UCLA
UCLA Previously Published Works

Title

An epigenome-wide association study of ambient pyrethroid pesticide exposures in California's central valley

Permalink

Authors

Furlong, Melissa A
Paul, Kimberly C
Yan, Qi
et al.

Publication Date

DOI

Peer reviewed

An Epigenome-Wide Association Study of Ambient Pyrethroid Pesticide Exposures In California's Central Valley

Melissa A. Furlong[1], **Kimberly C. Paul**[2], **Qi Yan**[2], **Yu-Hsuan Chuang**[2], **Myles G. Cockburn**[3], **Jeff M. Bronstein**[4], **Steve Horvath**[5,6], **Beate Ritz**[2,4]

[1] Department of Community, Environment, and Policy, University of Arizona Mel and Enid Zuckerman College of Public Health, Tucson, Arizona, USA

[2.] Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, CA, USA.

[3.] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, CA, USA

[4.] Department of Neurology, David Geffen School of Medicine, Los Angeles, CA, USA

[5.] Department of Human Genetics, David Geffen School of Medicine, Los Angeles, CA, USA

[6.] Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, CA, USA

Abstract

Background—Pyrethroid pesticide use is increasing worldwide, although the full extent of associated health effects is unknown. An epigenome-wide association study (EWAS) with exploratory pathway analysis may help identify potential pyrethroid-related health effects.

Methods—We performed an exploratory EWAS of chronic ambient pyrethroid exposure using control participants' blood in the Parkinson's Environment and Genes Study in the Central Valley of California (N=237). We estimated associations of living and working near agricultural pyrethroid pesticide applications in the past 5 years (binary) with site-specific differential methylation, and used a false discovery rate (FDR) cut off of 0.05 for significance. We controlled for age, sex, education, cell count, and an ancestral marker for Hispanic ethnicity. We normalized methylation values for Type I/II probe bias using Beta-Mixture Quantile (BMIQ) normalization, filtered out cross-reactive probes, and evaluated for remaining bias with Surrogate Variable Analysis (SVA). We also evaluated the effects of controlling for cell count and normalizing for Type I/II probe bias by comparing changes in effect estimates and p-values for the top hits across BMIQ and GenomeStudio normalization methods, and controlling for cell count. To facilitate broader interpretation, we annotated genes to the CpG sites and performed gene set overrepresentation analysis, using genes annotated to CpG sites that were associated with pyrethroids at a raw p<0.05, and controlling for background representation of CpG sites on the chip. We did this for both a biological process context (Gene Ontology terms) using missMethyl,

Corresponding Author: Melissa Furlong, mfurlong@email.arizona.edu PO Box 245210 Tucson AZ 85721.

and a disease set context using WebGestalt. For these gene set overrepresentation analyses we also used an FDR cut off of 0.05 for significance of gene sets.

Results—After controlling for cell count and applying BMIQ normalization, 4 CpG sites were differentially methylated in relation to pyrethroid exposures. When using GenomeStudio's Illumina normalization, 415 CpG sites were differentially methylated, including all four identified with the BMIQ method. In the gene set overrepresentation analyses, we identified 6 GO terms using BMIQ normalization, and 76 using Illumina normalization, including the 6 identified by BMIQ. For disease sets, we identified signals for Alzheimer's disease, leukemia and several other cancers, diabetes, birth defects, and other diseases, for both normalization methods. We identified minimal changes in effect estimates after controlling for cell count, and controlling for cell count generally weakened p-values. BMIQ normalization, however, resulted in different beta coefficients and weakened p-values.

Conclusions—Chronic ambient pyrethroid exposure is associated with differential methylation at CpG sites that annotate to a wide variety of disease states and biological mechanisms that align with prior research. However, this EWAS also implicates several novel diseases for future investigation, and highlights the relative importance of different background normalization methods in identifying associations.

Keywords

epigenetics; pyrethroids; pesticides; neurodevelopment; neurodegeneration

Introduction

As the use of organophosphate pesticides (OPs) worldwide has declined over the past decades, OPs have been replaced by an increasing use of alternate insecticides, including pyrethroids. Pyrethroids have become the most commonly used class of pesticides, both agriculturally and residentially, and as of 2010 they accounted for approximately 40% of all insecticide-treated acres in the United States (1).

Pyrethroids have been associated with a wide range of chronic health sequelae. These pesticides are designed to disrupt neurological functioning by targeting systems that are conserved across the animal kingdom, which primarily include voltage-gated sodium channels (2). More recent evidence shows that they also act on potassium, chloride, and calcium channels and/or currents (2–6). Pyrethroids also have been shown to have a wide range of downstream secondary effects, including genotoxicity and DNA damage (7), regulation of neuronal apoptosis (8, 9), and immunotoxicity (10, 11). Associated health effects include altered sperm and testis function (12–14), altered blood brain barrier functioning and oxidative damage in brain, liver, and kidney (15). A case report described a woman developing a motor neuron disorder resembling Amyotrophic Lateral Sclerosis after chronic inhalation of pyrethroids (16). Epidemiological studies have associated pyrethroid exposure during the prenatal period and early childhood with impaired neurodevelopment, emphasizing behavioral disturbances (17–20). Pyrethroid exposure has also been associated with Parkinson's disease (21, 22) and leukemia (23, 24) in some recent studies.

Thus, pyrethroid exposure may have subtle yet profound effects on a wide variety of health outcomes. Capturing this wide range of pyrethroid-associated health effects is difficult with traditional epidemiology studies that typically focus on a single outcome or exposure. Instead, multiple-outcome or exposome centered epidemiological approaches may be needed for shaping public health recommendations (25). An epigenome-wide association study, or EWAS, is one mechanism to study outcome-wide or exposome-wide epidemiology, since the epigenome theoretically may reflect any health outcome or exposure affected by or reflected in the epigenome. More importantly, epigenetic regulation has been proposed as a mechanism by which environmental exposures may influence biological responses and health. Exposures including air pollution (26, 27), organophosphate pesticides (28), and smoking (29) have been shown to induce epigenetic changes across the genome of peripheral blood cells. In these studies, methylated genes and areas overlap consistently with expected mechanisms of action of the exposure and/or with health outcomes previously associated with the exposure.

Since the field of EWAS is growing and methods are rapidly emerging, with diverging consensus on the most appropriate methods to use, we also use this opportunity to explore differences in results when considering two different controversial approaches. Since cell composition may be a mediator of the relationship between pyrethroid exposure and overall methylation status, and controlling for mediators violates a foundational epidemiological principle (30), we explored the effects of controlling for cell count on beta coefficients and p-values. We also explored the effects of two different background normalization methods, and compared beta mixture quantile normalization (BMIQ) against the standard Illumina normalization in Illumina's GenomeStudio software.

Here, we examine associations between chronic pyrethroid pesticide exposures and whole blood leukocyte DNA methylation in population control participants enrolled in a Parkinson's disease case-control study in central California, USA. We use an ambient exposure framework based on agricultural applications reported in the California Pesticide Use Registry. Living and working near these agricultural applications may result in exposures via inhalation, dermal absorption, hand-to-mouth transmission, and exposure to dust and dust reservoirs in the home (31, 32).

Methods

All subjects provided informed consent, and the study was approved by the UCLA Institutional Review Board.

Study Population

The Parkinson's Environment and Gene Study (PEG) is a case-control study that was designed to investigate the etiology of Parkinson's disease in California's central valley, a predominantly agricultural region. In the first phase, known as PEG1, the study recruited Parkinson's (PD) cases and population-based non-PD controls from three counties in central California (Kern, Fresno, Tulare) from 2001 to 2007. For this analysis, we only used the controls. Controls for PEG1 were identified from residential parcel listings from property tax assessor records. Controls were screened for eligibility by mail or telephone, and

participation was limited to one person per household. Eligibility criteria for controls included being over 35 years of age at enrollment, having lived in one of the three counties for at least the past 5 years, and not having received a Parkinsonism diagnosis. 121 of the original controls did not provide blood samples. More detailed recruitment information has been provided elsewhere (33).

For this study, we restricted participants to controls who self-identified as white race or Hispanic ethnicity among the PEG1 controls who provided venous blood samples (total n=237) and who did not have a PD diagnosis at baseline. PD causes fairly dramatic changes across the methylome. Since pyrethroids may influence these PD-methylomes differently from the methylomes of controls, and because some of these differences in PD methylomes might be due to pyrethroid exposure, we restricted this analysis to controls to limit the potential confounding introduced by the cases.

Interviewers were blinded to case/control status and conducted structured telephone interviews to obtain demographics and lifetime address histories.

DNA Methylation

DNA was extracted from peripheral whole blood, and samples were profiled using the Illumina Infinium 450k platform. All samples were prepared in accordance with manufacturer guidelines and the protocol for the Illumina 450k array, and are described in detail elsewhere (34). Samples were also randomized across well position and bisulfite conversion plate to reduce batch effects. This array measures 486k CpG sites, and covers 99% of RefSeq genes and 96% of CpG islands. In pre-processing, beta values were normalized using Beta-Mixture Quantile (BMIQ) normalization, we confirmed sex concordance, and we filtered out cross-reactive probes (35). In sensitivity analyses, we compared results from the BMIQ normalization with results using the background normalization method from GenomeStudio. We also evaluated residual batch effects and additional confounding using Surrogate Variable Analysis (SVA), with the SVA package in R (36). We first checked whether any surrogate variables were present in the primary model (BMIQ normalization and controlling for cell count). If SVs were present, we planned to include them in the final models.

Pesticide Exposure History

The California Pesticide Use Registry (PUR) is a state-wide registry of all commercial pesticide applications in California, with pesticide data from 1974 to present. We used a validated geospatial-based system (which has been described elsewhere (37)) to assign exposure status to all controls in PEG1. Briefly, the PUR reports pesticide applications at the level of the Public Land Survey System (PLSS), roughly a one-square mile resolution. We combined this PUR data with land-use and crop information from the state to improve the spatial resolution further. Then, we calculated total pounds of each active ingredient applied within 500 meters (38)of the participant's residence based on their reported address histories for each year from 1974 to the year of blood draw. For the five years prior to enrollment, the pounds per acre for each pyrethroid pesticide at each residential and occupational address were averaged over this five year period, using participant's entire residential and

occupational history. We then created a total pyrethroid count variable that reflected the number of pyrethroid pesticides a participant had been exposed to at greater than annual median levels (e.g., if a participant had been exposed to permethrin at greater than the median level for one year, and to zeta-cypermethrin at greater than the median level for that year, their total count would be 2). After examining the distribution of this continuous variable, we decided to dichotomize this count variable to indicate exposure to any pyrethroid pesticide at greater than the median level, over the past 5 years.

Statistical Analysis

All analyses were conducted in RV3.6.0 ("Planting of a Tree"). We first examined whether participants who provided blood were different from those who did not provide blood by age, sex, education, occupational exposure to pesticides, our pesticide exposure metric, and Hispanic ethnicity status. Then, we describe characteristics of the study population. To perform an epigenome-wide association study of the relationship between pyrethroid exposure and differential methylation, we used beta regression to evaluate associations between the binary pyrethroid exposure variable (any pyrethroid exposure > median for the last 5 years) and the beta values for each CpG site, with the package betareg in R. Beta regression has advantages for modeling associations with values that are distributed between 0 and 1; it allows coupled mean and variance terms to both depend on covariates, and has superior sensitivity and specificity for identifying significant associations, compared to t-tests and linear regressions with beta or M values (39). Since beta regression requires values to be between 0 and 1, we first 'squeezed' the data(39), using the formula (betavalue*(n obs-1) + 0.5)/n. This effectively maintains the rank ordering and values while slightly shifting values that are exactly equal to 0 or 1, thus facilitating beta regression. We used a log link to bound the proportions between 0 and 1. Thus, we present the exponentiated beta coefficients so that they represent the proportion change, e.g., the relative percentage change in methylation. We also used an lsmeans approach to generate the approximate absolute percent change in methylation expected from pyrethroid exposures (presented as P in Table 1). These estimates are approximately equivalent to what would be provided from a beta coefficient from a standard linear model. We a-priori selected the adjustment variables of age, sex, education, a genetic ancestry marker for Hispanic ancestry, and blood cell composition proportions (CD4T, granulocytes, B cells, and natural killer cells), based on evidence in the literature that these factors may confound relationships between environmental contaminants and health outcomes. In sensitivity analyses we additionally controlled for occupational exposure to pesticides (40). We used the Houseman method to estimate cell composition used in these models (41). We also evaluated EWAS models without controlling for cell composition. We then compared the results from the two models by plotting the relationships of these values in a scatterplot. We repeated this analysis to evaluate the effects of BMIQ normalization for Type I/II probe bias on beta coefficients and p-values. We present the CpG sites that were associated with pyrethroids at FDR q<0.05 from the method using the BMIQ normalization and controlling for cell count, and we also present the top 50 hits by p-value for the CpG sites associated with pyrethroids at FDR q<0.05 in the model using GenomeStudio's background normalization, and controlling for cell count. The complete list is presented in supplementary materials as an R dataset. We also estimated the associations between pyrethroids and the Houseman cell counts reported

above, as well as CD8 Naïve cells, CD4 Naïve cells, and CD8pCD28nCD45Ran (42), while controlling for age, education, sex, and Hispanic ethnicity.

We report Manhattan Plots, p-value histograms, q-q plots, and genomic inflation scores (lambda) for the BMIQ and Illumina normalization models. Since genomic inflation scores overestimate small p-value inflation in EWAS studies, particularly when the exposure may have many small effects throughout the epigenome (43), we employ the "bacon" method to estimate a Bayesian Genomic Inflation Factor (44). Briefly, this method constructs an empirical null distribution using a Gibbs Sampling algorithm by fitting a three-component normal mixture on z-scores. One component is forced to represent the null distribution with mean and standard deviation representing the bias and inflation. The other two components capture the amount of true associations present in the data. We implemented this method with the Bioconductor package "bacon".

To annotate CpG sites, we used UCSC gene annotations from the Illumina450k annotation database. If no gene was annotated to an identified CpG site, we identified the closest gene with the hiAnnotator package in R. In sensitivity analyses, we limited genes to genes within 100,000 basepairs of the CpG site. hiAnnotator interfaces with the UCSC web browser, downloads the appropriate tracks and tables from the web browser, creates GRange objects, and identifies the closest gene. We used the hg19 freeze, the "UCSC Genes" track and the "knownGene" table to identify the location of known genes. We also used the kgXref table to identify the gene symbols associated with the UCSC gene names. We used EntrezGene, GeneCard, and UniprotKB-Swissprot summaries reported on GeneCard to describe genes and gene proteins of these closest genes and their paralogs, and we preferentially reported associated diseases and biological functioning.

Gene Set Overrepresentation Analysis

To evaluate the possibility that the pyrethroid-associated CpG sites might reflect specific pyrethroid-associated biological pathways, we used a combination of resources. WebGestalt (WEB-Based GEne SeT AnaLysis Toolkit) and the missMethyl package (45) are freely available resources that provide biological and disease context to gene lists while allowing the user to control for background representation. We used WebGestalt to identify gene-associated diseases while comparing to the background gene list of the Illumina450K chip, and missMethyl to identify associated Gene-Ontology (GO) Biological Process terms, which account for the background representation of CpG sites across genes on the Illumina450K chip. Both protocols analyze groups of genes to estimate the observed vs expected number of genes in any given set. This generates a fold-enrichment value, which is a measure of the observed vs expected number of genes in any given pathway or disease set. Both processes require a list of genes to analyze for functionality. Since a large proportion of CpG sites that are above the 0.05 FDR-q value cutoff may be truly associated with pyrethroid exposure, and because overrepresentation analysis requires a large number of genes, we used the genes annotated to CpG sites that were associated with pyrethroids at the raw p cutoff of 0.05. In WebGestalt, we set the minimum number of genes for a category to 3. For all associated diseases and gene ontology terms, we report the FDR q-value cutoff, which represents a Fisher test of whether the biological set is overrepresented among the gene set. This q-value

is not related to the CpG-site pyrethroid-associated q-values. We assessed the OMIM disease states, GLAD4U disease states, and GO Biological Processes that were below the FDR-corrected q-value cut-off of 0.05 for the pathway, using the model that utilized BMIQ normalization and controlled for cell count. In sensitivity analyses, we also performed these gene set overrepresentation analyses with the lists that were generated from the models that utilized GenomeStudio's background normalization and controlled for cell count.

Results

Participants who did not provide blood samples were not different from participants who did provide blood for the characteristics of age, sex, ethnicity, education, or occupational exposure to pesticides. However, participants who did not provide blood (and were thus excluded from this study) were slightly more likely to be exposed to pyrethroid pesticides over the past 5 years (26% exposed, p<0.01). The 237 participants included in the study were on average approximately 67 years of age, and approximately half male (53%). Most (87%) of the participants identified as Non-Hispanic White, 19% were considered exposed to pyrethroids according to our metric of exposure, and 31% were exposed to OPs over the past 5 years (Table 1). Exposure to OPs and pyrethroids were moderately correlated (r= 0.56).

EWAS

Surrogate Variable Analysis did not identify any surrogate variables, and thus none were included as a covariate. The number of identified CpG sites that met an FDR q cutoff<0.05 varied based on model characteristics. After filtering out cross-reactive probes, applying BMIQ normalization to methylation values, and age, sex, education, a genetic ancestry marker for Hispanic ancestry, and blood cell composition proportions (CD8T, granulocytes, B cells, natural killer cells, and monocytes), we identified 4 CpG sites that met an FDR cutoff <0.05 (Table 2). These four identified CpG sites mapped to FAM20C, RUNX3, PIN4, and CDH11. FAM20C and CDH11 are both associated with calcium ion binding and skeletal deformities, while RUNX3 and PIN4 are associated with cancers.

However, when using Illumina's GenomeStudio's background normalization, we identified 415 CpG sites that met the FDR cutoff. All 4 of the CpG sites identified by the BMIQ models were included in these 415 sites. The functional annotations on the associated genes include a wide range of neurological disorders and functions, cancers, transcription factors, and developmental disorders.

When examining the comparisons for any CpG site that met an FDR q<0.05 for any of the 4 models tested (BMIQ with cells, BMIQ without cells, Illumina with cells, Illumina without cells), we plotted comparisons for 574 sites. Controlling for cell count had minimal effect on these beta coefficients (Figure 1 panel A, spearman correlation coefficient for cell counts vs no cell count control $\rho= 0.99$), although p-values displayed more variation (Figure 1 panel B, $\rho =0.89$), likely due to increasing the degrees of freedom in the models rather than changes in beta coefficients. This was generally true for both the models using BMIQ normalization (Figure 1 panels A & B) and the models using the standard Illumina GenomeStudio's background normalization (Figure 1 panels C & D, $\rho = 0.99$ and $\rho =0.62$,

respectively). However, a large cluster of CpG sites that were identified when using the background normalization had substantially different beta coefficients compared to the BMIQ normalization (Figure 1 panel E, $\rho=0.83$). The p-values from the two models also were substantially changed and weakened when using BMIQ normalization compared to the Illumina-normalized methylation values (Figure 1 panel F, $\rho=0.28$).

When evaluating the utility of controlling for cell count, the BMIQ model that did not control for cell count identified 6 CpG sites, including all 4 identified by the BMIQ model that did control for cell count, indicating a slightly greater sensitivity when not controlling for cell count. When comparing the differences for the beta coefficients (the log of the relative changes) for these 6 sites, the median difference in beta coefficients was 0.01, ranging from 0 to 0.05, suggesting that cell count is generally neither an important mediator nor important covariate for the pyrethroid – CpG relationship for these top hits. To put these in context in terms of relative differences in methylation proportions, for cg18234533, a 0.01 change in the beta coefficient would change the exp(B), a measure of relative effect, from 1.08 to 1.09. We also did not observe significant associations between pyrethroids and cell composition (Appendix Table 1).

Bayesian genomic inflation factors were excellent for all models, equaling 0.97 for the model using BMIQ normalization and controlling for cell count, 1.11 for the model using BMIQ normalization and not controlling for cell count, 1.06 for the model using Illumina normalization and controlling for cell count, and 1.03 for the model using Illumina normalization and not controlling for cell count.

We did not identify any individual chromosome with prominent and consistent associations (Figure 2). Instead, the p-value histograms indicated a mostly uniform distribution above 0.10, with a strong skew towards smaller values, possibly indicating widespread associations across the epigenome (Figure 2). QQ plots showed the least early departure for the BMIQ normalization models (Figure 2).

In sensitivity analyses, controlling for occupational exposure to pesticides did not materially change the beta coefficients, and thus does not confound the analysis we conducted.

Gene Set Overrepresentation—We next performed a gene set overrepresentation analysis for Gene Ontology terms, using the set of genes annotated to the CpG sites associated with pyrethroids at a raw p<0.05 in the BMIQ-normalized models that controlled for cell count (n=32,695 CpG sites), and also in the Illumina models that controlled for cell-count (N=69,180 CpG sites). The 6 GO terms that were identified using the BMIQ models included membrane, nucleoplasm, cytosol, protein binding, ATP binding, and transcription factor binding (Table 3). When using Illumina-normalized methylation models, we identified 67 GO terms, which included all the GO terms identified when using the BMIQ models. Notable additional pathways in this model included neural tube closure, in utero embryonic development, embryonic digit morphogenesis, glutamatergic synapse, dendrite, and several transcription related terms (Table 3).

In the disease set analysis from OMIM and GLAD4U (Figure 3), we observed several diseases related to cancer, metabolic disorders, and Neurological/ Musculoskeletal/ Developmental disorders. In OMIM, the smallest p-value associations were for Diabetes (Non-insulin dependent) and Acute Myeloid Leukemia, while the highest fold enrichments were for prostate cancer, Alzheimer's disease, and Mycobacterium tuberculosis. In GLAD4U, the disease with the lowest p-value was tracheoesophageal fistula (which was also represented in the OMIM hits), and the strongest fold enrichment ratios were for synostosis (adjacent bones become fused), hand deformities, and cluster headaches. When examining disease related gene set overrepresentation for the background-normalized CpG sites associated at $p<0.05$ with pyrethroids, this method identified many of the same diseases as when using BMIQ-normalization, and additionally identified a number of other cancers, lung diseases, and neurological disorders, including late onset Parkinson's disease, Prader-Willi Syndrome, Amyotrophic Lateral Sclerosis (ALS), neuropathy, neural tube defects, deafness, depression, psychomotor epilepsy, nervous system malformations, developmental disorder, Schizophrenia, neurodegenerative diseases, and brain diseases (Appendix 2). Results were similar when restricting to genes within 100,000 basepairs of the CpG site.

Discussion

In this study, pyrethroids were associated with a wide range of CpG sites that varied by chromosome, associated gene function, and disease state. Two of the top CpG sites are annotated to genes associated with calcium ion binding, which is consistent with reports of pyrethroids disrupting calcium ion flux and calcium ion channels (5, 46, 47). Several of the prominently identified gene ontologies and diseases have been previously implicated in pyrethroid exposure (e.g., Neurological, Developmental, Musculoskeletal, Cancer (16–20, 48–52)), while several are novel and have not been investigated in relation to pyrethroid exposure (e.g., Diabetes, Alzheimer's, birth defects, skeletal abnormalities). The related disease states reflect similar findings; Parkinson's disease (21, 53) and mental disorders (54, 55) have been associated with pyrethroids, and some cancers have been previously investigated with pyrethroid exposure (56), although the cancer literature tends to be mixed (57). Diabetes, Alzheimer's, and other cancers have not been thoroughly investigated. It is perhaps notable that Diabetes and Alzheimer's were both identified, as Alzheimer's disease has been hypothesized to be "Type 3 Diabetes", due to many shared biological mechanisms and the contributions of insulin to amyloid plaque formation and the formation of neurofibrillary tangles (58, 59). It is additionally notable that we identified several genes associated with Herpes Simplex Virus I (HSV1), which has also been associated with Alzheimer's disease (60–63), and several genes and pathways involved in WNT signaling, which has also recently been implicated for its importance in Alzheimer's disease(64). When using Illumina-normalized methylation values, we also identified overrepresentation of gene ontology terms that are consistent with the mode of action of pyrethroids (specifically disruption of the glutamatergic synapse, a known target of pyrethroids (65)). The epigenome may thus reflect both the utility of our exposure metric, as the EWAS captured the biological function of pyrethroids, as well as the potential for EWAS to function as an outcome-wide association study across disease states and biological processing.

Methodological Approaches

Interestingly, we identified fewer CpG sites and pathways when using BMIQ-normalized methylation compared to using the standard Illumina GenomeStudio background normalization. Prior studies are inconsistent regarding the purported value of different normalization procedures compared to Illumina normalization method (66, 67), and some studies suggest that concerns regarding normalization methods may be over inflated, since both type I and type II probes are highly reproducible (68). We did find that using background normalization identified the same sites as when using BMIQ normalization, although Illumina background normalization results in dramatically more hits. The Spearman correlation coefficient for the beta coefficients from the two methods was high, at 0.83, showing that the relative ordering of the effect estimates was maintained across normalization methods. Although these factors may appear to reflect a greater sensitivity when using the background Illumina standardization, prior studies have reported the opposite (67, 69). Thus, it also possible that many of the hits from the Illumina standardization method are false positives. However, the results from the disease analysis using BMIQ normalization is compatible with many of the diseases and functions related to the annotated genes from the top Illumina hits. Thus, exploring the implications of the annotated genes from the hits using the Illumina method is still worthwhile, with the obvious caveat that this is highly exploratory and needs replication in future studies.

We observed minimal confounding by cell count for our identified CpG sites. Instead, adding cell count variables to the models served to inflate standard errors and reduce the number of CpG sites that met a p-value threshold, which is an anticipated side effect of including more terms in the model. Despite this finding in our study, controlling for cell count likely still holds utility for other types of EWAS studies, particularly when the outcome might vary by cell count. Our models examined an environmental exposure in relation to methylation in this control population. Cell count may be an important consideration when including patients, or in other study designs, and may be critical for reducing bias in certain situations. However, unnecessarily adjusting for cell count across the board may lead to inflated standard errors, and researchers should assess the confounding or mediation potential on a study-by-study basis.

Capturing Exposure Relevance in EWAS

Despite their wide-ranging phenotypic effects observed in our study and the literature, pyrethroids are primarily designed to disrupt neurological functioning by interfering with voltage-gated sodium channels. However, they also disrupt voltage-gated chloride and calcium channels (70). NRXN1, MCC, FAM20C, and CDH11, which were all annotated to top hits identified in our EWAS, are all implicated in ion channel functioning, and have several downstream functions. For instance, alpha-type isoforms of NRXN1 play a role in the regulation of calcium channel activity and Ca(2+)-triggered neurotransmitter release at synapses and at neuromuscular junctions. NRXN1 is also associated with a wide variety of neurological and developmental conditions including autism, schizophrenia, nicotine dependence, hypotonia, mental retardation, and language delays (71–74). The identification of these genes in the EWAS suggest that our exposure metric may indeed be capturing true pyrethroid exposure, although it is unclear whether pyrethroids are influencing voltage gated

channels via methylation, or whether methylation occurs as a result of changes in voltage gated channel activity and/or morphology. Additionally, in our data, the direction of effect for CpG sites were mostly negative, indicating a hypomethylation effect. The clinical relevance of hypo- vs hypermethylation is variable, depending on the gene and outcome (75). However, in the developing mouse brain, broad hypomethylation of various genes (including ion channels) in neurons led to defects in neuronal morphology, excitability, and dendritic arborization(76). Additionally, in the same study, hypomethylation of these genes was associated with prolonged action potentials, and suggested an absence of the rapidly inactivating potassium current. This is consistent with reports that pyrethroids prolong the amount of time that sodium channels are kept open, leading to hyperexcitation of the nervous system (77).

Relevance of Identified CpG sites

The top CpG sites by p-value were mostly related to calcium ion binding, cancer, or neurological phenotypes. FAM20C, CDH11, USP24, INHBC, and CLEC2D are integral in calcium related processes and are associated with musculoskeletal disorders and bone growth. RUNX3 and PIN4 are related to cancer, and RUNX3 is also related to dendritic cells development. Some other notable sites identified by the model using the Illumina background normalization include cg15329860, associated with KDM5D, which is associated with spermatogenic failure, and its paralog is associated with x-linked mental retardation. cg21690979 in exon 1 in the promoter region of IRF2BP2 (Interferon Regulatory Factory 2 Binding Protein 2). IRF2BP2 has a number of functions, particularly as a corepressor of the transcriptional activity of IRF-2, which competes with IRF-1. In the brain, the IRF protein family act as transcription factors for genes involved in neuronal survival, neuroinflammation, and microglia activation (summarized in (78)).

Interestingly, IRF1 transcription was recently implicated in a twin study of Autism Spectrum Disorders (79), and the IRF2 gene was implicated in a case study of patients with intellectual disability (80). IRF2BP2 also regulates macrophage polarization and is involved in inflammatory response, and microglial IRF2BP2 may act to limit functional deficits after ischemic brain injury (81). Another study reports that IRF2BP2 interacts with enhanced postnatal care provided to reduce anxiety in mice, where mice with IRF2BP2-microglial knockout did not experience any benefits of enhanced postnatal care on anxiety(82). IRF2BP2 may thus be critical in responding to external insults, such as pyrethroid exposure. IRF2BPL (Interferon regulatory factor 2 binding protein like), an important paralog of IRF2BP2, is associated with neurodevelopmental disorder with regression, abnormal movements, loss of speech, and seizures (83). Another top hit, cg16275172, is in the promoter region of the 5'UTR 1st exon of TMOD3. TMOD3 blocks the elongation and depolymerization of actin filaments at the slow-growing, pointed end. The TMOD/TM complex contributes to the short actin protofilament, which defines the geometry of the membrane skeleton. TMOD2, an important paralog of TMOD3, encodes a neuronal-specific member of the tropomodulin family of actin-regulatory proteins. TMOD2 and TMOD1 are critical for dendritic branching, spine morphogenesis, and synapse formation, and TMOD2 was previously proposed as a candidate gene for Amyotrophic Lateral Sclerosis (84). ALS is a neurodegenerative disease characterized by motor and neuron degeneration, and is

regulated in part by actin dysregulation (85). Interestingly, a case report of pyrethroid inhalation intoxication describes a woman who developed ALS-like symptoms after chronic pyrethroid exposure, and subsequent amelioration of these symptoms after the pyrethroids were removed from her environment(16). Gene expression of TMOD1 and TMOD2 have also both been implicated in studies of Fragile X and Down Syndrome (86). These top hits reflect genes involved in bone mineralization, embryonic development, cancer, neurodegeneration and neurodevelopment. Interestingly, these trends were also observed in the disease-related gene set overrepresentation analysis, where we observed overrepresentation of Alzheimer's disease, several cancers, skeletal disorders, and mental disorders. In the expanded gene set analysis using the background normalization values, we identified a number of additional developmental and neurological diseases, including Parkinson's disease, developmental disorders, brain diseases, Prader-Willi syndrome, deafness, ALS, neural tube defects, and schizophrenia. Most of these diseases have not been investigated in relationship to pyrethroid exposure.

Challenges to Interpreting Pyrethroid-Disease Relationship in EWAS

A number of challenges exist in interpreting these results, including both the underlying assumptions for annotating genes to CpG sites for use in gene set overrepresentation analyses and for interpreting them in terms of causality and direction of effect. A primary consideration for gene annotation is that although we have identified associations with differential methylation of these CpG sites, we have not identified changes in gene expression or changes in methylation for the most relevant affected tissue. The existing literature that allows us to infer from CpG site methylation to gene expression, or assess relationships between blood CpG and other tissue methylation is still relatively scant. For instance, one of the top CpG sites, cg10065825, annotated to CDH11, has previously been associated with progression free survival in chronic lymphocytic leukemia (87), and is one of the only CpG sites with some prior information. Additionally, we explored correlations between blood methylation values and brain methylation values for several CpG sites annotated to genes with neurologic functions, using a freely available blood brain methylation tool(88). The blood methylation values for these sites were either not highly correlated or only moderately correlated with brain methylation (with the exception of cg15329860, which is correlated with brain methylation at approximately 0.98 in several brain regions) (88). This encourages caution when suggesting functional implications for identified CpG sites, and reflects gaps in the current literature. Whether our results reflect brain methylation is unclear and the functional implications of CpG site methylation is generally unknown for most CpGs. Still, our exploratory study did identify diseases and pathways that are consistent with the prior epidemiological literature on pyrethroids and health effects, suggesting at least some utility for this approach.

Pyrethroids may also be associated with pathways that are associated with the diseases we identify here, but may not cause those diseases. For instance, several of the disorders we identified are mostly genetically determined, and may be unlikely to be caused by an environmental agent. However, pyrethroids may perturb many pathways that are integral to the pathology of the disease. Many of these diseases are also developmental. However, our study population was predominantly older, and it is not possible to develop these diseases at

older ages. Additionally, it is unclear whether changes in the epigenome do truly cause phenotypic variation. Epigenomic changes may instead be a result of changes in disease or phenotypic state. It is also possible that the epigenome responds to pyrethroids with the same biological mechanisms that are involved in an unrelated disease. For instance, we identified several metabolic pathways and disorders in our GO Biological Processes and OMIM/GLAD4U disease set analyses. Pyrethroids may indeed be associated with metabolic disorders, which are then reflected in the epigenome. However, it is also possible that the human body uses the same metabolic pathways to metabolize and detoxify pyrethroids, which are up- or downregulated in response to pyrethroid exposures, and pyrethroids may not be related to metabolic disorders at all. Our results, and other EWAS results, should be interpreted as one part of a broader literature. For instance, although similar concerns can be applied to interpreting the pyrethroid- neurodevelopmental relationships, the supporting evidence in toxicological and epidemiological studies suggests there may indeed be an impact of pyrethroids on neurodevelopment.

This supporting evidence exists in both the epidemiological and animal literatures. Two prior epidemiological studies, also in California, showed that living near pyrethroid pesticide applications during pregnancy was associated with the development of autism in childhood (54, 55). Other epidemiological studies have shown associations between pyrethroids and Parkinson's disease (21, 53), behavioral problems (17–20), and ADHD (19, 49). Animal studies have reported similar findings (8, 89–93). To our knowledge, no studies have investigated the relationship between pyrethroids and Alzheimer's disease, although evidence is growing for a relationship between environmental exposures, in particular air pollution, and subsequent development of cognitive decline and/or Alzheimer's disease (94).

We also observed a number of CpG sites that mapped to genes involved in cancer pathways. There is some existing evidence of an association between pyrethroids and leukemia (23), and more evidence of associations between broad pesticide use and leukemia (23, 24, 95, 96). However, the overall literature on pesticide use and cancer tends to be mixed (57, 97).

Strengths and Weaknesses of the Study Design

There were several strengths to this study design. First, we were able to characterize ambient pesticide exposures over a long period in a general population cohort. Pyrethroids are rapidly metabolized, and persistent epigenetic changes in an adult population may require months or years in response to environmental insults. Thus, total ambient pesticide applications over a long period may be a more relevant measure of chronic exposure for methylome changes compared to other metrics, such as pesticide metabolites as biomarkers, which reflect exposure in the more recent past (hours to days). We did not explore differences in the time window for the exposure metric. Since pyrethroids were much less in agricultural use than OPs at the time of our study, a 5-year metric was chosen to achieve a balance between a relevant biological time frame and having enough "exposed" persons to conduct the EWAS. A one year window would have resulted in a much smaller exposed number of participants, and a 10 year window would not have added many exposed as exposures were much more common in the 5 years prior to enrollment/reference date. Thus,

future research is needed to assess whether there are more relevant time windows for these types of exposure.

Additionally, ambient pesticide exposure near agricultural fields assessed here does not capture diet or residential use, which is an important source of pesticide exposure in the general population (98). Still, if exposure from diet is not associated with ambient pesticide exposure from agricultural sources, omitting this source of variation is not likely to be a confounder of the ambient pesticide-epigenome relationship. Diet may be related to residence area and socioeconomic status, and we did not have such information. Yet, when we controlled for education and occupational exposure to pesticides the results did not change; thus, we did not observe confounding by these factors.

Additionally, this ambient-exposure based approach for pesticides has been successfully used to identify associations with a number of prior health outcomes, including Parkinson's disease (99), autism spectrum disorders (55), neural tube defects (100), and pancreatic cancer (101). Another strength to the study was the careful design for the 450k chips; all analyses were run in a single batch and samples were randomly assigned to well plate positions. Thus, laboratory artefacts from batch and well position are not likely to present any bias in this analysis, which was confirmed by the lack of identified components in the surrogate variable analysis.

Another consideration is that we used blood cells to measure methylation. Since this study was exploratory, blood cells may be the most relevant tissue to assess epigenome-wide associations for pyrethroid pesticides. However, the preponderance of identified neurological diseases and phenotypes, along with the brain-based mechanism of action for pyrethroids, may support an argument for using brain tissue or cerebrospinal fluid (CSF) to measure pyrethroid-associated methylation, since methylation patterns are highly tissue-specific. Unfortunately, extracting brain tissue or CSF, or other types of internal organ tissue, from healthy subjects is difficult and faces ethical concerns; thus, using blood cells may be the most feasible alternative for an EWAS study among healthy subjects. Certain challenges do exist: one of our top hits, for instance, was TMOD3. While TMOD2 is a candidate gene for ALS, TMOD2 is not expressed outside of the brain (84). Thus, the use of blood may preclude our ability to identify pyrethroid associations with CpG-associated genes that are not expressed or not differentially methylated outside of the brain. The results from the blood-based results may, however, provide some impetus for examining specific tissues or media for certain cancers, or in post-mortem brains of patients with neurodegenerative diseases.

Although our sample size was limited to 237 subjects, which likely limited our power, we were still able to identify several CpG sites that met FDR significance, and several pathways and diseases in gene set overrepresentation analysis. These results need to be independently replicated, and the dataset of CpG sites and associated p-values are available in the appendices for future researchers investigating pyrethroid EWAS studies. Thus, if our sample size precluded our ability to identify a CpG site with, for example, an FDR q value of 0.08 in our dataset, a replication analysis that identifies the same CpG site, should be able to note this and identify the CpG site as overlapping. The relatively small sample size did

limit our ability to assess associations by specific pyrethroid pesticide (e.g., deltamethrin, permethrin, allethrin, etc), since the numbers of exposed people for each pesticide would have been too small for evaluation. Different pyrethroid pesticides do display differing toxicities (102), and this may be critical to consider in future research. A limitation to our study is that we were unable to replicate our results in a replication dataset, and our sample size was sufficiently small that doing discovery and replication in our own dataset would have been challenging. Ambient chronic pesticide exposure is not an easily measured exposure metric; indeed, in the United States, only California and Arizona have pesticide use registries. Pesticide exposures are also not readily available in open source datasets. Interestingly, the participants included in this EWAS study were slightly less likely to be considered exposed to pyrethroids than those who refused a blood draw. However, participants did not differ on other metrics we examined and while we do not expect bias from exclusions, a higher number of exposed participants would have increased our statistical power to detect site specific methylation differences.

The exposed participants in our study were additionally comparatively highly exposed to other ambient agricultural pesticides, as this tri-county area lies in the Central Valley of California, which is one of the most agriculturally intensive regions in the United States. Ambient pyrethroid pesticide exposure is unlikely to be this high in most urban or non-agricultural rural areas. However, many cities and towns regularly spray pyrethroids in urban areas to control West Nile Virus and as a general form of mosquito control, and ambient pesticide exposure in these areas during mosquito season may be just as high, or higher, for residents living directly in the paths of these spray trucks. Additionally, home pesticide use is widespread (103), and exposures from foggers in the home may result in higher acute airborne exposures compared to living 500 meters from an agricultural pyrethroid application.

The wide range of phenotypes and gene functions that we observed may seem at first glance to display divergent functionality, but it is important to note that these symptoms are consistent with symptoms from pyrethroid pesticide poisoning reports. In one investigation of chronic sequelae following acute pyrethroid intoxications, a doctor in Germany documented a series of symptoms that regularly followed pyrethroid exposure, including mood disorders and reduced intellectual performance, visual and hearing problems, sensomotor polyneuropathies, vegetative nervous disorders, hypotonias, increased susceptibility to infectious diseases, autoimmune diseases, and skin problems (104). Although he observed alleviation of acute symptoms early on, the author documented residual effects up to several years after the incidents. In another study of Chinese reports of pyrethroid intoxications (105), authors observed commonalities among victims of poisoning incidents: abnormal facial sensations, nausea, headache, dizziness, anorexia, fatigue, weakness, parasthesia, and sweating were commonly noted. In severe cases, they also noted seizures, pulmonary edema, muscular fasciculation, and coma. Similar symptoms were noted in a study of occupational poisonings (106). We observed diseases and gene functions related to all of these symptoms. It is remarkable that the symptoms of pyrethroid pesticide poisoning, as reported in the literature, appear to be reflected in the epigenomes of otherwise healthy people who are likely unaware of their pyrethroid pesticide exposure. This may imply that the symptoms of pyrethroid pesticide poisonings are on the extreme end of a

continuous phenotypic spectrum, and that more minor health effects on the other end of the spectrum may be experienced at the lower end of pyrethroid exposures. Barring residual confounding or unaccounted-for problems with study design, the prior literature, along with our findings, appear to indicate that both low- and high-level chronic pyrethroid exposures disrupt a variety of critical biological processes, with a wide-ranging and potentially dramatic impact on several metrics of human health. Importantly, such effects may be reflected in the epigenome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Appendix

Appendix 1 Table 1

Associations of Pyrethroids with Cell Count (N=237)

	Beta (95% CI)
CD4 Naïve	−23.600 (−61.637, 14.436)
CD8 Naïve	8.033 (−6.119, 22.18)
CD8pCD28nCD45Ran	−0.967 (−2.209, 0.274)
CD4 T	0.005 (−0.014, 0.023)
CD8 T	−0.008 (−0.022, 0.006)
Granulocytes	0.016 (−0.014, 0.047)
Natural Killer Cells	−0.013 (−0.028, 0.002)
B cells	−0.003 (−0.012, 0.006)
Monocytes	−0.003 (−0.011, 0.004)

These models control for age, sex, and a Hispanic ancestry marker.

Appendix 2

Gene Set Overrepresentation Results for Diseases, OMIM and GLAD4U, from Illumina-Normalized Methylation Models

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
GLAD4U Database						
PA445752	Stress	616	506.36	1.13	3.8E–14	1.1E–10
PA445092	Nervous System Malformations	260	213.72	1.15	7.3E–09	1.0E–05
PA445914	Translocation, Genetic	504	414.29	1.11	1.7E–08	1.6E–05
PA443275	Adhesion	711	584.45	1.09	4.9E–08	3.4E–05

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
PA446836	Craniofacial Abnormalities	313	257.29	1.13	5.9E–08	3.4E–05
PA445000	Musculoskeletal Abnormalities	353	290.17	1.11	3.9E–07	1.9E–04
PA447208	Mental Disorders	637	523.62	1.08	1.5E–06	6.0C–04
PA4434S4	Ataxia	291	239.20	1.11	S.3E–06	1.8E–03
PA44616S	Leukemia, T-Cell	202	166.05	1.13	S.7E–06	1.8E–03
PA444944	Microcephaly	145	119.19	1.15	1.1E–05	3.1E–03
PA4436S3	Cell Transformation, Neoplastic	326	267.97	1.10	1.6E–05	4.1E–03
PA445100	Neuroblastoma	229	188.24	1.12	3.0E–05	6.6E–03
PA447278	Depression	221	181.66	1.12	3.0E–05	6.6E–03
PA4434S2	Astigmatism	52	42.74	1.22	3.7E–05	7.5E–03
PA444750	Leukemia	482	396.21	1.08	S.8E–05	LIE–02
PA445644	Shock	340	279.48	1.09	6.9E–05	1.2E–02
PA444833	Lymphatic Diseases	379	311.54	1.08	9.5E–05	1.5E–02
PA166114377	Hand-foot syndrome	60	49.32	1.20	1.1E–04	1.5E–02
PA444756	Leukemia, Lymphoid	292	240.03	1.10	1.1E–04	1.5E–02
PA165108377	Lymphoid leukemia NOS	292	240.03	1.10	1.1E–04	1.5E–02
PA444340	Lymphoma	363	298.39	1.09	1.1E–04	1.5E–02
PA165108622	Drug interaction with drug	480	394.56	1.07	1.2E–04	1–5E–02
PA4434SS	Ataxia Telangiectasia	185	152.07	1.12	1.3E–04	1.6E–02
PA443963	Dysostoses	100	82.20	1.16	1.4E–04	1.6E–02
PA165108317	Developmental disorder NOS	133	109.33	1.13	1.9E–04	2.2E–02
PA443842	Death	397	326.34	1.08	2.2E–04	2.4E–02
PA445538	Retinoblastoma	216	177.55	1.10	2.8E–04	2.7E–02
PA446858	Neurodegenerative Diseases	441	362.51	1.07	2.9E–04	2.7E–02
PA447230	HIV	826	678.98	1.05	3.0E–04	2.7E–02
PA446859	Loss of Heterozygosity	245	201.39	1.10	3.0E–04	2.7E–02
PA443269	Adenoma	185	152.07	1.11	3.1E–04	2.7E–02
PA443382	Anoxia	230	189.06	1.10	3.1E–04	2.7E–02
PA447322	Therapy-related acute myeloid leukemia (t-ML)	54	44.39	1.19	3.2E–04	2.7E–02
PA443553	Brain Diseases	480	394.56	1.07	3.2E–04	2.7E–02
PA44406S	Epilepsy	243	199.75	1.10	3.7E–04	3.0E–02
PA165108839	Psychomotor epilepsy	53	43.57	1.19	3.8E–04	3.06–02
PA443728	Chromosome Aberrations	355	291.81	1.08	4.1E–04	3.1E–02
PA446309	Lymphoma, T-Cell	190	156.18	1.11	4.1E–04	3.1E–02
PA444349	Lymphoproliferative Disorders	388	318.94	1.08	4.5E–04	3.3E–02
PA165108619	Peripheral neuroepithelioma	166	136.45	1.11	4.6E–04	3.3E–02
OMIM Database						
12S853	Diabetes mellitus, noninsulin-dependent	28	4.23	6.14	5.0E–15	5.0E–15
114480	Breast cancer	24	3.63	6.61	5.0E–15	5.0E–15

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
601626	Leukemia, acute myeloid	20	3.02	6.61	5.0E–15	5.0E–15
189960	Tracheoesophageal fistula with or without esophageal atresia	21	3.18	5.98	3.1E–14	1.2E–12
114500	Colorectal cancer	14	2.12	6.61	2.9E–12	9.1E–11
252010	Mitochondrial complex i deficiency	18	2.72	5.51	2.2E–10	5.1E–09
211980	Lung canceralveolar cell carcinoma, included	16	2.42	5.79	2.5E–10	5.1E–09
256000	Leigh syndrome	16	2.42	5.79	2.5E–10	5.1E–09
601665	Obesityleanness, included	15	2.27	5.73	1.5E–09	2.7E–08
176807	Prostate cancer	12	1.81	6.06	9.0E–09	1.4E–07
609423	Human immunodeficiency virus type 1, susceptibility	15	2.27	5.29	3.8E–08	5.2E–07
114550	Hepatocellular carcinoma	9	1.36	6.61	3.9E–08	5.2E–07
608446	Myocardial infarction, susceptibility tomyocardial infarction, susceptibility to, 1, included	13	1.97	5.60	5.1E–08	5.6E–07
600807	Asthma, susceptibility to	11	1.66	6.01	5.6E–08	5.6E–07
104300	AJiheimer disease	11	1.66	6.01	5.6E–08	5.6E–07
607948	Mycobacterium tuberculosis, susceptibility tomycobacterium tuberculosis, protection against	11	1.66	6.01	5.6E–08	5.6E–07
181500	Schizophrenia	14	2.12	5.20	2.0E–07	1.9E–06
180300	Rheumatoid arthritis	8	1.21	6.61	2.6E–07	2.3E–06
145500	Hypertension, essential	12	1.81	5.51	2.9E–07	2.4E–06
187500	Tetralogy of fallot	7	1.06	6.61	1.8E–06	1.3E–05
613065	Leukemia, acute lymphoblastic	7	1.06	6.61	1.8E–06	1.3E–05
268000	Retinitis pigmentosa	10	1.51	5.29	8.9E–06	6.5E–05
607174	Meningioma, familial, susceptibility to	6	0.91	6.61	1.2E–05	6.5E–05
144700	Renal cell carcinoma, nonpapillary	6	0.91	6.61	1.2E–05	6.5E–05
143890	Hypercholesterolemia, familial	6	0.91	6.61	1.2E–05	6.5E–05
209880	Central hypoventilation syndrome, congenital	6	0.91	6.61	1.2E–05	6.5E–05
130650	Beckwith-wiedemann syndrome	6	0.91	6.61	1.2E–05	6.5E–05
166710	Osteoporosis	6	0.91	6.61	1.2E–05	6.5E–05
171300	Pheochromocytoma	8	1.21	5.79	1.2E–05	6.5E–05
613659	Gastric cancergastric cancer, intestinal, included	8	1.21	5.79	1.2E–05	6.5E–05
611162	Malaria, susceptibility tomalaria, resistance to	17	2.57	3.89	4.1E–05	2.1E–04
217095	Conotruncal heart malformations	5	0.76	6.61	7.8E–05	3.0E–04
603932	Intervertebral disc disease	5	0.76	6.61	7.8E–05	3.0E–04
608089	Endometrial cancer	5	0.76	6.61	7.8E–05	3.0E–04
226650	Epidermolysis bullosa, junctional, non-herlitz type	5	0.76	6.61	7.8E–05	3.0E–04
240500	lmmunodeficiency, common variable, 2	5	0.76	6.61	7.8E–05	3.0E–04
609135	Aplastic anemiaaplastic anemia, susceptibility to	5	0.76	6.61	7.8E–05	3.0E–04
607785	Juvenile myelomonocytic leukemia	5	0.76	6.61	7.8E–05	3.0E–04
188470	Thyroid cancer, nonmedullary, 2	5	0.76	6.61	7.8E–05	3.0E–04

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
236670	Muscular dystrophy-dystroglycanopathy (congenital with brain and eyeanomalies), type a, 1	5	0.76	6.61	7.8E–05	3.0E–04
603896	Leukoencephalopathy with vanishing white matter	5	0.76	6.61	7.8E–05	3.0E–04
176270	Prader-willi syndrome	10	1.51	4.63	1.4E–04	5.1E–04
168600	Parkinson disease, late-onset	8	1.21	4.96	2.5E–04	8.7E–04
146110	Hypogonadotropic hypogonadism 7 with or without anosmia	14	2.12	3.78	3.4E–04	1.1E–03
255310	Myopathy, congenital, with fiber-type disproportion	6	0.91	5.51	4.1E–04	1.1E–03
254450	Myelofibrosismyelofibrosis with myeloid metaplasia	4	0.60	6.61	5.2E–04	1.1E–03
219700	Cystic fibrosis	4	0.60	6.61	5.2E–04	1.1E–03
273300	Testicular germ cell tumor	4	0.60	6.61	5.2E–04	1.1E–03
194190	WoK-Nrschhorn syndrome	4	0.60	6.61	5.2E–04	1.1E–03
136880	Fundus albipunctatusretinitis punctata albescens	4	0.60	6.61	5.2E–04	1.1E–03
209900	Bardet-biedl syndrome 1	7	1.06	4.72	1.3E–03	2.7E–03
415000	Spermatogenic failure, y–linked, 2	8	1.21	4.13	2.9E–03	4.5E–03
606963	Pulmonary disease, chronic obstructive	3	0.45	6.61	3.4E–03	4.5E–03
604571	Bare lymphocyte syndrome, type i	3	0.45	6.61	3.4E–03	4.5E–03
601410	Diabetes mellitus, transient neonatal, 1	3	0.45	6.61	3.4E–03	4.5E–03
607341	Focal cortical dysplasia, type ii	3	0.45	6.61	3.4E–03	4.5E–03
219000	Fraser syndrome 1	3	0.45	6.61	3.4E–03	4.5E–03
231200	Bernard-soulier syndrome	3	0.45	6.61	3.4E–03	4.5E–03
264350	Pseudohypoaldosteronism, type i, autosomal recessive	3	0.45	6.61	3.4E–03	4.5E–03
259500	Osteogenic sarcoma	3	0.45	6.61	3.4E–03	4.5E–03
614519	Hemorrhage, intracerebral, susceptibility to	3	0.45	6.61	3.4E–03	4.5E–03
268220	Rhabdomyosarcoma 2	3	0.45	6.61	3.4E–03	4.5E–03
608516	Major depressive disorder	3	0.45	6.61	3.4E–03	4.5E–03
231680	MULTIPLE ACYl–COa DEHYDROGENASE DEFICIENCY	3	0.45	6.61	3.4E–03	4.5E–03
109400	Basal cell nevus syndrome	3	0.45	6.61	3.4E–03	4.5E–03
604229	Anterior segment dysgenesis 5	3	0.45	6.61	3.4E–03	4.5E–03
163200	Schimmelpenning-feuerstein-mims syndrome	3	0.45	6.61	3.4E–03	4.5E–03
606864	Paraganglioma and gastric stromal sarcoma	3	0.45	6.61	3.4E–03	4.5E–03
613652	Clq deficiency	3	0.45	6.61	3.4E–03	4.5E–03
158810	Bethlem myopathy 1	3	0.45	6.61	3.4E–03	4.5E–03
602089	Hemangioma, capillary infantile	3	0.45	6.61	3.4E–03	4.5E–03
600791	Deafness, autosomal recessive 4, with enlarged vestibular aqueduct	3	0.45	6.61	3.4E–03	4.5E–03
248600	Maple syrup urine disease	3	0.45	6.61	3.4E–03	4.5E–03
304400	Deafness, x-linked 2	3	0.45	6.61	3.4E–03	4.5E–03

Gene Set	Description	Size	Expect	Ratio	P Value	FDR
145900	Hypertrophic neuropathy of dejerine-sottas	4	0.60	4.96	1.2E–02	1.5E–02
201300	Neuropathy, hereditary sensory and autonomic, type iia	4	0.60	4.96	1.2E–02	1.5E–02
105200	Amyloidosis, familial visceral	4	0.60	4.96	1.2E–02	1.5E–02
105400	Amyotrophic lateral sclerosis 1	4	0.60	4.96	1.2E–02	1.5E–02
188890	Tobacco addiction, susceptibility to	4	0.60	4.96	1.2E–02	1.5E–02
601634	Neural lube defects, folate-sensitive	4	0.60	4.96	1.2E–02	1.5E–02
60307S	Macular degeneration, age-related, 1	5	0.76	3.97	2.7E–02	3.3E–02

These are disease sets overrepresented among gene lists annotated to CpG sites that were associated with pyrethroids at a raw p<0.05. Models controlled for cell count, age, sex, education, and a genetic marker for Hispanic ancestry. Gene lists were derived from WebGestalt, with a minimum gene set size of 3. Gene set overrepresentations for diseases that use the BMIQ normalization are reported in the main findings (Figure 3).

References

1. Osteen CD, Fernandez-Cornejo J, Economic and policy issues of US agricultural pesticide use trends. Pest management science 69, 1001–1025 (2013). [PubMed: 23483682]

2. Bradberry SM, Cage SA, Proudfoot AT, Vale JA, Poisoning due to pyrethroids. Toxicological reviews 24, 93–106 (2005). [PubMed: 16180929]

3. Ray D, Sutharsan S, Forshaw P, Actions of pyrethroid insecticides on voltage-gated chloride channels in neuroblastoma cells. Neurotoxicology 18, 755 (1997). [PubMed: 9339822]

4. Hildebrand ME, McRory JE, Snutch TP, Stea A, Mammalian voltage-gated calcium channels are potently blocked by the pyrethroid insecticide allethrin. Journal of Pharmacology and Experimental Therapeutics 308, 805–813 (2004). [PubMed: 14634047]

5. Symington SB, Clark JM, Action of deltamethrin on N-type (Cav2. 2) voltage-sensitive calcium channels in rat brain. Pesticide biochemistry and physiology 82, 1–15 (2005).

6. Castellanos A. et al., Pyrethroids inhibit K2P channels and activate sensory neurons: basis of insecticide-induced paraesthesias. Pain 159, 92 (2018). [PubMed: 28937579]

7. Patel S, Pandey AK, Bajpayee M, Parmar D, Dhawan A, Cypermethrin-induced DNA damage in organs and tissues of the mouse: evidence from the comet assay. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 607, 176–183 (2006).

8. Hossain MM, Liu J, Richardson JR, Pyrethroid insecticides directly activate microglia through interaction with voltage-gated sodium channels. Toxicological Sciences 155, 112–123 (2016). [PubMed: 27655349]

9. Wu A, Li L, Liu Y, Deltamethrin induces apoptotic cell death in cultured cerebral cortical neurons. Toxicology and applied pharmacology 187, 50–57 (2003). [PubMed: 12628584]

10. Xu C. et al., Early life exposure of zebrafish (Danio rerio) to synthetic pyrethroids and their metabolites: a comparison of phenotypic and behavioral indicators and gene expression involved in the HPT axis and innate immune system. Environmental Science and Pollution Research 25, 12992–13003 (2018). [PubMed: 29480392]

11. Zhang Y. et al., Immunotoxicity of pyrethroid metabolites in an in vitro model. Environmental toxicology and chemistry 29, 2505–2510 (2010). [PubMed: 20853454]

12. Mani U. et al., Steroidogenic alterations in testes and sera of rats exposed to formulated fenvalerate by inhalation. Human & experimental toxicology 21, 593–597 (2002). [PubMed: 12507254]

13. Sakr S, Azab A, Effect of Pyrethroid Inhalation on Testis of Albino Rat. Pakistan Journal of Biological Sciences 4, 498–500 (2001).

14. Saillenfait A-M, Ndiaye D, Sabaté J-P, Pyrethroids: exposure and health effects–an update. International journal of hygiene and environmental health 218, 281–292 (2015). [PubMed: 25648288]

15. Gupta A, Nigam D, Gupta A, Shukla G, Agarwal A. (1999) Effect of pyrethroid-based liquid mosquito repellent inhalation on the blood–brain barrier function and oxidative damage in selected organs of developing rats in Journal of Applied Toxicology: An International Forum Devoted to Research and Methods Emphasizing Direct Clinical, Industrial and Environmental Applications (Wiley Online Library), pp 67–72.

16. Doi H. et al., Motor neuron disorder simulating ALS induced by chronic inhalation of pyrethroid insecticides. Neurology 67, 1894–1895 (2006). [PubMed: 17130437]

17. Furlong M, Barr DB, Wolff MS, Engel S, Prenatal exposure to pyrethroid pesticides and childhood behavior and executive functioning. Neurotoxicology 62, 231–238 (2017). [PubMed: 28811173]

18. Viel J-F et al., Behavioural disorders in 6-year-old children and pyrethroid insecticide exposure: the PELAGIE mother–child cohort. Occup Environ Med, oemed-2016–104035 (2017).

19. Dalsager L. et al., Maternal urinary concentrations of pyrethroid and chlorpyrifos metabolites and attention deficit hyperactivity disorder (ADHD) symptoms in 2–4-year-old children from the Odense Child Cohort. Environmental research, 108533 (2019). [PubMed: 31229776]

20. Oulhote Y, Bouchard MF, Urinary metabolites of organophosphate and pyrethroid pesticides and behavioral problems in Canadian children. Environmental Health Perspectives (Online) 121, 1378 (2013).

21. Furlong M. et al., Protective glove use and hygiene habits modify the associations of specific pesticides with Parkinson's disease. Environment international 75, 144–150 (2015). [PubMed: 25461423]

22. Ritz BR, Paul KC, Bronstein JM, Of pesticides and men: a California story of genes and environment in Parkinson's disease. Current environmental health reports 3, 40–52 (2016). [PubMed: 26857251]

23. Ding G. et al., Pyrethroid pesticide exposure and risk of childhood acute lymphocytic leukemia in Shanghai. Environmental science & technology 46, 13480–13487 (2012). [PubMed: 23153377]

24. Ferreira JD, Couto AC, Pombo-de-Oliveira MS, Koifman S, B. C. S. G. o. I. A. Leukemia, In utero pesticide exposure and leukemia in Brazilian children< 2 years of age. Environmental health perspectives 121, 269–275 (2012). [PubMed: 23092909]

25. VanderWeele TJ, Outcome-wide epidemiology. Epidemiology (Cambridge, Mass.) 28, 399 (2017).

26. Panni T. et al., Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. Environmental health perspectives 124, 983–990 (2016). [PubMed: 26731791]

27. Gruzieva O. et al., Epigenome-wide meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. Environmental health perspectives 125, 104–110 (2016). [PubMed: 27448387]

28. Paul KC et al., Organophosphate pesticide exposure and differential genome-wide DNA methylation. Science of the Total Environment 645, 1135–1143 (2018). [PubMed: 30248838]

29. Joubert BR et al., 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environmental health perspectives 120, 1425–1431 (2012). [PubMed: 22851337]

30. Rothman KJ, Greenland S, Lash TL, Modern epidemiology (Lippincott Williams & Wilkins, 2008).

31. Simcox NJ, Fenske RA, Wolz SA, Lee I-C, Kalman DA, Pesticides in household dust and soil: exposure pathways for children of agricultural families. Environmental health perspectives 103, 1126–1134 (1995). [PubMed: 8747019]

32. Lu C, Kedan G, Fisker-Andersen J, Kissel JC, Fenske RA, Multipathway organophosphorus pesticide exposures of preschool children living in agricultural and nonagricultural communities. Environmental research 96, 283–289 (2004). [PubMed: 15364595]

33. Kang GA et al., Clinical characteristics in early Parkinson's disease in a central California population-based study. Movement disorders: official journal of the Movement Disorder Society 20, 1133–1142 (2005). [PubMed: 15954133]

34. Horvath S, Ritz BR, Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. Aging (Albany NY) 7, 1130 (2015). [PubMed: 26655927]

35. Chen Y.-aet al., Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 8, 203–209 (2013). [PubMed: 23314698]

36. Leek JT et al., Tackling the widespread and critical impact of batch effects in high-throughput data. Nature Reviews Genetics 11, 733–739 (2010).

37. G. N. R. S. M. A. B. J. C. M. F. M. R. B, α-Synuclein Gene May Interact with Environmental Factors in Increasing Risk of Parkinson's Disease. Neuroepidemiology 35, 191–195 (2010). [PubMed: 20664293]

38. Rull RP, Ritz B, Historical pesticide exposure in California using pesticide use reports and land-use surveys: an assessment of misclassification error and bias. Environmental health perspectives 111, 1582 (2003). [PubMed: 14527836]

39. Triche TJ, Laird PW, Siegmund KD, Beta regression improves the detection of differential DNA methylation for epigenetic epidemiology. bioRxiv 10.1101/054643, 054643 (2016).

40. Liew Z, Wang A, Bronstein J, Ritz B, Job exposure matrix (JEM)-derived estimates of lifetime occupational pesticide exposure and the risk of Parkinson's disease. Archives of environmental & occupational health 69, 241–251 (2014). [PubMed: 24499252]

41. Houseman EA et al., DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics 13, 86 (2012). [PubMed: 22568884]

42. Horvath S, Levine AJ, HIV-1 infection accelerates age according to the epigenetic clock. The Journal of infectious diseases 212, 1563–1573 (2015). [PubMed: 25969563]

43. Yang J. et al., Genomic inflation factors under polygenic inheritance. European Journal of Human Genetics 19, 807 (2011). [PubMed: 21407268]

44. van Iterson M, van Zwet EW, Heijmans BT, Controlling bias and inflation in epigenome-and transcriptome-wide association studies using the empirical null distribution. Genome biology 18, 19 (2017). [PubMed: 28129774]

45. Mi H, Muruganujan A, Casagrande JT, Thomas PD, Large-scale gene function analysis with the PANTHER classification system. Nature protocols 8, 1551 (2013). [PubMed: 23868073]

46. Breckenridge CB et al., Evidence for a separate mechanism of toxicity for the Type I and the Type II pyrethroid insecticides. NeuroToxicology 30, S17–S31 (2009). [PubMed: 19766671]

47. Symington SB, Frisbie RK, Clark JM, Characterization of 11 commercial pyrethroids on the functional attributes of rat brain synaptosomes. Pesticide biochemistry and physiology 92, 61–69 (2008).

48. Wang N, Huang M, Guo X, Lin P, Urinary Metabolites of Organophosphate and Pyrethroid Pesticides and Neurobehavioral Effects in Chinese Children. Environmental Science & Technology 50, 9627–9635 (2016). [PubMed: 27524288]

49. Wagner-Schuman M. et al., Association of pyrethroid pesticide exposure with attention-deficit/ hyperactivity disorder in a nationally representative sample of US children. Environmental Health 14, 44 (2015). [PubMed: 26017680]

50. Domingues VF et al., Pyrethroid pesticide metabolite in urine and microelements in hair of children affected by autism spectrum disorders: A preliminary investigation. International journal of environmental research and public health 13, 388 (2016). [PubMed: 27482573]

51. del Pilar Navarrete-Meneses M, Pérez-Vera P, Pyrethroid pesticide exposure and hematological cancer: epidemiological, biological and molecular evidence. Reviews on environmental health 34, 197–210 (2019). [PubMed: 30903760]

52. Quist EM et al., Reevaluation of Hepatocellular Neoplasms in CD-1 Mice from a 2-year Oral Carcinogenicity Study with Permethrin. Toxicologic pathology 47, 11–17 (2019). [PubMed: 30384807]

53. Kannarkat G. et al., Common genetic variant association with altered HLA expression, synergy with pyrethroid exposure, and risk for Parkinson's disease: an observational and case–control study. NPJ Parkinson's disease 1, 15002 (2015).

54. von Ehrenstein OS et al., Prenatal and infant exposure to ambient pesticides and autism spectrum disorder in children: population based case-control study. bmj 364, l962 (2019). [PubMed: 30894343]

55. Shelton JF et al., Neurodevelopmental disorders and prenatal residential proximity to agricultural pesticides: the CHARGE study. Environmental Health Perspectives (Online) 122, 1103 (2014).

56. Alavanja MC et al., Non-hodgkin lymphoma risk and insecticide, fungicide and fumigant use in the agricultural health study. PLoS One 9, e109332 (2014). [PubMed: 25337994]

57. Boffetta P, Desai V, Exposure to permethrin and cancer risk: a systematic review. Critical reviews in toxicology 48, 433–442 (2018). [PubMed: 29687728]

58. De la Monte SM, Wands JR, Alzheimer's disease is type 3 diabetes—evidence reviewed. Journal of diabetes science and technology 2, 1101–1113 (2008). [PubMed: 19885299]

59. Kandimalla R, Thirumala V, Reddy PH, Is Alzheimer's disease a type 3 diabetes? A critical appraisal. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease 1863, 1078–1089 (2017). [PubMed: 27567931]

60. Haas JG, Lathe R, Microbes and Alzheimer's Disease: New Findings Call for a Paradigm Change. Trends in Neurosciences 41, 570–573 (2018). [PubMed: 30033181]

61. Lövheim H, Gilthorpe J, Adolfsson R, Nilsson L-G, Elgh F, Reactivated herpes simplex infection increases the risk of Alzheimer's disease. Alzheimer's & Dementia 11, 593–599 (2015).

62. Agostini S. et al., High avidity HSV-1 antibodies correlate with absence of amnestic Mild Cognitive Impairment conversion to Alzheimer's disease. Brain, behavior, and immunity 58, 254–260 (2016).

63. Readhead B. et al., Multiscale analysis of independent Alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. Neuron 99, 64–82. e67 (2018). [PubMed: 29937276]

64. Tapia-Rojas C, Inestrosa NC, Loss of canonical Wnt signaling is involved in the pathogenesis of Alzheimer's disease. Neural regeneration research 13, 1705 (2018). [PubMed: 30136680]

65. Wu J, Xia R, Shi N, Liu Y, Effects of pyrethroids on the glutamate uptake system in synaptic vesicle of rats. Wei sheng yan jiu= Journal of hygiene research 28, 261–262 (1999). [PubMed: 12712690]

66. Wang T. et al., A systematic study of normalization methods for Infinium 450K methylation data using whole-genome bisulfite sequencing data. Epigenetics 10, 662–669 (2015). [PubMed: 26036609]

67. Teschendorff AE et al., A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 29, 189–196 (2013). [PubMed: 23175756]

68. Wu MC et al., A systematic assessment of normalization approaches for the Infinium 450K methylation platform. Epigenetics 9, 318–329 (2014). [PubMed: 24241353]

69. Shiah Y-J, Fraser M, Bristow RG, Boutros PC, Comparison of pre-processing methods for Infinium HumanMethylation450 BeadChip array. Bioinformatics 33, 3151–3157 (2017). [PubMed: 28605401]

70. Soderlund DM, Molecular mechanisms of pyrethroid insecticide neurotoxicity: recent advances. Archives of Toxicology 86, 165–181 (2012). [PubMed: 21710279]

71. Ching MS et al., Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 153, 937–947 (2010).

72. Gauthier J. et al., Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. Human genetics 130, 563–573 (2011). [PubMed: 21424692]

73. Kirov G. et al., Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. Human molecular genetics 17, 458–465 (2007). [PubMed: 17989066]

74. Kirov G. et al., Neurexin 1 (NRXN1) deletions in schizophrenia. Schizophrenia bulletin 35, 851 (2009). [PubMed: 19675094]

75. Ouadid-Ahidouch H, Rodat-Despoix L, Matifat F, Morin G, Ahidouch A, DNA methylation of channel-related genes in cancers. Biochimica et Biophysica Acta (BBA)-Biomembranes 1848, 2621–2628 (2015). [PubMed: 25703813]

76. Hutnick LK et al., DNA hypomethylation restricted to the murine forebrain induces cortical degeneration and impairs postnatal neuronal maturation. Human molecular genetics 18, 2875–2888 (2009). [PubMed: 19433415]

77. Narahashi T, Frey J, Ginsburg KS, Roy M, Sodium and GABA-activated channels as the targets of pyrethroids and cyclodienes. Toxicology letters 64, 429–436 (1992). [PubMed: 1281937]

78. Furumai R, Tamada K, Liu X, Takumi T, UBE3A regulates the transcription of IRF, an antiviral immunity. Human Molecular Genetics 28, 1947–1958 (2019). [PubMed: 30690483]

79. Nguyen A, Rauch TA, Pfeifer GP, Hu VW, Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. The FASEB Journal 24, 3036–3051 (2010). [PubMed: 20375269]

80. Kashevarova AA et al., Array CGH analysis of a cohort of Russian patients with intellectual disability. Gene 536, 145–150 (2014). [PubMed: 24291026]

81. Cruz SA et al., Loss of IRF2BP2 in microglia increases inflammation and functional deficits after focal ischemic brain injury. Frontiers in cellular neuroscience 11, 201 (2017). [PubMed: 28769762]

82. Hari A. et al., IRF2BP2-deficient microglia block the anxiolytic effect of enhanced postnatal care. Scientific reports 7, 9836 (2017). [PubMed: 28852125]

83. Marcogliese PC et al., IRF2BPL is associated with neurological phenotypes. The American Journal of Human Genetics 103, 245–260 (2018). [PubMed: 30057031]

84. Cox PR, Zoghbi HY, Sequencing, expression analysis, and mapping of three unique human tropomodulin genes and their mouse orthologs. Genomics 63, 97–107 (2000). [PubMed: 10662549]

85. Hensel N, Claus P, The actin cytoskeleton in SMA and ALS: how does it contribute to motoneuron degeneration? The Neuroscientist 24, 54–72 (2018). [PubMed: 28459188]

86. McMillan EL, Kamps AL, Lake SS, Svendsen CN, Bhattacharyya A, Gene expression changes in the MAPK pathway in both Fragile X and Down syndrome human neural progenitor cells. American journal of stem cells 1, 154 (2012). [PubMed: 22822472]

87. Ronchetti D. et al., Distinct patterns of global promoter methylation in early stage chronic lymphocytic leukemia. Genes, Chromosomes and Cancer 53, 264–273 (2014). [PubMed: 24347044]

88. Hannon E, Lunnon K, Schalkwyk L, Mill J, Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. Epigenetics 10, 1024–1032 (2015). [PubMed: 26457534]

89. Singh AK et al., Long term exposure to cypermethrin induces nigrostriatal dopaminergic neurodegeneration in adult rats: postnatal exposure enhances the susceptibility during adulthood. Neurobiology of aging 33, 404–415 (2012). [PubMed: 20371137]

90. Hossain MM, Richardson JR, Mechanism of pyrethroid pesticide–induced apoptosis: role of Calpain and the ER stress pathway. Toxicological Sciences 122, 512–525 (2011). [PubMed: 21555338]

91. Hossain MM, DiCicco-Bloom E, Richardson JR, Hippocampal ER stress and learning deficits following repeated pyrethroid exposure. Toxicological Sciences 143, 220–228 (2014). [PubMed: 25359175]

92. Maurya SK, Mishra J, Abbas S, Bandyopadhyay S, Cypermethrin stimulates GSK3β-dependent Aβ and p-tau proteins and cognitive loss in young rats: Reduced HB-EGF signaling and downstream neuroinflammation as critical regulators. Molecular neurobiology 53, 968–982 (2016). [PubMed: 25575682]

93. Righi DA, Palermo-Neto J, Behavioral effects of type II pyrethroid cyhalothrin in rats. Toxicology and Applied Pharmacology 191, 167–176 (2003). [PubMed: 12946652]

94. Power MC, Adar SD, Yanosky JD, Weuve J, Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. Neurotoxicology 56, 235–253 (2016). [PubMed: 27328897]

95. Malagoli C. et al., Passive exposure to agricultural pesticides and risk of childhood leukemia in an Italian community. International journal of hygiene and environmental health 219, 742–748 (2016). [PubMed: 27693118]

96. Guha N. et al., Characterization of residential pesticide use and chemical formulations through self-report and household inventory: the Northern California Childhood Leukemia study. Environmental health perspectives 121, 276 (2013). [PubMed: 23110983]

97. Omidakhsh N. et al., Residential pesticide exposures in pregnancy and the risk of sporadic retinoblastoma: a report from the Children's Oncology Group. American journal of ophthalmology 176, 166–173 (2017). [PubMed: 28131887]

98. Lu C, Barr DB, Pearson MA, Waller LA, Dietary intake and its contribution to longitudinal organophosphorus pesticide exposure in urban/suburban children. Environmental health perspectives 116, 537 (2008). [PubMed: 18414640]

99. Wang A. et al., Parkinson's disease risk from ambient exposure to pesticides. European journal of epidemiology 26, 547–555 (2011). [PubMed: 21505849]

100. Rull RP, Ritz B, Shaw GM, Neural tube defects and maternal residential proximity to agricultural pesticide applications. American journal of epidemiology 163, 743–753 (2006). [PubMed: 16495467]

101. Clary T, Ritz B, Pancreatic cancer mortality and organochlorine pesticide exposure in California, 1989–1996. American journal of industrial medicine 43, 306–313 (2003). [PubMed: 12594778]

102. Soderlund DM et al., Mechanisms of pyrethroid neurotoxicity: implications for cumulative risk assessment. Toxicology 171, 3–59 (2002). [PubMed: 11812616]

103. Stout II DM et al., American Healthy Homes Survey: a national study of residential pesticides measured from floor wipes. Environmental science & technology 43, 4294–4300 (2009). [PubMed: 19603637]

104. Müller-Mohnssen H, Chronic sequelae and irreversible injuries following acute pyrethroid intoxication. Toxicology letters 107, 161–176 (1999). [PubMed: 10414793]

105. He F. et al., Clinical manifestations and diagnosis of acute pyrethroid poisoning. Archives of toxicology 63, 54–58 (1989). [PubMed: 2742502]

106. Chen S. et al., An epidemiological study on occupational acute pyrethroid poisoning in cotton farmers. Occupational and Environmental Medicine 48, 77–81 (1991).
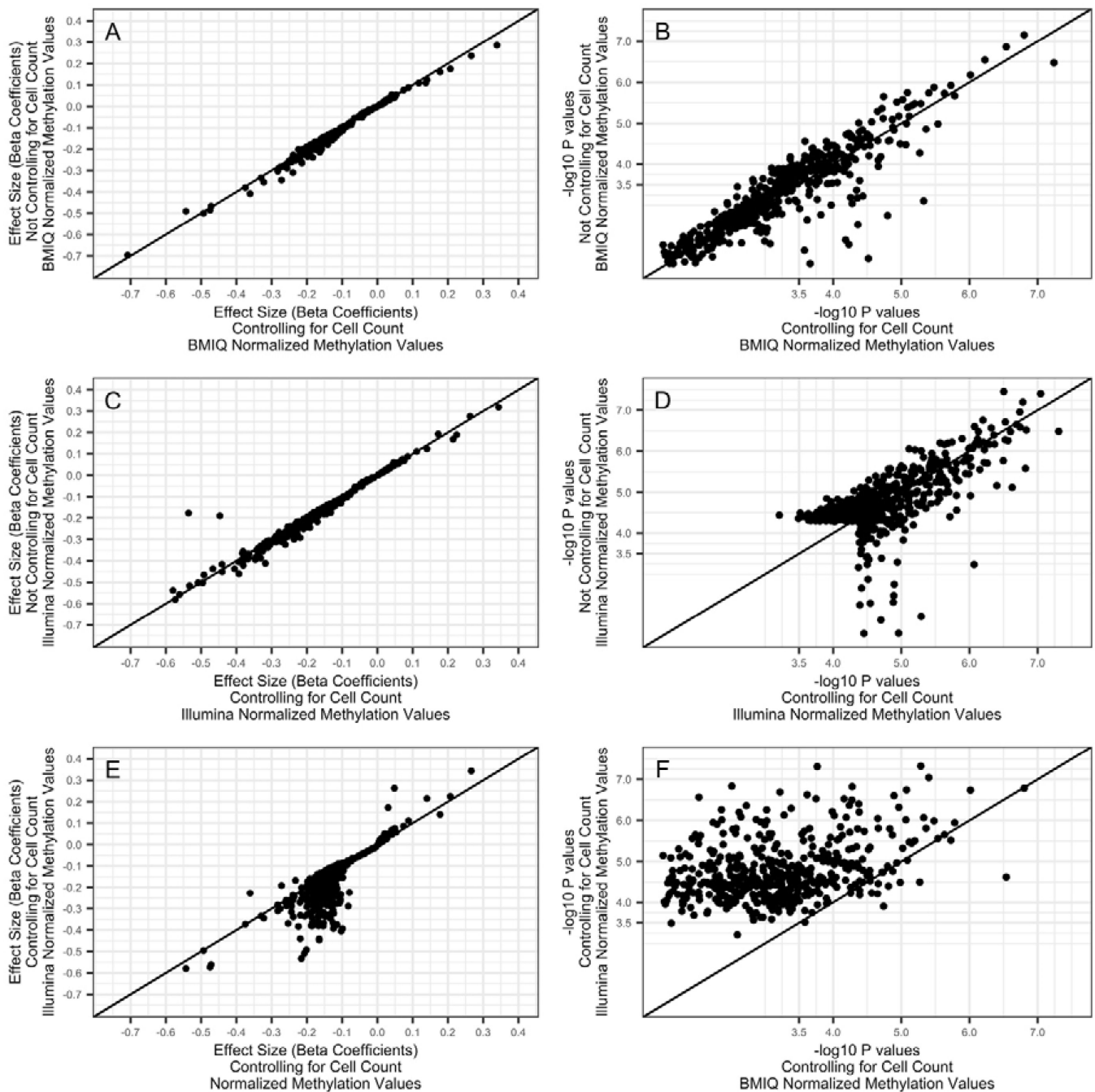
FIGURE 1. Comparing Effect Sizes and P-Values Before and After Controlling for Cell Count, and Before and After BMIQ Normalization for Type I/Type II Probe Bias

The first column (panels A, C, and E), compare beta coefficients across models, while the second column (panels B, D, and F) compare p-values. Beta coefficients represent the log(relative effect), and closely approximate an absolute change in methylation. The first row (panels A and B) compare models controlling for cell count, against models that do not control for cell count, both from models that use methylation values BMIQ normalized for Type I/II probe bias. The second row compares models controlling for cell count, against models that do not control for cell count, both from models that use the Illumina

GenomeStudio's normalized methylation values. The third row (panels E and F) compare results from models that use BMIQ methylation values, against models that use Illumina's normalized methylation values. The diagonal line in all panels falls along a slope of 1, e.g., perfect correlation. Only CpG sites that were associated at FDR p<0.05 in any model are plotted (total n=574). For visualization purposes, we exclude one CpG site that had a p-value of 2e-25 in the models using the Illumina normalized methylation values. All models control for age, sex, education, and a genetic marker for Hispanic ancestry.
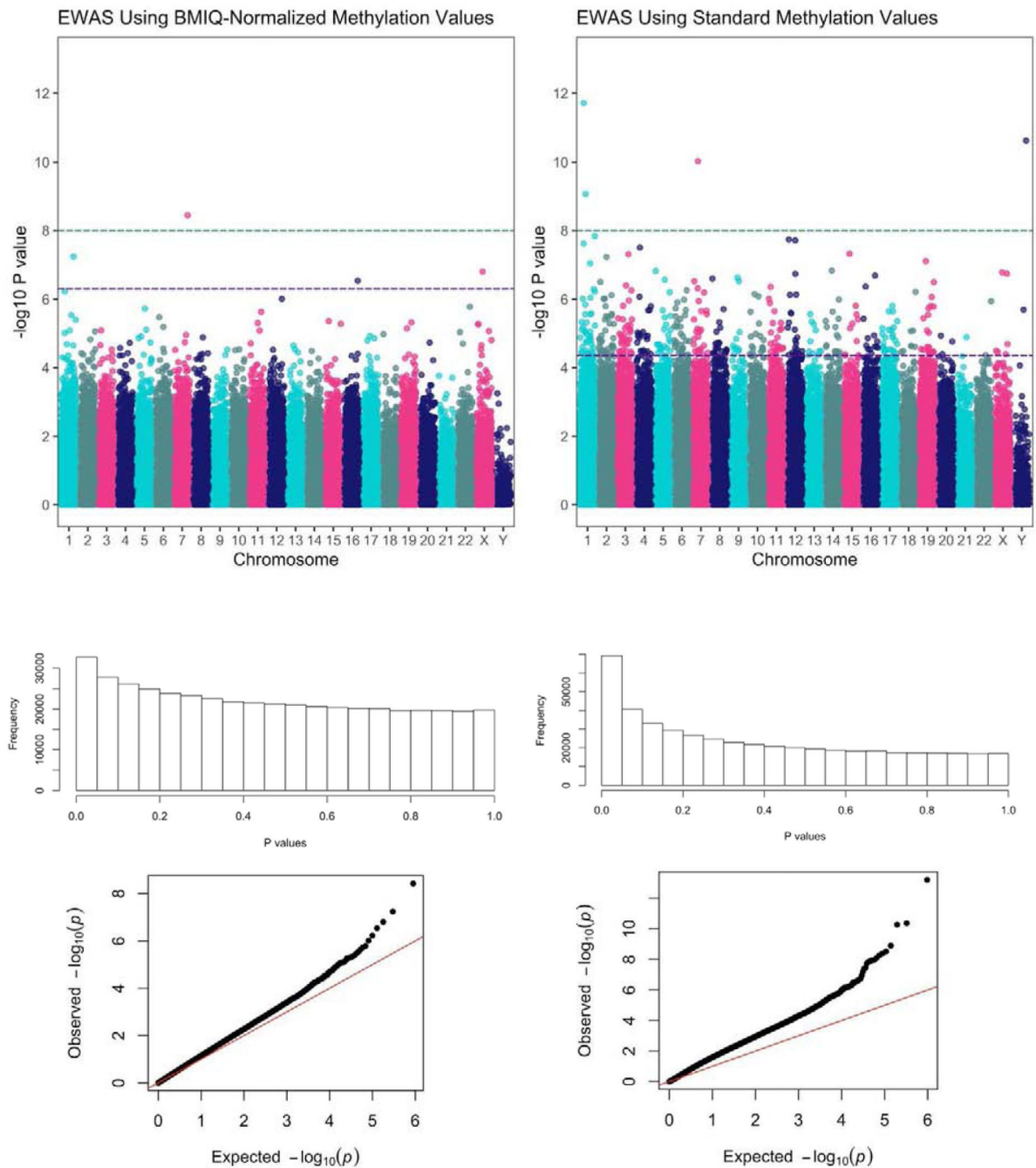
FIGURE 2. Characteristics of the EWAS: Manhattan Plots, P Value Histograms, and QQ Plots
The first column represents models using the methylation values BMIQ-normalized for Type I/II Probe Bias, and the second column represents models using the Illumina-normalized methylation values. The first row represents Manhattan plots, the second row are p value histograms, and the third row are QQ plots. In the Manhattan plot, the first dashed purple line represents the FDR cut off for significance, and the dashed green line represents the classic Bonferroni cut off for significance.
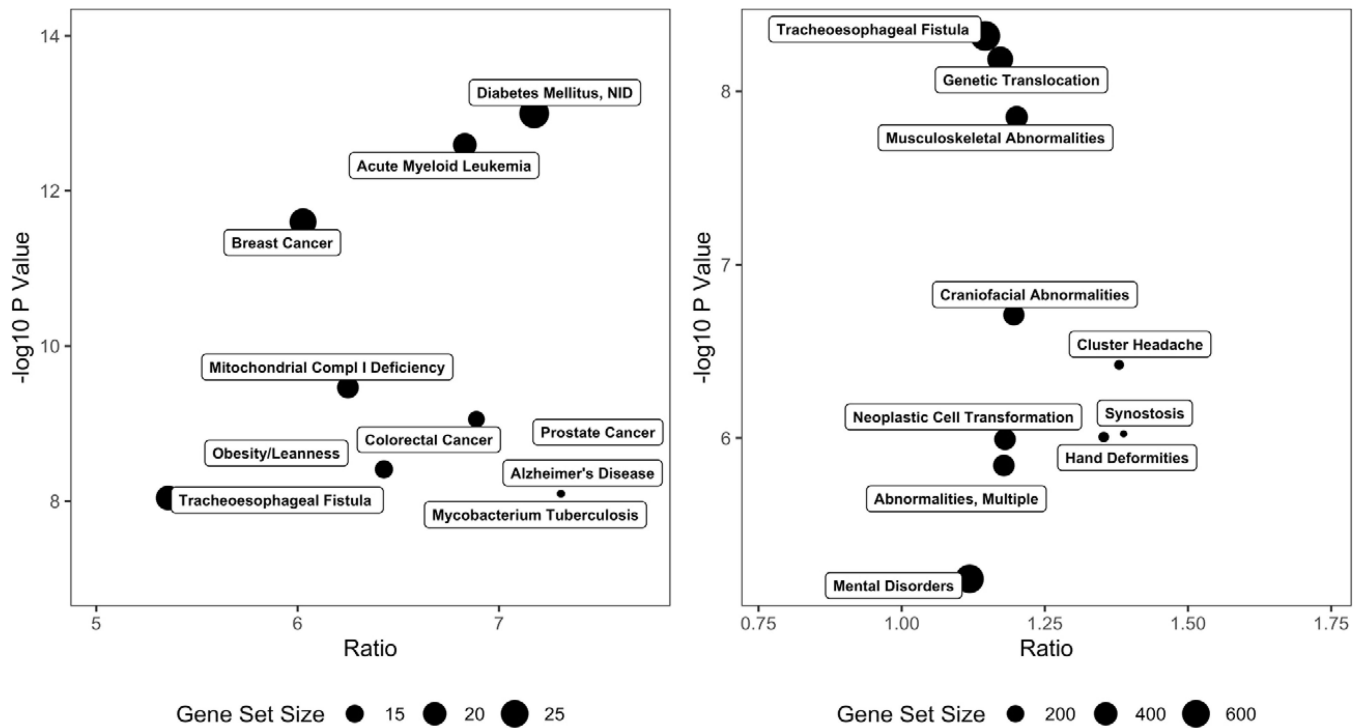
Figure 3. Volcano Plots of Disease Gene Sets, From OMIM (Left) and GLAD4U (Right) Databases

These plots show the Enrichment ratios and –log10 p values of disease gene sets from the OMIM database (left) and GLAD4U database (right). These sets are overrepresented at an FDR q<0.05 level, in the list of genes annotated to CpG sites that are associated with pyrethroid exposure at a raw p<0.05 (n=32,695 CpG sites, annotating to 5,782 unique gene names), using BMIQ normalization. The size of the markers indicates the relative size of the gene sets.

Table 1

Demographic Characteristics of the Study Participants (N = 237)

Characterstic	Study Participants
Age (mean, sd)	67.4 (12.8)
Sex (n, %)	
Male	126 (53.2)
Female	111 (46.8)
Ethnicity	
Non-Hispanic White	218 (87.3)
Hispanic White	19 (12.6)
Hispanic Ancestry Markers (mean, sd)	0.07 (0.17)
Pyrethroid Exposure	
<Median for All Pyrethroids in Past 5 Years	192 (81.0)
>Median for Any Pyrethroid in Past 5 Years	45 (19.0)
Organophosphate Exposure	
<Median for All OPs in Past 5 Years	164 (69.2)
>Median for Any OP in Past 5 Years	73 (30.8)

TABLE 2

List of CpG Sites Associated with Pyrethroid Exposures, Nearest Gene, and Description of Gene Function & Related Diseases

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	P	P	FDR q	Exp (β)	P	P	FDR q			
Associations of Pyrethroids with BMIQ Type1/Type2 Normalized Methylation Values at FDR q <0.05											
cg18234533	1.08	0.06	3.6e-9	<0.01	1.09	0.07	9.4e-11	1.1e-05		FAM20C, in gene	Binds calcium and phosphorylates proteins involved in bone mineralization. GO annotations include calcium ion binding and protein serine/threonine kinase activity. Associated diseases include Raine Syndrome and hypophosphatemia.
cg12459932	1.04	0.03	5.8e-8	0.01	1.05	0.05	8.5e-10	8.3e-05	Unclassifi ed	RUNX3, 35659	can either activate or suppress transcription. also interacts with other transcription factors. It functions as a tumor suppressor, and the gene is frequently deleted or transcriptionally silenced in cancer. GO annotations include Notch Signaling, Dendritic Cells Developmental Lineage Pathway. Diseases include testicular yolk Sac Tumor and Cleidocranial Dysplasia.
cg26578373	0.76	-0.05	1.6e-7	0.02	0.74	-0.05	1.6e-07	4.4e-03	Promoter	PIN4, 429	catalyzes the isomerization of peptidylprolyl bonds, and may play a role in the cell cycle, chromatin remodeling, and/or ribosome biogenesis. The encoded protein may play an additional role in the mitochondria. GO annotations include peptidyl-prolyl cis-trans isomerase activity. Diseases associated include Prostatic Acinar Adenocarcinoma.
cg10065825	0.70	-0.04	2.9e-7	0.03	0.80	-0.03	2.4e-05	4.3e-02		CDH11, 118278	integral membrane proteins that mediate calcium-dependent cell-cell adhesion. GO annotations include calcium ion binding. Diseases associated include Elsahy-Waters Syndrome and Aneurysmal Bone Cysts.
Associations of Pyrethroids with Illumina-Normalized Methylation Values at FDR q <0.05, Top 50 by FDR q value											
cg09545293	1.15	0.01	7.1E-02	0.71	2.40	0.06	1.9E-25	9.2E-20	Unclassifi ed Cell type specific	CNPY1 24228	expressed in the midbrain-hindbrain (MHB) boundary in zebrafish, binds FGFR1 (MIM 136350), and plays a role in FGF signaling, paralog CNPY2 regulator of neurite outgrowth by stabilizing myosin regulatory light chain (MRLC)
cg14709234	1.04	0.03	2.0E-05	0.20	1.07	0.06	1.9E-12	4.7E-07		USP24 3234	deubiquitinating enzymes. Associated with late-onset Parksinon's disease. Paralog USP32 GO terms include calcium ion binding and thiol-dependent ubiquitin-specific protease activity
cg15329860	1.0w al3	0.01	3.2E-01	0.85	1.19	0.07	2.4E-11	3.8E-06		KDM5D 216	Zinc finger domain associated with chromatin organization.. Diseases include spermatogenic failure, y-linked, y chromosome infertility. Paralog KDM5C

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	P	P	FDR q	Exp (β)	P	p	FDR q			
											associated with x linked mental retardation, alacrima, achalasia
cg18234533	1.08	0.06	3.6E-09	<0.01	1.09	0.07	9.4E-11	1.1E-05		FAM20C in gene	Binds calcium and phosphorylates proteins involved in bone mineralization. GO annotations include calcium ion binding and protein serine/threonine kinase activity. Associated diseases include Raine Syndrome and hypophosphatemia.
cg12459932	1.04	0.03	5.8E-08	0.01	1.05	0.05	8.5E-10	8.3E-05	Unclassified	RUNX3 35659	can either activate or suppress transcription. also interacts with other transcription factors. It functions as a tumor suppressor, and the gene is afrequently deleted or transcriptionally silenced in cancer. GO annotations include Notch Signaling, Dendritic Cells Developmental Lineage Pathway. Diseases include testicular yolk Sac Tumor and Cleidocranial Dysplasia.
cg21690979	0.81	−0.00	8.1E-06	0.15	0.71	−0.00	1.5E-08	1.2E-03	Promoter	IRF2BP2 in gene	Regulates macrophage polarization & inflammatory response. Associated with Immunodeficiency & Mesenchymal Condrosarcoma. Paralog IRF2BPL associated with Neurodevelopmental Disorder with Regression, Abnormal Movements, Loss of Speech, and Seizures
cg27526429	NA	NA	NA	NA	1.07	0.05	1.8E-08	1.2E-03		INHBC 48565	TGF-Beta superfamily, Inhibits activin a signaling. Transgenic mice exhibit deficits in testis, liver, prostate. Regulates secretion of follitropin by pituitary gland. Involved in hormone secretion, germ cell development, erythroid differentiation, nerve cell survival, embyonic axial development/bone growth.
cg02150435	0.80	−0.01	2.9E-05	0.23	0.64	−0.01	1.9E-08	1.2E-03	Promoter	LOC374443; CLEC2D 21861	Natural killer cell receptor c-type lectin family. Inhibits osteoclast formation, contains transmembrane domain near N-terminus. GO terms include transmembrane signaling receptor activity, carbohydrate binding. Paralog CLEC2A associated with malignant syringoma and eccrine sweat gland cancer.
cg11196904	0.83	−0.01	5.9E-07	0.05	0.78	−0.01	2.4E-08	1.3E-03	Promoter	APITD1 in gene	Read-through transcription between APITD1 and CORT genes. Paralog APITD1/CENPS, is in neuroblastoma tumor suppressor candidate region. Has a role in cell death pathway. Associated diseases include limited scleroderma, Fanconi anemia. CORT binds to all somatostatin receptors, depresses neuronal activity, regulates slow-wave sleep, locomotor activity. Diseases include Pituitary Adenoma, Intestinal Disease.
cg05991454	1.40	0.07	1.5E-04	0.34	1.73	0.16	3.1E-08	1.5E-03		POU4F2 1857	POU-domain transcription factor family and may be involved in maintaining visual system neurons in the retina. Also associated with breast cancer and

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	P	P	FDR q	Exp (β)	P	p	FDR q			
											accelerated tumor growth, differentiating neuroblastoma, and optic nerve disease. Also a DNA-binding transcription factor activity, and rNA polymerase II proximal promoter sequence-specific DNA binding. Paralog POU4F1 associated with growth of cervical tumors, developing sensory nervous system, optic neuropathy, and glaucoma.
cg16275172	0.78	−0.01	5.2E-06	0.13	0.77	−0.02	4.8E-08	2.0E-03	Promoter	TMOD3 21285	Homolog TMOD2 associated with Amyotrophic Lateral Sclerosis 1
cg26367456	1.02	0.01	1.7E-s04	0.35	1.03	0.03	4.9E-08	2.0E-03		ALCAM in gene	activated leukocyte cell adhesion molecule. Associated with stork bite, melanoma. Cell adhesion molecules, L1CAM interactions. GO annotations include signaling receptor binding. Paralog BCAM associated with Lutheran Null and Sickle Cell disease.
cg15037358	NA	−0.01	NA	NA	1.05	0.04	5.9E-08	2.2E-03		NRXN1 123537	Neurexin 1 encodes a single-pass type I membrane protein. Cell-surface receptors bind neuroligins to form Ca(2+)-dependent neurexin/neuroligin complexes at synapses in the central nervous system. Required for efficient neurotransmission. Involved in the formation of synaptic contacts. Associated with Pitt-Hopkins-like syndrome-2 (severe mental retardation, speech development, autistic behavior, breathing abnormalities, decreased reflexes, protruding tongue with drooling), Chromosome 2p16.3 deletion syndrome (Schizophrenia 17). Paralog NRXN2 associated with Leukocyte Adhesion Deficiency and Pancreatic Gastrinoma. Pathways include transmission across chemical synapses, and muscular dystrophies and dystrophin-glycoprotein complex. GO annotations include transmembrane signaling receptor activity and calcium channel regulator activity.
cg22460206	0.90	−0.01	1.4E-01	0.76	0.67	−0.02	7.8E-08	2.7E-03	Unclassifi ed	ZNF527 3393	associated with HSV1, nucleic acid binding.
cg06501070	1.02	0.02	3.9E-06	0.13	1.03	0.02	9.0E-08	2.9E-03		LPAR3 96	G protein-coupled receptor family, as well as the EDG family of proteins. This protein functions as a cellular receptor for lysophosphatidic acid and mediates lysophosphatidic acid-evoked calcium mobilization. Associated with ovarian cancer, triple-receptor negative breast cancer. paralog LPAR2 associated with pulmonary fibrosis and spinal stenosis.
cg10537269	0.84	−0.01	3.1E-03	0.50	0.73	−0.02	1.5E-07	4.3E-03	Promoter	KIAA0586; TIMM9 569	Associated with Joubert syndrome 23 (includes cerebellar ataxia, oculomotor apraxia, hypotonia, psychomotor delay, neonatal breathing abnormalities, also renal disease and polydactyly). Also associated with Short-rib thoracic dysplasia 14 with polydactyl:

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	P	P	FDR q	Exp (β)	P	p	FDR q			
											short ribs & tubular bones, polydactyly, cleft palate, major organ anomalies. Sometimes lethal neonatally.
cg23958684	1.05	0.04	5.3E–05	0.25	1.06	0.05	1.5E–07	4.3E–03		MCC 1444	candidate colorectal tumor suppressor gene that is thought to negatively regulate cell cycle progression. Associated with familial adenomatous polyposis. GO annotations include calcium ion binding. Paralog USHBP1 associated with Usher Syndrome.
cg26578373	0.76	–0.05	1.6E–07	0.02	0.74	–0.05	1.6E–07	4.4E–03	Promoter	PIN4 429	catalyzes the isomerization of peptidylprolyl bonds, and may play a role in the cell cycle, chromatin remodeling, and/or ribosome biogenesis. The encoded protein may play an additional role in the mitochondria. GO annotations include peptidyl-prolyl cis-trans isomerase activity. Diseases associated include Prostatic Acinar Adenocarcinoma.
cg22009709	0.49	–0.14	8.4E–06	0.15	0.46	–0.05	1.8E–07	4.4E–03		PIN4 in gene	catalyzes the isomerization of peptidylprolyl bonds, and may play a role in the cell cycle, chromatin remodeling, and/or ribosome biogenesis. The encoded protein may play an additional role in the mitochondria. GO annotations include peptidyl-prolyl cis-trans isomerase activity. Diseases associated include Prostatic Acinar Adenocarcinoma.
cg07069406	0.76	–0.01	9.7E–07	0.07	0.74	–0.02	1.8E–07	4.4E–03		ZNF664;CDC92 29722	CCDC92 associated with lipdystrophy (abnormal distribution of fat and cardiovascular problems), and Senior-Loken Syndrome (cycstic kidney disease, also causes retinal dystrophy)
cg00843885	0.83	–0.00	6.0E–04	0.42	0.83	–0.01	2.0E–07	4.7E–03	Promoter	TAOK2 2595	Paralog TAOK1 associated with chromosome 17q11.2 deletion syndrome, "NF1 microdeletion syndrome": facial dysmorphism, mental retardation, early-onset neurofibromas, risk of malignant peripheral nerve sheath tumors
cg13593530	1.05	0.04	2.4E–04	0.37	1.08	0.06	2.4E–07	5.2E–03		AGPAT2 9425	converts lysophosphatidic acid to phosphatidic acid, the second step in de novo phospholipid biosynthesis. Mutations in this gene have been associated with congenital generalized lipodystrophy (CGL), or Berardinelli-Seip syndrome, a disease characterized by a near absence of adipose tissue and severe insulin resistance. GO annotations include transferase activity, transferring acyl groups, 1-acylglycerol-3-phosphate o-acyltransferase activity.
cg13551243	0.83	–0.01	1.3E–05	0.17	0.85	–0.01	2.5E–07	5.3E–03	Promoter	ZNF7 1730	associated with HSV1. gO annotations include nucleic acid binding and dNA-binding transcription factor activity.

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	β	P	FDR q	Exp (β)	β	p	FDR q			
cg25575520	1.02	0.02	9.3E-03	0.56	1.04	0.03	2.7E-07	5.5E-03		MIER3 3286	transcriptional repressor. GO terms include chromatin binding. Paralog associated with hairy tongue and amelogenesis imperfecta (disorder characterized by hypoplastic enamal, delayed/failed secondary dentition). Amelogenesis imperfecta also associated with breast cancer. FAM20A also associated with this gene.
cg14010814	1.02	0.02	1.8E-04	0.36	1.04	0.03	3.0E-07	5.5E-03		DBC1 in gene	associated with bladder cancer.
cg21442556	1.05	0.03	6.9E-05	0.27	1.07	0.05	3.0E-07	5.5E-03		MIER3 10150	actin polymerization in cells, mediates formation of branched actin networks, Paralog MIER1 has associated diseases include fanconi anemia, Scapuloperoneal myopathy.
cg23975564	NA	NA	NA	NA	1.05	0.04	3.1E-07	5.5E-03		GCFC2 161704	transcriptional repressor. Associated with dyslexia, reading disorder. GO annotations include DNA binding transcription factor activity and proximal promoter DNA binding transcription repressor activity, RNA polymerase II-specific.
cg0262189	0.87	-0.01	5.3E-05	0.25	0.89	-0.01	3.2E-07	5.5E-03	Promoter	SERTAD1 1382	Developmental and pathological neuron death; required to initiate apoptotic cell cycle pathway; mood disorders, prostate cancer progression by binding androgen receptor ligand, cancer. Paralog SERTAD3 also associated with tumor growth
cg09726198	0.85	-0.11	4.2E-05	0.25	0.64	-0.01	4.0E-07	6.6E-03	Unclassified	PDHB 29433	mitochondrial multienzyme complex that catalyzes the overall conversion of pyruvate to acetyl-CoA and carbon dioxide, and provides the primary link between glycolysis and the tricarboxylic acid (TCA) cycle. Associated with Pyruvate dehydrogenase E1-Beta deficiency (a metabolic disease characterized by buildup of lactic acid, and variety of neurological problems)
cg02507033	0.86	-0.01	5.8E-05	0.25	0.88	-0.01	4.3E-07	6.8E-03	Promoter	C16orf75/RMI2 in gene	essential for genome stability. Associated diseases include microcephaly, increased sister chromatid exchange 2, malignant ovarian cyst.
cg27207764	1.03	0.02	1.9E-03	0.47	1.04	0.03	4.4E-07	6.8E-03		RTN3 in gene	Reticulon family expressed in neuroendocrine tissues. Associated with extraskeletal chondrome and Alzheimer's disease. Associated with protein-protein interactions at synapses and transmission across chemical synapses.
cg08876591	0.82	-0.01	1.1E-05	0.16	0.79	-0.02	4.8E-07	7.3E-03	Promoter	ADAM22 in gene	Essential for correct myelination in peripheral nervous system. Early infantile epileptic encephalopathy, brachydactyly
cg13071672	0.82	-0.01	4.1E-04	0.40	0.78	-0.02	5.1E-07	7.5E-03	Promoter	DYRK3 1169	catalyze autophosphorylation on serine/threonine and tyrosine residues. Associated diseases include neuroaspergillosis and mental retardation.

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	β	P	FDR q	Exp (β)	β	p	FDR q			
cg22226973	0.82	−0.02	4.4E−03	0.52	0.61	−0.03	5.5E−07	7.7E−03	Promoter	CDKAL1 11823	associated with gene expression, trNA processing, GO annotations include transferase activity and iron-sulfur cluster binding. Associated diseases include Diabetes.
cg18853071	0.84	−0.00	9.5E−04	0.44	0.83	−0.01	5.5E−07	7.7E−03	Promoter	UBE2E1 1266	ubiquitin is an important cellular mechanism for targeting abnormal or short-lived proteins for degradation. Associated with spinocerebellar Ataxia. GO annotations include ligase activity, acid-amino acid ligase activity. Paralog UBE2E2 associated with late onset Parkinson's disease.
ch.1.792811 77F	NA		NA	NA	0.83	−0.01	6.0E−07	7.9E−03		ADGRL4 36250	Endothelial orphan receptor that acts as a key regulator of angiogenesis. Associated with epilepsy. GO annotations include G protein-coupled receptor activity and transmembrane signaling receptor activity
cg00572356	0.81	−0.02	1.0E−04	0.31	0.78	−0.02	6.1E−07	7.9E−03	Promoter	ESCO1 12442	acetyltransferases involved in sister chromatid cohesion. Associated diseases include Sc Phocomelia syndrome and Roberts Syndrome (bone disorders characterized by limb and facial abnormalities, some mental retardation). GO annotations include transferase activity, transferring acyl groups.
cg23290343	0.85	−0.01	4.1E−05	0.25	0.86	−0.01	6.2E−07	7.9E−03		FBXL17 in gene	protein-ubiquitin ligases. Paralog FBXL19 associated with psoriasis.
cg26308909	1.01	0.01	2.4E−03	0.48	1.02	0.02	6.3E−07	7.9E−03		MAD1L1 in gene	MAD1L1 may play a role in cell cycle control and tumor suppression. Associated with prostate cancer, tertroperitoneum carcinoma.
cg16006296	NA		NA	NA	1.03	0.02	7.0E−07	8.6E−03		LOC399744, ZNF37A 330885	gene expression, HSV1. GO annotations include nucelic acid binding and DNA binding transcription factor activity.
cg20162066	NA		NA	NA	1.04	0.03	7.3E−07	8.6E−03		OSBPL8 in gene	Lipid transporter between endoplasmic reticulum and plasma membrane. It binds mutliple lipid-containing molecules, including phosphatidylserine, phosphatidylinositol 4-phosphate (PI4P) and oxysterol, and promotes their exchange between the endoplasmic reticulum and the plasma membrane. Associated with short-rib thoracic dysplasia (skeletal cilipathy). GO annotations include cholesterol binding. Paralog OSBPL5 associated with Beckwith-wiedemann syndrome (Growth disorder with higher risk of tumors).
cg14950855	0.87	−0.00	5.2E−04	0.41	0.76	−0.00	7.4E−07	8.6E−03		DNAJC22 1874	heat shock protein. Associated with sialolithiasis (disorder of salivary gland) and patent foramen ovale (cardiovascular defect). GO annotations include unfolded protein binding and chaperone binding.

CpG Site	BMIQ Normalized Methylation Values				Illumina GenomeStudio Normalized Methylation Values				Reg Element	Nearest Gene & nearest |bp distance|	Description of nearest Gene from Gene Card
	Exp (β)	P	P	FDR q	Exp (β)	P	p	FDR q			
cg13522118	1.03	0.02	9.7E–05	0.31	1.04	0.03	7.6E–07	8.6E–03		RALB in gene	GTP-binding protein. Associated with pancreatic cancer and epileptic encephalopathy. Required for suppression of apoptosis. GO annotations include GTP binding and ubiquitin protein ligase binding. Paralog RALA associated with laryngeal tuberculosis and pancreatic cancer.
cg13525697	1.02	0.02	2.3E–03	0.48	1.03	0.03	8.2E–07	9.0E–03		SLC16A9 91538	Solute Carrier Family 16 Member 9. Proton linked monocarboxylate transporter. Associated with gout. Paralog SLC16A14 associated with allan-herndon-dudley syndrome (neurological disorder resulting in severe intellectual disability and problems with movement)
cg03373646	0.77	–0.02	6.4E–05	0.26	0.76	–0.03	8.4E–07	9.0E–03	Unclassified Cell type specific	CRIPAK 14620	negative regulator of PAK1. Associated with infantile hypophosphatasia (severe bone disease characterized by defective skeletal mineralition), and laryngeal benign neoplasm.
cg10130497	0.98	–0.02	4.7E–06	0.13	0.98	–0.02	8.6E–07	9.0E–03		ARRDC5 90	Paralog ARRDC4 associated with mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes. Plays a role in endocytosis of activated G protein-coupled receptors
cg12742796	0.83	–0.03	2.2E–05	0.20	0.75	–0.03	9.0E–07	9.3E–03	Unclassified	TTF2 128	Transcription termination factor. dsDNA-dependent ATPase activity and RNA polymerase II termination activity. Associated diseases include sexual sadism and cornea plana. GO annotations include hydrolase activity and DNA-dependent ATP-ase activity. Paralog HLTF associated with schinke immunoosseous dysplasia (short stature, kidney disease, weakened immune system) and colorectal cancer.
cg12218368	0.87	–0.00	1.4E–03	0.46	0.84	–0.01	9.5E–07	9.3E–03	Promoter	USP39 224	ubiquitin specific peptidase 39. associated with miyoshi muscular dystrophy. GO annotations include thiol-dependent ibiquitinyl hydrolase activity. Paralog USP46 associated with tracheal cancer.
cg05019221	1.01	0.01	8.6E–06	0.15	1.02	0.02	9.7E–07	9.3E–03		OTOS 104362	Otospiralin is synthesized by nonsensory cells (fibrocytes) of the inner ear, and downregulation of otospiralin in guinea pigs leads to deafness. Associated with deafness.
cg15429732	0.88	–0.01	4.3E–03	0.52	0.69	–0.03	9.7E–07	9.3E–03	Promoter	ZFPL1;CDCA5 213	zinc finger protein. Required for cis-Golgi integrity and efficient ER to Golgi transport. Involved in the maintenance of the integrity of the cis-Golgi, possibly via its interaction with GOLGA2/GM130

These results are from the model that controls for sex, age, education, and Hispanic ancestry. Exp(B) reflects the relative effect of pyrethroid exposure derived from using a log link in the beta regression model. P is the absolute difference derived from lsmeans of the beta regression model. P is the raw p-value. FDR Q refers to the FDR adjusted p-value. Nearest gene is either the annotation from Illumina,

or in the event a gene is not provided from Illumina, it is the annotation from the hiAnnotator package for the nearest gene. The gene description and associated diseases from gene card include a short synopsis of the GeneCard description and associated diseases for the identified gene and for paralogs, if relevant.

Table 3

Gene Set Overrepresentation Analysis of Pyrethroid-Associated CpG Sites, GO Biological Processes

ONTOLOGY	TERM	N	DE	P.DE	FDR
BMIQ Normalized Methylation					
CC	membrane	1853	1313.3	8.0E–08	7.0E–04
CC	nucleoplasm	3123	2164.8	1.4E–06	7.9E–03
CC	cytosol	4871	3288.1	1.3E–05	3.9E–02
MF	protein binding	9664	6483.3	2.0E–12	3.6E–08
MF	ATP binding	1451	1041.5	4.0E–06	1.8E–02
MF	transcription factor binding	264	213.5	5.3E–06	1.9E–02
Illumina Normalized Methylation					
BP	negative regulation of transcription by RNA polymerase II	707	655.5	9.2E–12	1.8E–08
BP	positive regulation of transcription, DNA-templated	535	497.0	4.8E–10	6.5E–07
BP	positive regulation of transcription by RNA polymerase II	975	872.0	7.2E–8	6.3E–05
BP	protein ubiquitination	419	384.0	1.3E–07	1.1E–04
BP	cellular response to DNA damage stimulus	224	212.5	3.2E–07	2.1E–04
BP	viral process	366	337.0	3.7E–07	2.4E–04
BP	proteasome-mediated ubiquitin-dependent protein catabolic process	133	129.5	1.5E–06	9.0E–04
BP	cell division	344	316.0	1.8E–06	9.3E–04
BP	mRNA splicing, via spliceosome	224	210.0	1.8E–06	9.3E–04
BP	neural tube closure	76	76.0	4.1E–06	2.0E–03
BP	in utero embryonic development	168	160.0	1.1E–05	4.6E–03
BP	protein polyubiquitination	206	192.5	1.8E–05	7.1E–03
BP	mitotic cell cycle	90	88.0	1.9E–05	7.5E–03
BP	protein deubiquitination	236	218.5	3.3E–05	1.2E–02
BP	DNA repair	194	179.5	6.8E–05	2.2E–02
BP	translational initiation	130	122.5	1.0E–04	3.0E–02
BP	embryonic digit morphogenesis	56	56.0	1.4E–04	3.8E–02
BP	cilium assembly	152	142.0	1.4E–04	3.8E–02
BP	RNA splicing	143	134.5	1.5E–04	4.2E–02
BP	Golgi organization	65	64.0	1.6E–04	4.2E–02
CC	nucleoplasm	3123	2860.8	2.2E–47	1.9E–43
CC	cytosol	4871	4293.7	1.6E–31	9.3E–28
CC	membrane	1853	1692.8	4.8E–26	2.1E–22
CC	nucleus	4561	3946.8	6.1E–19	1.8E–15
CC	mitochondrion	1106	996.0	4.9E–14	1.2E–10
CC	centrosome	447	424.5	5.4E–14	1.2E–10
CC	nuclear speck	386	363.5	1.6E–10	2.6E–07
CC	Golgi apparatus	820	741.0	3.7E–10	5.4E–07
CC	protein-containing complex	592	535.0	6.2E–08	5.8E–05

ONTOLOGY	TERM	N	DE	P.DE	FDR
CC	mitochondrial matrix	349	321.0	1.7E–07	1.3E–04
CC	cytoplasm	3433	2918.7	7.1E–07	4.5E–04
CC	perinuclear region of cytoplasm	684	611.2	1.5E–06	9.0E–04
CC	microtubule	254	238.0	1.7E–06	9.3E–04
CC	glutamatergic synapse	346	321.5	1.8E–06	9.3E–04
CC	focal adhesion	397	365.5	2.5E–06	1.2E–03
CC	late endosome	120	116.5	4.8E–06	2.2E–03
CC	lysosomal membrane	285	263.0	5.2E–06	2.3E–03
CC	nuclear body	277	256.5	6.9E–06	3.0E–03
CC	endoplasmic reticulum	850	743.6	8.3E–06	3.6E–03
CC	nuclear chromatin	167	159.0	1.2E–05	4.8E–03
CC	Golgi membrane	585	522.0	2.0E–05	7.7E–03
CC	intracellular membrane-bounded organelle	579	514.0	5.1E–05	1.8E–02
CC	dendrite	331	303.0	5.9E–05	2.0E–02
CC	spindle	114	109.0	9.7E–05	3.0E–02
CC	nuclear membrane	234	215.5	1.1E–04	3.3E–02
CC	microtubule cytoskeleton	122	116.0	1.4E–04	3.8E–02
CC	lamellipodium	175	164.0	1.4E–04	3.8E–02
CC	cilium	187	172.0	1.6E–04	4.3E-02
MF	protein binding	9664	8427.8	1.0E–58	1.8E–54
MF	RNA binding	1334	1229.6	3.4E–24	1.2E–20
MF	ATP binding	1451	1295.2	3.0E–11	5.3E–08
MF	cadherin binding	280	267.2	1.5E–09	1.8E–06
MF	protein kinase binding	401	375.5	3.5E–09	4.1E–06
MF	DNA-binding transcription factor activity	480	444.0	8.6E–09	9.4E–06
MF	enzyme binding	345	323.5	1.3E–08	1.4E–05
MF	transcription factor binding	264	251.5	6.3E-08	5.8E–05
MF	identical protein binding	1041	919.1	2.2E–07	1.7E–04
MF	protein C-terminus binding	186	179.0	2.4E–07	1.7E–04
MF	RNA polymerase II proximal promoter sequence-specific DNA binding	379	352.5	2.4E–07	1.7E–04
MF	DNA-binding transcription repressor activity, RNA polymerase II-specific	224	212.0	2.2E–06	1.1E–03
MF	histone deacetylase binding	102	99.0	2.1E–05	7.7E–03
MF	protein domain specific binding	255	235.0	3.4E–05	1.2E–02
MF	double-stranded DNA binding	72	71.0	4.2E–05	1.5E–02
MF	DNA-binding transcription activator activity, RNA polymerase II-specific	390	354.0	6.5E–05	2.2E–02
MF	transcription corepressor activity	205	191.5	8.4E–05	2.7E–02
MF	DNA-binding transcription factor activity, RNA polymerase II-specific	1537	1330.5	9.0E–05	2.8E–02
MF	ubiquitin-protein transferase activity	224	206.0	1.7E–04	4.5E–02

These overrepresentation pathways were derived from the gometh function in the missMethyl R package, after accounting for prior probabilities. The input list includes the CpG sites that were associated with pyrethroids at raw p<0.05 (N=32,695 CpG sites for BMIQ normalized models and

N=69,180 for Illumina normalized models). These estimates were from the EWAS beta regression model that controlled for age, sex, education, cell count, and an ancestral marker for Hispanic ancestry. Ontology codes are BP=biological process, CC=cellular component, MF= molecular function. N is number of genes in the GO term, DE is the number of genes that are differentially methylated, P.DE is the p-value for over-representation of the GO term, and FDR is the False discovery rate q value.