ENVIRONMENTAL EPIGENETICS

Environmental Epigenetics, 2019, 1–10

doi: 10.1093/eep/dvz014
Research article

RESEARCH ARTICLE

Cadmium exposure and *MEG3* methylation differences between Whites and African Americans in the NEST Cohort

John S. House [1,2], Jonathan Hall[1,3], Sarah S. Park[1,3], Antonio Planchart[1,3], Eric Money[4,5], Rachel L. Maguire[1,3], Zhiqing Huang[6], Carolyn J. Mattingly[1,3], David Skaar[1,3], Jung Ying Tzeng[1,7], Thomas H. Darrah[8], Avner Vengosh[9], Susan K. Murphy[6], Randy L. Jirtle[1,3] and Cathrine Hoyo[1,3,*]

[1]Center for Human Health and the Environment, North Carolina State University, Raleigh, NC, USA; [2]Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA; [3]Department of Biological Sciences, North Carolina State University, Raleigh, NC, USA; [4]Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA; [5]Department of Natural Resources, North Carolina State University, Raleigh, NC, USA; [6]Department of Obstetrics and Gynecology, Duke University School of Medicine, Durham, NC, USA; [7]Department of Statistics, North Carolina State University, Raleigh, NC, USA; [8]Division of Climate, Water, and Environment, School of Earth Sciences, The Ohio State University, Columbus, OH, USA; [9]Nicholas School of the Environment, Duke University, Durham, NC, USA

*Correspondence address. Center for Human Health and the Environment, North Carolina State University, Raleigh, NC, USA and Department of Biological Sciences, North Carolina State University, Raleigh, NC, USA. Tel: 919 515 0540; Fax: 919-515-7169; E-mail: choyo@ncsu.edu

Managing Editor: Mike Skinner

Abstract

Cadmium (Cd) is a ubiquitous environmental pollutant associated with a wide range of health outcomes including cancer. However, obscure exposure sources often hinder prevention efforts. Further, although epigenetic mechanisms are suspected to link these associations, gene sequence regions targeted by Cd are unclear. Aberrant methylation of a differentially methylated region (DMR) on the *MEG3* gene that regulates the expression of a cluster of genes including *MEG3, DLK1, MEG8, MEG9* and *DIO3* has been associated with multiple cancers. In 287 infant–mother pairs, we used a combination of linear regression and the Getis-Ord Gi* statistic to determine if maternal blood Cd concentrations were associated with offspring CpG methylation of the sequence region regulating a cluster of imprinted genes including *MEG3*. Correlations were used to examine potential sources and routes. We observed a significant geographic co-clustering of elevated prenatal Cd levels and *MEG3* DMR hypermethylation in cord blood ($P = 0.01$), and these findings were substantiated in our statistical models ($\beta = 1.70$, se $= 0.80$, $P = 0.03$). These associations were strongest in those born to African American women ($\beta = 3.52$, se $= 1.32$, $P = 0.01$) compared with those born to White women ($\beta = 1.24$, se $= 2.11$, $P = 0.56$) or Hispanic women ($\beta = 1.18$, se $= 1.24$,

Received 22 March 2019; **revised** 21 June 2019; **accepted** 19 July 2019

$P = 0.34$). Consistent with Cd bioaccumulation during the life course, blood Cd levels increased with age ($\beta = 0.015$ μg/dl/year, $P = 0.003$), and Cd concentrations were significantly correlated between blood and urine ($\rho > 0.47$, $P < 0.01$), but not hand wipe, soil or house dust concentrations ($P > 0.05$). Together, these data support that prenatal Cd exposure is associated with aberrant methylation of the imprint regulatory element for the *MEG3* gene cluster at birth. However, neither house-dust nor water are likely exposure sources, and ingestion via contaminated hands is also unlikely to be a significant exposure route in this population. Larger studies are required to identify routes and sources of exposure.

Key words: maternal exposures; imprinted genes; offspring epigenetics; cadmium; CpG methylation; racial epigenetic outcome differences

Introduction

Due to large industrial activity in the last two centuries, cadmium (Cd) is now a ubiquitous environmental contaminant that ranks in the top 10 chemicals of concern [1]. Once in the body, toxicity is amplified by bioaccumulation due to slow excretion rates [2]. This leads to its accumulation in target tissues such as the liver and pancreas, and in long-term storage compartments, including the kidney, neural tissues and bone [3]. The half-life of Cd in humans is estimated to be between 10 and 30 years [4, 5]. Emerging data suggest that ethnic minorities may be at higher risk of non-occupational exposure [6, 7], and co-exposure to other toxic metals, including arsenic (As) and lead (Pb) may interact with Cd to exacerbate health effects [1]. Although cigarette smoke is the main source of non-occupational exposure, environmental health agencies also point to contaminated dietary staples, water and pollution-contaminated air from which Cd can be ingested or inhaled. Recent data based on residential addresses of pregnant women also revealed geographic clustering of Cd and other toxic metals in an inner-city neighborhood, suggesting contaminated soils of some urban landscapes may be a source [6]. However, empirical data linking sources and routes to the body burden are limited.

Reported health effects of high occupational levels of Cd are consistent with target tissue injury (e.g. nephrotoxicity, neurotoxicity and carcinogenicity). In contrast, effects of exposure to lower non-occupational Cd and co-occurring toxic metals has been difficult to discern. The International Agency of Research on Cancer (IARC) has classified Cd as a probable carcinogen based on years of *in vitro* and *in vivo* evidence and inconsistent human data. However, in a recent comprehensive review, only half of the 57 human studies found that Cd exposure at levels experienced by the general population increased the risk of components of metabolic syndrome, an established risk factor for multiple cancers [8]. These inconsistencies may partially be due to an inability of prior studies to measure and account for co-occurring metals. Recent advances in inductively coupled plasma (ICP)-mass spectrometry (MS) have enabled the simultaneous measurement of multiple metals in smaller specimen quantities, characteristic in otherwise healthy humans [9].

The paucity of mechanistic data linking molecular intermediates to Cd exposure and adverse health outcomes in humans has also limited our understanding of the effects of lower Cd doses, common in the general population, on health outcomes. Such mechanistic insights have the potential to inform early intervention strategies. Epigenetic regulation, a means by which gene expression is altered in response to environmental cues, including exposure to Cd during the prenatal period, provides flexibility in adapting to changing environments [10]. However, the sequence regions and signaling pathways affected by Cd exposure have been difficult to replicate. In part, this results from the majority of available data being measured using methylation arrays and other targeted regions for which the temporal stability of DNA methylation alterations are unknown. This temporal ambiguity between Cd exposure and methylation alterations complicates causal inference, especially for adult-onset diseases with suspected fetal origins, such as cancer. In contrast, methylation of CpG dinucleotides at regions that control genomic imprinting is differentially established in a sex-dependent manner in gametes and resists the wave of global demethylation that occurs after fertilization. As such, methylation marks of imprint regulatory sequences are temporally stable and have been proposed to serve as an archive of early exposures [11–14]. Recent research using whole-genome bisulfite sequencing with 12× to 15× coverage suggests that imprinted gene control regions are over-represented in differentially methylated regions (DMRs) of offspring and mothers exposed to Cd [15]. However, these findings have not been validated using more quantitatively accurate methods such as pyrosequencing.

Among the most studied cancer-associated imprinted gene, DMRs are those regulating the Delta-like 1 (*DLK1*)-Maternally expressed gene 3 (*MEG3*) domain at chromosome 14q32.2. This imprinted domain includes *MEG3*-IG (methylation established on the paternal allele in sperm) and *MEG3* (methylation established on the paternal allele post-fertilization) DMRs. These DMRs regulate the imprinting and expression of a gene cluster that includes paternally expressed *DLK1* and maternally expressed *MEG3* and *MEG8* (Fig. 1). The *MEG3*-IG DMR is located intergenically ~13kb upstream of the *MEG3* transcription start site, while the *MEG3* DMR is intragenic and overlaps with the *MEG3* promoter. While the *MEG3*-IG is the main imprint control region that functions in the placenta to control proper allelic expression of transcripts in this region, the *MEG3* DMR maintains the active (unmethylated) status at the IG-DMR on the maternal allele and functions in somatic tissues, allowing for expression of downstream genes. Hypermethylation of the maternally imprinted *MEG3* DMR has been associated with multiple forms of cancer, many in organs targeted by Cd (reviewed in [16, 17]); however, an association between Cd and altered methylation of the *MEG3* DMR has not been described.

In these analyses, we evaluate whether early prenatal Cd exposure associates with differential methylation of the *MEG3* DMR control region in offspring. To inform public health policy decisions, we also assessed correlations between environmental routes of exposure including house dust, soil and water with Cd concentrations found in individuals in a geographic cluster with high Cd exposure.

Results

The characteristics of $n = 287$ study participants included in these analyses are summarized in Table 1. African American

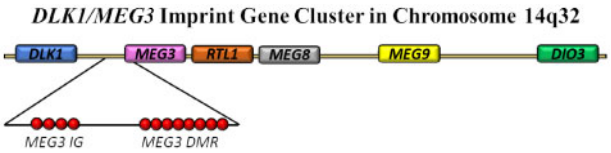

Figure 1: *DLK1/MEG3* imprint gene cluster in chromosome 14q32. Representation of the imprinted control regions for *DLK1/MEG3* locus on chromosome 14q32. Red circles represent where assessed methylation percentages at four CpGs in the intergenic DMR (*MEG3 IG*) and eight CpGs were obtained in the *MEG3* DMR

Table 1: characteristics of the 287 study participants

Characteristic	Frequency (%) or (sd)
Maternal	
Maternal age (years)	28.2 (sd = 5.8)
Race/ethnicity	
Black/African American	103 (35.9%)
White	85 (29.6%)
Hispanic	88 (30.7)
Other	11 (3.8%)
Parity	
None	96 (33.5%)
One	102 (35.5%)
Two or more	89 (31.0%)
Maternal pre-pregnancy BMI (kg/m^2)	27.33 (sd = 7.2)
Smoking during pregnancy	
Yes	41 (14.6%)
No	239 (85.4%)
Offspring	
Weeks of gestation	38 (sd = 2.0)
Males (%)	144 (50.2%)
Birth weight (g)	3264 (sd = 613.0)
Metals (Min, 25%, 50%, 75%, Max)	
Cadmium (ng/g)	0.000, 0.101, 0.226, **0.513**, 3.951
Lead (μg/dl)	0.038, 0.167, 0.358, 0.831, 259.4
Arsenic (μg/dl)	0.006, 0.039, 0.044, 0.051, 0.437
Methylation% (Min, 25%, 50%, 75%, Max)	
***MEG3* [mean (sd) = 72.86 (5.50)]**	56.60, 69.01, 72.26, 77.09, 85.21

Cadmium was dichotomized at the cut-point between third and fourth quartiles and is indicated in bold.

women comprised 36% of the participants, while Hispanics and Whites comprised ~30% each. Mean pre-pregnancy body mass index (BMI) was 27 kg/m^2 (sd = 7.2) and ranged from 16 to 67 kg/m^2; however, it varied by ethnicity—obese women (i.e. BMI > 30 kg/m^2) were more likely to be African American (52%) than White (19%). Cigarette smoking during pregnancy also varied by ethnicity as 29% of African Americans, 10% of Whites and 4% of Hispanics reported smoking. Gestational age at birth averaged 38.5 weeks, and mean birth weight was 3264 g (sd = 613 g). This was lower in African Americans. The mean age at delivery was 28 years (range, 18–43 years), and the majority of women (69%) had one or more children before the index pregnancy; these factors did not vary by ethnicity. From blood obtained during the first trimester (mean gestational age, 12 weeks), the geometric mean for maternal Cd concentrations (n = 287) was 0.24 ng/g [interquartile range (IQR) = 0.10–0.51], and that of Pb and As were 0.41 μg/dl (IQR = 0.17–0.83) and 0.05 μg/dl (IQR = 0.03–0.05), respectively (Table 1). The highest levels were in African Americans.

Table 2: maternal blood cadmium concentrations and offspring *MEG3* DMR hypermethylation

Characteristic[a]	β, se	P-value
Cadmium (ng/g) (Referent = lower 75%)	**1.70, 0.80**	**0.03**
Maternal pre-pregnancy obesity	0.40, 0.78	0.61
Maternal gestational smoking	−0.77, 1.01	0.45
Gestational age (weeks)	−0.14, 0.16	0.39
Lead (μg/dl)	−0.00, 0.02	0.79
Arsenic (μg/dl)	−2.13, 7.02	0.76
Sex (vs. males)	0.06, 0.66	0.93
Black/African American (vs. White)	**1.84, 0.89**	**0.04**
Hispanic (vs. White)	0.53, 1.00	0.60
Other (vs. White)	−1.37, 1.76	0.44
Education (Some college vs. no college)	−1.30, 0.78	0.10
Ethnicity/race stratified—cadmium (ng/g)[b]		
White (n = 83)	1.24, 2.11	0.56
Black/African American (n = 101)	**3.52, 1.32**	**0.01**
Hispanic (n = 83)	1.18, 1.24	0.34

[a]Full model multiple linear regression. Complete observations from the n = 287.
[b]Multiple linear regression adjusted for all covariates except race.

Associations between Cd Exposure and *MEG3* DMR Methylation

We measured the methylation of the eight CpG sites in the *MEG3* DMR (Fig. 1), and the mean methylation fraction for the *MEG3* DMR was 72.8% (sd = 5.5%). Regression models adjusted for co-exposure to Pb, As, pre-pregnancy maternal obesity, sex and gestational age at delivery and socioeconomic status and cigarette smoking showed that higher prenatal Cd concentrations [defined as upper quartile Cd exposed mothers (0.513 ng/g)] were associated with a 1.70% higher CpG methylation in all offspring at birth (β = 1.70, se = 0.80, P = 0.03) (Table 2). Additional adjustment for parity and maternal age at delivery did not alter these findings. In this fully adjusted model, being born to an African American mother was independently associated with higher methylation of the *MEG3* DMR (β = 1.84, se = 0.89, P = 0.04).

When stratified by ethnicity, associations between high maternal Cd exposure and *MEG3* DMR hypermethylation were most apparent in offspring of African Americans (β = 3.52, se = 1.32, P = 0.01) compared with Whites (β = 1.24, se = 2.11, P = 0.56) and Hispanics (β = 1.18, se = 1.24, P = 0.34) (Table 2). Tests for heterogeneity of regression coefficients for each ethnic group revealed that associations of Cd exposure and *MEG3* hypermethylation were significantly higher in African Americans compared to both Whites (P = 0.02) and Hispanics (P = 0.03). Scatterplots of mean *MEG3* methylation and maternal Cd exposure shown by race in Fig. 2 support that offspring of African American women had significantly higher methylation levels of the *MEG3* DMR. Because the skewed Cd distribution that was not improved by log-transformation necessitated Cd to be dichotomized at the top quartile, we considered models where Cd was dichotomized at the top tertile or quintile, to ensure that our findings were not unduly influenced by how Cd was dichotomized. Our findings that Cd exposure and being born to an African American mother were independently associated with higher MEG3 DMR, remained unchanged, and heterogeneity test to determine whether the associations differed by race remained unaltered (data not shown).

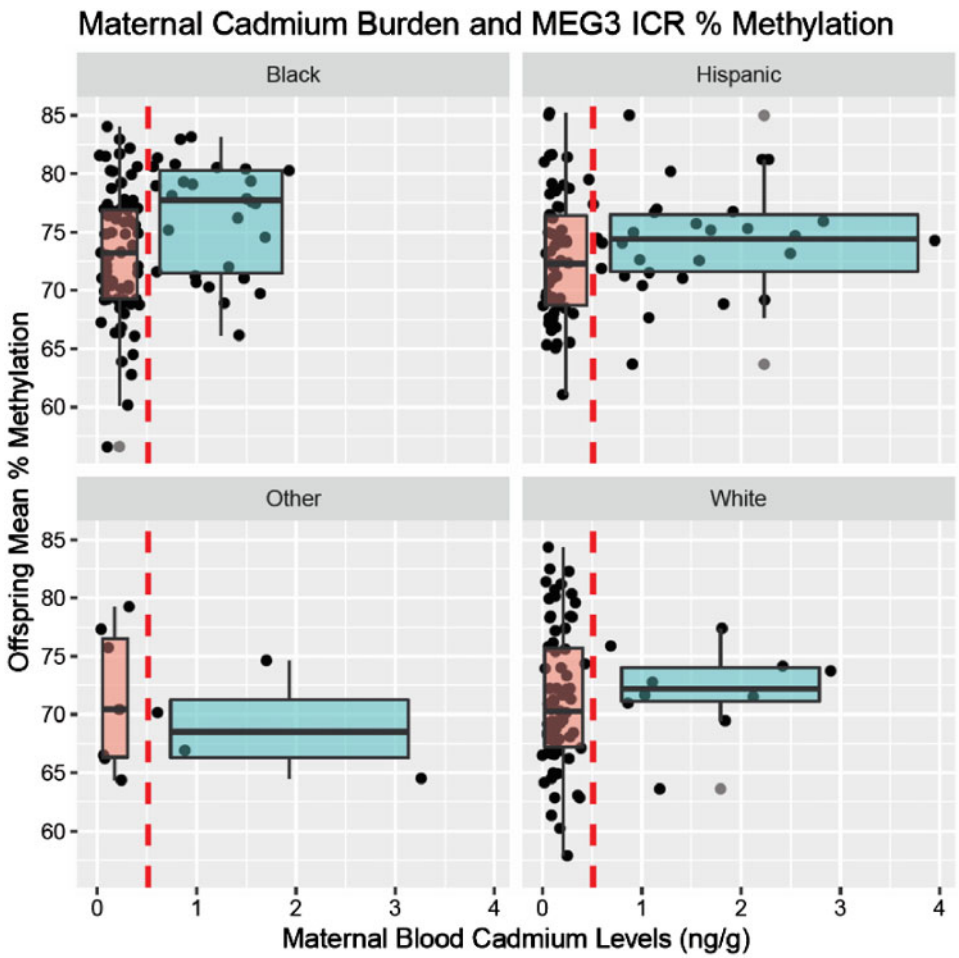

Figure 2: maternal cadmium exposure and offspring *MEG3* imprint control region (ICR) methylation. Mean offspring methylation of eight CpGs in the ICR of *MEG3* are plotted against maternal blood levels of cadmium and stratified by race. Dotted red line represents overall Q3/Q4 cut-point for high exposure (0.513 ng/g). Box and whisker plots represent the IQR and median for referent exposure (red box) and high exposure (blue)

Mediation by *MEG3* DMR Methylation of Associations between Cd and Lower Birth Weight

Despite limited statistical power, given previously reported inverse associations of maternal Cd exposure and birthweight [18], we conducted mediation analyses to assess whether *MEG3* methylation mediates this association. The direct effect of the highest quartile of maternal Cd burden on offspring birthweight was a loss of 174 g of birthweight [$\beta = -174$; 95% CI = (-310, -47.0); $P = 0.01$]. The tested mediation effect of *MEG3* methylation was not significant in the full model with all ethnic groups [$\beta = -10.9$; 95% CI = (-37.0, -6.39); $P = 0.23$], or when examined only for African Americans. In the fully adjusted model with all races, the proportion of this association explained by *MEG3* methylation was 5.16% ($P = 0.24$). When examined only on African American mothers, the proportion increased to 28.4%, but was still not significant ($P = 0.18$).

MEG3 DMR Methylation and Gene Expression Cluster

To determine the functional significance of the *MEG3* DMR hypermethylation that co-clustered with Cd and Pb exposure, we measured and correlated gene expression of three of the genes in the imprint gene cluster regulated by the *MEG3* DMR (i.e. *DLK1*, *MEG3* and *DIO3*) with DNA methylation in the cord blood of a subset of 22 offspring. We found negative correlations between both *MEG3* imprint control regions and the expression of *DIO3* and *DLK1* (Supplementary Table S3), but these results were not significant and difficult to interpret due to small sample size.

Geospatial Cluster of *MEG3* Hypermethylation

Figure 3 shows results of the geospatial cluster analysis (Getis-Ord Gi*) of offspring *MEG3* hypermethylation dichotomized at the top quartile (>77% methylation), together with maternal blood concentration for Cd and Pb based on street address during the prenatal period of gestation. Among the 287 offspring, 32% of African Americans were in the top quartile, compared with 21% of Whites and 22% of Hispanics. We found evidence for geospatial clustering of *MEG3* DMR hypermethylation with high confidence ($P < 0.001$). Intriguingly, this geographic clustering of *MEG3* DMR hyper-methylation coincided with elevated geographical Cd and Pb concentrations previously reported (Fig. 3) [6]. This spatial co-location of *MEG3* DMR hypermethylation with elevated concentrations of Cd in the population supports the correlation between aberrant methylation and exposure.

Potential Sources of Environmental Chemical Exposure

To identify potential sources and routes of toxic metal exposure, we also examined correlations between levels of Cd in

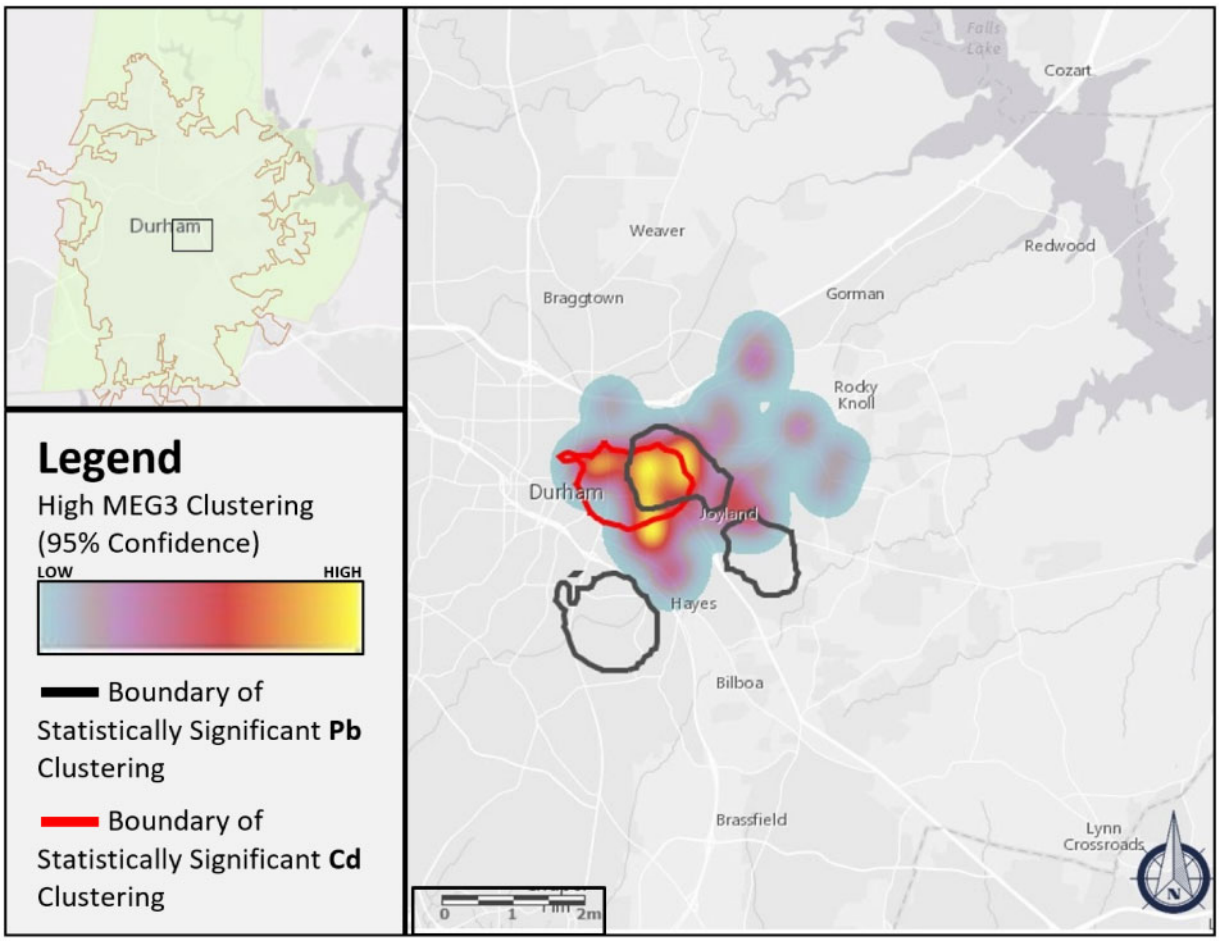

Figure 3: elevated blood cadmium levels in Durham County coincide with *DLK1/MEG3* hypermethylation. Geospatial clustering of offspring in the upper quartile of *MEG3* hypermethylation (yellow) are shown superimposed on the geographical clustering of those mothers with elevated levels of Cd and Pb in blood. Statistical significance is based on the Getis-Ord Gi* *P*-value and z-value. This statistic examines data points in proximity to one another to determine geographic areas where high *MEG3* methylation values are clustered together compared to the overall values across the study area. Statistical significance in this case means there is a 95% confidence that a cluster exists and is not due to random chance in the 'yellow' areas (*MEG3*) or within the 'black' boundary (Pb) and 'red' boundary (Cd) on the map

blood, urine, soil, tap water and house dust in an independent sample ($n = 37$) of randomly selected children and adult residents in the geographic cluster. The range of measurements are summarized in Supplementary Figure S1. We found that concentrations of Cd in blood and urine were positively correlated with each other ($\rho = 0.47$, $P < 0.01$), and Cd blood concentrations were associated with age, as expected ($\beta = 0.015$ µg/dl/year, $P = 0.003$ and Table 3). However, Cd levels in soil, house dust and water were not significantly correlated with urine or blood Cd concentrations (Table 3). Removing two outliers (Supplementary Figure S1) did not materially alter these correlations (Supplementary Table S1).

Discussion

Exposure to non-occupational levels of Cd may be carcinogenic, and other common trace metals may exacerbate the effects [19]; however, until recently, the effects of other toxic metals were not easily measurable in limited quantities of source material, characteristic of human specimens. Although epigenetic responses to Cd exposure are the hypothesized mechanisms of action, the affected genomic regions are still unclear. We evaluated whether prenatal Cd exposure was associated with hypermethylation of the *MEG3* DMR in offspring as similar patterns of *MEG3* DMR methylation have been linked to the development or progression of multiple cancers (reviewed in [20]). Our key findings were that after adjusting for co-occurring toxic metals, As and Pb and other prenatal factors previously shown to alter CpG methylation at birth, exposure to Cd *in utero* was associated with hypermethylation of the *MEG3* DMR and the magnitude of this effect was comparable to that of being born to an African American woman. These findings were corroborated by geospatial analysis demonstrating that hypermethylation of the *MEG3* DMR (defined as quartile 4 or >77% mean methylation) coincided geospatially with elevated Cd exposure. The *MEG3* DMR effects are unlikely to be influenced by co-exposure to As or Pb. To our knowledge, we present the first evidence linking ethnic-specific hypermethylation of a region that controls a large imprinted gene domain to non-occupational Cd exposure, a common environmental toxicant in humans.

We found striking ethnic differences in the association between Cd exposure and methylation of *MEG3*. While these findings may be artifactual, and should be confirmed in larger studies, other studies have demonstrated epigenetic variability for both ethnicity and environmental exposures by ethnicity [21, 22]. Methylation pattern alterations in sequence regions

Table 3: environmental and body burden of Cd concentrations

N	Blood	Urine	Soil	Water	Dust (W)	Dust (V)	Age
Blood	36	0.47 (<0.01)	−0.12 (0.49)	0.18 (0.32)	−0.04 (0.83)	0.17 (0.54)	0.58 (<0.01)
Urine		36	0.03 (0.85)	0.31 (0.08)	−0.06 (0.75)	0.42 (0.10)	0.16 (0.37)
Soil			35	−0.14 (0.43)	0.01 (0.94)	0.12 (0.66)	−0.09 (0.62)
Water				35	−0.06 (0.74)	0.14 (0.60)	0.08 (0.63)
Dust (W)					37	−0.36 (0.16)	−0.17 (0.31)
Dust (V)						17	0.26 (0.32)
Age							37

Spearman correlations (*P*-values) are shown for the levels of Cd measured in blood (ng/g), urine (ng/g), soil (μg/g), household water (μg/l), house dust wipe (W, μg/l), house dust vacuumed (V, μg/g) and age for youngest household occupant. Two measurements were removed due to value > mean + 5 standard deviations (see Supplementary Methods).

that regulate genomic imprinting in response to environmental exposures have been reported previously in humans. For example, both severe and moderate *in utero* caloric restriction during the 'Dutch Famine' and the Gambia [23, 24] was associated with hypomethylation of the imprint control region that regulates expression of the *IGF2/H19* imprinted gene cluster in offspring six decades after birth [25]. Pb exposure *in utero* and in early life is associated with hypomethylation of regions regulating the *IGF2/H19*, *PLAGL1/HYMA1* and *PEG3* imprinted gene clusters among 30-year-old adults [26]. As adults, these individuals displayed decreased brain volume [27] and increased risk of neurological disorders and juvenile delinquency [28, 29].

MEG3 was first identified in mice as Gene Trap Locus 2 (*Gtl2*). Imprinting of the human ortholog of *Gtl2*, now referred to as Maternally Expressed Gene 3, or *MEG3*, was first reported by our group [30]. *MEG3* encodes a long noncoding RNA (lncRNA) that is normally expressed in multiple tissues including placenta, brain, adrenal gland, pancreas, spleen and liver, and is reciprocally imprinted with paternally expressed *DLK1*. The expression of these genes is controlled by CpG methylation at the germline-derived intergenic differentially methylated DMR (*MEG3-IG*) located 13 kb upstream of *MEG3*, and at the *MEG3* DMR that overlaps the promoter (Fig. 1) [30]. In mice, deletion of the *MEG3-IG* DMR on the maternal chromosome results in an epigenotype switch, where the domain starts exhibiting the transcriptional profile of the paternal chromosome [31, 32]. Aberrant hyper-methylation of the *MEG3-IG* or *MEG3* DMRs is correlated with greatly decreased *MEG3* transcripts, which in turn is associated with increased expression of the reciprocally imprinted *DLK1*, cell proliferation *in vitro*, more aggressive and higher grade human tumors including meningioma, colon, nasopharyngeal, liver as well as leukemia [20]. Restoration of *MEG3* expression suppresses tumor growth [33]. Since one function of the *MEG3* DMR is to maintain an active (unmethylated) status in the *MEG3* IG region, hypermethylation of *MEG3* DMR in response to Cd has the potential to alter the expression of a large number of genes in this imprinted domain, including *MEG3* and other imprinted noncoding RNA genes such as *MEG8* and *MEG9*. These observations are consistent with our findings that hypermethylation of the *MEG3* DMR was correlated, albeit not significantly, with the expression of *DLK1*, *DIO3* and *MEG3* (Supplementary Table S3).

While mechanisms elucidating how *MEG3* DMR hypermethylation alters the trajectories of many cancers for the worse are still unclear, previous studies have indicated that *MEG3* functions as a tumor suppressor through activation of p53. p53 is a crucial tumor suppressor involved in preventing cancer through its response to cellular stressors, acting as a transcription factor and inducing the expression of genes involved in senescence, apoptosis and cell cycle arrest. *MEG3* has be shown to interact with the p53 DNA-binding domain at various p53 target genes [34]. Ectopic expression of *MEG3* in human cancer cell lines restores tumor suppressive activities and results in accumulation of p53 protein, activation of p53 downstream targets and inhibition of cell proliferation by inducing cell cycle arrest and apoptosis [33, 35]. It is therefore plausible that Cd-induced decreases in *MEG3* transcripts could result in loss of *MEG3* tumor suppressive function and increase susceptibility to cancer.

The weak correlation observed between *MEG3* DMR methylation and expression of the *MEG3* transcripts may be because in humans, the *MEG3* gene contains 10 exons and a number of RNA isoforms transcribed from human *MEG3* [36, 37]. All isoforms contain common exons 1, 2, 3, 8 and 10, but exhibit variation in expression of exons 4–7 and 9 [36]. The ability to regulate p53 and cell proliferation varies among *MEG3* isoforms and is dose dependent [34, 35], likely depending on the degree or pattern of methylation in this promoter region. Factors that influence *MEG3* splicing could affect the ability of *MEG3* RNAs to regulate p53 and subsequent cell cycle arrest, apoptosis and cancer suppression. Interestingly, 12 *MEG3* isoforms are present in human fetal liver, but only a few are present in adult tissue, suggesting a difference in regulation of splicing mechanisms between adult and fetal tissues [36]. Thus, understanding the relationship between Cd, *MEG3* DMR methylation and health outcomes will require an assessment of alternatively spliced *MEG3* transcription in addition to CpG methylation.

We attempted to elucidate what environmental exposures are driving elevated Cd levels with home visits to 35 individuals in the geographical region identified in Fig. 3. The body burden of Cd measured in blood increased in a linear fashion with age. We had expected to identify routes of environmental exposures for body burden via correlations between environmental sources and human body fluid but were unable to do so. However, we cannot exclude the possibility that the lack of significant correlations may be due to the limited number of samples we assessed with home visits. Large studies are needed to identify sources and routes of Cd exposure that can guide intervention efforts.

This study has several strengths. The longitudinal design, with Cd measured in first trimester blood specimens and *MEG3* DMR methylation measured in cord blood leukocytes enhances causal inference. Most Cd (~99%) attaches to erythrocytes which have a lifespan of ~120 days, such that first trimester measurements are likely to provide an accurate representation of the levels of Cd present during epigenetic reprogramming that occurs during the periconceptional period. We have implemented multiplexed measurements of Cd, Pb and As with the

use of ICP-MS, which enabled statistical adjustment for these co-occurring toxic metals in statistical analyses. Further, we utilized geospatial analyses to assess the overlap of geographical distribution of MEG3 DMR hypermethylation with environmental exposure to corroborate our statistical models. These analyses support that levels of Cd, and not those of Pb or As, associate with *MEG3* DMR hypermethylation. The multi-ethnic composition of our cohort enabled stratification by ethnicity, revealing considerably large effects in offspring of African Americans that were not apparent in Whites or Hispanics. While the significance of these ethnic differences is unclear, they suggest that, at a minimum, race-specific analyses are warranted in human epigenetic studies as well in the assessment of Cd exposure and imprinted gene methylation. The latter will require a concerted effort in characterizing the human imprint regulatory network of genes. Although this study solely focused on the regulatory region of *MEG3*, this region regulates a large number of imprinted genes, including *DLK1, MEG8, MEG9, MEG11* and *DIO3* and similar patterns of methylation have been associated with multiple cancers. The follow-up of the cohort is short, limited only to age five years for a small number of children. Thus, the associations between Cd exposure, *MEG3* DMR methylation and phenotypes that are likely to be maintained into adulthood have as yet to be shown in humans.

Despite these limitations, our geospatial data and linear regression analyses provide novel evidence for Cd, and not for Pb or As, associating with *MEG3* DMR hypermethylation, with the potential to alter gene expression for a large number of paternally and maternally expressed genes in the imprinted domain. Similar methylation in this regulatory region has been associated with the initiation of aggressiveness of multiple human tumors. Larger studies are needed to identify routes of exposure and to verify associations in order to inform public health intervention efforts, and to identify the array of genes and specific pathways perturbed by Cd via altered CpG methylation that elevate risk of chronic diseases including cancer.

Methods

Study Participants

Participants derived from the Newborn Epigenetics STudy (NEST), a cohort of pregnant women and their children whom, at enrollment, intended to receive obstetric care at one of two facilities in Durham County, NC, between 2009 and 2011. All participants provided informed consent. Accrual protocol details have been provided elsewhere [38]. Briefly, 1700 pregnant women were enrolled consecutively. Inclusion criteria were pregnancy, aged 18+ years or older and English or Spanish speaking. We excluded those planning to relinquish custody of the child and those whom, despite intent, did not deliver at one of the two facilities due to our inability to obtain umbilical cord blood. These analyses are restricted to 287 infant–mother pairs in whom at least Cd, As and Pb were measured in maternal blood (i.e. the first 310 enrolled) and for whom CpG methylation was measured in umbilical cord blood leukocytes. This study was approved by Duke University School of Medicine's Institutional Review Board.

A total of 36 randomly selected households in the geographic clusters, where the Cd geospatial cluster was identified, were re-contacted between December 2014 and February 2016, to confirm the body burdens of Cd, As and Pb and identify potential sources of exposure. Blood and urine samples were obtained from the youngest member of the household, for a total of 10 children under the age of 20 years and 26 aged 20–64 years. Hands and floor samples were obtained using hand and dust wipes on wooden floors and vacuums on carpeted floors. This study was approved by North Carolina State University Institutional Review Board.

Measurements

MEG3 DMR Methylation

Trained phlebotomists drew 10 ml of pregnant women's blood into ethylenediaminetetraacetic acid (EDTA)-containing vacutainer tubes at enrollment. DNA for pyrosequencing was extracted from umbilical cord blood buffy coat using Puregene reagents. Pyrosequencing was performed using a Qiagen Pyromark Q96 MD Pyrosequencer. Primers and PCR conditions have been previously described in detail [39, 40] as well as in Supplementary Methods. The methylation fraction for each CpG dinucleotide was calculated using PyroQ CpG Software (Qiagen). The methylation fraction was analysed at multiple CpG sites for the *MEG3* DMRs.

Blood Cd, Pb and As

From the 10 ml of peripheral blood obtained from pregnant women in EDTA-containing vacutainer tubes, 1 ml of whole blood was removed and stored. About 200 µl aliquots were used to measure Cd, Pb and As concentrations using a Perkin Elmer Dynamic Reaction Cell (DRC) II axial field ICP-MS at the University of Massachusetts, Boston as previously described [9, 41, 42]. To clean sample lines and reduce memory effects, sample lines were sequentially washed with 18.2 MΩ cm resistance water (Milli-Q water purification system, Millipore, Bedford, MA, USA) for 90 s and a 2% nitric acid solution for 120 s between analyses. Procedural blanks were analysed within each block of 10 samples, to monitor and correct for instrumental and procedural backgrounds. Calibration standards used to determine metal in blood included aliquots of Milli-Q water, and NIST 955c SRM spiked with known quantities of each metal in a linear range from 0.025 to 10 µg/dl. Standards were prepared from 1000 mg/l single element standards (SCP Science, USA). Method detection limits (MDLs) were calculated according to the two-step approach using the $t_{99}S_{LLMV}$ method (USEPA, 1993) at 99% CI ($t = 3.71$) [43]. The MDLs yielded values of 0.006, 0.005 and 0.071 µg/dl for Cd, Pb and As, respectively. Limits of detection (LOD) were 0.002, 0.002 and 0.022 µg/dl for Cd, Pb and As, respectively, and limits of quantification (LOQ) (according to Long and Winefordner, 1983) were 0.0007, 0.0006 and 0.0073 µg/dl for Cd, Pb and As, respectively. The number of samples below the LOD for Cd, Pb and As were 2, 1 and 1, respectively. Correlations between each metal are reported in Supplementary Table S2.

House Dust and Hand Wipe Cd and Pb

In water samples, trace metals were analysed after calibration using serial dilutions of NIST 1643e standard spiked with U and Th. Ghost wipe samples were digested with ultra-pure HNO_3 (Fisher Optima) and double-distilled water (resistivity >18.2 MU*cm). For measurement of total Pb, Cd and other trace elements, 0.034 g of soil and dust floor samples were digested. Reference materials used for instrument calibration and standardization included USGS certified geochemical reference materials G2, W2, BIR, SCO, SDO, AGV, RGM and SDC. All digestion processes of the dust samples were prepared using ultra-pure HNO_3 and HF reagents (Fisher Optima) and double-distilled water. All plastic polyethylene bottles and tubes were soaked

and cleaned with 1 N HCl and 1 N HNO_3, and then rinsed with deionized water. Trace elements were analysed via a VG Plasmaquad 3 ICP-MS at Duke's Nicholas School of the Environment.

Expression Data

Expression of *DLK1*, *DIO3* and *MEG3* were assessed using Taqman® gene expression assays (Hs00171584_m1, Hs00956431_s1 and Hs00292028_m1 respectively) using 18S (HS99999901_s1) as the housekeeping gene. Protocols were followed according the manufacturer's instructions. In brief, delta Ct values net of 18S were then subtracted from mean delta Ct value for each measured transcript by individual to get an index of low to high expression individuals (Supplementary Data S1).

Covariate Data

In addition to other toxic metals, factors that may influence the association between Cd exposure and CpG methylation include maternal obesity, parity, gestational age, ethnicity and offspring sex. Maternal obesity was estimated from self-reported maternal pre-pregnancy anthropometric measurements, including current height (m) and usual pre-pregnancy weight (kg) and BMI that was expressed as kg/m^2. Participants were defined as non-obese (BMI $<30\,kg/m^2$) and obese (BMI $\geq 30\,kg/m^2$) according to the World Health Organization classifications. Ethnicity was self-reported as African American, White, Hispanic or Other. Parity, gestational age at delivery and sex of offspring were reported and verified using medical records.

Statistical Analyses

Cluster Analysis

MEG3 DMR methylation percentages were normally distributed, and ranged from 56 to 85%, IQR = 69–77%, with mean and median of 72.3 and 72.9%, respectively. A *MEG3* DMR methylation fraction at or exceeding the top quartile (Q3/Q4 = 77% methylation) was defined as *MEG3* DMR hypermethylation. To ensure that an appropriate denominator was used, these analyses were limited to the subset of women living in Durham County at the time of blood draw used for metal measurements. To minimize the effects of cigarette smoking in this cluster analysis, smokers were also excluded.

The Hot Spot Analysis and Kernel Density Tools within ArcGIS 10.2.2 for cluster mapping and analysis was used to generate geographical heat maps from infant *MEG3* DMR methylation at birth, yet de-identify the specific prenatal address of study participants. The *MEG3* DMR hypermethylation density was calculated over a $1\,km^2$ grid and reported as $ng/g/km^2$ and the Getis-Ord Gi* statistic was calculated for each offspring. This statistic is based on the *MEG3* DMR hypermethylation value, a spatial weight between neighboring features and the total number of features. The Gi* statistic generated a Z-score and corresponding *P*-value that tested the hypothesis that a given pattern was the result of randomness in *MEG3* DMR methylation fractions. For this analysis, a fixed-distance band of 1500 m was used to calculate the spatial influence of neighboring points. Points falling outside the 1500 m band were considered non-influential in the calculation of the Gi*. The resulting significance levels (*P*-value) were then plotted using a heat map to show high ($P < 0.01$) and low cluster probabilities for a given methylation fraction.

Linear Regression Models

We used general linear regression models to evaluate the relationship between Cd exposure during the first trimester and *MEG3* DMR methylation at birth. Because of the severely right-skewed distribution of Cd concentrations that could not be improved by log-transformation, Cd exposure was dichotomized as upper quartile of blood concentration (0.513 ng/g) compared with all others. These models were adjusted for maternal co-exposure to Pb and As, race/ethnicity (African American, White, Hispanic and Other), cigarette smoking during gestation (yes or no), maternal obesity before pregnancy (continuous BMI), weeks of gestation (continuous) as a general indicator of fetal distress, offspring sex and maternal education, a stable indicator of socioeconomic status. Refined models were stratified separately by race/ethnicity and estimates for race were tested for homogeneity in SAS 9.4.

Mediation analysis

We used the *mediation* [44] package in R to test if mean *MEG3* methylation acted as a mediator on previous work showing associations between *MEG3* hypermethylation and low birthweight.

Supplementary Data

Supplementary data are available at *EnvEpig* online.

Conflict of Interest Statement. None declared.

References

1. ATSDR. Agency for Toxic Substances and Disease Registry. http://www.atsdr.cdc.gov/. 2011 (1 February 2019, date last accessed).
2. Nawrot T, Plusquin M, Hogervorst J, Roels HA, Celis H, Thijs L, Vangronsveld J, Van Hecke E, Staessen JA. Environmental exposure to cadmium and risk of cancer: a prospective population-based study. *Lancet Oncol* 2006;**7**:119–26.
3. Solenkova NV, Newman JD, Berger JS, Thurston G, Hochman JS, Lamas GA. Metal pollutants and cardiovascular disease: mechanisms and consequences of exposure. *Am Heart J* 2014; **168**:812–22.
4. Ishizaki M, Suwazono Y, Kido T, Nishijo M, Honda R, Kobayashi E, Nogawa K, Nakagawa H. Estimation of biological half-life of urinary cadmium in inhabitants after cessation of environmental cadmium pollution using a mixed linear model. *Food Addit Contam Part A* 2015;**32**: 1273–6.
5. Cheng X, Niu Y, Ding Q, Yin X, Huang G, Peng J, Song J. Cadmium exposure and risk of any fracture: a PRISMA-compliant systematic review and meta-analysis. *Medicine* 2016;**95**:e2932.
6. King KE, Darrah TH, Money E, Meentemeyer R, Maguire RL, Nye MD, Michener L, Murtha AP, Jirtle R, Murphy SK *et al*. Geographic clustering of elevated blood heavy metal levels in pregnant women. *BMC Public Health* 2015;**15**: 1035.
7. McKelvey W, Gwynn RC, Jeffery N, Kass D, Thorpe LE, Garg RK, Palmer CD, Parsons PJ. A biomonitoring study of lead, cadmium, and mercury in the blood of New York City adults. *Environ Health Perspect* 2007;**115**:1435–41.
8. Green AJ, Hoyo C, Mattingly CJ, Luo Y, Tzeng JY, Murphy SK, Buchwalter DB, Planchart A. Cadmium exposure increases

the risk of juvenile obesity: a human and zebrafish comparative study. *Int J Obes* 2018;**42**:1285–95.
9. de Angelis P, Miller RK, Darrah TH, Katzman PJ, Pressman EK, Kent TR, O'Brien KO. Elemental content of the placenta: a comparison between two high-risk obstetrical populations, adult women carrying multiples and adolescents carrying singletons. *Environ Res* 2017; **158**:553–65.
10. Park SS, Skaar DA, Jirtle RL, Hoyo C. Epigenetics, obesity and early-life cadmium or lead exposure. *Epigenomics* 2017;**9**: 57–75.
11. Skaar DA, Li Y, Bernal AJ, Hoyo C, Murphy SK, Jirtle RL. The human imprintome: regulatory mechanisms, methods of ascertainment, and roles in disease susceptibility. *ILAR J* 2012; **53**:337–54.
12. Woodfine K, Huddleston JE, Murrell A. Quantitative analysis of DNA methylation at all human imprinted regions reveals preservation of epigenetic stability in adult somatic tissue. *Epigenet Chromatin* 2011;**4**:1.
13. Ollikainen M, Craig JM. Epigenetic discordance at imprinting control regions in twins. *Epigenomics* 2011;**3**:295–306.
14. Hoyo C, Murphy SK, Jirtle RL. Imprint regulatory elements as epigenetic biosensors of exposure in epidemiological studies. *J Epidemiol Community Health* 2009;**63**:683–4.
15. Cowley M, Skaar DA, Jima DD, Maguire R, Hudson KM, Park SS, Sorrow P, Hoyo C. Imprinting control regions are hotspots for perturbation of DNA methylation by cadmium exposure in mothers and newborn children. *Environ Mol Mutagen* 2017; **58**:S39–S39.
16. Balas MM, Johnson AM. Exploring the mechanisms behind long noncoding RNAs and cancer. *Noncoding RNA Res* 2018;**3**: 108–17.
17. Zhou Y, Zhang X, Klibanski A. MEG3 noncoding RNA: a tumor suppressor. *J Mol Endocrinol* 2012;**48**:R45–53.
18. Vidal AC, Semenova V, Darrah T, Vengosh A, Huang Z, King K, Nye MD, Fry R, Skaar D, Maguire R *et al.* Maternal cadmium, iron and zinc levels, DNA methylation and birth weight. *BMC Pharmacol Toxicol* 2015;**16**:20.
19. Tilley SK, Reif DM, Fry RC. Incorporating ToxCast and Tox21 datasets to rank biological activity of chemicals at superfund sites in North Carolina. *Environ Int* 2017;**101**:19–26.
20. Li J, Shen H, Xie H, Ying Y, Jin K, Yan H, Wang S, Xu M, Wang X, Xu X *et al.* Dysregulation of ncRNAs located at the DLK1DIO3 imprinted domain: involvement in urological cancers. *Cancer Manage Res* 2019;**11**:777–87.
21. Galanter JM, Gignoux CR, Oh SS, Torgerson D, Pino-Yanes M, Thakur N, Eng C, Hu D, Huntsman S, Farber HJ *et al.* Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *Elife* 2017;**6**:1–24.
22. Rahmani E, Shenhav L, Schweiger R, Yousefi P, Huen K, Eskenazi B, Eng C, Huntsman S, Hu D, Galanter J *et al.* Genome-wide methylation data mirror ancestry information. *Epigenet Chromatin* 2017;**10**:1.
23. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr* 2007;**27**: 363–88.
24. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, Zhang W, Torskaya MS, Zhang J, Shen L *et al.* Season of conception in rural Gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet* 2010;**6**:e1001252.
25. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE, Lumey LH. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA* 2008;**105**:17046–9.
26. Li Y, Xie C, Murphy SK, Skaar D, Nye M, Vidal AC, Cecil KM, Dietrich KN, Puga A, Jirtle RL *et al.* Lead exposure during early human development and DNA methylation of imprinted gene regulatory elements in adulthood. *Environ Health Perspect* 2016;**124**:666–73.
27. Cecil KM, Brubaker CJ, Adler CM, Dietrich KN, Altaye M, Egelhoff JC, Wessel S, Elangovan I, Hornung R, Jarvis K *et al.* Decreased brain volume in adults with childhood lead exposure. *PLoS Med* 2008;**5**:e112.
28. Dietrich KN, Krafft KM, Shukla R, Bornschein RL, Succop PA. The neurobehavioral effects of early lead exposure. *Monogr Am Assoc Ment Defic* 1987;**8**:71–95.
29. Dietrich KN, Ris MD, Succop PA, Berger OG, Bornschein RL. Early exposure to lead and juvenile delinquency. *Neurotoxicol Teratol* 2001;**23**:511–8.
30. Wylie AA, Murphy SK, Orton TC, Jirtle RL. Novel imprinted DLK1/GTL2 domain on human chromosome 14 contains motifs that mimic those implicated in IGF2/H19 regulation. *Genome Res* 2000;**10**:1711–8.
31. Rogers ED, Ramalie JR, McMurray EN, Schmidt JV. Localizing transcriptional regulatory elements at the mouse Dlk1 locus. *PLoS One* 2012;**7**:e36483.
32. Steshina EY, Carr MS, Glick EA, Yevtodiyenko A, Appelbe OK, Schmidt JV. Loss of imprinting at the Dlk1-Gtl2 locus caused by insertional mutagenesis in the Gtl2 5′ region. *BMC Genet* 2006;**7**:44.
33. Zhang X, Zhou Y, Mehta KR, Danila DC, Scolavino S, Johnson SR, Klibanski A. A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *J Clin Endocrinol Metab* 2003;**88**:5119–26.
34. Zhu J, Liu S, Ye F, Shen Y, Tie Y, Zhu J, Wei L, Jin Y, Fu H, Wu Y *et al.* Long noncoding RNA MEG3 interacts with p53 protein and regulates partial p53 target genes in hepatoma cells. *PLoS One* 2015;**10**:e0139790.
35. Zhou Y, Zhong Y, Wang Y, Zhang X, Batista Dl, Gejman R, Ansell PJ, Zhao J, Weng C, Klibanski A. Activation of p53 by MEG3 non-coding RNA. *J Biol Chem* 2007;**282**:24731–42.
36. Zhang X, Rice K, Wang Y, Chen W, Zhong Y, Nakayama Y, Zhou Y, Klibanski A. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. *Endocrinology* 2010;**151**:939–47.
37. Sekita Y, Wagatsuma H, Irie M, Kobayashi S, Kohda T, Matsuda J, Yokoyama M, Ogura A, Schuster-Gossler K, Gossler A *et al.* Aberrant regulation of imprinted gene expression in Gtl2lacZ mice. *Cytogenet Genome Res* 2006;**113**:223–9.
38. Liu Y, Murphy SK, Murtha AP, Fuemmeler BF, Schildkraut J, Huang Z, Overcash F, Kurtzberg J, Jirtle R, Iversen ES *et al.* Depression in pregnancy, infant birth weight and DNA methylation of imprint regulatory elements. *Epigenetics* 2012; **7**:735–46.
39. Murphy SK, Huang Z, Hoyo C. Differentially methylated regions of imprinted genes in prenatal, perinatal and postnatal human tissues. *PLoS One* 2012;**7**:e40924.
40. Nye MD, Hoyo C, Huang Z, Vidal AC, Wang F, Overcash F, Smith JS, Vasquez B, Hernandez B, Swai B *et al.* Associations between methylation of paternally expressed gene 3 (PEG3), cervical intraepithelial neoplasia and invasive cervical cancer. *PLoS One* 2013;**8**:e56325.
41. Luo Y, McCullough LE, Tzeng JY, Darrah T, Vengosh A, Maguire Rl, Maity A, Samuel-Hodge C, Murphy SK, Mendez MA *et al.* Maternal blood cadmium, lead and arsenic levels,

nutrient combinations, and offspring birthweight. *BMC Public Health* 2017;**17**:354.

42. DeLoid G, Cohen JM, Darrah T, Derk R, Rojanasakul L, Pyrgiotakis G, Wohlleben W, Demokritou P. Estimating the effective density of engineered nanomaterials for in vitro dosimetry. *Nat Commun* 2014;**5**:3514.

43. Darrah TH, Prutsman-Pfeiffer JJ, Poreda RJ, Ellen Campbell M, Hauschka PV, Hannigan RE. Incorporation of excess gadolinium into human bone from medical contrast agents. *Metallomics* 2009;**1**:479–88.

44. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. *J Stat Softw* 2014;**59**:1–38.