danish cohort (P = 4.8 times 10-3) and in a US cohort (P = 3.3 times 10-9). Compared with non-carriers, heterozygous and homozygous carriers of the at-risk alleles (38% and 7% of the population, respectively) have relative risks of 1.45 and 2.41. This corresponds to a population attributable risk of 21%. The TCF7L2 gene product is a high mobility group box–containing transcription factor previously implicated in blood glucose homeostasis. It is thought to act through regulation of proglucagon gene expression in enteroendocrine cells via the Wnt signaling pathway2 (Grant et al., 2006).

3 Summary of biomarkers of early effects and susceptibility at the individual level

Table 1. Early effect and susceptibility biomarkers associated with the health effects relevant to HEALS

Health outcome	Biomarker	Type	Matrix	Study
	phosphatidylcholines	Metabolomics	Urine	(Ho et al., 2013)
	diglycerides			
	triglycerides			
	cholesterol			
	cortol			
	cholic acid			
Allergic Asthma	IL-5	Transcriptomics	PBMC	(Raedler et al., 2015)
	IL-13			
	IFN-γ			
	CLIC4			
	TSC1			
	IL37			
	PSTPIP2			
	TREM1			
	RGS13			
	RORC			
	RHEB			
	EIF4B			
	FCRL5			
Asthma	dodecanoic acid	Metabolomics	Plasma	(Yu et al., 2016)
	myristic acid			
	phytosphingosine			
	sphinganine			
	inosine			
	taurocholic acid			
	uric acid			
	lysoPC (22:6)			
	lysoPC (18:2)			
	lysoPC (20:4)			
	l-tryptophan			
	PS (18:2(9Z,12Z)/18:0)			

Health outcome	Biomarker	Type	Matrix	Study
	succinate	Metabolomics	Serum	(Chang et al., 2015)
	inosine			
	3,4-dihydroxybenzoic acid			
	phenylalanine			
Asthma	methionine	Metabolomics	Serum	(Amaral, 2014)
	glutamine			
	histidine			
	formate			
	methanol			
	acetate			
	choline			
	arginine			
	glucose			
	phosphatidylcholines			
	polyunsaturated phosphatidylcholines			
	Acetylated products		Exhaled breath condensate	
	oxidized compounds			
	urocanic acid			
	methyl-imidazoleacetic acid			
	isoleucyl-proline			
	ammonium ions			
	retinoic acid			
	deoxyadenosine-related metabolites			
Asthma	methionine	Metabolomics	Serum	(Jung et al., 2013)
	glutamine			
	histidine			
	formate			
	methanol			
	acetate			
	choline			
	O-phosphocholine			
	arginine			
	glucose			
Asthma	CL20	Transcriptomics	Nasal epithelial cells	(Castro-Nallar et al., 2015)
	IL1A			

Health outcome	Biomarker	Type	Matrix	Study
	RAK2			
	TNF			
	C8orf4			
	FAM46A			
	BIRC3			
	ATF3			
	JUN			
	CCL20			
	TIFA			
	TNFAIP3			
	CSF3			
	IL8			
	ARRDC3			
	NFKBIA			
Asthma	IL13	Transcriptomics	Nasal epithelial cells	(Poole et al., 2014)
	IL5			
	POSTN			
	CLCA1			
	SERPINB2			
	DPP4			
	CST1			
	TPSAB1			
	MSA42			
Childhood asthma	1-(methylsulfanyl)propane	Metabolomics	Exhaled breath condensate	(Amaral, 2014)
	ethylbenzene			
	1,4-dichlorobenzene			
	4-isopropenyl-1-methylcyclohexene			
	2-octenal			
	1-isopropyl-3-methylbenzene			
	1,7-dimethylnaphtalene			
	1-(methylsulfanyl)propane		Exhaled breath condensate	(Gahleitner et al., 2013)
	ethylbenzene			
	1,4-dichlorobenzene			
	4-isopropenyl-1-methylcyclohexene			
	2-octenal			
	octadecyne			

Health outcome	Biomarker	Type	Matrix	Study
	1-isopropyl-3-methylbenzene			
	1,7-dimethylnaphtalene			
	urocanic acid		Urine	(Mattarucchi et al., 2012)
	methyl-imidazoleacetic acid			
Obesity-related asthma	methane	Metabolomics	Exhaled breath condensate	(Maniscalco et al., 2016)
	pyruvate			
	glyoxylate			
	dicarboxylate			
Autism	sphingosine 1-phosphate	Metabolomics	Serum	(Wang et al., 2016)
	docosahexaenoic acid			
	Glycine/serine	Metabolomics	Urine	(Stein et al., 2015)
	Arginine			
	Glutamate			
	Aspartate			
	Glutamine			
	Alanine			
	indoxyl sulfate	Metabolomics	Urine	(Diémé et al., 2015)
	N-α-acetyl-l-arginine			
	methyl guanidine			
	phenylacetylglutamine			
Autism	*β*-alanine			(Mavel et al., 2013)
	glycine			
	taurine			
	succinate			
	creatine			
	3-methylhistidine			
	succinate	Metabolomics	Urine	(Emond et al., 2013)
	glycolate			
	hippurate			
	3-hydroxyphenylacetate			
	vanillylhydracrylate			
	3-hydroxyhippurate			
	4-hydroxyphenyl-2-hydroxyacetate			
	1*H*-indole-3-acetate			
	phosphate			
	palmitate			

Health outcome	Biomarker	Type	Matrix	Study
	stearate			
	3-methyladipate			
Autism	citric acid	Metabolomics	Urine	(West et al., 2014)
	glutaric acid			
	hippuric acid			
	phenylacetylglutamine			
	dimethylamine			
	succinic acid		Plasma	
	methylhexa-decanoic acids			
	hepta-decanoic acids			
	tetra-decanoic acids			
	3-aminoisobutyric acid			
	Creatinine			
	Aspartate			
	glutamate			
	DHEAS			
	isoleucine			
	homocitruline			
Autism	3-(3-hydroxyphenyl)-3-hydroxypropanoic acid	Metabolomics	Urine	(Noto et al., 2014)
	3,4-dihydroxybutyric acid			
	glycolic acid			
	glycine			
	cis-aconitic acid			
	phenylalanine			
	tyrosine			
	p-hydroxyphenylacetic acid			
	homovanillic acid			
	glycine	Metabolomics	Plama	(Heberling and Dhurjati, 2015)
	serine			
	threonine			
	alanine			
	histidine			
	glutamine			
	glutamate			
	taurine			

Health outcome	Biomarker	Type	Matrix	Study
	glutathione			
	thioredoxins			
Autism	Chd8	Transcriptomics	Cells	(Durak et al., 2016)
	PRC2			
	TSHZ3		cerebral cortical projection neurons	(Caubit et al., 2016)
	hsa-miR-21-3p		post-mortem brains	(Wu et al., 2016)
	hsa_can_1002-m			
	MSNP1AS		bacterial artificial chromosome	(Inoue and Inoue, 2016)
	MAL		human brain cortical tissue	(Gupta et al., 2014)
	C11orf30			
Autism	*A2BP1*	Transcriptomics	Brain Tissue	(Voineagu et al., 2011)
	M12			
	CADPS2			
	AH11			
	CNTNAP2			
	SLC25A12			
Bipolar Disorder	pyruvate	Metabolomics	Serum	(Yoshimi et al., 2016)
	N-acetylglutamic acid			
	α-ketoglutarate			
	arginine			
	β-alanine			
	serine			
Fetal alcohol spectrum disorders	L-thyroxine	Metabolomics	human embryonic stem cell-derived neural lineages	(Palmer et al., 2012)
	5'-methylthioadenosine			
	indoleacetaldehyde			
	L-kynurenine			
	Succinyladenosin			
Global Developmental Delay	methionine	Metabolomics	Dried blood stain	(Guo et al., 2011)
	tyrosine			
	glycine			
	citrulline			
	carnitine			
	palmitate			
	hydroxypalmitate			

Health outcome	Biomarker	Type	Matrix	Study
motor and language retardation	glutarate		Urine	
	3-hydroxyglutarate			
white matter lesions	glutamine	Metabolomics	?	(Wisnowski et al., 2013)
	glutamate			
	lactate			
Neurodevelopmental disorders	InSyn1	Transcriptomics	frozen brain sections	(Uezu et al., 2016)
	VGF			(Sha et al., 2012)
	NPAS3			
	SOX			
	dihydroxyacetone phosphate	Metabolomics		
	fructose-1,6-bisphosphate			
	succinate			
	α-ketoglutarate			
	NAD(+)			
Schizophrenia	2-hydroxyglutaric acid	Metabolomics	Plasma	(Napoli et al., 2015)
	cholesterol			
	fatty acid			
	lactate			
	pyruvate			
	ammonia			
	2-hydroxyvaleric acid			
	1,3-propanediol			
	citrulline			
	2,3-butanediol			
	HIF-1α	Transcriptomics		
	IGF1			
	MYC			
Neurodegenerative Disorders	phenylethylamine	Metabolomics	Plasma	(Giulivi et al., 2016)
	aconitate			
	isocitrate			
	oleamide			
Parkinson	myoinositol	Metabolomics	Brain areas	(Rango et al., 2013)
Parkinson	coumaric acid	Metabolomics	Urine	(Luan et al., 2015)
	tryptophan			
	tyrosine			

Health outcome	Biomarker	Type	Matrix	Study
	succinic acid			
	pimelic acid			
	lysine			
	hypoxanthine			
	pyridoxic acid			
	glutaric acid			
	hexanoylglycine			
	Indoleacetic acid			
	Aspartic acid			
	acetylphenylalanine			
	hydroxytryptophan			
	kynurenine			
	furoylglycine			
	cortisol			
	hydroxyphenylacetic acid			
	tiglylglycine			
	aminobutyric acid			
	xanthurenic acid			
	hydroxyprogesterone			
	isoleucine			
	alanine			
	leucine			
	phenylacetylglutamine			
	dihydrocortisol			
	phenylalanine			
Parkinson	cortisol	Metabolomics	Urine	(Luan et al., 2015)
	dihydrocortisol			
	hydroxyprogesterone			
	21-deoxycortisol			
	kynurenine			
	hydroxytryptophan			
	xanthurenic acid			
	hydroxyphenylacetic acid			
	acetylphenylalanine			
	phenylacetyglutamine			
	phenylalanine			

Health outcome	Biomarker	Type	Matrix	Study
	tyrosine	Metabolomics	Cerebrospinal fluid	(Luan et al., 2015)
	phenylalanine			
Parkinson	α-synuclein	Metabolomics	zebrafish	(Ren et al., 2016)
	ubiquitin			
	dodecanoic acid			
	hexadecanoic acid			
	octadecenoic acid			
	arachidonic acid			
	eicosanoic acid			
	propanoic acid			
	benzenedicarboxylic acid			
	butanoic acid			
	docosenoic acid			
	phthalic acid			
Parkinson	α synuclein	Metabolomics	Blood	(Mehta and Adler, 2015)
	DJ-1			
	uric acid			
	HNE-DJ-1)			
	polyamine			
	purine			
	pyruvate			
	redox			
	-hydroxykynurenine			
	N-acetylated amino acids			
	tryptophan			
	bilirubin			
	ergothioneine			
	levodopa			
	biliverdin			
	caffeine			
Parkinson	Lewy type α-synucleinopathy	Metabolomics	submandibular gland biopsy	(Mehta and Adler, 2015)
	α synuclein		Tissue	
	Aβ42		cerebrospinal fluid	
	tau protein			
	α synuclein			

Health outcome	Biomarker	Type	Matrix	Study
Parkinson	Glutathione	Metabolomics	Plasma	I(Ibanez et al., 2015)
	Uric acid			
	3-Hydroxykynurenine		cerebrospinal fluid	
	Oxidized glutathione			
	3-hydroxyisovaleric acid			
	tryptophan			
	creatinine			
	N^8-acetylspermidine		Serum	(Ibanez et al., 2015)
	Myo-inositol	Metabolomics	Brain	
	18:2 fatty acid			
	total fatty acid			
	N-acetylaspartate			
	phosphocreatine			
	phosphocholine			
Parkinson	Amino acids	Metabolomics	cerebrospinal fluid and plasma	(Ibanez et al., 2015)
	pyroglutamate			
	2-oxoisocaproate			
	3-Hydroxyisovaleric acid			
	tryptophan			
	creatinine			
	C16–C18 saturated			
Parkinson	454 metabolites	Metabolomics	neuroblastoma cell line	(Roede et al., 2014)
	ERK	Transcriptomics		
	heme oxygenase-1			
	sulfiredoxin-1			
	Slc7a11			
	Nrf2			
	PPAR-γ			
Parkinson	purine	Metabolomics	cerebrospinal fluid	(Andersen et al., 2016)
	tryptophan			
	NfL			
	amyloid			
	tau-proteins			
	noradrenaline			
	ApoE			
	purine metabolite xanthine over HVA			

Health outcome	Biomarker	Type	Matrix	Study
	3-hydroxykynurenine			
	UCHL-1			
	Aβ1-42			
	kynurenic acid			
	urate			
	copper			
	flt3-ligand			
	3-HK/KYNA			
	hydroxyindoleacetic acid			
	GSSG			
	Neuromodulin			
	8-hydroxyguanosine			
	silicic acid			
	BDNF			
	Lipid peroxidation			
	interleukins			
	Clusterin			
	GFAP			
	YKL-40			
	Ubiquitin			
	UCHL-1			
	Apolipoprotein A1 and A2			
	Tryptophan			
	Transthyretin			
	Albumin			
	Creatinine			
	Xylitol			
	GAP-43			
	vitamin D-binding protein			
	Haptoglobin			
	Corticosterone			
	8-hydroxyguanosine			
	β2-microglobulin			
Parkinson	DOPAC	Metabolomics	catecholaminergic neurons	(Andersen et al., 2016)
	homovanillic acid			
	DHPG			

Health outcome	Biomarker	Type	Matrix	Study
	(MHP			
	DJ-1	Transcriptomics	cerebrospinal fluid	(Andersen et al., 2016)
	a-synuclein			
	glutathione S-transferase P			
	D-growth factors			
	sCD14			
Parkinson	GFP-LC3 puncta	Transcriptomics	*PINK1*-deficient cells	(Gómez-Sánchez et al., 2016)
	LC3-II			
Parkinson	tryptophan	Metabolomics	Serum	(Hatano et al., 2016)
	caffeine			
	bilirubin			
	ergothioneine			
Parkinson	*N*8-acetyl spermidine	Metabolomics	Serum	(Botas et al., 2015)
	polyamine			
	8-Hydroxy-2-deoxyguanosine		Plasma	
	glutathione			
	Uric acid			
	3-hydroxykynurenine		CSF	
	glutathione			
	urate			
Alzheimer Disease	Glycine	Metabolomics	Brain	(Ibanez et al., 2015)
	lactate			
	aspartate			
	leucine			
	Taurine			
	N-acetyl-l-aspartate			
	glutamate			
	glutamine			
	valine			
Alzheimer Disease	Beta-alanine	Metabolomics	cerebrospinal fluid	(Ibanez et al., 2015)
	aspartate			
	asparagine			
	alanine			
	l-cysteine			
	Carnitine			

Health outcome	Biomarker	Type	Matrix	Study
	l-methionine			
	methionine-cysteine-glutamate			
	Uridine			
	Methylsalsolinol			
	taurine			
	Creatinine			
	dopamine			
	quinone			
	Vanillylmandelic acid			
	xanthine			
Alzheimer Disease	Glutamine	Metabolomics	cerebrospinal fluid and plasma	(Ibanez et al., 2015)
	2-aminoadipic acid			
	Linoleic acid			
	butedioic acid			
	arachidonic acid			
	N,*N*-dimethylglycine			
	oxoproline			
	glutamic acid			
	proline			
Alzheimer Disease	Choline	Metabolomics	Serum	(Ibanez et al., 2015)
	creatinine			
	dimethyl-arginine			
	cysteine disulfide			
	Ceramides			
	diacylglycerols			
	fatty acids			
	Lipidic molecules			
	Asparagine			
	methionine			
	histidine			
	acetyl-spermidine			
	acyl carnitines			
	phenylalany-phenylalanine			
	Methionine			
Alzheimer Disease	5-Hydroxyindoleacetic acid	Metabolomics	Urine	(Ibanez et al., 2015)
	N^1-acetylspermidine			

Health outcome	Biomarker	Type	Matrix	Study
	desaminotyrosine			
Alzheimer Disease	Lysophosphatidyl choline	Metabolomics	Plasma	(Ibanez et al., 2015)
	beta-alanine			
	l-cysteine			
	l-methionine			
	aspartate			
	asparagine			
	methionine-cysteine-glutamate			
	arachidonic acid			
	N,*N*-dimethylglycine			
	thymine			
	glutamine			
	glutamic acid			
	cytidine			
Alzheimer Disease	desmosterol	Metabolomics/Lipodomics	Plasma	(Sato et al., 2012)
	cholesterol			
Alzheimer Disease	Choline	Metabolomics	cerebrospinal fluid	(Kang et al., 2015)
	dimethylarginine			
	arginine			
	valine			
	serine			
	histidine			
	creatine			
	carnitine			
	suberylglycine			
	Aβ42			
	phosphorylated tau			
	total tau			
	methionine			
	5-HIAA			
	vanillylmandelic acid			
	xanthosine			
Alzheimer Disease	Norepinephrine	Metabolomics	cerebrospinal fluid	(Trushina and Mielke, 2014)
	methionine			
	alpha-tocopherol			
	5-hydroxytryptophan			

Health outcome	Biomarker	Type	Matrix	Study
	tyramine			
	Choline			
	valine,			
	carnitine			
	serine			
Alzheimer Disease	GCA	Metabolomics	Plasma	(Trushina and Mielke, 2014)
	GCDCA			
	GDCA			
	N16:0			
	N21:0			
	Lysophospolipid 18:1			
Alzheimer Disease	phosphatidylcholine		Serum	(Trushina and Mielke, 2014)
	plasmalogens			
	sphingomyelins			
	sterols			
	Dihydroxybutanoic acid			
Alzheimer Disease	sphingolipid	Lipodomics	Plasma	(Trushina and Mielke, 2014)
	Amyloid β-protein	Metabolomics		(Irizarry, 2004)
	Brain-plasma Aβ flux			
	Aβ autoantibodies			
	Platelet APP isoforms			
	Cholesterol			
	24S-Hydroxycholesterol			
	Homocysteine			
Alzheimer Disease	sphingolipids	Lipidomics	Plasma	(Han et al., 2011b)
	N16:0			
	N21:0			
	N22:1			
	N20:0			
	N23:0			
	N24:2			
	N17:1			
	ceramide			
	20:2 SM			
	OH-N25:0 ceramide			

Health outcome	Biomarker	Type	Matrix	Study
	OH-N24:2			
	OH-N24:1			
	N28:2			
Alzheimer Disease	desmosterol	Metabolomics	Plasma	(Botas et al., 2015)
	cholesterol			
	sphingomyelin lipids			
	ceramide species			
Alzheimer Disease	choline		CSF	(Botas et al., 2015)
	dimethylarginine			
	arginine			
	valine			
	proline			
	serine			
	histidine			
	creatine			
	carnitine			
	cortisol			
	cysteine			
	methionine			
	5-hydroxyindoleacetic acid			
	vanillylmandelic acid			
	xanthosine			
	tyrosine			
	tryptophan			
	purine			
	norepinephrine			
Alzheimer Disease	glutamate	Metabolomics	Hippocampus, cortex, midbrain and cerebellum tissues	(Botas et al., 2015)
	glutamine			
	taurine			
	gamma-amino butyric acid			
	choline			
	succinate			
	lactate			
	aspartate			
	glycine			
	alanine			

Health outcome	Biomarker	Type	Matrix	Study
	iso-leucine			
Alzheimer Disease	phospholipids	Metabolomics	Serum	(Botas et al., 2015)
	2,4-dihydroxybutanoic acid			
	phosphatidylcholines			
ALS	Lactate	Metabolomics	cerebrospinal fluid	(Ibanez et al., 2015)
	acetate			
	sum of glutamine and glutamic acid			
	lactate		brain	
	N-acetyl-aspartate			
	creatine			
	choline			
	myo-inositol			
ALS	Phosphate	Metabolomics	Plasma	(Lawton et al., 2014)
	Cortisone			
	3-methylxanthine			
	Delta-tocopherol			
	Creatine			
	Theobromine			
	Iminodiacetate			
	Palmitoyl sphingomyelin			
	Octadecanedioate			
	7-methylxanthine			
	3-dehydrocarnitine			
	Urate			
	1,2-propanediol			
	4-vinylphenol sulfate			
	Serine			
	Cysteine			
	Proline			
	Hexadecanedioate			
	2-hydroxybutyrate			
	Alpha-ketobutyrate			
	1-methylurate			
	Pyroglutamine			
	Dodecanedioate			
	Paraxanthine			

Health outcome	Biomarker	Type	Matrix	Study
	Creatinine			
	1-stearoyl-GPI (18:0)			
	Arachidonate (20:4n6)			
	Glutamine			
ALS	Glu	Metabolomics	Plasma	(Patin et al., 2016)
	val			
	aminoadipic acid			
	spermidine			
	creatinine			
	met			
	Glutamic acid			
	ADMA			
	L-DOPA			
	histidine			
	polyamine			
	dimethylarginine			
	putrescine			
	Asn			
	Phe			
	Ser			
	spermidine			
	dimethylargnine			
	arginine			
	proline			
	tryptophan			
	branched amino acid			
	N-acetyl-ornithine			
	ornithine			
	putrescine			
	hydroxyproline			
	2-Aminoadipate			
	Carnosine			
	lysophosphatidylcholine		cerebral cortex	
ALS	creatine kinase	Transcriptomics	Plasma	(Patin et al., 2016)
	proline 4 hydroxylase			
	spermidine synthase			

Health outcome	Biomarker	Type	Matrix	Study
	pyruvate dehydrogenase			
	beta (Pdhb)			
	dihydrolipoamide dehydrogenase			
	pyruvate deshydrogenase complex (PDC)			
	arylalkylamine N-acetyltransferase			
	3-hydroxyanthranilate			
	3,4-dioxygenase			
	kynureninase			
ALS	3-Methylxanthine	Metabolomics	Plasma	(Blasco et al., 2016)
	theobromine			
	theophylline			
	paraxanthine			
	1-methyluric acid			
	L-serine			
	L-histidine			
	Pyruvic acid			
	alpha-ketobutyrate			
	2-oxoisovaleric acid			
	Pantothenate			
	Pyruvic acid			
	L-glutamine			
	D-glutamic acid			
	L-serine			
	L-tyrosine			
	sarcosine			
	Formic acid			
	Acetic acid			
	Acetone			
	Pyruvic acid			
	Ascorbic acid			
	pyroglutamic acid			
	pantothenic acid			
ALS	glutamate	Metabolomics	Serum	(Botas et al., 2015)
	beta-hydroxybutyrate			
	acetate			

Health outcome	Biomarker	Type	Matrix	Study
	acetone			
	formate			
	glutamine			
	histidine			
	N-acetyl derivatives			
	acetate	Metabolomics	CSF	(Botas et al., 2015)
	acetone			
	pyruvate			
	ascorbate			
	glutamine			
	glutamate			
ALS	Glucose	Metabolomics	CSF	(Gray et al., 2015)
	lactate			
	citric acid			
	ethanol			
	glycolytic			
	citric acid			
	creatine		Plasma	
	ascorbate			
	acetone			
	glutamate			
	β-hydroxybutyrate			
	glutamate		Serum	
Multiple Sclerosis	choline	Metabolomics	CSF	(Botas et al., 2015)
	myo-inositol			
	threonate			
	3-hydroxybutyrate			
	citrate			
	phenylalanine			
	2-hydroxyisovalerate			
	formate			
	acetate			
	lactate			
	creatinine			
	phenylalanine			
	sorbitol			

Health outcome	Biomarker	Type	Matrix	Study
	fructose			
	arginine			
	alanine			
	O-phosphoethanolamine			
	putrescine			
	mannose	Metabolomics	Plasma	
	fatty acid			
	bile acid			
	taurine			
	tryptophan			
	histidine			
	linoleic			
	d-arginine			
Huntington Disease	NAA	Metabolomics	Serum and CSF	(Botas et al., 2015)
	glutamine			
	succinic acid			
	glucose			
	lactate			
	glucose			
	malonate	Metabolomics	Serum	
	procatabolic			
	NAA		Striatum	
	γ-aminobutyric acid			
	aspartate			
	alanine			
	acetate			
	choline			
	phosphocholine			
	glycerophosphocholine			
	glutamine			
	taurine			
	myo-inositol			
	lactate			
	creatine			
Huntington Disease	4-aminobutyrate	Metabolomics	post-mortem human brain	(Graham et al., 2016)
	glycine			

Health outcome	Biomarker	Type	Matrix	Study
	formate			
	l-glutamic acid			
	tyrosine			
	l-phenylalanine			
	aspartate			
	inosine			
	myo-inositol			
	taurine			
	urea			
	uracil			
	l-leucine			
	valine			
	niacinamide			
	tyrosine			
	l-phenylalanine			
	ethanolamine			
	homocitrulline			
	N-acetylaspartic acid			
	niacinamide			
Obesity	glycine	Metabolomics	Serum	(Zhang et al., 2013)
	glutamine			
	glycero-phosphatidylcholine 42:0			
	PCaa 32:0			
	PCaa 32:1			
	PCaa 40:5			
	lysoPC C14:0		Plasma	
	lysoPC C18:0			
	lysoPC C18:1			
	trigonelline		Gut flora	
	2-hydroxyisobutyrate			
	hippuric acid			
	xanthine			
Obesity	C14:0	Metabolomics	Serum	(Newgard et al., 2009)
	C16:0		Urine	
	C16:1			
	C18:1			

Health outcome	Biomarker	Type	Matrix	Study
	C20:4			
	palmitoleate			
	glycine			
	ethylmalonate			
	isobutyryl glycine			
	isovaleryl glycine			
	α-ketoglutarate			
	BCAA			
	methionine			
	Glx			
	phenylalanine			
	tyrosine			
	C5 acylcarnitines			
Obesity	lysoPC C14:0	Metabolomics	Plasma	(Kim et al., 2010)
	lysoPC C18:0			
	lysoPC C18:1			
	valine			
	leucine			
	propionyl carnitine			
	butyryl carnitine			
	tryptophan			
	hexanoyl carnitine			
	L-carnintine			
Obesity	C1qr1	Transcriptomics	Adipose tissue	(Morton et al., 2011)
	Np3r			
	Thbs1			
	Ppp1r3d			
	Tmepai			
	Trp53inp2			
	Ttc7b			
	Tuba1a			
	Fgf13			
	Fmr			
	PFKFB3		Abdominal subcutaneous adipose tissue biopsies	(Jiao et al., 2008)

Health outcome	Biomarker	Type	Matrix	Study
Obesity-related asthma	methane	Metabolomics	Exhaled breath condensate	(Maniscalco et al., 2016)
Childhood obesity	unsaturated lipids	Metabolomics	Serum	(Zhang et al., 2013)
	myoinositol			
	pyruvate			
	threonine			
	tyrosine			
	glucose			
	urea			
T1D	Azacitidine	Metabolomics	Plasma	(Dutta et al., 2012)
	4-Oxo-norfloxacin			
	Hydroxyhydroquinone			
	Pirenzepine			
	Phenylethylamine			
	2,3,4-Trihydroxybenzylhydrazide			
	4-(diaminomethylideneamino)butanoic acid			
	3-Indoleacetic Acid			
	Tyramine			
	Cinnamic acid			
	COX5B	Transcriptomics	Skeletal Muscle cells	
	COX10			
	ubiquinol cytochrome c reductase			
	ATP5F1			
	cAMP-dependent protein kinase (PKA)			
	protein kinase B (Akt)			
	prostaglandin 12 synthase			
	X-linked inhibitor of apoptosis			
	mitogen-activated protein kinase kinase kinase 7 (TAK1)			
T2D	3-hydroxybutyrate	Metabolomics	Serum	(Zhang et al., 2014)
	altered bile acids			
	PI C18:0/22:6			
	SM dC18:0/20:2			
	α-hydroxybutyrate			Xu et al. 2013
	proline			
	phenylalanine			

Health outcome	Biomarker	Type	Matrix	Study
	myristic			
	palmitic			
	stearic acid			
T2D	ornithine	Metabolomics	Urine	(Salek et al., 2007)
	histidine			
	leucine			
	isoleucine			
	2-oxoisovalerate			
	N-methylnicotinamide			
	tryptophan/tryptamine			
	N-methyl-2-pyridone-5-carboxamide			
	non-esterified fatty acids (NEFAs)	Metabolomics	Plasma	(Han et al., 2011b)
	esterified fatty acids (EFAs)			
	decanoyl carnitine			(Ha et al., 2012)
	ox-LDL			
	8-epi-PGF2α			
	lysoPC (C14:0)			
T2D	ADCY5	Transcriptomics	HapMap CEU sample data	(Dupuis et al., 2010)
	ADCY5			
	GCK			
	DGKB-TMEM195			
	CRY2			
	FADS1			
	GLIS3			
	SLC2A2			
	PROX1			
	C2CD4B			
	HOMA-IR			
	FG/T2D-associated locus MTNR1B			
	G6PC2			
	TCF7L2			
	SLC30A8			
	MADD			
	FAM148B			
	Pro12Ala - PPARG gene	Transcriptomics		(Prokopenko et al., 2008)
	Glu23Lys - KCNJ11 gene			

Health outcome	Biomarker	Type	Matrix	Study
	DG10S478			(Grant et al., 2006)
T2D	Mannosidase Alpha Class 1A Member 2	Transcriptomics	beta-cells, pancreatic islets, arterial tissue, peripheral blood mononuclear cells, liver, and skeletal muscle	(Calimlioglu et al., 2015)
	miR-335			
	miR-16-5p			
	ITIH4	Transcriptomics	pancreatic islets, arterial tissue, and beta cells	
	AKAP13			
	ZFP36L2-EIF4E-MTRR			
	SMAD3		beta cells	
	SKP1			
	YWHAQ			
	CTNNB1			
	CSNK2A1			
	ESR1	Transcriptomics	pancreatic islets	
	CSNK2A1			
	MYC			
	SMAD3			
	RELA		PBMC	
	TP53			
	EGR1			
	NCOR1			
	BCL6			
	NR3C1			
	SMAD1			
	PPARA		liver	
	CSNK2A1			
	MDM2		skeletal muscle	
T2DM	*m/z* 802.5	Metabolomics	Plasma	(Zhu et al., 2011)
	m/z 716.5			
	m/z 885.5			
	m/z 540.3			
	m/z 790.5			
	m/z 738.5			
	m/z 909.5			
	m/z 568.3			

Health outcome	Biomarker	Type	Matrix	Study
	m/z 854.5			
	PI [M−H]$^-$ *m/z* 909			
	SM [M−HCOO]$^-$ *m/z* 801			
	valine	Metabolomics	serum	(Bao et al., 2009)
	maltose			
	urate			
	octadecanoate			
	arachidonate			
	glutamate			
	butanoate			
	hippuric acid			(Zhao et al., 2010)
	methylxanthine			
	methyluric acid			
	3-hydroxyhippuric acid			
	glucose	Metabolomics	plasma	(Li et al., 2009)
	2-hydroxyisobutyric acid			
	linoleic acid			
	palmitic acid			
	phosphate			
	arginine		Serum	(Drabkova et al., 2015)
	asparagine			
	serine			
	threonine			
	valine			
	tryptophan			
T2D and Obesity	leucine	Metabolomics	Plasma	(Fiehn et al., 2010)
	2-ketoisocaproate			
	valine			
	cysteine			
	histidine			
	2-hydroxybutanoate			
	carbohydrate derivatives			
	glutamate			
	lactate			
	phosphoric acid			

Health outcome	Biomarker	Type	Matrix	Study
	alcohol dehydrogenase 1B (ADH1B)	Transcriptomics	abdominal subcutaneous adipose tissue biopsies	(Jenkinson et al., 2016)

4 Statistics for biomarkers clustering and selection

4.1 Hierarchical clustering analysis

The cluster analysis of the variables provides the ability of arrange numerical variables to hierarchical clusters (Berkhin, 2006). In particular, this analysis identifies cluster to homogenous hierarchical groups of cases that are not always previously and initially observable. Moreover, the procedure of hierarchical clustering relies on a discriminant analysis that examines redundancy between variables and the statistical significant of the groups and it is used for assessing collinearity of variables (Anderberg, 2014).

The principle approach of Hierarchical Clustering (AHC) is an iterative classification method that perform the following steps:

1. Initially it is assigned every object to a separate cluster.
2. Then it is calculated the dissimilarity between the N objects.
3. Then a cluster of two object is created if two objects minimize the agglomeration criterion
4. Finally, the dissimilarity between this class and the N-2 other objects is calculated using the agglomeration criterion.

An appropriate *metric-distance* can be a criterion to measure the similarity in shape between two profiles, but can also capture inverse relationships. This process is repeated until all the objects have been clustered. The results of the procedure is a binary clustering tree whose root is the class that contains all the observations. Thought this procedure the tree can been divided to levels according the user objective criteria. A very common metric is Euclidean distance (Danielsson, 1980) which is applying based on the formula:

$$\|a-b\|_2 = \sqrt{\sum_i (a_i + b_i)^2}$$

In this study the Hmisc R package (Harrell Jr and Dupont, 2006) was used to apply the Hoeffding D statistic and squared Pearson correlations. The results of the analysis was illustrated in a full cluster tree (dendrogram) depicturing variables in each grouping as well the distance between groupings.

The Pearson's Correlation is a linear correlation metric that measures the similarity in shape between two profiles while the Hoeffding D has the advantage of being sensitive to many types of dependence, including highly non-monotonic relationships (Hoeffding, 1948).

4.2 Environment Wide Associations Study (EWAS)

Environment-Wide Association Study (EWAS) is a population-based data analysis approach that correlates multiple environmental factors to disease. The EWAS approach has been introduced by Patel et al. (2010) and it is based on Genome-Wide Association Study (GWAS). In GWAS multiple

genetic factors are assayed along with phenotypic information on each individual. In particular, GWAS enhance the discovery of genetic variants that they can postulate about the function of the pathways discovered in diseased individuals. In other words, the genetic factors are the independent variables, and the phenotype is the dependent variable. The GWAS have been widely and effectively applied to emerge the missing heritability of complex diseases (Manolio et al., 2009). Especially, among a wide range of studies and researches, GWAS have used to provide new insights into type 2 diabetes aetiology (Frayling, 2007), to investigate genomic characteristics of trait/disease-associated SNPs (TASs) (Hindorff et al., 2009) , to map chromatin marks across cell types to systematically characterize regulatory elements, their cell-type specificities and their functional iterations (Ernst et al., 2011) and to examined the role of common genetic variation in schizophrenia (Ripke et al., 2011).

EWAS has been conducted to evaluated the hypotheses regarding the broad contribution of the environment to disease (Patel et al., 2010). In EWAS, the place of the genome domain is replaced by the envirome domain. The Envirome consists of the environmental factors which are the quantity of the individual exposures that has directly measured. Particularly, the measures can be the amount of a chemical substances in human tissues and organs, a self-report historical exposure and common well–being characteristics such as family and social economic status (SES). Hence, Patel et al. (2012 conducted EWAS to comprehensively and systematically explore and associate multiple environmental factors discovering and replicating robust correlations with serum lipid levels. Also, EWAS was used to evaluate possible additive role of both contaminants and lifestyle factors regarding Metabolic Syndrome revealing the existent of associations in the examined elderly population (Lind et al., 2013).

The EWAS framework can carried out a systematic sensitivity analyses, whereby validated factors are modeled under different assumptions or with additional covariates. Also, using a pair-wise validation method it is computed the correlation of dependence between the factors, revealing potential evidence for exposure or confounding route. The EWAS framework and analysis approach is based on the GWAS and the framework of EWAS were introduced by Patel et al. (2010). Firstly, in the proposed framework it is conducted an initial scan for environmental factors associated with the observed effects/variables through general linear modeling (e.g logistic regression). The model takes into consideration the factor which have been adjusted for known confounder considering the existent environmental association. Secondly, the false discovery rate is applied on multiple hypotheses. Finally, factors that it is deemed that significantly associated with the observed factors beyond the region of false discovery are “validated” in independent cohorts.

5 Usefulness and feasibility assessment

5.1 Application on the REPRO_PL cohort

5.1.1 Introduction

The Polish Mother and Child Cohort (REPRO_PL, www.repropl.com), is a multicenter prospective cohort study of environmental factors contributing to pregnancy outcomes, children's health and neurodevelopment, established in 2007. Briefly, women were recruited during the first trimester of pregnancy at maternity units or clinics in selected regions of Poland if they fulfilled the following inclusion criteria:

- single pregnancy up to 12 weeks of gestation,
- no assisted conception,
- no pregnancy complications and
- no chronic diseases as specified in the study protocol.

The women were interviewed 3 times during the pregnancy (once in each trimester) in order to collect and update demographic and socio-economic data, medical and reproductive history, and information about occupational exposure and lifestyle factors. After delivery, detailed information regarding pregnancy outcomes and children's health was obtained. Two years after birth, assessment of exposure, health status as well as neurodevelopment of each child was performed by a pediatrician as well as a psychologist/child development specialist. At that time also information concerning socio-demographic, environmental and lifestyle factors have been updated by conducting interviews with mothers. Current analysis was restricted to 149 mother-child pairs from Lodz (central Poland) district.

The overall analysis consists of

5.1.2 Data used in the analysis

5.1.2.1 Exposure factors

For the association, several exposure factors have been investigated, including:

- Prenatal phthalate exposure, by measuring phthalate metabolites in the urine collected from the mothers during the third trimester of pregnancy (range 30–34 weeks)

- postnatal exposure was determined by measuring phthalate metabolites in the urine from children at around the 24th month of age (range 23– 28 months). In both cases the following 11 phthalate metabolites were measured: a) five Low-MWP: monoethyl phthalate (MEP) (metabolite of diethyl phthalate; DEP), mono-iso-butyl phthalate (MiBP) (metabolite of di-iso-butyl phthalate; DiBP), mono-n-butyl phthalate (MnBP) (metabolite of di-n-butyl phthalate; DnBP), 3OH-mono-n-butyl phthalate (3OH-MnBP) (secondary metabolite of DnBP), and monobenzyl phthalate (MBzP) (metabolite of butyl-benzyl phthalate; BBzP), and b) six High-MWP: mono(2- ethylhexyl) phthalate (MEHP) (metabolite of di(2-ethylhexyl) phthalate; DEHP), 5OH-mono(2-ethylhexyl) phthalate (5OH-MEHP) and 5oxo-mono(2-ethylhexyl) phthalate

(5oxo-MEHP) (secondary metabolites of DEHP), 7OH-mono-methyloctyl phthalate (OH-MiNP) and 7oxo-mono-methyloctyl phthalate (oxo-MiNP) (secondary metabolites of di-iso-nonyl phthalate; DiNP), and mono-n-octyl phthalate (MnOP) (metabolite of di-n-octyl phthalate; DnOP).

- Prenatal exposure to tobacco constituents was assessed based on the cotinine level in saliva collected from the women during the third trimester of pregnancy
- Child exposure to environmental tobacco smoke (ETS) was assessed based on the cotinine level in the urine collected during the visit scheduled for assessment of child health and neurodevelopment
- Pb in the mother plasma during during pregnancy, delivery and in cord blood
- Hg in maternal hair during the 30-34th week of pregnancy
- 1-HP in mother urine during the 20-24th week of pregnancy

5.1.2.2 Exposure and effect modifiers

Additional factors considered as exposure and effects modifiers were included as well. These included:

- Sociodemographic parameters such as
  - Socioeconomic status
  - Mother education
  - Marital status
  - Day-care attendance at one and two years of age
- Mother and child anthropometric parameters
  - Mother bodyweight
  - Mother height
  - Mother age at birth
  - Child birth weight
  - Child gender
- Other gestational and post-delivery factors
  - Child birth weight
  - Mode of delivery
  - Breastfeeding
- Presence of micronutrients, minerals and vitamins
  - Zn in the mother plasma during pregnancy, delivery and in cord blood
  - Se in the mother plasma during pregnancy, delivery and in cord blood

  - Cu in the mother plasma during pregnancy, delivery and in cord blood
  - Vitamin D2
  - Vitamin D3

- Enzymatic activity
  - Thiobarbituric acid reactive substances in the mother plasma during pregnancy
  - GPX3 glutathione peroxidase 3 in the mother plasma during pregnancy
  - GPX1 glutathione peroxidase 1 in the mother plasma during pregnancy
  - Superoxide dismutase in the mother plasma during pregnancy
  - Ceruloplasmin in the mother plasma during pregnancy

5.1.2.3 Health outcomes

The health outcomes investigated in this study included children neurodevelopment, assessed at around the 24th (±1.2) month of age (range, 23 to 28 months) using the Bayley Scales of Infant and Toddler Development, third edition (Bayley-III). Examination was performed at pediatric medical centers at two University Hospitals in Lodz. The testing was done in the presence of the mother or a relative by a psychologist or a child development specialist. Before commencement of the examination the training of examiners in the field of the Bayley-III was performed by a professional. Bayley-III is an individually applied examination that assesses the developmental functioning of children up to 42 months of age. The test presents child with situations and tasks designed to produce an observable set of behavioral responses. In the current analysis, focus was on child cognitive, motor and language development. The child psychomotor development measured by row score/chronologic age was yielded with each subtest, and composite scores for language, motor scales, and composite score equivalent for cognitive scale were generated based on such data.

5.1.3 EWAS analysis of exposure, socioeconomic, biochemical, and health outcomes

5.1.3.1 Data clustering

For clustering the various exposure related data, the two different clustering techniques described in Chapter 4 were used. The results are graphicaly illustrated in Figure 2 and Figure 3 respectivelly.

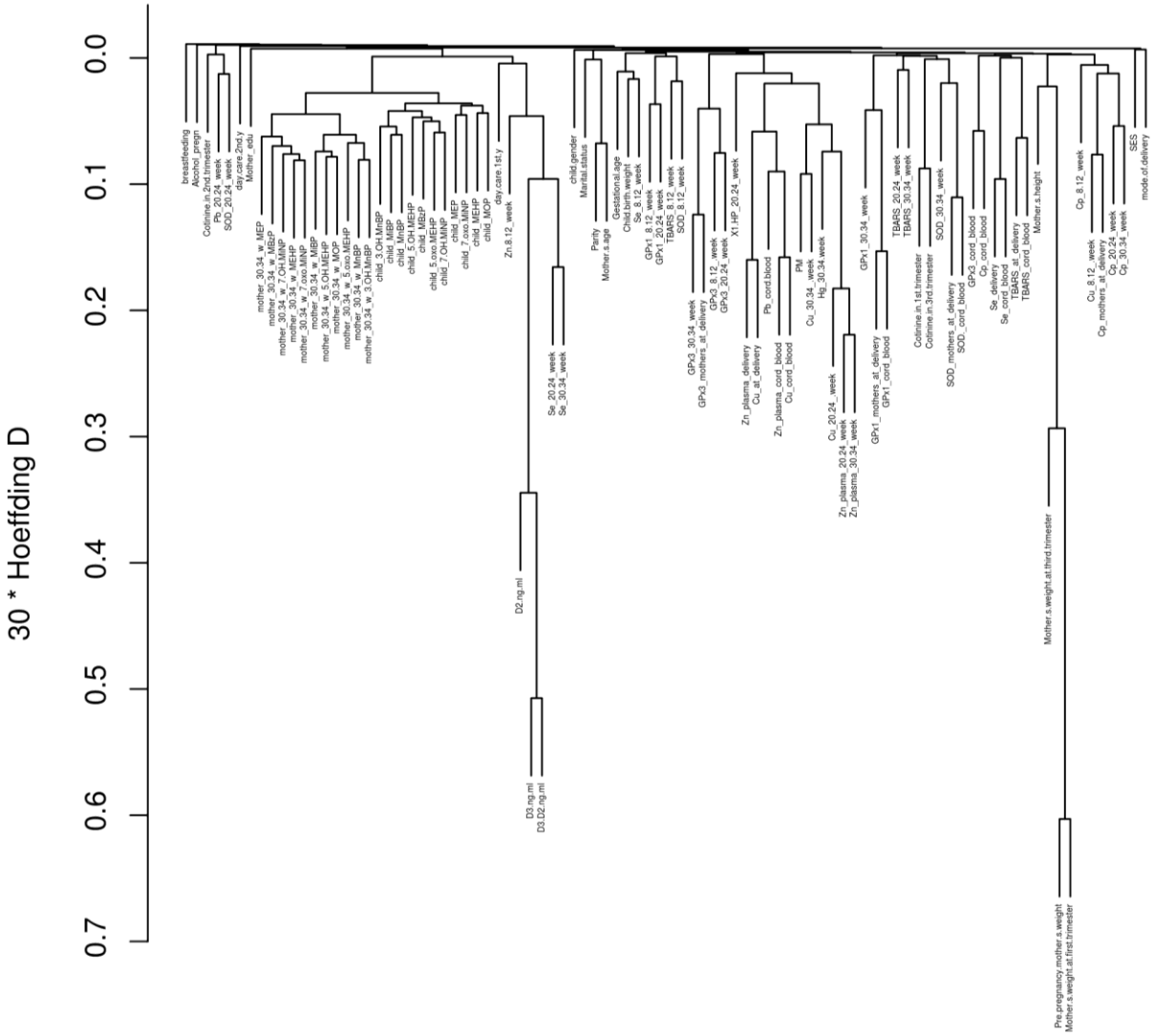

Figure 2. Hierarchical clustering using the Hoeffding D method

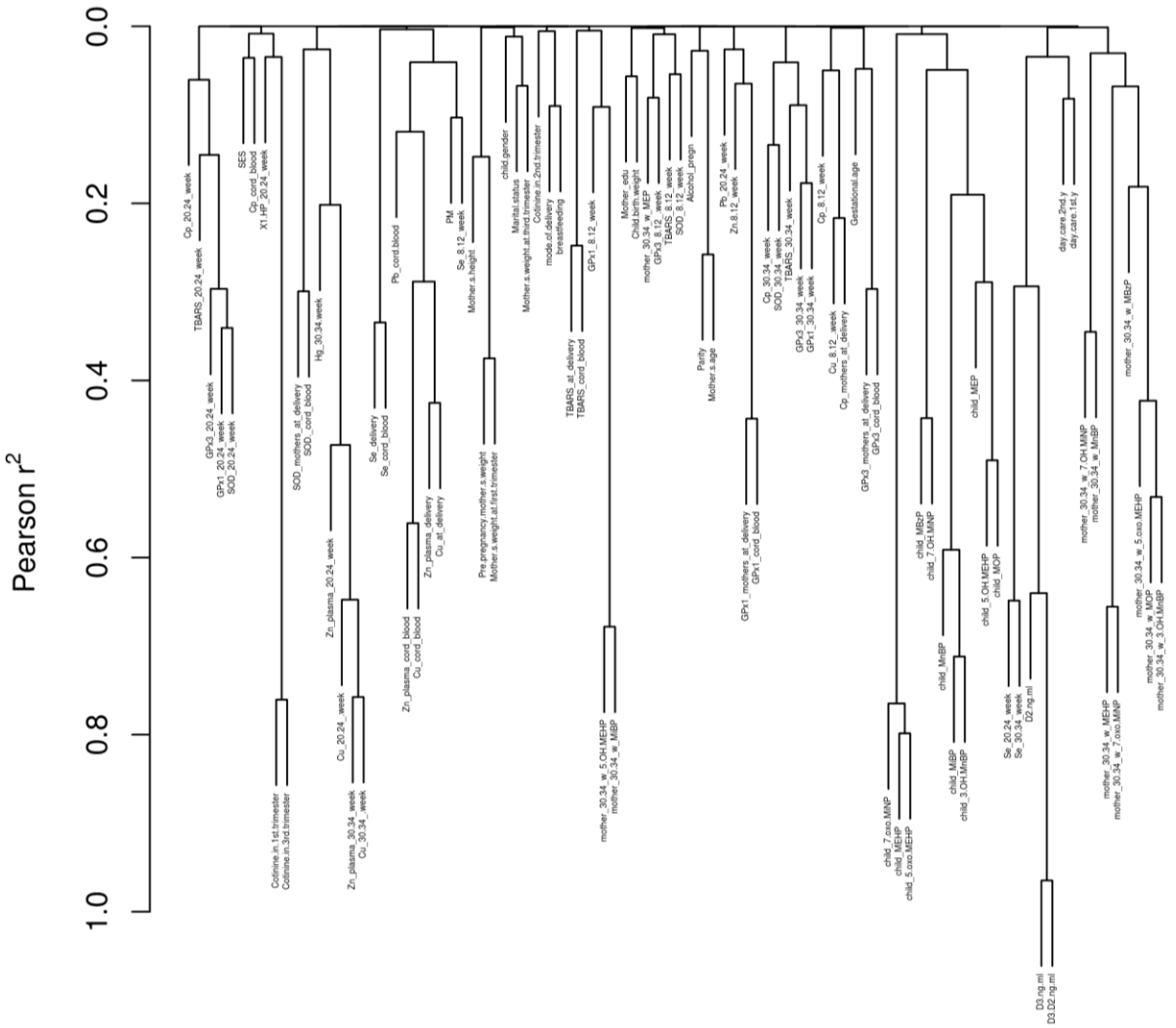

Figure 3. Hierarchical clustering using the Pearson correlation

The auto-correlations of the various parameters, are illustrated in both the heatmap (Figure 4) and the correlation globe (Figure 5).

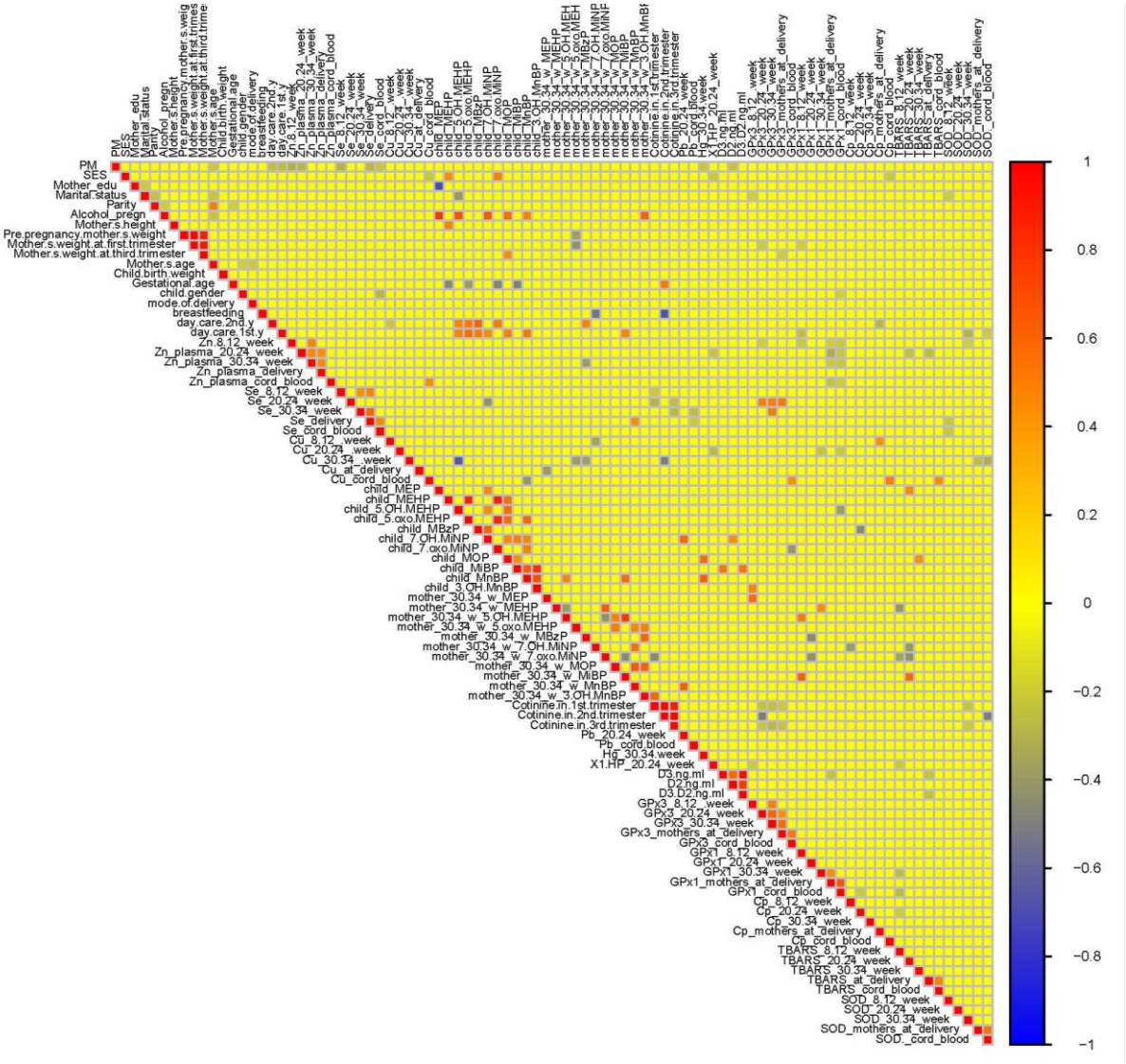

Figure 4. Heatmap of the exposure parameters of the REPRO_PL study

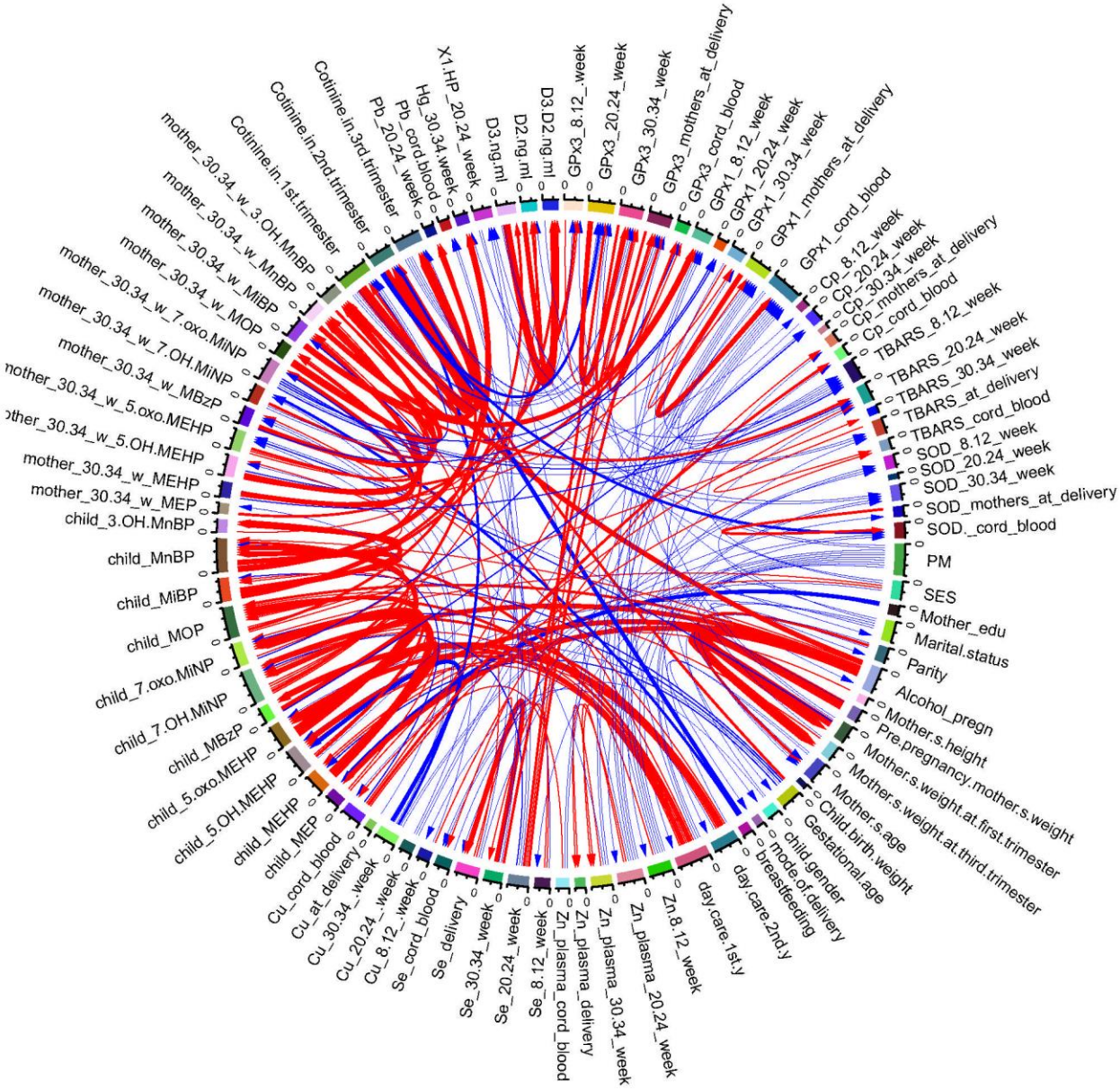

Figure 5. Correlation globe of the environmental and exposure factors of the REPRO_PL study

5.1.3.2 *Results and interpretation of EWAS analysis*

Investigation of the associations between the various neurodevelopment indices evaluated by the Bailey test revealed that different factors might have a negative, or a beneficial contribution.

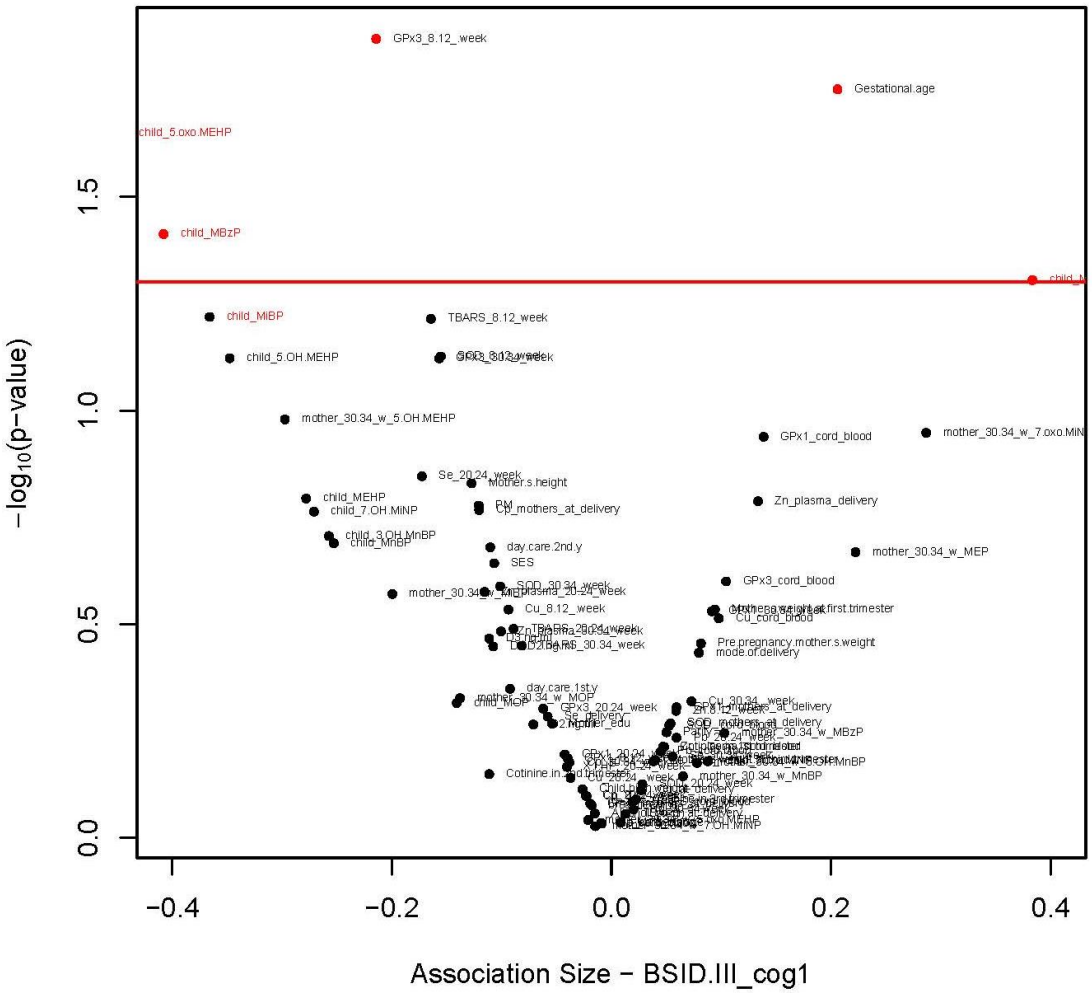

Figure 6. Association of the cognitive development at one year of age with exposure and modifiers

With regard to cognitive development at one year of life, it was found that child exposure to DEHP and BBzP result in lower cognitive scores, while the same occurs for the levels of glutathione peroxidase 3 in the mother plasma during pregnancy. On the other hand, the gestational age of the mother has beneficial effect, however the causality (socio-developmental or physiology) has to be further investigated.

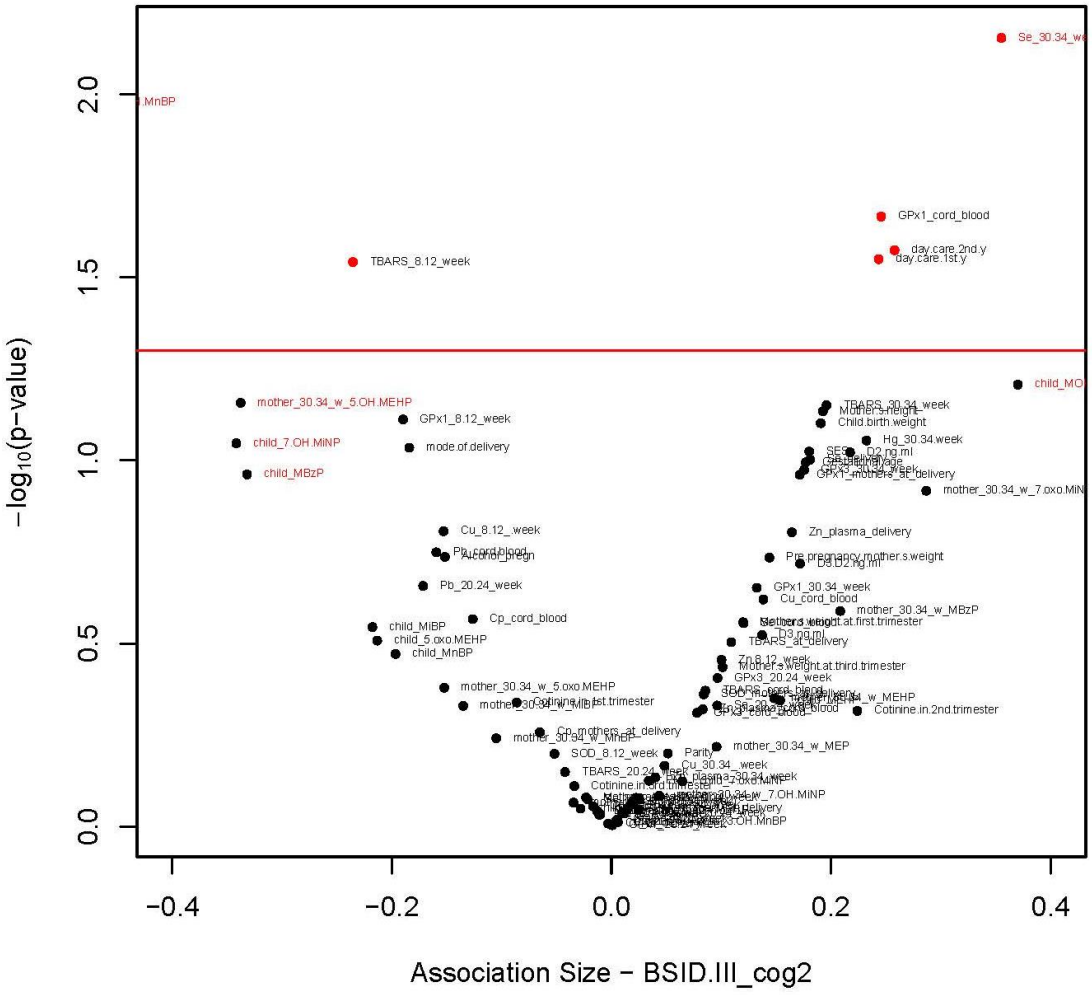

Figure 7. Association of the cognitive development at two years of age with exposure and modifiers

In the second year of age (Figure 7), cognitive development is negatively affected by child exposure to DnBP and the presence of thiobarbituric acid reactive substances in the mother plasma during pregnancy. On the contrary, the presence of selenium during the third semester of pregnancy seem to be a significant beneficial parameter, followed by glutathione peroxidase 1 in the mother plasma during pregnancy. It is also very important to highlight that cognitive development is strongly depends on the attendance of day care center by the child; children hosted in day care centers present higher cognitive skills.

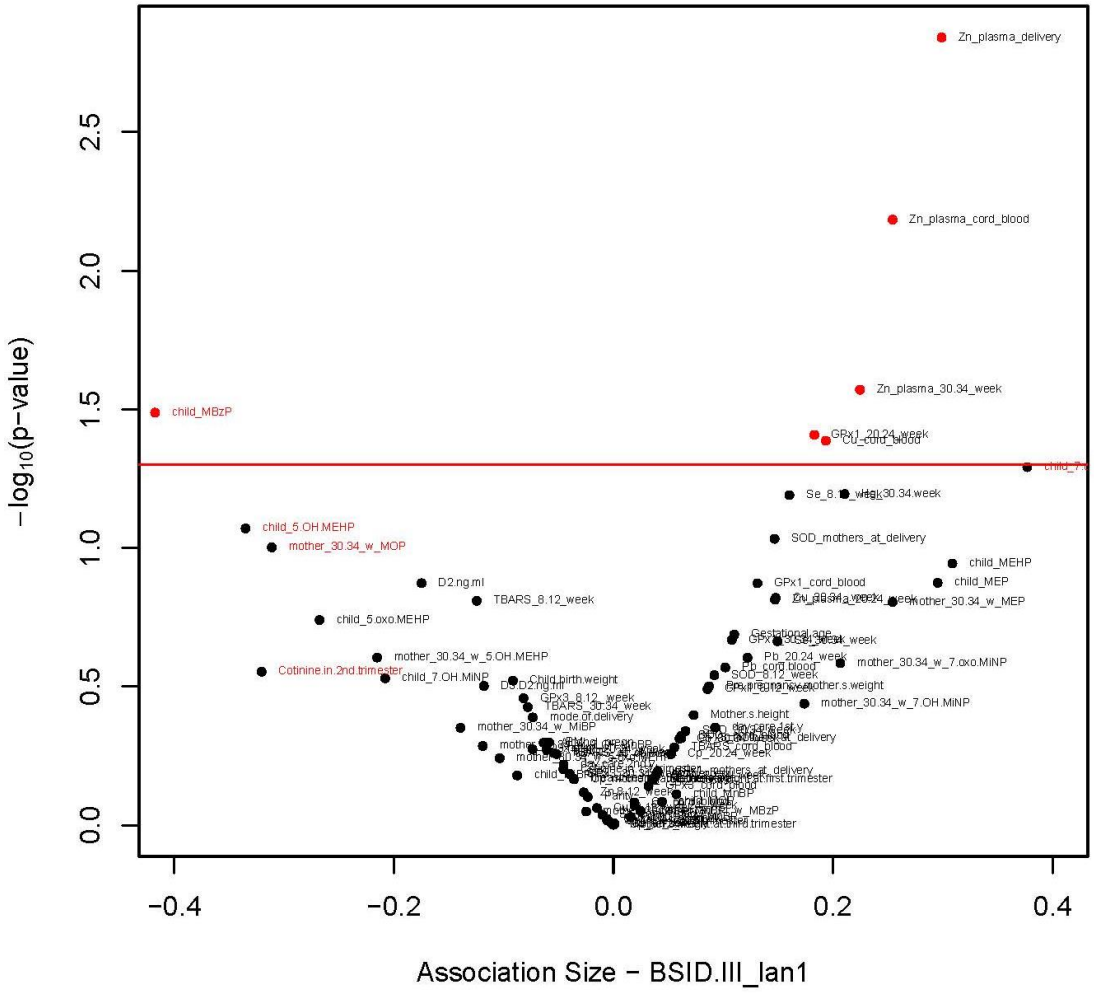

Figure 8. Association of the language development at one year of age with exposure and modifiers

With regard to language development at one year of age (Figure 8), child exposure to phthalates (BBzP) seem to affect negatively. Interestingly, the presence in maternal plasma of Zn (during all the trimesters of pregnancy), glutathione peroxidase 1 (during the second trimester) and Cu act beneficially in the linguistic development.

Language development at two years of age is affected negatively by child exposure to DEP, DnBP and DinP, indicating once more the detrimental effects of children exposure to both low and high molecular phthalates. However, prenatal exposure to thiobarbituric acid reactive substances, glutathione peroxidase 3, glutathione peroxidase 1 and selenium during the last trimester of pregnancy, indicate the importance of this developmental period for the future language development of the child.

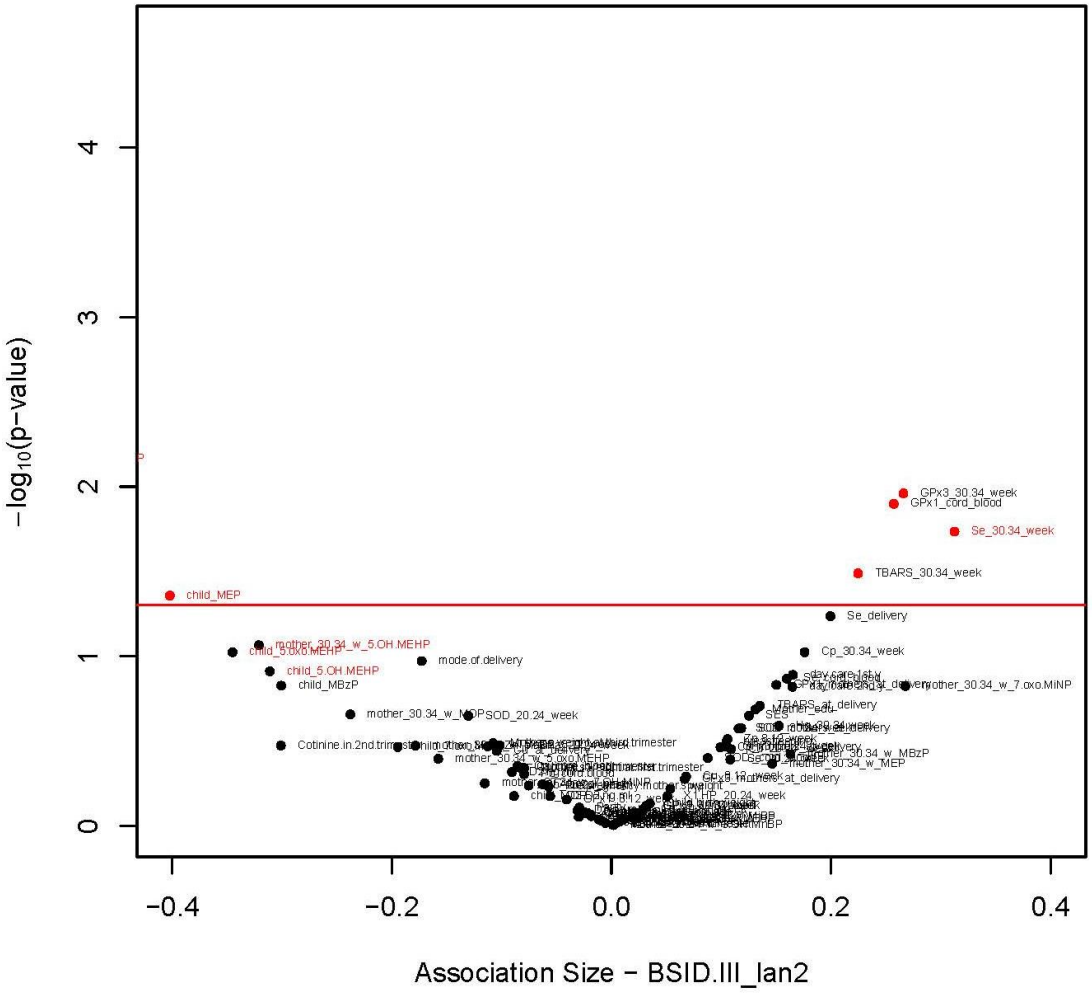

Figure 9. Association of the language development at two years of age with exposure and modifiers

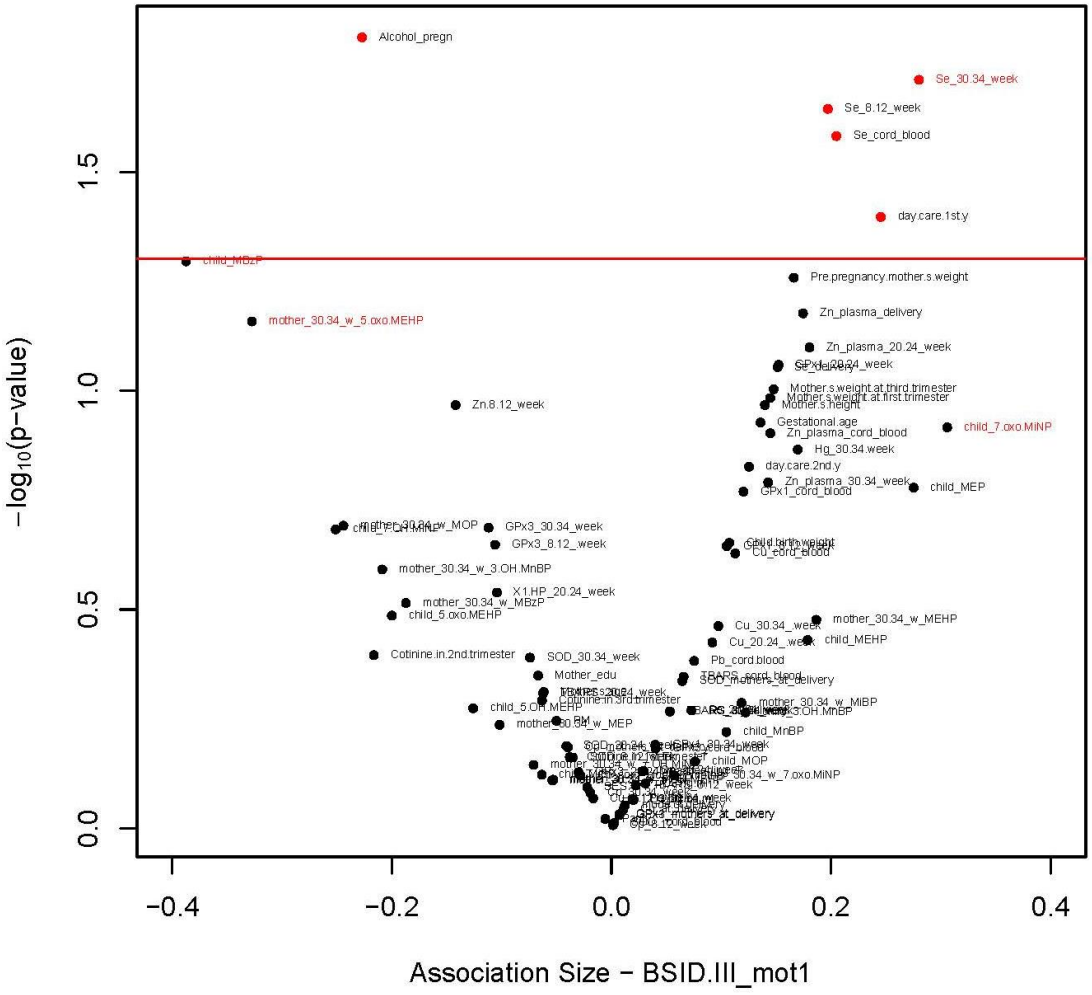

Figure 10. Association of the motor development at one year of age with exposure and modifiers

With regard to motor development, none of the environmental exposure factors seem to hold a key role. At year one (Figure 10), maternal alcohol consumption during pregnancy results in lower scores of motor development, while the presence of selenium (during the 1st and the 3rd trimester), contributes beneficially. In addition, child attendance in a day care center contributes beneficially as well.

Regarding the motor development in the second year (Figure 11), a very interesting finding is that increased maternal bodyweight during the 1st and the 3rd trimester result in lower motor development. On the other hand, the presence of thiobarbituric acid reactive substances and Hg in the mother plasma during the 3rd trimester result in higher motor development scores. It has to be noted that the presence of Hg does not act (in terms of physiology) positively in motor development;

it is rather an indicator of high consumption of food items (e.g. fish), rich in Hg, but also in other nutrients (e.g. omega 3 fatty acids) that are beneficial for child development.

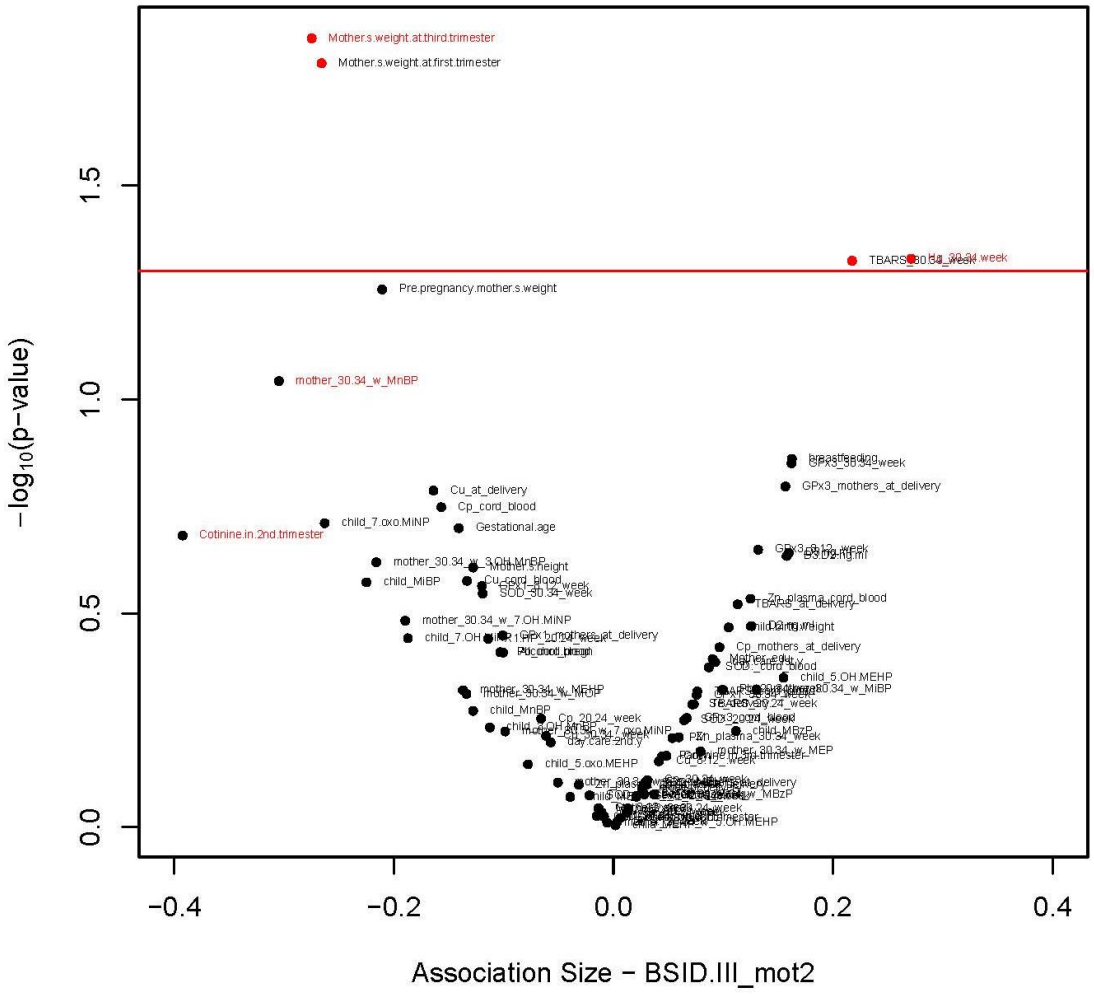

Figure 11. Association of the motor development at two years of age with exposure and modifiers

5.1.4 Toxicity pathway Analysis

Various statistical techniques can be combined in a proper way to identify potential biomarkers and demonstrate how useful the selected biomarkers are. A key issue for metabolomics studies is to avoid over-fitting the data. Because of the large number of metabolites and the relatively small sample size, a complex model can over-utilize (over-fit) the data specific information and show very good performance, but that good result is useless if it cannot be duplicated using a new set of test data. Proper model evaluation and validation is therefore a necessary step to understand the true

performance of a model and the potential biomarkers. So, preprocessing methods, such as filtering, normalization and mean-centering, are crucial to pathway analysis (Bowei et al., 2014).

Mass Profiler Professional, an Agilent's software, provides the necessary tools for preprocessing steps, statistical analysis and pathway analysis.

5.1.4.1 Clustering Analysis

In Hierarchical Clustering Analysis (HCA), the multidimensional data obtained by untargeted metabolomics analysis is reduced to a correlation/distance matrix. Then, the samples are clustered together in a single dendogram and the distances between the samples reflect the similarities and differences between their metabolomic patterns.

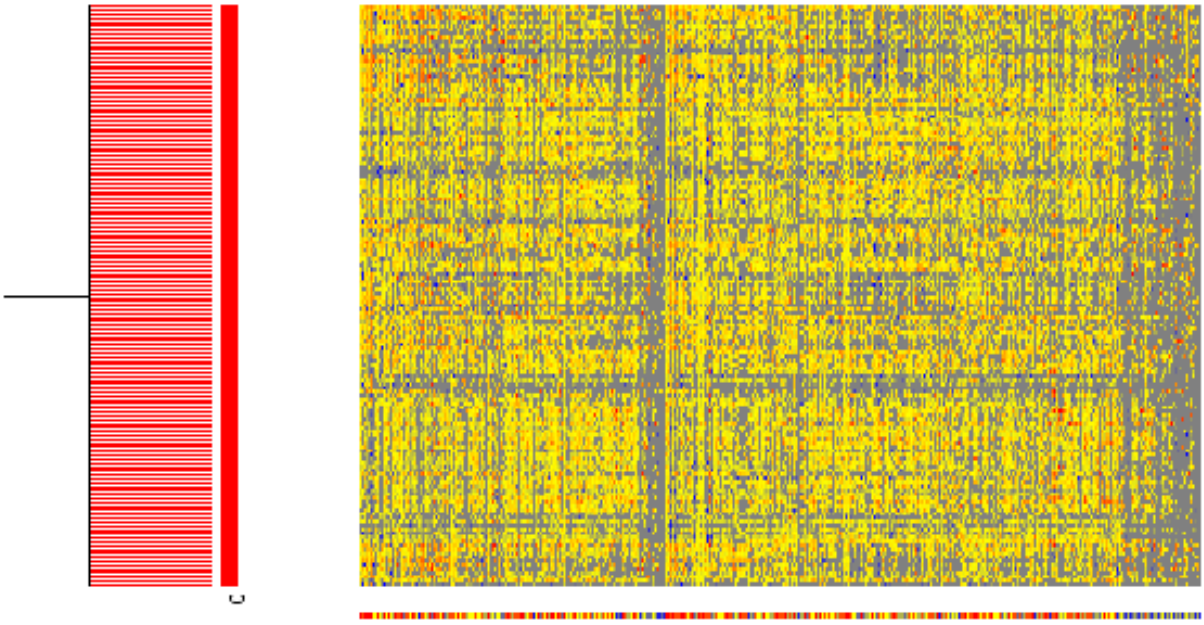

Figure 12. HCA, Similarity Measure: Euclidean, Number of clusters: 3, Maximum number of iterations: 50

Different distance measures (e.g. Euclidean, Pearson, Spearman) and clustering algorithms (e.g. average linkage, single linkage, complete linkage, ward, centroids) can be used in the development of HCA models (Caraveras et al., 2014). For this data set Euclidean will be used as distance metric and Wards as linkage rule. After that step have been created an entity list, which will be the sample data file for pathway analysis.

5.1.4.2 Pathway Analysis

The simple algorithms have present defaults to quickly create a pathway view for further investigation. On the other hand, the advanced algorithms provide the ability to specify several filter criteria for input data and relations.

5.1.4.2.1 Simple Analysis

As input data will be used the identified biomarkers and metabolic pathways will be created with NLP Network discovery:

- Workflow: NLP Network Discover
  - Direct Interpretations
  - Matching statistics (Matched and redundant)
- Analysis Filters: Relation score ≥ 9
- Relation types: member, transport, expression, regulation, binding, promoter binding, metabolism, protein modification.
- Pathways from: WikiPathways, KEGG, BioCyc.

Mass Profiler Professional first tries to find direct matches between the pathway entities (pathways from databases) and the entities in the selected entity list (input data). A direct match can only occur when entities from both pathways and entity lists have identifiers from the same annotation.

Table 2. Identified biomarkers

Isocitrate dehydrogenase (NADP(+))	IDP	Riboflavin
IDH1	Istidina	Rize
IDH2	LNAC	Saccharin
1-Methylhistidine	Mefloquine	Salicylurate
2'-deoxyinosine	Methoxyhydroxyphenylglycol	Sirup
3'-SL	Mimosine	Sorbo
3-Chlorotyrosine	N-benzoylglycine	Sulfalene
3008-B	NARIGENIN	Thiothixene
4'-Hydroxyflurbiprofen	Nevirapine	Tinidazole
4-Pyridoxic acid	Glumin	Tobrex
5-HTP	Guaifenesin	Tryptophan
5-acetamido-6-formamido-3-methyluracil	H-Asp-OH	Uridine
5-oxoproline	HPPA	Usaf hc-1
6-Mercaptopurine	Hgln	Vasc

6-Sulfatoxymelatonin	GENOP	WSQR DQ CVQ
7,9-dihydro-1H-purine-2,6,8(3H)-trione	Glumin	Xanthurenate
734571A	Beflavin	Xylose
AAMU	CATECHOL	Zeranol
APAP	CF3COO(-)	acetaminophen O-glucuronide
Ac-Try	CF3COOH	androsterone 3-glucosiduronic acid
Acetaminophen	CRESOL	cis-Aconitate
Acetylcysteine	Catechin	estrone 3-glucosiduronic acid
Adrin	Cerulenin	ketoglutarate
Afalanine	Clavulanic Acid	meletin
Alna	Clonidine	testosterone 3-glucosiduronic acid
Amifostine	D-glucono-1,5-lactone	Niflumic Acid
Andros-S	DHLA	Noname
Asp-Phe	DOPAC	Nystatin
Azadc	Depas	O-Due
Dexfenfluramine	Dicid	Edetic Acid
Dihydroxymandelic acid	Ecgonine methyl ester	Enalaprilat
Epinephrine	Ethacrynic acid	Flunarizine

The **Pathway entities of Experiment Type** columns contain the number of entities in the pathway that are of the experiment type: the number of genes and proteins in the pathway if the experiment type is Genomics/Proteomics/Transcriptomics and the number of metabolites if the experiment type is Metabolomics, regardless of whether the entities are present in the technology or not.

Table 3. Identified pathways

Pathway list	Pathway Experiment Entities
methylthiopropionate biosynthesis	1
acyl carrier protein metabolism	1
L-dopa degradation	1
ceramide degradation	1
thiosulfate disproportionation III (rhodanese)	1

Pathway list	**Pathway Experiment Entities**
lanosterol biosynthesis	1
melatonin degradation III	1
glutamate removal from folates	1
2-amino-3-carboxymuconate semialdehyde degradation to glutaryl-CoA	1
acetyl-CoA biosynthesis III (from citrate)	1
glutathione biosynthesis	1
citrulline degradation	1
threonine degradation II	1
L-dopachrome biosynthesis	1
alanine biosynthesis III	1
sulfite oxidation IV	1
D-mannose degradation	1
1,25-dihydroxyvitamin D3 biosynthesis	1
asparagine biosynthesis I	1
L-cysteine degradation II	1
adenine and adenosine salvage VI	1
L-cysteine degradation III	1
oxidized GTP and dGTP detoxification	1
UDP-N-acetyl-D-galactosamine biosynthesis I	1
adenine and adenosine salvage I	1
4-hydroxybenzoate biosynthesis	1
geranylgeranyldiphosphate biosynthesis	1
thio-molybdenum cofactor biosynthesis	1
MAP kinase cascade	1
4-hydroxyproline degradation I	1
4-hydroxyphenylpyruvate biosynthesis	1
histamine biosynthesis	1
Alpha-tocopherol degradation	1
sorbitol degradation I	1
methylglyoxal degradation VI	1
phosphatidylethanolamine biosynthesis III	1
(S)-reticuline biosynthesis II	1
thiamin salvage III	1
diphthamide biosynthesis	1
Hs_Colchicine_Metabolic_Pathway_WP2536_74479	1

Pathway list	Pathway Experiment Entities
tyrosine biosynthesis IV	2
guanosine nucleotides degradation III	2
L-serine degradation	2
glycerol-3-phosphate shuttle	2
glutamate dependent acid resistance	2
lactose degradation III	2
L-glutamine biosynthesis II (tRNA-dependent)	2
purine ribonucleosides degradation to ribose-1-phosphate	2
alanine biosynthesis II	2
D-glucuronate degradation I	2
glutamate degradation X	2
taurine biosynthesis	2
glycine biosynthesis I	2
formaldehyde oxidation II (glutathione-dependent)	2
flavin biosynthesis IV (mammalian)	2
Beta-alanine degradation I	2
epoxysqualene biosynthesis	2
NADH repair	2
glycine degradation (creatine biosynthesis)	2
D-myo-inositol (3,4,5,6)-tetrakisphosphate biosynthesis	2
alanine degradation III	2
L-cysteine degradation I	2
glycine biosynthesis III	2
pyridoxal 5'-phosphate salvage pathway	2
4-aminobutyrate degradation I	2
guanine and guanosine salvage I	2
cardiolipin biosynthesis II	2
anandamide degradation	2
fatty acid Alpha-oxidation III	2
uridine-5'-phosphate biosynthesis	2
asparagine degradation I	2
thyroid hormone biosynthesis	2
UDP-D-xylose and UDP-D-glucuronate biosynthesis	2
glutamine biosynthesis I	2
putrescine biosynthesis III	2

Pathway list	Pathway Experiment Entities
spermine biosynthesis	2
cysteine biosynthesis/homocysteine degradation	2
GDP-L-fucose biosynthesis I (from GDP-D-mannose)	2
spermidine biosynthesis I	2
GDP-L-fucose biosynthesis II (from L-fucose)	2
glutamate biosynthesis II	2
sulfate activation for sulfonation	2
Hs_Butyrate-induced_histone_acetylation_WP2366_88590	2
Hs_Evolocumab_Mechanism_WP3408_88049	2
dolichol and dolichyl phosphate biosynthesis	3
inosine-5'-phosphate biosynthesis II	3
methylglyoxal degradation I	3
fatty acid Beta-oxidation III (unsaturated, odd number)	3
tetrahydrobiopterin biosynthesis I	3
sucrose degradation V (mammalian)	3
acyl-CoA hydrolysis	3
N-acetylglucosamine degradation I	3
PRPP biosynthesis I	3
molybdenum cofactor biosynthesis	3
histamine degradation	3
mitochondrial L-carnitine shuttle pathway	3
lysine degradation II	3
5-aminoimidazole ribonucleotide biosynthesis I	3
glutamine degradation I	3
methylmalonyl pathway	3
coenzyme A biosynthesis	3
fatty acid biosynthesis initiation II	3
D-myo-inositol (1,4,5,6)-tetrakisphosphate biosynthesis	3
glutamate degradation II	3
L-carnitine biosynthesis	3
tetrahydrobiopterin biosynthesis II	3
1D-myo-inositol hexakisphosphate biosynthesis V (from Ins(1,3,4)P3)	3
aspartate biosynthesis	3
S-adenosyl-L-methionine biosynthesis	3
serine biosynthesis	3

Pathway list	**Pathway Experiment Entities**
phenylethylamine degradation I	3
methylglyoxal degradation III	3
acetate conversion to acetyl-CoA	3
tetrahydrofolate salvage from 5,10-methenyltetrahydrofolate	4
NAD phosphorylation and dephosphorylation	4
proline biosynthesis I	4
leukotriene biosynthesis	4
dermatan sulfate degradation (metazoa)	4
galactose degradation I (Leloir pathway)	4
heme biosynthesis from uroporphyrinogen-III I	4
trehalose degradation II (trehalase)	4
pentose phosphate pathway (oxidative branch)	4
2-oxobutanoate degradation I	4
dermatan sulfate biosynthesis (late stages)	4
ketogenesis	4
arginine degradation I (arginase pathway)	4
ethanol degradation II	4
myo-inositol biosynthesis	4
dTMP de novo biosynthesis	4
catecholamine biosynthesis	4
adenine and adenosine salvage III	4
trans, trans-farnesyl diphosphate biosynthesis	4
S-methyl-5-thio-Alpha-D-ribose 1-phosphate degradation	4
NAD biosynthesis III	4
glutamate degradation III (via 4-aminobutyrate)	4
eumelanin biosynthesis	4
N-acetylglucosamine degradation II	4
glycerol degradation I	4
CMP-N-acetylneuraminate biosynthesis I (eukaryotes)	4
Hs_Caffeine_and_Theobromine_metabolism_WP3633_85154	4
Hs_Felbamate_Metabolism_WP2816_88575	4
Hs_Amino_acid_conjugation_of_benzoic_acid_WP521_88588	4
Hs_eIF5A_regulation_in_response_to_inhibition_of_the_nuclear_export_system_WP3302_87966	4
pyruvate fermentation to lactate	5
superpathway of serine and glycine biosynthesis I	5

Pathway list	Pathway Experiment Entities
folate polyglutamylation	5
tryptophan degradation X (mammalian, via tryptamine)	5
tryptophan degradation to 2-amino-3-carboxymuconate semialdehyde	5
creatine-phosphate biosynthesis	5
glutaryl-CoA degradation	5
proline biosynthesis II (from arginine)	5
calcium transport I	5
NAD biosynthesis from 2-amino-3-carboxymuconate semialdehyde	5
tetrapyrrole biosynthesis II	5
NAD salvage pathway III	5
glucose and glucose-1-phosphate degradation	5
ketolysis	5
UDP-N-acetyl-D-glucosamine biosynthesis II	5
aspartate degradation II	5
Rapoport-Luebering glycolytic shunt	5
salvage pathways of pyrimidine deoxyribonucleotides	5
BMP Signalling Pathway	5
tRNA splicing	5
superoxide radicals degradation	5
citrulline-nitric oxide cycle	5
protein citrullination	5
serotonin and melatonin biosynthesis	5
4-hydroxy-2-nonenal detoxification	5
urate biosynthesis/inosine 5'-phosphate degradation	5
ceramide biosynthesis	5
GDP-glucose biosynthesis	5
Hs_Catalytic_cycle_of_mammalian_Flavin-containing_MonoOxygenases_(FMOs)_WP688_84769	5
Hs_Diclofenac_Metabolic_Pathway_WP2491_73891	5
androgen biosynthesis	6
zymosterol biosynthesis	6
selenocysteine biosynthesis II (archaea and eukaryotes)	6
spermine and spermidine degradation I	6
arginine degradation VI (arginase 2 pathway)	6
GDP-mannose biosynthesis	6
pentose phosphate pathway (non-oxidative branch)	6

Pathway list	Pathway Experiment Entities
oxidative ethanol degradation III	6
ethanol degradation IV	6
arginine biosynthesis IV	6
chondroitin sulfate degradation (metazoa)	6
tyrosine degradation I	6
phosphatidylcholine biosynthesis I	6
phenylalanine degradation I (aerobic)	6
phosphatidylethanolamine biosynthesis II	6
urea cycle	6
Fatty acid biosynthesis	6
dopamine degradation	7
adenosine nucleotides degradation II	7
inositol pyrophosphates biosynthesis	7
histidine degradation III	7
glycogen degradation II	7
Hs_Acetylcholine_Synthesis_WP528_79855	7
Hs_DDX1_as_a_regulatory_component_of_the_Drosha_microprocessor_WP2942_87380	7
Hs_Arylamine_metabolism_WP694_88582	7
Hs_Arachidonate_Epoxygenase_-_Epoxide_Hydrolase_WP678_71506	7
thymine degradation	8
fatty acid activation	8
sphingomyelin metabolism	8
oleate biosynthesis II (animals)	8
serotonin degradation	8
dolichyl-diphosphooligosaccharide biosynthesis	8
uracil degradation II (reductive)	8
Hs_Effects_of_Nitric_Oxide_WP1995_88595	8
Hs_Folate-Alcohol_and_Cancer_Pathway_WP1589_87508	8
Hs_FTO_Obesity_Variant_Mechanism_WP3407_87514	8
Hs_Aflatoxin_B1_metabolism_WP699_70509	8
Hs_Dual_hijack_model_of_Vif_in_HIV_infection_WP3300_87896	8
fatty acid Alpha-oxidation II	9
noradrenaline and adrenaline degradation	9
choline biosynthesis III	9
purine nucleotides de novo biosynthesis II	9

Pathway list	Pathway Experiment Entities
UDP-N-acetyl-D-galactosamine biosynthesis II	9
glycogen degradation III	9
heme biosynthesis II	9
putrescine degradation III	9
glutathione redox reactions I	9
Hs_Benzo(a)pyrene_metabolism_WP696_85085	9
Hs_Codeine_and_Morphine_Metabolism_WP1604_74317	9
sphingosine and sphingosine-1-phosphate metabolism	10
NAD biosynthesis II (from tryptophan)	10
pentose phosphate pathway	10
purine nucleotides degradation II (aerobic)	10
Hs_Gastric_acid_production_WP2596_87520	10
Hs_Composition_of_Lipid_Particles_WP3601_87388	10
tryptophan degradation III (eukaryotic)	11
Gamma-glutamyl cycle	11
mevalonate pathway I	12
pyrimidine ribonucleotides interconversion	12
Hs_Bone_Morphogenic_Protein_(BMP)_Signalling_and_Regulation_WP1425_86082	12
D-myo-inositol (1,4,5)-trisphosphate degradation	13
NAD salvage pathway II	13
cholesterol biosynthesis II (via 24,25-dihydrolanosterol)	13
stearate biosynthesis I (animals)	13
colanic acid building blocks biosynthesis	13
cholesterol biosynthesis III (via desmosterol)	13
cholesterol biosynthesis I	13
Hs_Estrogen_Receptor_Pathway_WP2881_88032	13
Hs_EV_release_from_cardiac_cells_and_their_functional_effects_WP3297_88048	13
Hs_Dopamine_metabolism_WP2436_71387	13
pyrimidine ribonucleotides de novo biosynthesis	14
triacylglycerol degradation	14
superpathway of geranylgeranyldiphosphate biosynthesis I (via mevalonate)	14
salvage pathways of pyrimidine ribonucleotides	15
D-myo-inositol (1,3,4)-trisphosphate biosynthesis	15
1D-myo-inositol hexakisphosphate biosynthesis II (mammalian)	15
Hs_Fatty_Acid_Omega_Oxidation_WP206_85418	15

Pathway list	Pathway Experiment Entities
Hs_Biogenic_Amine_Synthesis_WP550_85678	15
Hs_Cholesterol_Biosynthesis_WP197_81059	15
fatty acid Beta-oxidation I	16
Hs_Deregulation_of_Rab_and_Rab_Effector_Genes_in_Bladder_Cancer_WP2291_85341	16
Hs_Drug_Induction_of_Bile_Acid_Pathway_WP2289_88593	17
Hs_ACE_Inhibitor_Pathway_WP554_84372	17
Hs_Ganglio_Sphingolipid_Metabolism_WP1423_79812	18
Hs_Estrogen_metabolism_WP697_86903	18
CDP-diacylglycerol biosynthesis I	19
D-myo-inositol-5-phosphate metabolism	19
Hs_Farnesoid_X_Receptor_Pathway_WP2879_88056	19
superpathway of D-myo-inositol (1,4,5)-trisphosphate metabolism	20
glycolysis I	20
phosphatidylglycerol biosynthesis II (non-plastidic)	20
Hs_Apoptosis_Modulation_by_HSP70_WP384_67054	20
gluconeogenesis I	21
3-phosphoinositide degradation	21
Hs_Complement_Activation_WP545_87260	22
Hs_Blood_Clotting_Cascade_WP272_84876	22
Hs_Fatty_Acid_Biosynthesis_WP357_85181	22
Hs_Cori_Cycle_WP1946_79691	23
Hs_Estrogen_signaling_pathway_WP712_78491	23
Hs_EBV_LMP1_signaling_WP262_86888	24
Hs_Angiogenesis_WP1539_79949	24
D-myo-inositol (1,4,5)-trisphosphate biosynthesis	25
Hs_Eicosanoid_Synthesis_WP167_82702	25
Hs_Differentiation_of_white_and_brown_adipocyte_WP2895_87889	25
3-phosphoinositide biosynthesis	26
glutathione-mediated detoxification I	26
Hs_EPO_Receptor_Signaling_WP581_86897	26
triacylglycerol biosynthesis	27
superpathway of cholesterol biosynthesis	27
Hs_Follicle_Stimulating_Hormone_(FSH)_signaling_pathway_WP2035_86080	27
Hs_Gastric_Cancer_Network_1_WP2361_86831	29
Hs_Cytokines_and_Inflammatory_Response_WP530_79331	30

Pathway list	Pathway Experiment Entities
Hs_Dopaminergic_Neurogenesis_WP2855_87239	30
Hs_Extracellular_vesicle-mediated_signaling_in_recipient_cells_WP2870_88052	30
Hs_Bladder_Cancer_WP2828_87377	31
Hs_Constitutive_Androstane_Receptor_Pathway_WP2875_79890	32
Base excision repair	33
Hs_Endothelin_Pathways_WP2197_74852	33
Hs_Alpha_6_Beta_4_signaling_pathway_WP244_85199	33
Hs_Factors_and_pathways_affecting_insulin-like_growth_factor_(IGF1)-Akt_signaling_WP3850_88165	33
Hs_Fatty_Acid_Beta_Oxidation_WP143_79783	34
Hs_Fluoropyrimidine_Activity_WP1601_84700	34
DNA replication	36
tRNA charging	38
phospholipases	38
Hs_Amyotrophic_lateral_sclerosis_(ALS)_WP2447_85186	38
Hs_G13_Signaling_Pathway_WP524_72112	38
Hs_Alanine_and_aspartate_metabolism_WP106_74147	40
Hs_ATM_Signaling_Pathway_WP2516_84490	41
Hs_Eukaryotic_Transcription_Initiation_WP405_73594	41
Hs_DNA_Replication_WP466_79981	42
Porphyrin and chlorophyll metabolism	43
Hs_Fas_Ligand_(FasL)_pathway_and_Stress_induction_of_Heat_Shock_Proteins_(HSP)_regulation_WP314_87381	43
Hs_Aryl_Hydrocarbon_Receptor_Pathway_WP2873_79696	46
Amino sugar and nucleotide sugar metabolism	47
Hs_Aryl_Hydrocarbon_Receptor_WP2586_85335	48
Hs_Diurnally_Regulated_Genes_with_Circadian_Orthologs_WP410_84683	48
N-Glycan biosynthesis	49
Hs_Differentiation_Pathway_WP2848_87893	50
Hs_Apoptosis-related_network_due_to_altered_Notch3_in_ovarian_cancer_WP2864_79278	53
Hs_Cardiac_Progenitor_Differentiation_WP2406_73324	53
Hs_Cardiac_Hypertrophic_Response_WP2795_85087	54
Hs_Cell_Differentiation_-_Index_WP2029_82903	54
Hs_Copper_homeostasis_WP3286_87382	55
Hs_ErbB_Signaling_Pathway_WP673_86901	55
Hs_Complement_and_Coagulation_Cascades_WP558_85997	61
Hs_Endochondral_Ossification_WP474_87977	64

Pathway list	Pathway Experiment Entities
Hs_AGE-RAGE_pathway_WP2324_78487	66
Glycolysis / Gluconeogenesis	66
superpathway of inositol phosphate compounds	67
Hs_Folate_Metabolism_WP176_85063	67
Hs_DNA_Damage_Response_WP707_82937	68
Hs_AMP-activated_Protein_Kinase_(AMPK)_Signaling_WP1403_86073	68
Hs_G1_to_S_cell_cycle_control_WP45_80001	68
Hs_Arrhythmogenic_Right_Ventricular_Cardiomyopathy_WP2118_71265	78
Hs_Apoptosis_WP254_88580	87
Hs_Cytoplasmic_Ribosomal_Proteins_WP477_67139	88
Hs_Androgen_receptor_signaling_pathway_WP138_79958	89
Hs_G_Protein_Signaling_Pathways_WP35_88601	92
Hs_Corticotropin-releasing_hormone_signaling_pathway_WP2355_79973	92
Hs_Apoptosis_Modulation_and_Signaling_WP1772_87370	95
Hs_B_Cell_Receptor_Signaling_Pathway_WP23_79985	98
Hs_Allograft_Rejection_WP2328_84648	100
Hs_Cell_Cycle_WP179_70629	103
Hs_Electron_Transport_Chain_WP111_86860	104
Hs_DNA_Damage_Response_(only_ATM_dependent)_WP710_79974	114
Hs_Adipogenesis_WP236_80209	131
Oxidative phosphorylation	133
Hs_Brain-Derived_Neurotrophic_Factor_(BDNF)_signaling_pathway_WP2380_87427	144
Hs_Ectoderm_Differentiation_WP2858_87931	145
Hs_Endoderm_Differentiation_WP2853_88152	146
Hs_Alzheimers_Disease_WP2059_87372	149
Hs_Calcium_Regulation_in_the_Cardiac_Cell_WP536_80211	150
Hs_EGF-EGFR_Signaling_Pathway_WP437_79266	162
Hs_Focal_Adhesion_WP306_80308	191
Hs_Circadian_rythm_related_genes_WP3594_87161	210
MAPK signaling pathway	257
Pathways in cancer	327

5.1.4.2.2 Advanced Analysis

Single Experiment Analysis (SEA) as the name implies, the SEA only identifies matching pathways for one experiment. In addition, the curated pathways options (for example WikiPathways, BioCyC pathways, KEGG, or BioPAX pathways) as well as NLP created pathways can also be individually selected as sources for pathway analysis.

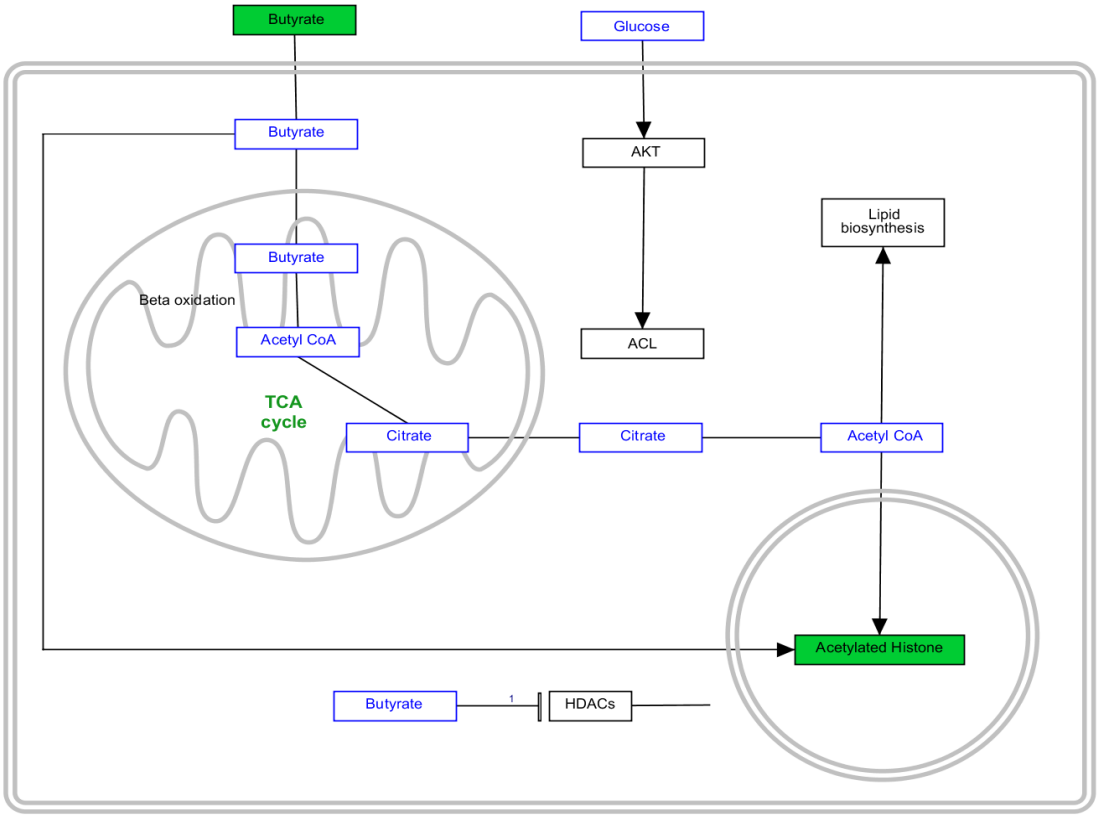

Figure 13. Pathway from single experiment analysis: Butyrate-induced histone acetylation

At this point we have to mention that no p-values are computed for entities from Metabolomics experiments during an SEA workflow to avoid a misrepresentation of the significance of matching pathways caused by the fact that Mass Profiler Professional uses the technology (All Entities list) as a reference for p-value computation. The technology of a metabolomics experiment is limited to only the measured metabolites with an observable abundance in the experiment. Pathways on the other hand are likely to contain many other metabolites that may not be present in the technology. This results in a pathway p-value computed with the technology as reference to be higher than a more realistic p-value computed with a comprehensive reference set of global entities.

5.2 Application on the HERACLES Greek cohort

5.2.1 Introduction

The HERACLES Greek cohort, is a cohort study aiming at assessing the contribution of heavy metals environmental contamination associated to children neurodevelopment. The study has been established in 2012. Around 350 children aged 3 to 8 living in the proximity between 0.5 to 12 km were enroled.

5.2.2 Data used in the analysis

5.2.2.1 Exposure factors

For the association, several exposure factors have been investigated, including:

- Exposure to heavy metals, including:
  - Cd, Hg and As in urine
  - Pb in blood
  - Mn and Hg in hair
- Additional proxies of exposure, such as
  - Distance from the contaminated sites
  - Concentration of heavy metals in the soil of the child address

5.2.2.2 Exposure and effect modifiers

Additional factors considered as exposure and effects modifiers were included as well. These included:

- Sociodemographic parameters such as
  - Socioeconomic status
  - Mother education
  - Father education
  - Stress events
- Child anthropometric parameters and post-delivery factors
  - Child body mass index
  - Child gender
  - Breastfeeding
- Presence of micronutrients, minerals and vitamins
  - Se in the mother plasma during pregnancy, delivery and in cord blood
- Detailed dietary habits
  - Consumption of meat products (pork meat, beef, lamb, sausages)
  - Consumption of fish
  - Consumption of sea food

- Consumption of poultry (eggs, chicken)
- Consumption of dairy products (milk, yogurt)
- Consumption of nuts
- Consumption of fruits
- Consumption of vegetables
- Consumption of snacks (biscuits, chocolates)

5.2.2.3 Health outcomes investigated

The health outcomes considered in this study are relevant to the neurodevelopmental disorders in children estimated following the administration to the children or their parents and teachers the following four test batteries.

The **Child Behavior Checklist** (Achenbach and Rescorla, 2001), also called the Achenbach System of Empirically Based Assessment, is a report form to screen for emotional, behavioral, and social problems. The CBCL's questions are associated with problems on a syndrome scale in eight different categories: anxious/depressed, withdrawn/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behavior, and aggressive behavior. The CBCL also has a scale set to show scores associated with disorders from the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association., 2000): anxiety, oppositional defiant disorder, conduct problems, somatic problems, affective problems, and attention deficit disorder. Many studies have demonstrated a high rate of reliability between the scales of the CBCL and actual psychological diagnosis (Warnick et al., 2008).

The **Cambridge Neuropsychological Test Automated Battery** (CANTAB); it has been used to assess neurocognitive performance in modeling studies of Chronic Fatigue Syndrome (CFS) (Capuron et al., 2001; Robbins and Sahakian, 2002). CANTAB has modules for several neurocognitive functions and processes including psychomotor and motor speed, reasoning and planning abilities, memory and attention, and frontal, temporal and hippocampal dysfunctions. Thus, it allows assessment of neuro-cognitive dysfunctions associated with neurologic disorders, phannacologic manipulations, and neuro-cognitive syndromes.

The **Social Responsiveness Scale** (SRS); it is often used to measure Autism Spectrum Disorders (ASD) severity. The Social Responsiveness Scale (SRS) is a parent and teacher-completed screening questionnaire measures social ability of children from 4 years to 18 years old. It is used primarily to measure Autism Spectrum Disorders (ASD) severity. Although SRS is frequently referred to as a measure of “social impairment,” many SRS items describe other core features of ASD, including communication deficits and repetitive behaviors (Constantino et al., 2000), as well as symptoms not exclusively related to ASD diagnostic criteria (Grzadzinski et al., 2011).

The **Wechsler Intelligence Scale for children – Fourth Edition** (Wechsler, 2003); it is an individually administered measure of intelligence intended for children aged six years to 16 years and 11 months. WISC-IV yields measures of general intelligence as reflected in both verbal and nonverbal

(performance) abilities and specific indices including verbal comprehension, perceptual reasoning, working memory and processing speed.

5.2.3 EWAS analysis

5.2.3.1 Data clustering

For clustering the various exposure related data, the two different clustering techniques described in Chapter 4 were used. The results are graphically illustrated in Figure 14 and Figure 15 respectively.

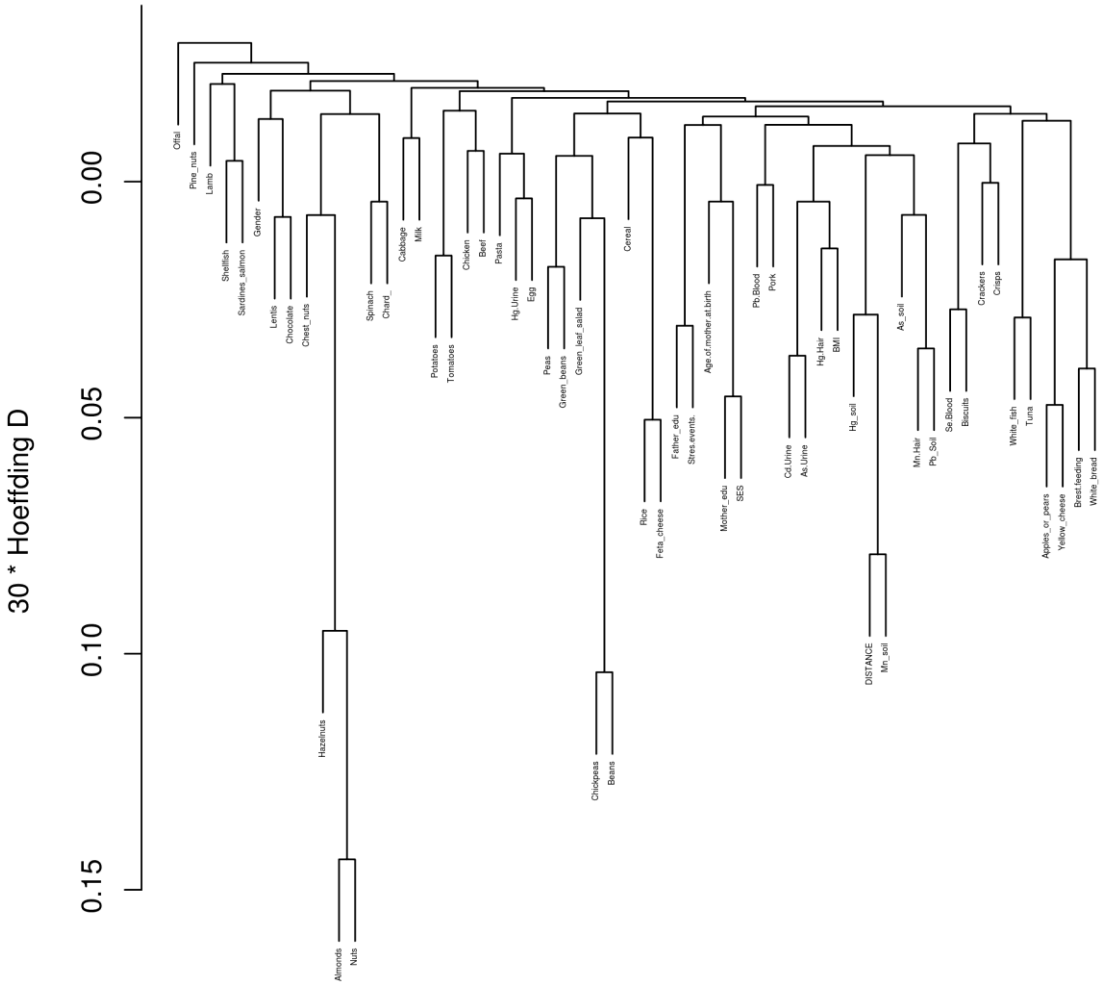

Figure 14. Hierarchical clustering using the Hoeffding D method

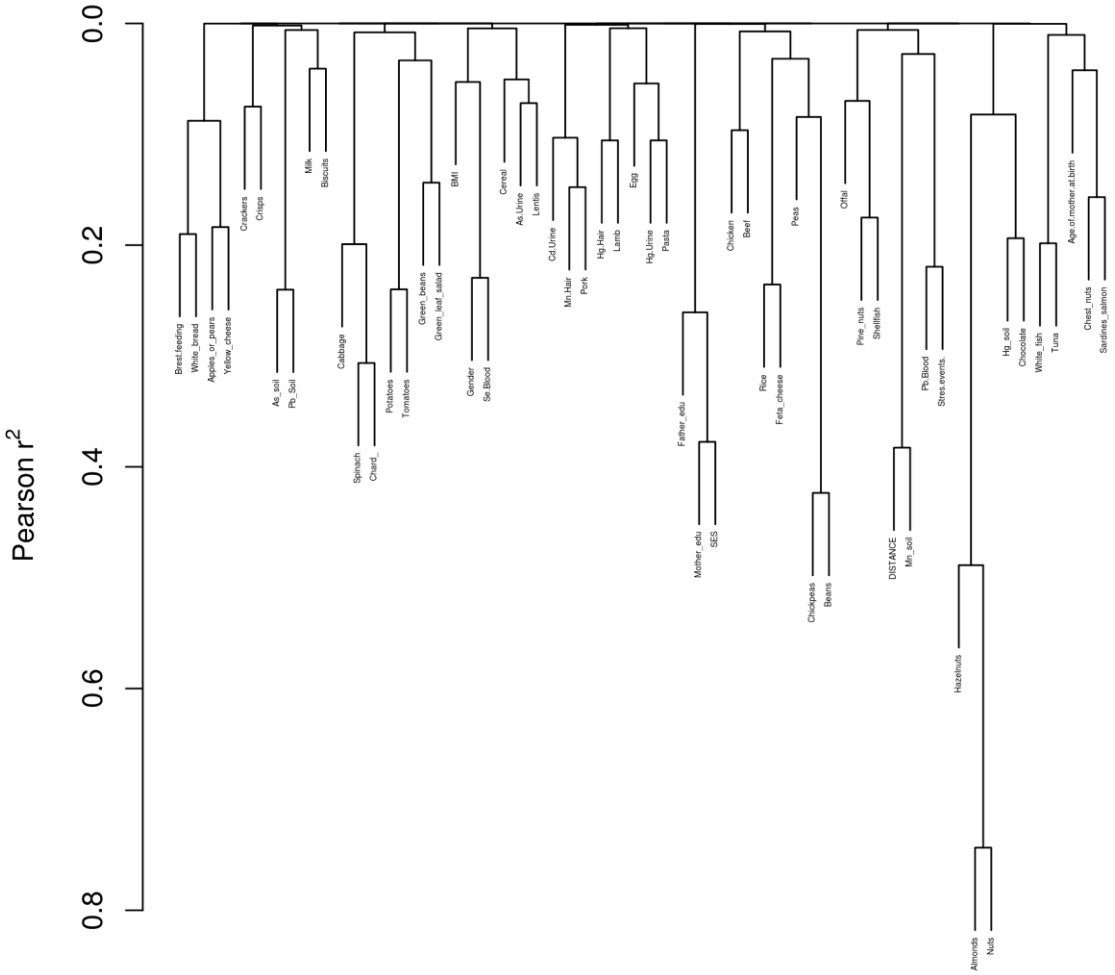

Figure 15. Hierarchical clustering using the Pearson correlation

The auto-correlations of the various parameters, are illustrated in both the heatmap (Figure 16) and the correlation globe (Figure 17).

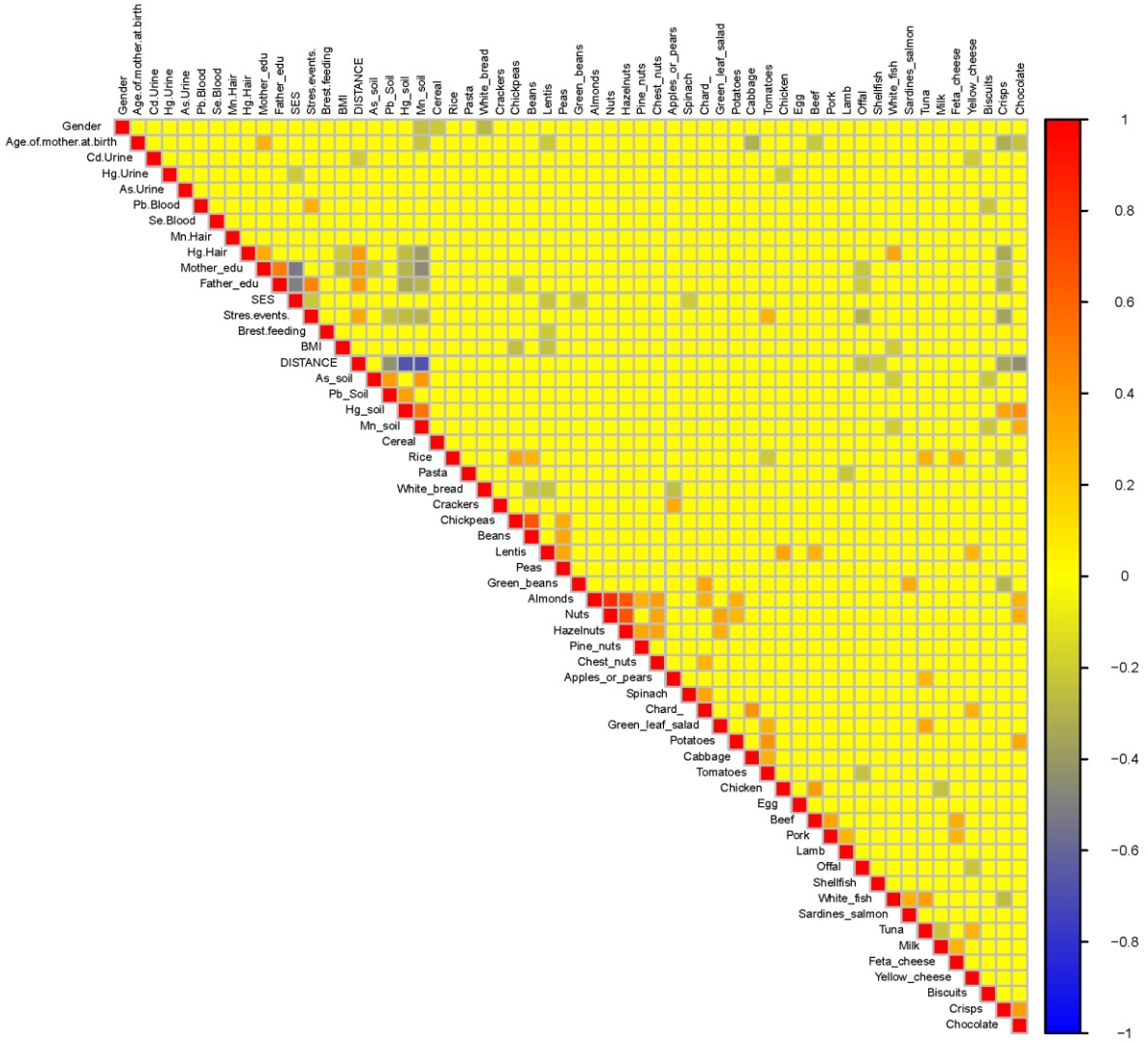

Figure 16. Heatmap of the exposure parameters of the HERACLES study

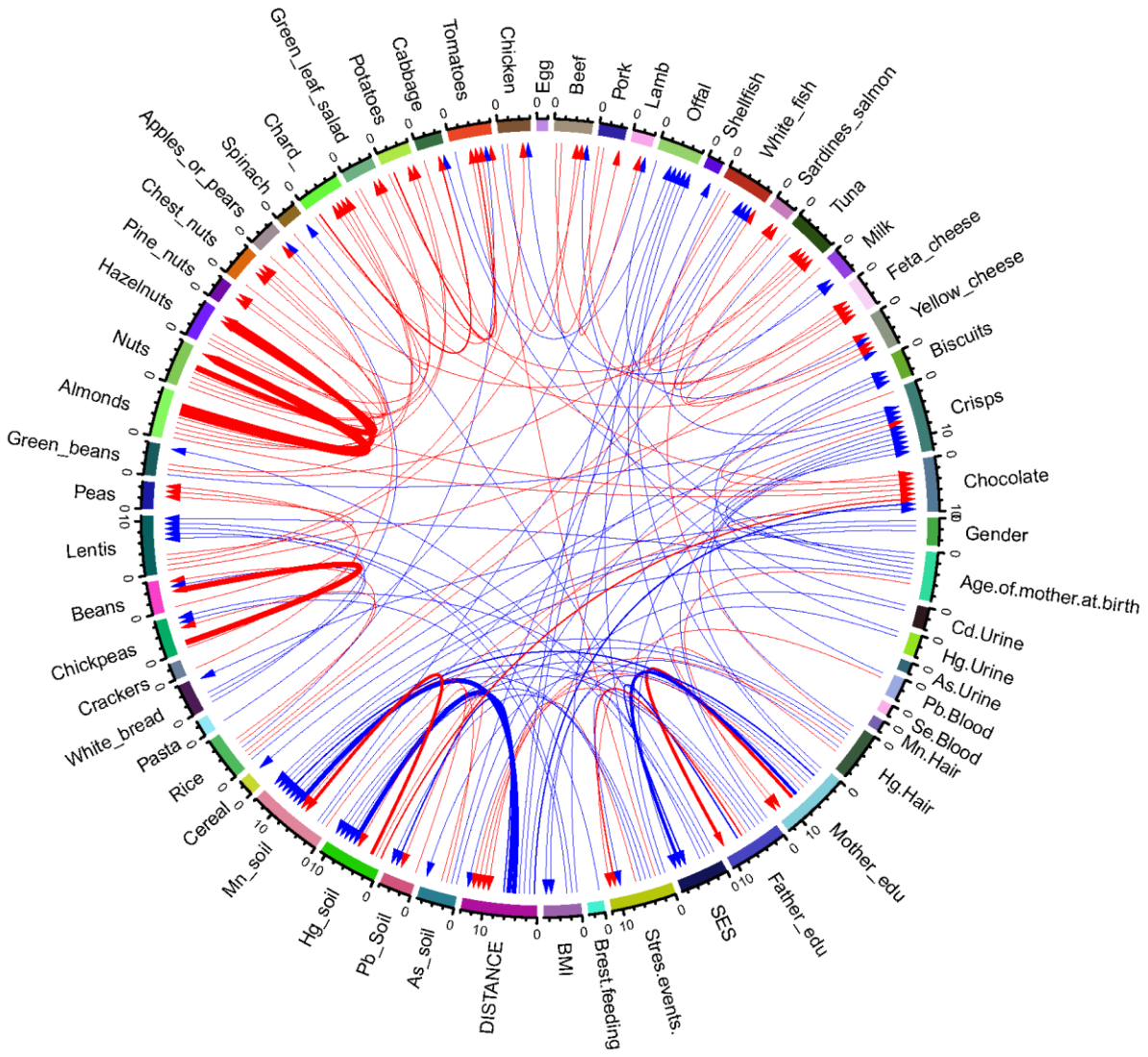

Figure 17. Correlation globe of the environmental, dietary and exposure factors of the HERACLES study

5.2.3.2 Results and interpretation of EWAS analysis

The outcomes of the associations among the various exposure and sociodemographic factors for one selected indicative outcome of each test battery are illustrated in the following Figures.

EWAS analysis results relevant to the **Child Behavioral CheckList** (CBCL) test battery results show that socio-cultural factors are strongly associated with children behavior. More specifically the *mother school title* and the *age of the mothers at birth* show both a robust statistical association (p-value<0.05 and in some cases p-value <0.01) with most of the CBCL indices considered. Looking at the volcano plots both parameters show a negative association with the CBCL scores indicating that lower educational level of the mothers as well as a lower age of the mother at the children birth may have negative impact on the children behavior.

The *stress index* was derived by merging the total number of stressful events detected by the mother and their average intensity is also playing an important role on the children behavior (p-value<0.05 and in some cases p-value<0.01) showing a negative effect on both internalizing and externalizing problems indices such as anxiety and depression, withdrawal and depression and somatic complaints, aggressive and rule-breaking behavior.

The *concentration of lead in blood* shows a strong statistical significance (p-value <0.05) with most of the CBCL indices analyzed. In this case the association shows a positive direction revealing a negative impact of higher blood concentration of lead on the on cognitive functions in children. This result is confirmed by a number of research studies which indicate exposure to lead as one of the most environmental determinants of neurodevelopmental disorders in children. On this subject the National Toxicology Program (NTP) has concluded that childhood lead exposure is associated with reduced cognitive function, including lower intelligence quotient (IQ) and reduced academic achievement (National Toxicology Program, 2012). The NTP has also concluded that childhood lead exposure is associated with attention-related behavioral problems (including inattention, hyperactivity, and diagnosed attention-deficit/hyperactivity disorder (ADHD)) and increased incidence of problem behaviors including delinquent, criminal, or antisocial behavior (National Toxicology Program, 2012).

Of opposite sign but still with robust statistical significance is the association of the concentration of selenium in blood which appears to act as beneficial element especially with regard to Internalizing Problems and ADHD as measured by CBCL battery indices. These results confirm the antioxidant properties of selenium which is a well-known regulator of brain function (Dominiak et al., 2016). These positive properties that selenium possesses are attributed to its ability to be incorporated into selenoproteins as an amino acid. Several selenoproteins are expressed in the brain, in which some of them, e.g. glutathione peroxidases (GPxs), thioredoxin reductases (TrxRs) or selenoprotein P (SelP), are strongly involved in antioxidant defense and in maintaining intercellular reducing conditions. Since increased oxidative stress has been implicated in neurological disorders higher levels of selenium in blood may be among the important factors protecting against those pathologies.

Breast feeding in the first months of children life is another parameter that shows a significant statistical association (p-value <0.05) especially with the internalizing problems as measured by CBCL battery indices. Also in this case the association shows a negative sign indicating that breastfeeding and especially its duration during the first year of life results in a beneficial effect on anxiety/depression, withdrawal/depression and somatic complaints as reported by the CBCL indices.

The concentration of mercury in hair reveals a strong association (p-value <0.05) with many CBCL indices considered, however its effect appears to have a controversial behavior as witnessed by its negative sign reported in the volcano plots indicating that higher concentration levels of Hg in hair may results in potential positive effect on the problem behavior in children.

CBCL indices as measured by teachers reveal slightly different patterns. Even though the socio cultural factors such as *mother school title* still show robust associations with most of the Child Behavioral CheckList test battery outcomes, other variables appear to play an important role. Among them the *distance of the residence address from the waste management site* shows a strong association especially with the internalizing problems. The negative sign of the association corroborates the negative impact of living in areas close to the waste management site especially on anxiety/depression, withdrawal/depression and somatic complaints.

Concentration of lead in blood is yet another significant variable (p-value <0.05) associated with Attention Deficit Hyperactivity Disorder while *Breast feeding* shows a strong association with Oppositional Defiant Disorder (ODD). Among the various food items considered, some of them show significant statistical association with CBCL indices. Consumption of *pork* (e.g. pork dishes, lard, bacon, salami) appears to be inversely associated (p-value <0.05) with the CBCL indices related to externalizing problems such as aggressive and rule-breaking behavior as well as with association with Oppositional Defiant Disorder and with Conduct Problems. High consumption of *chicken* reveals a strong association with Attention Deficit Hyperactivity Disorder measured by the teachers. Consumption of cabbage and lentils appears to influence negatively Attention Deficit Hyperactivity Disorder, Oppositional Defiant Disorder and with Conduct Problems too. High consumption of *coffee* is associated with externalizing problems such as aggressive and rule-breaking behavior and with Attention Deficit Hyperactivity Disorder and Conduct Problems measured by the teachers. *Fish consumption* reveals opposite effects on the basis of the fish type: while higher consumption of sardines and/or salmon (also canned) appears to be statistically associated with ADHD, consumption of herring, mackerel and trout (also canned) indicates a beneficial effect on the CBCL indices related to Conduct Problems. Higher consumption of *pine nuts and nuts* as well as of *white and wheat bread* indicates a beneficial effect on externalizing problems such as aggressive and rule-breaking behavior. Finally, higher consumption of

eggs and of beans are also associated to beneficial effects on internalizing problems (i.e. anxiety/depression, withdrawal/depression and somatic complaints) indices.

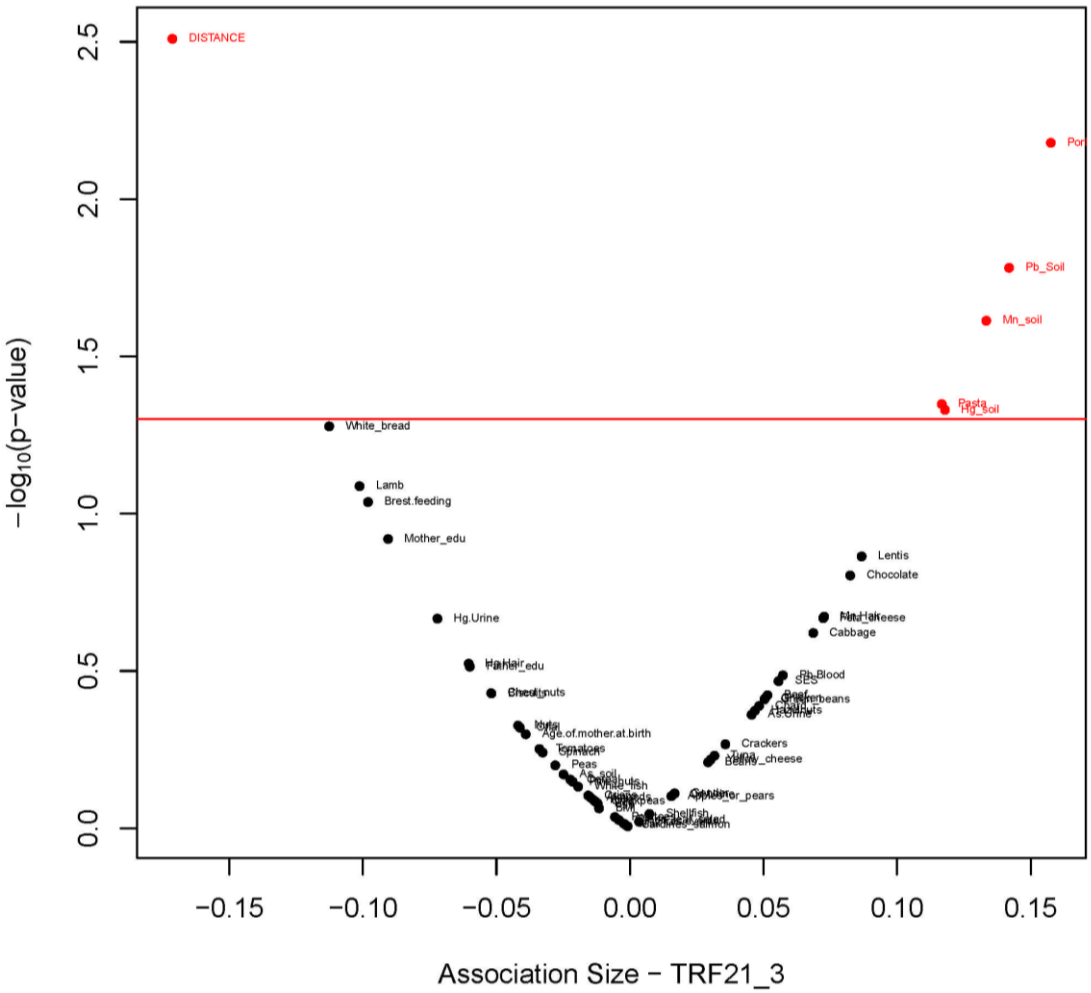

Figure 18. Associations of attention deficit / hyperactivity problems (from the CBCL test battery) with the environmental, dietary and exposure factors

EWAS analysis results relevant to the **Cambridge Neuropsychological Test Automated Battery (CANTAB)** test battery results show that concentration of Manganese in the hair is associated (p-value <0.05) with the Spatial Working Memory (SWM) with a positive sign revealing that higher Manganese levels in the hair increase error generation. Socio-cultural factors show again significant statistic association with the outcomes of the CANTAB test battery.

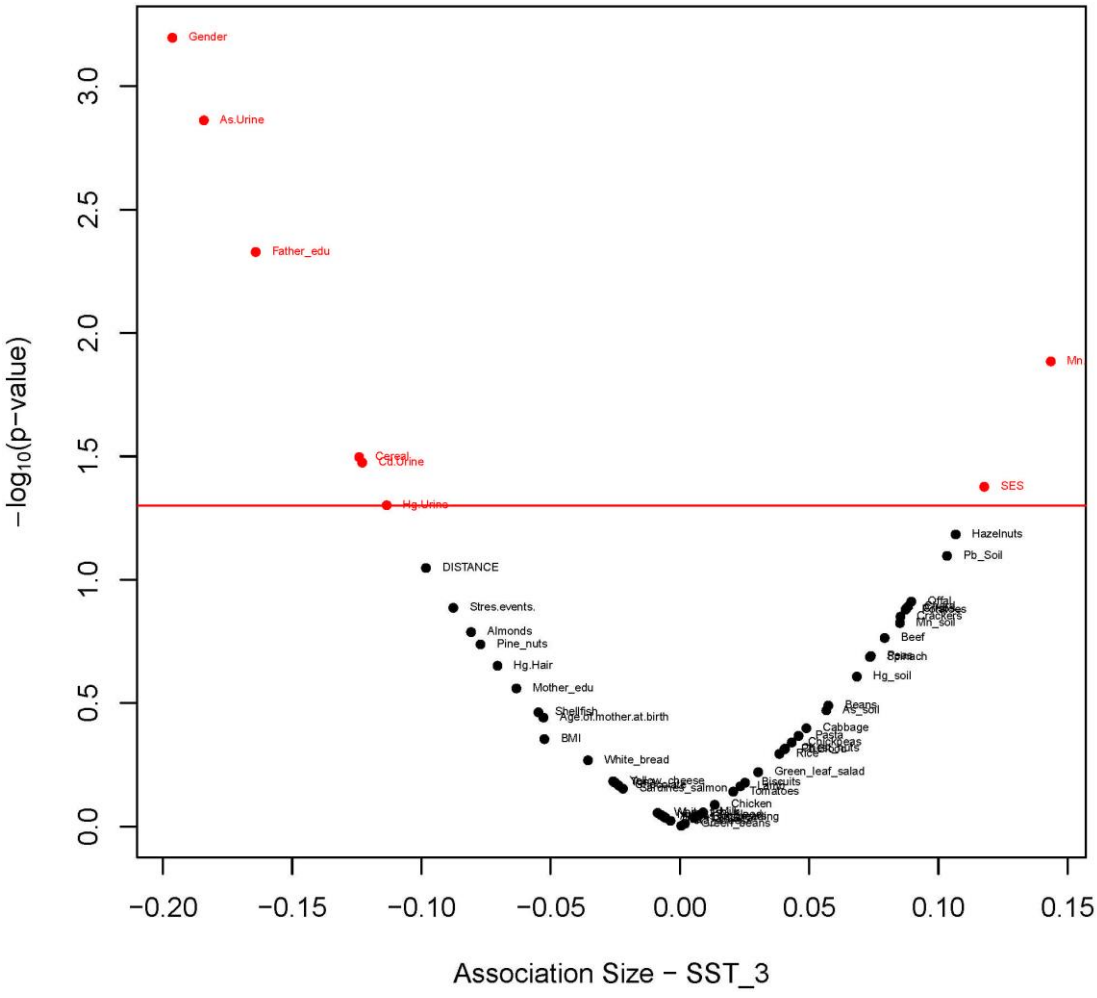

Figure 19. Associations of Stop Signal Task Mean correct RT on GO trials (from the CANTAB test battery) with the environmental, dietary and exposure factors

More in detail *Mother School title* appears to have a beneficial effect (p-value < 0.05) on the Stop Signal Task while *Father School title* on the spatial Working Memory Strategy index (p-value < 0.05). The *stress index* is also strongly associated (p-value <0.01) with the Spatial Working Memory (SWM) with a negative sign showing that higher stress levels decrease the error production.

EWAS analysis results relevant to the **Social Responsiveness Scale (SRS)** test battery show that also in this case socio-cultural factors are strongly associated with the Social Responsiveness Scale outcomes considered. *Mother school title* (p-value <0.000) and to a lower extent *Father school title* (p-value < 0.05) show both a robust statistical association the T scores of both the parents and teachers.

Moreover, the associations have a negative direction demonstrating that lower educational level of the parents may have negative impact on the Autism Spectrum Disorder (ASD) impairments of children.

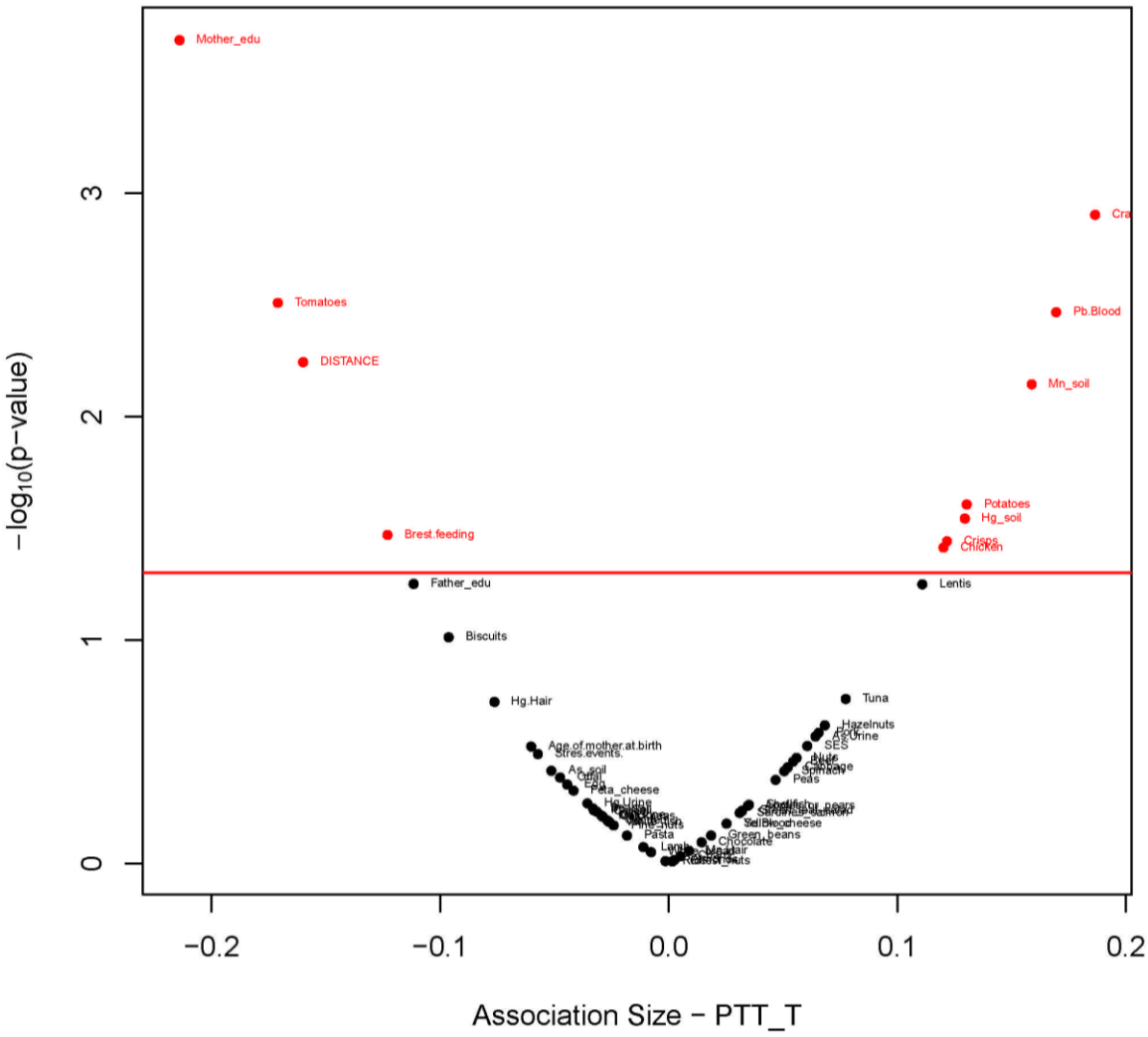

Figure 20. Associations of Total score T / Teachers (from the Social Responsiveness Scale test battery) with the environmental, dietary and exposure factors

Distance of the residence address from the waste management site shows a good association (p-value <0.001) with the T scores of the teachers (Figure 20). The negative sign of the association confirms the potential negative impact of living in the areas close to the waste management site on ASD impairments of children

Breastfeeding in the first months of children life also shows a good statistical association (p-value <0.05) with T scores as reported by teachers. The association shows a negative sign indicating the positive effect of breastfeeding, and especially its duration during the first year of life, on ASD impairments of children.

Among the biomonitoring data *selenium* in blood appears to be inversely associated (p-value <0.05) with the T scores of the parents. The negative sign of the association supports the positive impact of selenium on the neurodevelopmental disorders. Mercury concentration in hair shows a significant statistical association (p-value <0.05) with SRS battery indices and its effect appears to result in potential positive effect on ASD impairments.

Among the different food items higher consumption of pork dishes (lard, bacon, sausage) (p <0.01), coffee (p-value < 0.01), chicken (p < 0.05), breadsticks, crackers and rusks (p-value <0.05) and lentils (p-value <0.05) are associated with higher T scores of the SRS test battery indicating a potential negative effect on ASD impairments. On the contrary higher consumption of tomatoes (p-value < 0.001), white fish (e.g. codfish, rumble fish) (p-value < 0.01) and soft cheese (p-value <0.05)) are related with lower T scores of the SRS test battery signifying a potential positive effect on ASD impairment of children.

EWAS analysis results relevant to the **WISC-IV** test results show that the variable *distance of the residence address from the waste management site* is a key factor associated with almost all the indices of the WISC IV test. More specifically this variable shows a robust statistical association (p-value <0.001) with the Intelligence Quotient (IQ), Verbal Comprehension index, Perceptual Reasoning index, Working Memory index. Analysis of the results show a positive association with the WISC IV scores indicating that living far from the waste management site has a positive impact on the children cognitive functions. Some interesting conclusions can be drawn from the analysis of food consumption patterns. *Tomatoes* consumption appears to be statistically (p-value <0.05) associated to QI, Verbal Comprehension index and Working Memory index while cereal consumption reveals a strong association (p-value < 0.01) with the Perceptual Reasoning index. Both these food items show a positive sign meaning that their consumption has potential positive effects on the cognitive functions of the children. Epidemiological evidence suggests that consumption of lycopene, natural antioxidant presents in tomatoes, is able to reduce the risk of chronic diseases such as cancer, cardiovascular diseases as well as psychiatric syndromes (Story et al., 2010). In another study (Li and Zhang, 2007) reported that low serum levels of lycopene have been associated with increased risk of psychiatric disorders. One review of 22 studies examining the association of breakfast cereal consumption and academic performance in children and adolescents concluded that breakfast consumption may

improve cognitive function related to memory, test grades, and school attendance (Rampersaud et al., 2005).

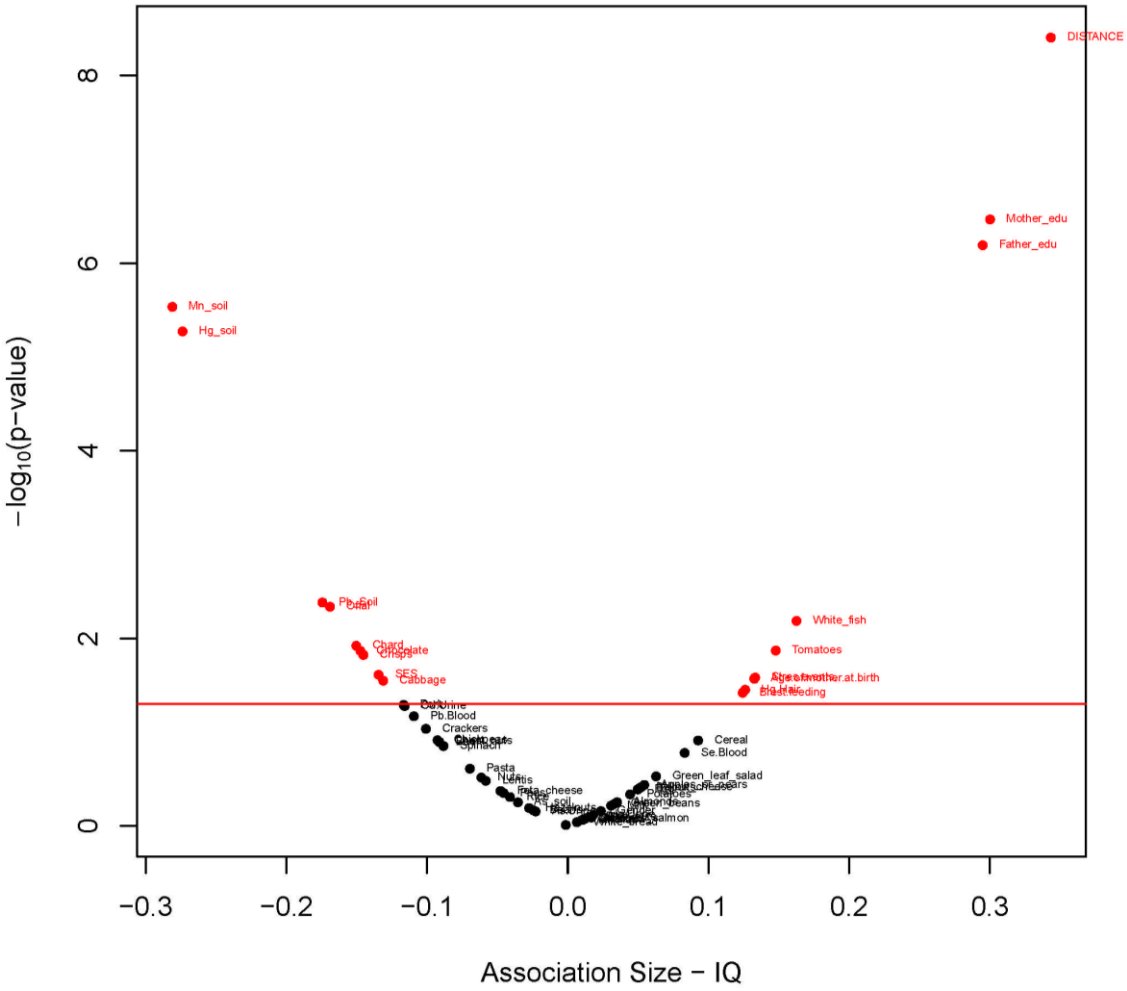

Figure 21. Association of intelligence quotient (from the WISC-IV test battery) with the environmental, dietary and exposure factors

Like for the CBCL test battery *consumption of fish* shows a dual behavior depending on the type of fish considered: higher consumption of *white fish* appears to have positive effects on the IQ and Verbal Comprehension index (p-value <0.001) while higher consumption of white fish reveals a negative effect (p-value <0.01) on Perceptual Reasoning index and Working Memory index.

6 Common nodes of toxicity pathways among the investigated endpoints

HEALS deals with 3 categories of health endpoints of high scientific and regulatory concern. Although in the past in epidemiology theses outcomes were treated as completely independent health states, recent findings have identified synergies among them.

Up to now, there is limited evidence on the common nodes. Forno et al. (2015) performed a cross-sectional study of 1429 adolescents aged 12 to 17 years in the 2007-2010 National Health and Nutrition Examination Survey. Adjusted regression was used to assess the relationships among obesity, insulin sensitivity/resistance, metabolic syndrome (MS), and lung function in children with and without asthma. Insulin resistance was negatively associated with FEV1 and forced vital capacity (FVC) in adolescents with and without asthma, whereas MS was associated with lower FEV1/FVC ratios, with a more pronounced decrease found among asthmatic patients; these associations were driven by overweight/obese adolescents. Higher body mass index was associated with a decrease in FEV1/FVC ratios among adolescents with insulin resistance. Compared with healthy participants, adolescents with MS had an approximately 2% decrease in FEV1/FVC ratios, adolescents with asthma had an approximately 6% decrease, and those with MS and asthma had approximately 10% decreased FEV1/FVC ratios ($P < .05$). Insulin resistance and MS were associated with worsened lung function in overweight/obese adolescents. Thus, it was concluded that asthma and MS synergistically decrease lung function, as do obesity and insulin resistance. These factors might contribute to the pathogenesis of asthma severity in obese patients and warrant further investigation.

Additional evidence on obesity as a major risk factor for asthma was provided by Leiria et al. (2015). Obesity is known to increase disease severity in asthmatic subjects and also to impair the efficacy of first-line treatment medications for asthma, worsening asthma control in obese patients. This concept is in agreement with the current understanding that some asthma phenotypes are not accompanied by detectable inflammation, and may not be ameliorated by classical anti-inflammatory therapy. There are growing evidences suggesting that the obesity-related asthma phenotype does not necessarily involve the classical TH2-dependent inflammatory process. Hormones involved in glucose homeostasis and in the pathogeneses of obesity likely directly or indirectly link obesity and asthma through inflammatory and non-inflammatory pathways. Furthermore, the endocrine regulation of the airway-related pre-ganglionic nerves likely contributes to airway hyperreactivity (AHR) in obese states. Clinical and basic studies highlighted the association between adipokines or insulin with pro-inflammatory mediators and showed that in accordance with other obesity-associated diseases, lowgrade inflammation may be determinant for the pathogeneses of asthma in obese patients. One key factor in clarifying the association between asthma and obesity is to understand the intercommunication between the innate immune system signals in the lung and the adipose tissue, and to know how insulin and adipokines mediate this conversation. Furthermore, central pathways regulating airway contractility are likely to modulate bronchial contractility in an obese state, as insulin and leptin signaling pathways are impaired in hypothalamic and extra-hypothalamic nuclei in this condition, influencing other peripheral functions, such as respiratory functions. it is not obesity per se, but metabolic changes (insulin resistance, leptin resistance and reduced circulating levels of adiponectin)

that accompany obesity, which potentiate AHR directly or through cross-talk with inflammatory pathways.

In addition, recent advances on the identification of disease causality (Stein et al., 2008)have identified common nodes among the involved toxicity pathways (Figure 22). Oxidative stress can trigger or increase inflammation through the activation of nuclear factor kappa B (NFkB), which is known to be sensitive to oxidative stress. Similarly, a variety of evidence suggests that amyloid-beta, a key factor in Alzheimer's disease, interacts in complex ways with inflammation and oxidative stress. While there is evidence of complex interactions—with amyloid-beta causing ROS/inflammation, and ROS/inflammation causing amyloid-beta production, evidence increasingly suggests that oxidative stress and inflammation commonly initiate this process.

Some critical questions that will be further studied in HEALS that will be greatly facilitated by multi-omics pathway analysis and biomarker discovery are the following ones:

- Are other common nodes of disease onset beyond oxidative stress and inflammation?
- What we can obtain more from regulatory pathway analysis
- What if we start hypothesis building and from reconstruction of Pathways of Toxicity?
- What can we learn from compounds associated to multiple endpoints (e.g. BPA has been associated both to neurodevelopmental disorders, as well as obesity and diabetes)?
- Do these different outcomes share a common mechanism of disease, or one environmental stressor might activate two different PoT?
- If two different PoT are activated, to what extent do these pathways converge in terms of gene expression and metabolomics signatures?

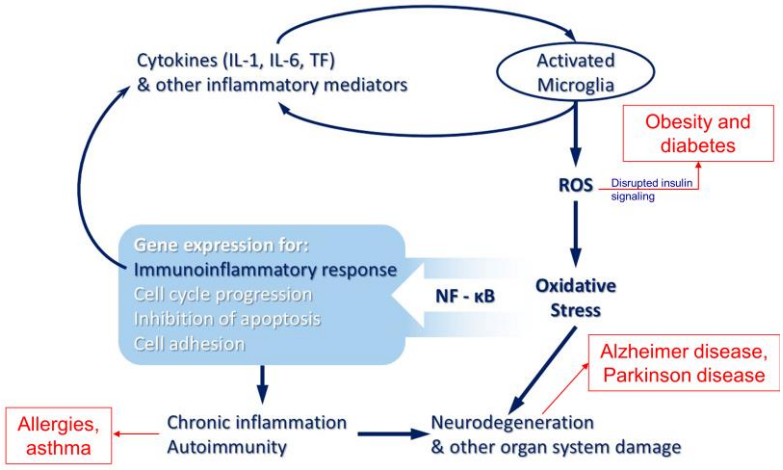

Figure 22. Mechanistic links of the various health endpoints of relevance in HEALS (Stein et al., 2008)

7 Conclusions

From the review process it was identified that several transcriptional and metabolite profiles have been to be associated with the health endpoints of interest in HEALS. What is really important is to identify the biomarkers that are really relevant to potential pathways of toxicity, especially in relation to the presence of environmental stressors. Thus, toxicity pathway analysis is a key process in the overall analysis for providing (a) a biologically plausible content and (b) for identifying nodes of toxicity that could eventually trigger more than one health endpoint. Given the above, pathways related to mitochondrial respiration and oxidative stress both in prenatal as well as post-natal exposure, seem to hold a key role in the evolvement of disease states. Once more it needs to be highlighted, that by simply using transcriptome and metabolome profiles without putting them in a biological contest, would result in data driven and eventually study-specific associations that will not provide real explanation on the causality between exposure and disease.

The Environment Wide Associations Study (EWAS) was found to provide the necessary framework for an unbiased stance towards the derivation of initial hypotheses associating exposure and disease, as well as the use of multiple types of data. In HEALS a novel approach towards defining causal associations between health status and environmental stressors is introduced, by do not considering confounders but examining all related factors as co-variates, which are in “linkage disequilibrium” with each other. REPRO_PL study and HERACLES study were using different data for associating different type of environmental, exposure sociodemographic data and biospecimens, as well as different test batteries. However, despite the differences in the availability of the data, the strength of the methodology allowed us to better identify associations between different exposure factors and neurodevelopment. In the case of Poland, it was clearly illustrated that several gestation factors have a beneficial (e.g. the concentration of selenium in maternal blood) or a negative influence (e.g. maternal bodyweight) in child neurodevelopment. On the other hand, child exposure to phthalates itself has a stronger negative influence in child neurodevelopment than maternal exposure. From the HERACLES study, it was found that children exposure to heavy metals and proximity to waste management sites have a negative influence in child neurodevelopment, however these effects are significantly modified by sociodemographic parameters (such as children SES and parents educational level), as well as dietary components such as nuts, hazelnuts and almonds; these food items provide a beneficial contribution, related to their high content in in omega 3 fatty acids. Direct actions of omega-3 polyunsaturated fatty acids on neuronal composition, neurochemical signaling and cognitive function constitute a multidisciplinary rationale for classification of dietary lipids as "brain foods."

To complete the analysis in both cohorts, further steps include the results of the toxicity pathways identified to be perturbed. The identification of perturbed pathways, will allow us to use them as additional variables in the EWAS analysis and in this way additional findings on the causality of neurodevelopmental levels will be elucidated.

References

Abu Bakar MH, Sarmidi MR, Cheng KK, Ali Khan A, Suan CL, Zaman Huri H, Yaakob H. Metabolomics - the complementary field in systems biology: a review on obesity and type 2 diabetes. Mol Biosyst 2015; 11: 1742-74.

Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms & profiles: an integrated system of mult-informant assessment. University of Vermont, Research Center for Children, Youth & Families, Burlington, 2001.

Amaral AFS. Metabolomics of asthma. Journal of Allergy and Clinical Immunology 2014; 133: 1497-1499.e1.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. American Psychiatric Association, Washington, 2000.

Anderberg MR. Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks. Vol 19: Academic press, 2014.

Andersen AD, Binzer M, Stenager E, Gramsbergen JB. Cerebrospinal fluid biomarkers for Parkinson's disease - a systematic review. Acta Neurol Scand 2016.

Bao Y, Zhao T, Wang X, Qiu Y, Su M, Jia W, Jia W. Metabonomic variations in the drug-treated type 2 diabetes mellitus patients and healthy volunteers. J Proteome Res 2009; 8: 1623-30.

Berkhin P. A survey of clustering data mining techniques. Grouping multidimensional data. Springer, 2006, pp. 25-71.

Blasco H, Patin F, Madji Hounoum B, Gordon PH, Vourc'h P, Andres CR, Corcia P. Metabolomics in amyotrophic lateral sclerosis: how far can it take us? Eur J Neurol 2016; 23: 447-54.

Botas A, Campbell HM, Han X, Maletic-Savatic M. Metabolomics of neurodegenerative diseases. Int Rev Neurobiol 2015; 122: 53-80.

Brandt EB, Biagini Myers JM, Acciani TH, Ryan PH, Sivaprasad U, Ruff B, LeMasters GK, Bernstein DI, Lockey JE, LeCras TD, Khurana Hershey GK. Exposure to allergen and diesel exhaust particles potentiates secondary allergen-specific memory responses, promoting asthma susceptibility. Journal of Allergy and Clinical Immunology 2015; 136: 295-303.e7.

Broix L, Jagline H, L Ivanova E, Schmucker S, Drouot N, Clayton-Smith J, Pagnamenta AT, Metcalfe KA, Isidor B, Louvier UW, Poduri A, Taylor JC, Tilly P, Poirier K, Saillour Y, Lebrun N, Stemmelen T, Rudolf G, Muraca G, Saintpierre B, Elmorjani A, Deciphering Developmental Disorders s, Moise M, Weirauch NB, Guerrini R, Boland A, Olaso R, Masson C, Tripathy R, Keays D, Beldjord C, Nguyen L, Godin J, Kini U, Nischke P, Deleuze J-F, Bahi-Buisson N, Sumara I, Hinckelmann M-V, Chelly J. Mutations in the HECT

domain of NEDD4L lead to AKT-mTOR pathway deregulation and cause periventricular nodular heterotopia. Nat Genet 2016; advance online publication.

Calimlioglu B, Karagoz K, Sevimoglu T, Kilic E, Gov E, Arga KY. Tissue-Specific Molecular Biomarker Signatures of Type 2 Diabetes: An Integrative Analysis of Transcriptomics and Protein-Protein Interaction Data. Omics 2015; 19: 563-73.

Capuron L, Ravaud A, Dantzer R. Timing and specificity of the cognitive changes induced by interleukin-2 and interferon-alpha treatments in cancer patients. Psychosom Med 2001; 63: 376-86.

Carlsson M, Thorell L, Sjolander A, Larsson-Faria S. Variability of total and free IgE levels and IgE receptor expression in allergic subjects in and out of pollen season. Scand J Immunol 2015; 81: 240-8.

Castro-Nallar E, Shen Y, Freishtat RJ, Pérez-Losada M, Manimaran S, Liu G, Johnson WE, Crandall KA. Integrating microbial and host transcriptomics to characterize asthma-associated microbial communities. BMC Medical Genomics 2015; 8: 50.

Caubit X, Gubellini P, Andrieux J, Roubertoux PL, Metwaly M, Jacq B, Fatmi A, Had-Aissouni L, Kwan KY, Salin P, Carlier M, Lieden A, Rudd E, Shinawi M, Vincent-Delorme C, Cuisset J-M, Lemaitre M-P, Abderrehamane F, Duban B, Lemaitre J-F, Woolf AS, Bockenhauer D, Severac D, Dubois E, Zhu Y, Sestan N, Garratt AN, Le Goff LK, Fasano L. TSHZ3 deletion causes an autism syndrome and defects in cortical projection neurons. Nat Genet 2016; advance online publication.

Chang C, Guo Z-g, He B, Yao W-z. Metabolic alterations in the sera of Chinese patients with mild persistent asthma: a GC-MS-based metabolomics analysis. Acta Pharmacologica Sinica 2015; 36: 1356-1366.

Chen A, Dietrich KN, Huo X, Ho SM. Developmental neurotoxicants in e-waste: An emerging health concern. Environmental Health Perspectives 2011; 119: 431-438.

Constantino JN, Przybeck T, Friesen D, Todd RD. Reciprocal social behavior in children with and without pervasive developmental disorders. J Dev Behav Pediatr 2000; 21: 2-11.

Dervola KSN, Johansen EB, Walaas SI, Fonnum F. Gender-dependent and genotype-sensitive monoaminergic changes induced by polychlorinated biphenyl 153 in the rat brain. NeuroToxicology 2015; 50: 38-45.

Diémé B, Mavel S, Blasco H, Tripi G, Bonnet-Brilhault F, Malvy J, Bocca C, Andres CR, Nadal-Desbarats L, Emond P. Metabolomics Study of Urine in Autism Spectrum Disorders Using a Multiplatform Analytical Methodology. Journal of Proteome Research 2015; 14: 5273-5282.

Dominiak A, Wilkaniec A, Wroczynski P, Adamczyk A. Selenium in the Therapy of Neurological Diseases. Where is it Going? Curr Neuropharmacol 2016; 14: 282-99.

Drabkova P, Sanderova J, Kovarik J, kandar R. An Assay of Selected Serum Amino Acids in Patients with Type 2 Diabetes Mellitus. Adv Clin Exp Med 2015; 24: 447-51.

Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JRB, Egan JM, Lajunen T, Grarup N, Sparsø T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proença C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan Ja, Elliott A, McCarroll SA, Payne F, Roccasecca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen Y-DI, Chines P, Clarke R, Coin LJM, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day INM, de Geus E, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen A-L, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PRV, Jørgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martínez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orrù M, Pakyz R, Palmer CNA, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AFH, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurðsson G, Sijbrands EJG, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvänen A-C, Tanaka T, Thorand B, Tichet J, Tönjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JCM, Yarnell JWG, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Consortium D, Consortium G, Global BC, Borecki IB, Loos RJF, Meneton P, Magnusson PKE, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Ríos M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WHL, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Hamsten A, on behalf of Procardis c, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BWJH, Boomsma D, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin M-R, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I, for the Mi. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nature genetics 2010; 42: 105-116.

Durak O, Gao F, Kaeser-Woo YJ, Rueda R, Martorell AJ, Nott A, Liu CY, Watson LA, Tsai L-H. Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. Nat Neurosci 2016; advance online publication.

Dutta T, Chai HS, Ward LE, Ghosh A, Persson X-MT, Ford GC, Kudva YC, Sun Z, Asmann YW, Kocher J-PA, Nair KS. Concordance of Changes in Metabolic Pathways Based on Plasma Metabolomics and Skeletal Muscle Transcriptomics in Type 1 Diabetes. Diabetes 2012; 61: 1004-1016.

Emond P, Mavel S, Aïdoud N, Nadal-Desbarats L, Montigny F, Bonnet-Brilhault F, Barthélémy C, Merten M, Sarda P, Laumonnier F, Vourc'h P, Blasco H, Andres CR. GC-MS-based urine metabolic profiling of autism spectrum disorders. Analytical and Bioanalytical Chemistry 2013; 405: 5291-5300.

Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 2011; 473: 43-49.

Fiehn O, Garvey WT, Newman JW, Lok KH, Hoppel CL, Adams SH. Plasma Metabolomic Profiles Reflective of Glucose Homeostasis in Non-Diabetic and Type 2 Diabetic Obese African-American Women. PLoS ONE 2010; 5: e15234.

Forno E, Han Y-Y, Muzumdar RH, Celedón JC. Insulin resistance, metabolic syndrome, and lung function in US adolescents with and without asthma. Journal of Allergy and Clinical Immunology 2015; 136: 304-311.e8.

Frayling TM. Genome–wide association studies provide new insights into type 2 diabetes aetiology. Nature Reviews Genetics 2007; 8: 657-662.

Gahleitner F, Guallar-Hoyas C, Beardsmore CS, Pandya HC, Thomas CP. Metabolomics pilot study to identify volatile organic compound markers of childhood asthma in exhaled breath. Bioanalysis 2013; 5: 2239-47.

Giulivi C, Napoli E, Tassone F, Halmai J, Hagerman R. Plasma Biomarkers for Monitoring Brain Pathophysiology in FMR1 Premutation Carriers. Front Mol Neurosci 2016; 9: 71.

Gómez-Sánchez R, Yakhine-Diop SMS, Bravo-San Pedro JM, Pizarro-Estrella E, Rodríguez-Arribas M, Climent V, Martin-Cano FE, González-Soltero ME, Tandon A, Fuentes JM, González-Polo RA. PINK1 deficiency enhances autophagy and mitophagy induction. Molecular & Cellular Oncology 2016; 3: e1046579.

Graham SF, Kumar PK, Bjorndahl T, Han B, Yilmaz A, Sherman E, Bahado-Singh RO, Wishart D, Mann D, Green BD. Metabolic signatures of Huntington's disease (HD): 1H NMR analysis of the polar metabolome in post-mortem human brain. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 2016; 1862: 1675-1684.

Grant SFA, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 2006; 38: 320-323.

Grarup N, Sandholt CH, Hansen T, Pedersen O. Genetic susceptibility to type 2 diabetes and obesity: from genome-wide association studies to rare variants and beyond. Diabetologia 2014; 57: 1528-1541.

Gray E, Larkin JR, Claridge TDW, Talbot K, Sibson NR, Turner MR. The longitudinal cerebrospinal fluid metabolomic profile of amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration 2015; 16: 456-463.

Grzadzinski R, Di Martino A, Brady E, Mairena MA, O'Neale M, Petkova E, Lord C, Castellanos FX. Examining autistic traits in children with ADHD: does the autism spectrum extend to ADHD? J Autism Dev Disord 2011; 41: 1178-91.

Guo L, Li BX, Deng M, Wen F, Jiang JH, Tan YQ, Song YZ, Liu ZH, Zhang CH, Kobayashi K, Wang ZN. Etiological analysis of neurodevelopmental disabilities: single-center eight-year clinical experience in south China. J Biomed Biotechnol 2011; 2011.

Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, West AB, Arking DE. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. Nature Communications 2014; 5: 5748.

Ha CY, Kim JY, Paik JK, Kim OY, Paik Y-H, Lee EJ, Lee JH. The association of specific metabolites of lipid metabolism with markers of oxidative stress, inflammation and arterial stiffness in men with newly diagnosed type 2 diabetes. Clinical Endocrinology 2012; 76: 674-682.

Han L-D, Xia J-F, Liang Q-L, Wang Y, Wang Y-M, Hu P, Li P, Luo G-A. Plasma esterified and non-esterified fatty acids metabolic profiling using gas chromatography–mass spectrometry and its application in the study of diabetic mellitus and diabetic nephropathy. Analytica Chimica Acta 2011a; 689: 85-91.

Han X, Rozen S, Boyle SH, Hellegers C, Cheng H, Burke JR, Welsh-Bohmer KA, Doraiswamy PM, Kaddurah-Daouk R. Metabolomics in Early Alzheimer's Disease: Identification of Altered Plasma Sphingolipidome Using Shotgun Lipidomics. PLoS ONE 2011b; 6: e21643.

Harrell Jr FE, Dupont MC. The Hmisc Package. R package version 2006; 3: 0-12.

Hatano T, Saiki S, Okuzumi A, Mohney RP, Hattori N. Identification of novel biomarkers for Parkinson's disease by metabolomic technologies. J Neurol Neurosurg Psychiatry 2016; 87: 295-301.

Hawi Z, Cummins TD, Tong J. Rare DNA variants in the brain-derived neurotrophic factor gene increase risk for attention-deficit hyperactivity disorder: a next-generation sequencing study. 2016.

Heberling C, Dhurjati P. Novel systems modeling methodology in comparative microbial metabolomics: identifying key enzymes and metabolites implicated in autism spectrum disorders. Int J Mol Sci 2015; 16: 8949-67.

Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proceedings of the National Academy of Sciences of the United States of America 2009; 106: 9362-9367.

Hinks TSC, Zhou X, Staples KJ, Dimitrov BD, Manta A, Petrossian T, Lum PY, Smith CG, Ward JA, Howarth PH, Walls AF, Gadola SD, Djukanović R. Innate and adaptive T cells in asthmatic patients: Relationship to severity and disease mechanisms. Journal of Allergy and Clinical Immunology 2015; 136: 323-333.

Ho WE, Xu YJ, Xu F, Cheng C, Peh HY, Tannenbaum SR, Wong WS, Ong CN. Metabolomics reveals altered metabolic pathways in experimental asthma. Am J Respir Cell Mol Biol 2013; 48: 204-11.

Hoeffding W. A class of statistics with asymptotically normal distribution. The annals of mathematical statistics 1948: 293-325.

Ibanez C, Cifuentes A, Simo C. Recent advances and applications of metabolomics to investigate neurodegenerative diseases. Int Rev Neurobiol 2015; 122: 95-132.

Inoue YU, Inoue T. Brain enhancer activities at the gene-poor 5p14.1 autism-associated locus. Scientific Reports 2016; 6: 31227.

Irizarry MC. Biomarkers of Alzheimer Disease in Plasma. NeuroRX 2004; 1: 226-234.

Jenkinson CP, Göring HHH, Arya R, Blangero J, Duggirala R, DeFronzo RA. Transcriptomics in type 2 diabetes: Bridging the gap between genotype and phenotype. Genomics Data 2016; 8: 25-36.

Jiao H, Kaaman M, Dungner E, Kere J, Arner P, Dahlman I. Association analysis of positional obesity candidate genes based on integrated data from transcriptomics and linkage analysis. Int J Obes (Lond) 2008; 32: 816-25.

Jung J, Kim SH, Lee HS, Choi GS, Jung YS, Ryu DH, Park HS, Hwang GS. Serum metabolomics reveals pathways and biomarkers associated with asthma pathogenesis. Clin Exp Allergy 2013; 43: 425-33.

Kang J, Lu J, Zhang X. Metabolomics-based promising candidate biomarkers and pathways in Alzheimer's disease. Pharmazie 2015; 70: 277-82.

Kim CS, Park JR, Yu SH, Kang JG, Ryu OH, Lee SJ, Hong EG, Kim D-M, Yoo JM, Ihm SH, Choi MG, Yoo HJ. Impact of Serum Adiponectin Concentration on Progression of Carotid Atherosclerosis in Patients with Type 2 Diabetes Mellitus. Endocrinol Metab 2012; 27: 31-38.

Kim JY, Park JY, Kim OY, Ham BM, Kim H-J, Kwon DY, Jang Y, Lee JH. Metabolic Profiling of Plasma in Overweight/Obese and Lean Men using Ultra Performance Liquid Chromatography and Q-TOF Mass Spectrometry (UPLC−Q-TOF MS). Journal of Proteome Research 2010; 9: 4368-4375.

Larché M, Robinson DS, Kay AB. The role of T lymphocytes in the pathogenesis of asthma. Journal of Allergy and Clinical Immunology 2003; 111: 450-463.

Lawton KA, Brown MV, Alexander D, Li Z, Wulff JE, Lawson R, Jaffa M, Milburn MV, Ryals JA, Bowser R, Cudkowicz ME, Berry JD. Plasma metabolomic biomarker panel to distinguish patients with amyotrophic lateral sclerosis from disease mimics. Amyotroph Lateral Scler Frontotemporal Degener 2014; 15: 362-70.

Leiria LOS, Martins MA, Saad MJA. Obesity and asthma: beyond TH2 inflammation. Metabolism 2015; 64: 172-181.

Lelieveld SH, Reijnders MRF, Pfundt R, Yntema HG, Kamsteeg E-J, de Vries P, de Vries BBA, Willemsen MH, Kleefstra T, Lohner K, Vreeburg M, Stevens SJC, van der Burgt I, Bongers EMHF, Stegmann APA, Rump P, Rinne T, Nelen MR, Veltman JA, Vissers LELM, Brunner HG, Gilissen C. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat Neurosci 2016; 19: 1194-1196.

Li N, Buglak N. Convergence of air pollutant-induced redox-sensitive signals in the dendritic cells contributes to asthma pathogenesis. Toxicology Letters 2015; 237: 55-60.

Li X, Xu Z, Lu X, Yang X, Yin P, Kong H, Yu Y, Xu G. Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for diabetes mellitus. Analytica Chimica Acta 2009; 633: 257-262.

Li Y, Zhang J. Serum concentrations of antioxidant vitamins and carotenoids are low in individuals with a history of attempted suicide. Nutr Neurosci 2007; 10: 51-8.

Lind PM, Risérus U, Salihovic S, Bavel Bv, Lind L. An environmental wide association study (EWAS) approach to the metabolic syndrome. Environment International 2013; 55: 1-8.

Luan H, Liu L-F, Tang Z, Zhang M, Chua K-K, Song J-X, Mok VCT, Li M, Cai Z. Comprehensive urinary metabolomic profiling and identification of potential noninvasive marker for idiopathic Parkinson's disease. Scientific Reports 2015; 5: 13888.

Maniscalco M, Paris D, Melck DJ, D'Amato M, Zedda A, Sofia M, Stellato C, Motta A. Coexistence of Obesity and Asthma Determines a Distinct Respiratory Metabolic Phenotype. J Allergy Clin Immunol 2016.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A. Finding the missing heritability of complex diseases. Nature 2009; 461: 747-753.

Mattarucchi E, Baraldi E, Guillou C. Metabolomics applied to urine samples in childhood asthma; differentiation between asthma phenotypes and identification of relevant metabolites. Biomedical Chromatography 2012; 26: 89-94.

Mavel S, Nadal-Desbarats L, Blasco H, Bonnet-Brilhault F, Barthélémy C, Montigny F, Sarda P, Laumonnier F, Vourc′ h P, Andres CR, Emond P. 1H-13C NMR-based urine metabolic profiling in autism spectrum disorders. Talanta 2013; 114: 95-102.

McGeachie MJ, Dahlin A, Qiu W, Croteau-Chonka DC, Savage J, Wu AC, Wan ES, Sordillo JE, Al-Garawi A, Martinez FD, Strunk RC, Lemanske RF, Liu AH, Raby BA, Weiss S, Clish CB, Lasky-Su JA. The metabolomics of asthma control: a promising link between genetics and disease. Immunity, Inflammation and Disease 2015; 3: 224-238.

Mehta SH, Adler CH. Advances in Biomarker Research in Parkinson's Disease. Current Neurology and Neuroscience Reports 2015; 16: 7.

Morton NM, Nelson YB, Michailidou Z, Di Rollo EM, Ramage L, Hadoke PWF, Seckl JR, Bunger L, Horvat S, Kenyon CJ, Dunbar DR. A Stratified Transcriptomics Analysis of Polygenic Fat and Lean Mouse Adipose Tissues Identifies Novel Candidate Obesity Genes. PLoS ONE 2011; 6: e23944.

Napoli E, Tassone F, Wong S, Angkustsiri K, Simon TJ, Song G, Giulivi C. Mitochondrial Citrate Transporter-dependent Metabolic Signature in the 22q11.2 Deletion Syndrome. J Biol Chem 2015; 290: 23240-53.

National Toxicology Program. NTP Monograph on Health Effects of Low-Level Lead. National Institute of Environmental Health Sciences, Research Triangle Park, NC, 2012.

Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, Haqq AM, Shah SH, Arlotto M, Slentz CA, Rochon J, Gallup D, Ilkayeva O, Wenner BR, Yancy WE, Eisenson H, Musante G, Surwit R, Millington DS, Butler MD, Svetkey LP. A Branched-Chain Amino Acid-Related Metabolic Signature that Differentiates Obese and Lean Humans and Contributes to Insulin Resistance. Cell metabolism 2009; 9: 311-326.

Noto A, Fanos V, Barberini L, Grapov D, Fattuoni C, Zaffanello M, Casanova A, Fenu G, De Giacomo A, De Angelis M, Moretti C, Papoff P, Ditonno R, Francavilla R. The urinary metabolomics profile of an Italian autistic children population and their unaffected siblings. J Matern Fetal Neonatal Med 2014; 27 Suppl 2: 46-52.

Palmer JA, Poenitzsch AM, Smith SM, Conard KR, West PR, Cezar GG. Metabolic Biomarkers of Prenatal Alcohol Exposure in Human Embryonic Stem Cell-derived Neural Lineages. Alcoholism, clinical and experimental research 2012; 36: 1314-1324.

Patel CJ, Bhattacharya J, Butte AJ. An environment-wide association study (EWAS) on type 2 diabetes mellitus. PloS one 2010; 5: e10746.

Patel CJ, Cullen MR, Ioannidis JP, Butte AJ. An Environment-wide Association Study Identifies Novel and Replicated Nutrients and Pollutants Associated with Serum Lipid Levels. Circulation 2012; 125: A013.

Patin F, Corcia P, Vourc'h P, Nadal-Desbarats L, Baranek T, Goossens JF, Marouillat S, Dessein AF, Descat A, Madji Hounoum B, Bruno C, Leman S, Andres CR, Blasco H. Omics to Explore Amyotrophic Lateral Sclerosis Evolution: the Central Role of Arginine and Proline Metabolism. Mol Neurobiol 2016.

Poole A, Urbanek C, Eng C, Schageman J, Jacobson S, O'Connor BP, Galanter JM, Gignoux CR, Roth LA, Kumar R, Lutz S, Liu AH, Fingerlin TE, Setterquist RA, Burchard EG, Rodriguez-Santana J, Seibold MA. Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. Journal of Allergy and Clinical Immunology 2014; 133: 670-678.e12.

Prokopenko I, McCarthy MI, Lindgren CM. Type 2 diabetes: new genes, new understanding. Trends in Genetics 2008; 24: 613-621.

Putschögl FM, Gaum PM, Schettgen T, Kraus T, Gube M, Lang J. Effects of occupational exposure to polychlorinated biphenyls on urinary metabolites of neurotransmitters: A cross-sectional and longitudinal perspective. International Journal of Hygiene and Environmental Health 2015; 218: 452-460.

Raedler D, Ballenberger N, Klucker E, Böck A, Otto R, Prazeres da Costa O, Holst O, Illig T, Buch T, von Mutius E, Schaub B. Identification of novel immune phenotypes for allergic and nonallergic childhood asthma. Journal of Allergy and Clinical Immunology 2015; 135: 81-91.

Rampersaud GC, Pereira MA, Girard BL, Adams J, Metzl JD. Breakfast habits, nutritional status, body weight, and academic performance in children and adolescents. J Am Diet Assoc 2005; 105: 743-60; quiz 761-2.

Rango M, Arighi A, Marotta G, Ronchi D, Bresolin N. PINK1 parkinsonism and Parkinson disease: Distinguishable brain mitochondrial function and metabolomics. Mitochondrion 2013; 13: 59-61.

Ren C, Hu X, Li X, Zhou Q. Ultra-trace graphene oxide in a water environment triggers Parkinson's disease-like symptoms and metabolic disturbance in zebrafish larvae. Biomaterials 2016; 93: 83-94.

Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, Lin DY, Duan J, Ophoff RA, Andreassen OA, Scolnick E, Cichon S, St. Clair D, Corvin A, Gurling H, Werge T, Rujescu D, Blackwood DHR, Pato CN, Malhotra AK, Purcell S, Dudbridge F, Neale BM, Rossin L, Visscher PM, Posthuma D, Ruderfer DM, Fanous A, Stefansson H, Steinberg S, Mowry BJ, Golimbet V, De Hert M, Jönsson EG, Bitter I, Pietiläinen OPH, Collier DA, Tosato S, Agartz I, Albus M, Alexander M, Amdur RL, Amin F, Bass N, Bergen SE, Black DW, Børglum AD, Brown MA, Bruggeman R, Buccola NG, Byerley WF, Cahn W,

Cantor RM, Carr VJ, Catts SV, Choudhury K, Cloninger CR, Cormican P, Craddock N, Danoy PA, Datta S, De Haan L, Demontis D, Dikeos D, Djurovic S, Donnelly P, Donohoe G, Duong L, Dwyer S, Fink-Jensen A, Freedman R, Freimer NB, Friedl M, Georgieva L, Giegling I, Gill M, Glenthøj B, Godard S, Hamshere M, Hansen M, Hansen T, Hartmann AM, Henskens FA, Hougaard DM, Hultman CM, Ingason A, Jablensky AV, Jakobsen KD, Jay M, Jürgens G, Kahn RS, Keller MC, Kenis G, Kenny E, Kim Y, Kirov GK, Konnerth H, Konte B, Krabbendam L, Krasucki R, Lasseter VK, Laurent C, Lawrence J, Lencz T, Lerer FB, Liang KY, Lichtenstein P, Lieberman JA, Linszen DH, Lönnqvist J, Loughland CM, MacLean AW, Maher BS, Maier W, Mallet J, Malloy P, Mattheisen M, Mattingsdal M, McGhee KA, McGrath JJ, McIntosh A, McLean DE, McQuillin A, Melle I, Michie PT, Milanova V, Morris DW, Mors O, Mortensen PB, Moskvina V, Muglia P, Myin-Germeys I, Nertney DA, Nestadt G, Nielsen J, Nikolov I, Nordentoft M, Norton N, Nöthen MM, O'Dushlaine CT, Olincy A, Olsen L, O'Neill FA, Ørntoft TF, Owen MJ, Pantelis C, Papadimitriou G, Pato MT, Peltonen L, Petursson H, Pickard B, Pimm J, Pulver AE, Puri V, Quested D, Quinn EM, Rasmussen HB, Réthelyi JM, Ribble R, Rietschel M, Riley BP, Ruggeri M, Schall U, Schulze TG, Schwab SG, Scott RJ, Shi J, Sigurdsson E, Silverman JM, Spencer CCA, Stefansson K, Strange A, Strengman E, Stroup TS, Suvisaari J, Terenius L, Thirumalai S, Thygesen JH, Timm S, Toncheva D, Van Den Oord E, Van Os J, Van Winkel R, Veldink J, Walsh D, Wang AG, Wiersma D, Wildenauer DB, Williams HJ, Williams NM, Wormley B, Zammit S, Sullivan PF, O'Donovan MC, Daly MJ, Gejman PV. Genome-wide association study identifies five new schizophrenia loci. Nature Genetics 2011; 43: 969-978.

Robbins TW, Sahakian BJ. Computer Methods of Assessment of Cognitive Function. Principles and Practice of Geriatric Psychiatry. John Wiley & Sons, Ltd, 2002, pp. 147-151.

Roede JR, Uppal K, Park Y, Tran V, Jones DP. Transcriptome–metabolome wide association study (TMWAS) of maneb and paraquat neurotoxicity reveals network level interactions in toxicologic mechanism. Toxicology reports 2014; 1: 435-444.

Salek RM, Maguire ML, Bentley E, Rubtsov DV, Hough T, Cheeseman M, Nunez D, Sweatman BC, Haselden JN, Cox RD, Connor SC, Griffin JL. A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. Physiological Genomics 2007; 29: 99-108.

Sato Y, Suzuki I, Nakamura T, Bernier F, Aoshima K, Oda Y. Identification of a new plasma biomarker of Alzheimer's disease using metabolomics technology. Journal of Lipid Research 2012; 53: 567-576.

Schmutz J, Wheeler J, Grimwood J, Dickson M, Yang J, Caoile C, Bajorek E, Black S, Chan YM, Denys M, Escobar J, Flowers D, Fotopulos D, Garcia C, Gomez M, Gonzales E, Haydu L, Lopez F, Ramirez L, Retterer J, Rodriguez A, Rogers S, Salazar A, Tsai M, Myers RM. Quality assessment of the human genome sequence. Nature 2004; 429: 365-368.

Sha L, MacIntyre L, Machell JA, Kelly MP, Porteous DJ, Brandon NJ, Muir WJ, Blackwood DH, Watson DG, Clapcote SJ, Pickard BS. Transcriptional regulation of neurodevelopmental and metabolic pathways by NPAS3. Mol Psychiatry 2012; 17: 267-79.

Stein J, Responsibility GBPfS, Science, Network EH. Environmental Threats to Healthy Aging: With a Closer Look at Alzheimer's & Parkinson's Diseases: Greater Boston Physicians for Social Responsibilty, 2008.

Stein TP, Schluter MD, Steer RA, Guo L, Ming X. Bisphenol A Exposure in Children With Autism Spectrum Disorders. Autism Res 2015; 8: 272-83.

Stephenson ST, Brown LAS, Helms MN, Qu H, Brown SD, Brown MR, Fitzpatrick AM. Cysteine oxidation impairs systemic glucocorticoid responsiveness in children with difficult-to-treat asthma. Journal of Allergy and Clinical Immunology 2015; 136: 454-461.e9.

Story EN, Kopec RE, Schwartz SJ, Harris GK. An update on the health effects of tomato lycopene. Annu Rev Food Sci Technol 2010; 1: 189-210.

Suhre K, Shin S-Y, Petersen A-K, Mohney RP, Meredith D, Wägele B, Altmaier E, CardioGram, Deloukas P, Erdmann J, Grundberg E, Hammond CJ, de Angelis MH, Kastenmüller G, Köttgen A, Kronenberg F, Mangino M, Meisinger C, Meitinger T, Mewes H-W, Milburn MV, Prehn C, Raffler J, Ried JS, Römisch-Margl W, Samani NJ, Small KS, Wichmann HE, Zhai G, Illig T, Spector TD, Adamski J, Soranzo N, Gieger C. Human metabolic individuality in biomedical and pharmaceutical research. Nature 2011; 477: 10.1038/nature10354.

Trushina E, Mielke MM. Recent advances in the application of metabolomics to Alzheimer's Disease. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 2014; 1842: 1232-1239.

Uezu A, Kanak DJ, Bradshaw TW, Soderblom EJ, Catavero CM, Burette AC, Weinberg RJ, Soderling SH. Identification of an elaborate complex mediating postsynaptic inhibition. Science 2016; 353: 1123-9.

Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, Mill J, Cantor R, Blencowe BJ, Geschwind DH. Transcriptomic Analysis of Autistic Brain Reveals Convergent Molecular Pathology. Nature 2011; 474: 380-384.

Wang H, Liang S, Wang M, Gao J, Sun C, Wang J, Xia W, Wu S, Sumner SJ, Zhang F, Sun C, Wu L. Potential serum biomarkers from a metabolomics study of autism. J Psychiatry Neurosci 2016; 41: 27-37.

Warnick EM, Bracken MB, Kasl S. Screening Efficiency of the Child Behavior Checklist and Strengths and Difficulties Questionnaire: A Systematic Review. Child and Adolescent Mental Health 2008; 13: 140-147.

Wechsler D. Wechsler Intelligence Scale for Children. Psychological Corporation, San Antonio TX, 2003.

West PR, Amaral DG, Bais P, Smith AM, Egnash LA, Ross ME, Palmer JA, Fontaine BR, Conard KR, Corbett BA, Cezar GG, Donley EL, Burrier RE. Metabolomics as a tool for discovery of biomarkers of autism spectrum disorder in the blood plasma of children. PLoS One 2014; 9: e112445.

Wild CP. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiology Biomarkers and Prevention 2005; 14: 1847-1850.

Winnier DA, Fourcaudot M, Norton L, Abdul-Ghani MA, Hu SL, Farook VS, Coletta DK, Kumar S, Puppala S, Chittoor G, Dyer TD, Arya R, Carless M, Lehman DM, Curran JE, Cromack DT, Tripathy D, Blangero J, Duggirala R, Göring HHH, DeFronzo RA, Jenkinson CP. Transcriptomic Identification of <italic>ADH1B</italic> as a Novel Candidate Gene for Obesity and Insulin Resistance in Human Adipose Tissue in Mexican Americans from the Veterans Administration Genetic Epidemiology Study (VAGES). PLoS ONE 2015; 10: e0119941.

Wisnowski JL, Bluml S, Paquette L, Zelinski E, Nelson MD, Jr., Painter MJ, Damasio H, Gilles F, Panigrahy A. Altered glutamatergic metabolism associated with punctate white matter lesions in preterm infants. PLoS One 2013; 8: e56880.

Wu YE, Parikshak NN, Belgard TG, Geschwind DH. Genome-wide, integrative analysis implicates microRNA dysregulation in autism spectrum disorder. Nat Neurosci 2016; advance online publication.

Xia M, Viera-Hutchins L, Garcia-Lloret M, Noval Rivas M, Wise P, McGhee SA, Chatila ZK, Daher N, Sioutas C, Chatila TA. Vehicular exhaust particles promote allergic airway inflammation through an aryl hydrocarbon receptor-notch signaling cascade. Journal of Allergy and Clinical Immunology 2015a; 136: 441-453.

Xia M, Viera-Hutchins L, Garcia-Lloret M, Noval Rivas M, Wise P, McGhee SA, Chatila ZK, Daher N, Sioutas C, Chatila TA. Vehicular exhaust particles promote allergic airway inflammation through an aryl hydrocarbon receptor–notch signaling cascade. Journal of Allergy and Clinical Immunology 2015b; 136: 441-453.

Xu F, Tavintharan S, Sum CF, Woon K, Lim SC, Ong CN. Metabolic Signature Shift in Type 2 Diabetes Mellitus Revealed by Mass Spectrometry-based Metabolomics. The Journal of Clinical Endocrinology & Metabolism 2013; 98: E1060-E1065.

Yoshimi N, Futamura T, Kakumoto K, Salehi AM, Sellgren CM, Holmén-Larsson J, Jakobsson J, Pålsson E, Landén M, Hashimoto K. Blood metabolomics analysis identifies abnormalities in the citric acid cycle, urea cycle, and amino acid metabolism in bipolar disorder. BBA Clinical 2016; 5: 151-158.

Yu M, Cui F-X, Jia H-M, Zhou C, Yang Y, Zhang H-W, Ding G, Zou Z-M. Aberrant purine metabolism in allergic asthma revealed by plasma metabolomics. Journal of Pharmaceutical and Biomedical Analysis 2016; 120: 181-189.

Zhang A-h, Qiu S, Xu H-y, Sun H, Wang X-j. Metabolomics in diabetes. Clinica Chimica Acta 2014; 429: 106-110.

Zhang A, Sun H, Wang X. Power of metabolomics in biomarker discovery and mining mechanisms of obesity. Obes Rev 2013; 14: 344-9.

Zhao X, Fritsche J, Wang J, Chen J, Rittig K, Schmitt-Kopplin P, Fritsche A, Häring H-U, Schleicher ED, Xu G, Lehmann R. Metabonomic fingerprints of fasting plasma and spot urine reveal human pre-diabetic metabolic traits. Metabolomics 2010; 6: 362-374.

Zhu C, Liang Q-l, Hu P, Wang Y-m, Luo G-a. Phospholipidomic identification of potential plasma biomarkers associated with type 2 diabetes mellitus and diabetic nephropathy. Talanta 2011; 85: 1711-1720.