Environment International 98 (2017) 212–218

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

Genetic and methylation variation in the CYP2B6 gene is related to circulating *p,p′*-dde levels in a population-based sample

Lars Lind [a,*,1], Esther Ng [b,1], Erik Ingelsson [c,d], Cecilia Lindgren [b,e], Samira Salihovic [c,f], Bert van Bavel [f], Anubha Mahajan [b], Erik Lampa [a], Andrew P. Morris [b,g], P. Monica Lind [h]

[a] Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden
[b] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[c] Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[d] Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA
[e] Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, USA
[f] MTM Research Centre, School of Science and Technology, Örebro University, Sweden
[g] Department of Biostatistics, University of Liverpool, Liverpool, UK
[h] Department of Medical Sciences, Occupational and Environmental Medicine, Uppsala University, Sweden

ARTICLE INFO

Article history:
Received 1 July 2016
Received in revised form 7 November 2016
Accepted 8 November 2016
Available online 11 November 2016

Keywords:
GWAS
CYP2B6
DDE
Metabolism
Methylation

ABSTRACT

Objectives: Since the metabolism of the organochlorine pesticide dichlorodiphenyltrichloroethane (DDT) is not fully known in humans, we evaluated if circulating levels of a major breakdown product of DDT, *p,p′*-DDE, were related to genome-wide genetic and methylation variation in a population-based sample.
Methods: In the population-based Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study (1016 subjects all aged 70), circulating levels of *p,p′*-DDE were analyzed by high-resolution chromatography coupled to high-resolution mass spectrometry (HRGC/HRMS). Genetic variants were genotyped and imputed (1000 Genomes reference, March 2012 release). Methylation sites were assayed using the Illumina HumanMethylation450 array in whole blood. A genome-wide association study (GWAS) approach was applied.
Results: Evidence for genome-wide significant association with *p,p′*-DDE levels was observed only for a locus at chromosome 19 corresponding to the *CYP2B6* gene (lead SNP rs7260538). Subjects being homozygote for the G allele showed a median level of 472 ng/g lipid, while the corresponding level for those being homozygote for the T allele was 192 ng/g lipid ($p = 1.5 \times 10^{-31}$). An analysis conditioned on the lead SNP disclosed a distinct signal in the same gene (rs7255374, position chr19:41520351; $p = 2.2 \times 10^{-8}$).
A whole-genome methylation analysis showed one significant relationship vs. *p,p′*-DDE levels ($p = 6.2 \times 10^{-9}$) located 7 kb downstream the *CYP2B6* gene (cg27089200, position chr19:41531976). This CpG-site was also related to the lead SNP ($p = 3.8 \times 10^{-35}$), but mediated only 4% of the effect of the lead SNP on *p,p′*-DDE levels.
Conclusion: Circulating levels of *p,p′*-DDE were related to genetic variation in the *CYP2B6* gene in the general elderly population. DNA methylation in this gene is not closely linked to the *p,p′*-DDE levels.

1. Introduction

Dichlorodiphenyltrichloroethane (DDT) is an insecticide heavily used since the second half of World War II. Due to reproductive problems observed in wild animals, DDT was banned in the 1970s and 1980s in most high-income countries. Despite its toxic properties, DDT is still used in many developing countries, mainly to fight malaria (Eskenazi et al., 2009). A major metabolite of DDT is 2, 2-*bis* (4-chlorophenyl)-1, 1-dichloroethene (*p,p′*-DDE). *p,p′*-DDE is highly lipophilic and thereby accumulates in adipose tissue and has been estimated to have a half-life of 10–15 years. Although DDT is not used as a pesticide any longer in high-income countries, due to its persistence in the environment and accumulation in the food chain in fish and meat, there is still exposure taking place. However, this exposure is not as high as before the ban. An accumulation occurs with ageing, so higher levels of DDT/*p,p′*-DDE are seen in older subjects than in younger individuals as a result from a continuous cumulative exposure (Ye et al., 2015). We have recently reported that high levels of *p,p′*-DDE are related to prevalent obesity, diabetes and hypertension (Lee et al., 2011; Lind et al., 2014b; Ronn et al., 2011).

* Corresponding author at: Department of Medical Sciences, Occupational and Environmental Medicine, Uppsala University, 751 85 Uppsala, Sweden.
E-mail addresses: lars.lind@medsci.uu.se (L. Lind), elengss@gmail.com (E. Ng), erik.ingeslsson@medsci.uu.se (E. Ingelsson), celi@broadinstitute.org (C. Lindgren), samira.salihovic@medsci.uu.se (S. Salihovic), Bert.VanBavel@oru.se (B. van Bavel), anubha@well.ox.ac.uk (A. Mahajan), erik.lampa@ucr.uu.se (E. Lampa), amorris@well.ox.ac.uk (A.P. Morris), monica.lind@medsci.uu.se (P.M. Lind).
[1] Contributed equally to this work.

http://dx.doi.org/10.1016/j.envint.2016.11.010

Several studies have shown that DDT alters the activity of many microsomal enzymes, including those involved in phase I and phase II metabolism of xenobiotics (Lubet et al., 1992; Madhukar and Matsumura, 1979) in different species (Abernathy et al., 1971; Bunyan et al., 1972; Henneman et al., 1994; Li et al., 1995; Lubet et al., 1990). Pharmacodynamic studies on CYP2B induction indicated no important differences between the isomer *p,p′*-DDT and its metabolite DDE (Nims et al., 1998). The effects consisted mainly of an induction of the CYP2B subfamily, a lesser induction on CYP3A, and minimal or no induction of CYP1A. On this basis, DDT has been considered a phenobarbiturate-type of inducer (Nims et al., 1998; Okey, 1972).

Toxicokinetic studies of persistent organic pollutants, such as DDT/DDE, have mainly been conducted in an experimental setting (as cited above), and therefore such results have to be validated in humans. One way to perform such studies in humans is to relate functional genetic variations in genes known to be involved in the kinetics of environmental contaminants to levels of the contaminant of interest, like relating levels of polychlorinated biphenyls to single nuclear polymorphisms (SNP) in the CYP1A1 gene (Lind et al., 2014a). However, such an approach demands a detailed prior knowledge of the metabolism of the compound of interest. Another approach is to use a genome-wide association study (GWAS), which test a great number of SNP across the genome without a prior hypothesis. We have previously used the GWAS approach and found that circulating levels of several of the PCBs were related to variation in the CYP2B6 gene (Ng et al., 2015b), and whole blood manganese levels to be related to variation in the SLC39A8 and SLC30A10 genes, while mercury and cadmium levels shared associations with variation in other genes (Ng et al., 2015a).

The expression of a protein is not only governed by the variation in base-pairs in the genes, but also by epigenetic mechanisms, like methylation. For example, it has been shown that *p,p′*-DDE levels are linked to alterations in global DNA methylation (Rusiecki et al., 2008).

Since the metabolism of DDT and *p,p′*-DDE has mainly been studied in the experimental setting, we used the GWAS approach to relate genetic variation to circulating levels of *p,p′*-DDE. For this purpose, we used data from the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study (Lind et al., 2005), in which extensive genotyping has been performed, together with measurements of circulating *p,p′*-DDE levels. We also investigated if *p,p′*-DDE levels were related to differential methylation using a whole-genome approach.

2. Material and methods

2.1. Subjects

The PIVUS study was originally designed to study markers of subclinical cardiovascular disease as risk factors for incident cardiovascular diseases (Lind et al., 2005). Eligible subjects were all aged 70 and lived in the community of Uppsala, Sweden, with a total population of approximately 175,000 individuals. The subjects were randomly chosen from the register of community living kept by the City council of Uppsala. Of a total of 2027 invited individuals (50% being females), 1016 subjects participated, giving a participation rate of 50.1%. The study was approved by the Ethics Committee of the University of Uppsala and all the participants gave their informed consent prior to the study.

All subjects were investigated in the morning after an overnight fast. No medication or smoking was allowed after midnight. The participants were asked to answer a questionnaire about their medical history, smoking habits and regular medication. Blood samples for determinations of *p,p′*-DDE levels, genotyping and DNA methylation were drawn at the same time in the fasting state at the age of 70 years. *p,p′*-DDE levels were measured in serum and DNA was prepared from leukocytes in whole blood specimens.

2.1. p,p′-DDE analyses

p,p′-DDE levels were measured in stored serum samples using a Micromass Autospec Ultima (Waters, Milford, MA, USA) high-resolution gas chromatography coupled to a high resolution mass spectrometry (HRGC/HRMS) system based on the method by Sandau and colleagues (Barr et al., 2003) with some modifications. A more detailed description of the analysis in this sample has previously been presented (Salihovic et al., 2012).

2.2. Genotyping and imputation

Genotyping was performed on all participants using the Illumina Metabochip together with the Illumina OmniExpress chip. Samples were excluded based on call rate <95%, extreme heterozygosity (>3 SD from the mean), gender discordance, duplicated samples, close relatives or ethnic outliers. Variants with exact Hardy-Weinberg equilibrium (HWE) p-value $< 1 \times 10^{-6}$, call rate < 0.99, SNPs with minor allele frequency [MAF] <5% were excluded from the scaffold prior to imputation. The cleaned genotype data were imputed up to the 1000 Genomes, March 2012 release reference panel (multi-ethnic panel on NCBI build 37 [b37]) using IMPUTE v.2.2.2.

2.3. Regional DNA methylation

Methylation sites across the genome were assayed using the Illumina HumanMethylation450k Beadchip, which detects methylation based on genotyping of bisulfite-converted genomic DNA, covering 482,421 CpG-sites and 3091 non-CpG sites. Samples were excluded based on call rate (98.5% probes with detection p-value < 0.01), leukocyte count > 10 ($\times 10^9$ cells/L), bisulfite conversion efficiency outliers, or more than one mismatch when comparing the SNPs on the methylation chip and the Omni/Metabochip genotyping chips. Data on the X and Y chromosomes were not used in the analysis. A quantile normalization of the signal intensities was performed per individual and undertaken separately for type-I and type-II probes of the chip. Beta-values were then calculated as the percentage methylation at a site, denoted degree of methylation in the text. A total of 20,522 methylation sites were excluded from the analysis since their probes mapped to multiple locations in the genome with at least two mismatches, in accordance with methods used by other investigators (Grundberg et al., 2013).

2.4. Statistical analyses

Since not normally distributed, *p,p′*-DDE was natural log-transformed before analysis, and used in the further analysis as a continuous variable. Since no priori hypotheses regarding which genes that could be related to *p,p′*-DDE levels were given in the present study, a GWAS with *p,p′*-DDE as dependent variable was performed using the score-based test in SNPTEST 2.4.1 (Marchini et al., 2007), accounting for imputation uncertainty in a missing data likelihood and assuming an additive genetic effect. Only SNPs with MAF >0.05 and IMPUTE2 info >0.4 were included in the analysis. Gender and two principal components (based on the genetic structure in the sample, here used to adjust for any genetic heterogeneity within the sample) were included as covariates and a p-value $< \times 10^{-8}$ defined a genome-wide significant finding.

The genomic inflation factor lambda (Devlin and Roeder, 1999) was calculated as a quality control to assess the evidence for residual population structure that was not accounted for in the association analysis. Even studies with relatively homogeneous populations are susceptible to residual confounding by population stratification.

The associations between SNPs within and around *CYP2B6* (10 kb up and downstream of the transcript boundaries on b37 obtained by the UCSC Table Browser [http://genome.ucsc.edu/cgi-bin/hgTables]) were visualized using the program LocusZoom 1.1 (Pruim et al., 2010). Following identification of the most significantly associated SNP in the