bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Long-term effects of wildfire smoke exposure during early-life on the nasal epigenome in rhesus macaques

Anthony P. Brown[1], Lucy Cai[1], Benjamin I. Laufer[2], Lisa A. Miller[1,3], Janine M. LaSalle[2], Hong Ji[1,3*]

Affiliations:

[1]California National Primate Research Center, Davis, CA 95616, USA; [2]Department of Medical Microbiology and Immunology, MIND Institute, Genome Center, University of California, Davis, CA 95616, USA. [3]Department of Anatomy, Physiology and Cell biology, School of Veterinary Medicine, University of California, Davis, CA 95616, USA

***Corresponding author**: Hong Ji, PhD, Department of Anatomy, Physiology and Cell Biology, School of Veterinary Medicine, California National Primate Research Center, University of California, Davis, CA, USA. Phone: 530-754-0679. Email: hgji@ucdavis.edu

Running title: DNA methylation and gene expression changes associated with early-life exposure to wildfire smoke in rhesus monkeys

Abstract Word Counts: 271; Total Word Count: 5442

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Background

Wildfire smoke is responsible for around 20% of all particulate emissions in the U.S. and affects millions of people worldwide. Children are especially vulnerable, as ambient air pollution exposure during early childhood is associated with reduced lung function. Most studies, however, have focused on the short-term impacts of wildfire smoke exposures. We aimed to identify long-term baseline epigenetic changes associated with early-life exposure to wildfire smoke. We collected nasal epithelium samples for whole genome bisulfite sequencing (WGBS) from two groups of adult female rhesus macaques: one group born just before the 2008 California wildfire season and exposed to wildfire smoke during early-life (n = 8), and the other group born in 2009 with no wildfire smoke exposure during early-life (n = 14). RNA-sequencing was also performed on a subset of these samples.

Results

We identified 3370 differentially methylated regions (DMRs) (difference in methylation ≥ 5% empirical $p < 0.05$) and 1 differentially expressed gene (*FLOT2*) (FDR< 0.05, fold of change ≥ 1.2). The DMRs were annotated to genes significantly enriched for synaptogenesis signaling, protein kinase A signaling, and a variety of immune processes, and some DMRs significantly correlated with gene expression differences. DMRs were also significantly enriched within regions of bivalent chromatin (top odds ratio = 1.46, q-value < 3×10^{-6}) that often silence key developmental genes while keeping them poised for activation in pluripotent cells.

Conclusions

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

These data suggest that early-life exposure to wildfire smoke leads to long-term changes in the methylome over genes impacting the nervous and immune systems, but follow-up studies will be required to test whether these changes influence transcription following an immune/respiratory challenge.

Keywords: Wildfire smoke, whole genome bisulfite sequencing, RNA-sequencing, rhesus macaques, early life

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

BACKGROUND

According to the National Interagency Fire Center, there were 50,477 wildfires (4.7 million acres burned) in the United States in 2019. In total, 212 million Americans lived in counties affected by wildfires in 2011 (1). These wildfires have contributed to levels of air pollution in the United States that have been linked to premature death (2-5). About 20% of all fine particulate emissions in the U.S. are from wildfire smoke, while half of all particulate matter less than 2.5 μm in diameter ($PM_{2.5}$) in California resulted from wildfires (2). $PM_{2.5}$ are especially harmful, as these particles are able to penetrate the respiratory system and the lungs (2). Exposure to these particles has been associated with asthma, bronchitis, lung cancer, and cardiovascular disease (3-5). Young children are especially vulnerable to these negative health effects, as studies have linked air pollution exposure in children to reduced lung function (6, 7), reduced height-for-age (8), increased blood pressure (9), and an increased risk of developing asthma and eczema (10). Most of these studies, however, focused on the short-term effects of exposures to wildfire smoke or polluted air and none have performed an unbiased assessment of gene pathways impacted by wildfire smoke exposure.

A cohort of rhesus macaques (*Macaca mulatta*) that were exposed in their first three months of life to a harsh wildfire season in 2008 in California was previously studied to understand some of the long-term effects of wildfire smoke exposure (11). Peripheral blood mononuclear cells (PBMCs) were cultured and challenged with either LPS or flagellin, and secretions of IL-8 and IL-6 were compared to macaques that were born in 2009 ($PM_{2.5}$ and ozone levels were much lower in 2009 compared to 2008) (11). Lung function was also compared between exposed and control macaques. Compared to control macaques, wildfire smoke-exposed macaques had significantly reduced lung volume. Female wildfire-exposed macaques showed

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

reduced production of IL-8 compared to controls, while male wildfire-exposed macaques showed reduced production of IL-6 compared to controls (11). This study implied that early-life exposure led to a difference in IL-8 and IL-6 production following an immune challenge, but it was still unclear to what degree these macaques exhibited baseline differences at the level of epigenetics and gene expression.

The epigenetic mark of DNA methylation has the potential to reflect past exposures with long-lived marks on genes, while the transcriptome reflects current levels of gene expression in a sampled tissue. To test the hypothesis that early-life wildfire smoke exposure would result in detectable epigenetic differences to gene pathways reflecting cellular function, we performed the integrated unbiased approaches of whole genome bisulfite sequencing (methylome) and RNA-sequencing (transcriptome) from nasal epithelial samples collected the same cohorts of female macaques examined a decade earlier for lung functions and immune responses. We identified a large number of genes associated with early-life exposure-related differential methylation involved in neuronal and immune signaling. In contrast, only one differentially expressed gene (*FLOT2*) was stably associated with early-life wildfire smoke exposure.

RESULTS

Exposure to wildfire smoke during infancy is associated with long-lasting changes to DNA methylation patterns in nasal epithelial cells.

To test the effects of early-life wildfire smoke-exposure on methylation status throughout the genome, we performed whole genome bisulfite sequencing on nasal epithelial samples collected from 22 adult rhesus macaques in 2019 (8 born in 2008 and exposed to high levels of $PM_{2.5}$ and

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ozone due to wildfires, 14 born in 2009 and therefore has relatively low levels of exposure to $PM_{2.5}$ and ozone; Figure 1, Table 1). Though there were several shared exposures to high levels of wildfire smoke $PM_{2.5}$ (> 35ug/m^3, the 24-hour $PM_{2.5}$ National Ambient Air Quality Standard) and ozone after the 2009 cohort was born (especially in the year of 2019), there was one high exposure event that only the 2008 cohort was exposed to in early-life (10 days above 35ug/m^3, Figure 1, Table 1). We assessed 26,609,677 CpG sites and identified 3370 differentially methylated regions (DMRs) between exposed and non-exposed samples (Figure 2, empirical $p < 0.05$, differences in methylation>5%). The majority of these DMRs were hypermethylated in exposed animals (2899, ~86%). A total of 114 (3.38%) of these DMRs were primarily located in CpG islands (12), 287 (8.52%) were located in CpG shores (0-2kb from island), 205 (6.08%) were located in CpG shelves (2-4kb from island), and 2764 (82.02%) were in the open sea (>4kb from island). This distribution was significantly different than expected by chance, with an enrichment towards CpG islands, shores, and shelves compared to regions assayed (Supplementary Figure 1). These 3370 DMRs were annotated to 2139 genes (Supplementary Table 1), of which 1852 genes were associated with DMRs hypermethylated in the exposed group, while 376 genes were associated with DMRs hypomethylated in the exposed group, and 89 genes were associated with both hypermethylated and hypomethylated DMRs (examples of DMRs shown in Figure 3). The DMRs were significantly more associated with promoters and exons than expected by chance, while they were less associated with intergenic regions than expected by chance (Supplementary Figure 2). The genes associated with DMRs as a whole were significantly enriched (FDR < 0.05) for 186 IPA canonical pathways, including *axonal guidance signaling*, *synaptogenesis signaling pathway*, *protein kinase A signaling*, *IL-15 production*, *CXCR4 signaling*, and *Th1 and Th2 activation pathway* (Figure 4, Supplementary Table 2).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Genes associated with hypermethylated DMRs were enriched for 187 IPA pathways, 168 of which were also enriched in genes associated with DMRs as a whole. The 19 unique IPA pathways enriched in hypermethylated DMRs include *14-3-3-mediated signaling, LPS-stimulated MAPK Signaling,* and *NF-κB activation by viruses* (Supplementary Table 3). Genes associated with hypomethylated DMRs were enriched for 41 IPA pathways, 23 of which were also enriched in genes associated with DMRs as a whole. The 18 unique IPA pathways enriched in hypomethylated DMRs include *dermatan sulfate biosynthesis, xenobiotic metabolism PXR signaling pathway*, and *HOTAIR regulatory pathway* (Supplementary Table 4).

Impact of wildfire smoke-associated DNA methylation changes on TF binding.

As the binding of transcription factors (TFs) are often influenced by DNA methylation, we performed a HOMER analysis to determine whether any transcription factor binding sites were enriched in these wildfire smoke-associated DMRs (13). A total of 131 transcription factor motifs were enriched in all DMRs ($q < 0.05$; Supplementary Table 5). Eight of the top ten most highly enriched TF motifs are part of the bZIP TF family (shown in Table 2). When testing for TF binding site enrichment in only DMRs that were hypermethylated in exposed macaques, six of the top ten were part of the bZIP TF family, while none of the top ten enriched TF binding sites in hypomethylated DMRs were part of the bZIP TF family (five out of ten contained homeobox motifs). Interestingly, the TFs whose binding sites were most enriched in all wildfire smoke-associated DMRs were primarily unmethylated (Table 2) in other ChIP-seq datasets (14), so the differential methylation could theoretically have a large impact on transcription factor binding and expression (15). In support of this, DNA methylation generally inhibits binding of bZIP TF members to DNA (15, 16).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Regions with hypomethylated DMRs are enriched for bivalent chromatin marks across tissue types.

In order to understand the gene regulatory role of regions with wildfire smoke-DMRs, we searched for the enrichment of 15 pre-defined chromatin states across 127 epigenomes from multiple tissues and cell types in the Roadmap Epigenomics project (17). After converting the *M. maculatta* coordinates into human (hg38) coordinates and using LOLA, the DMRs as a whole were enriched for bivalent chromatin marks (top odds ratio for any mark = 1.46, q-value < 3 x 10^{-6}; Figure 5A). Bivalent chromatin marks represent co-existing activating and repressing marks, which often silence key developmental genes while keeping them poised for activation in pluripotent cells (18). Hypomethylated DMRs seemed to drive this enrichment (top odds ratio for any mark = 2.05, q-value < 0.02; Figure 5B), though hypermethylated DMRs showed enrichment (top odds ratio for any mark = 1.51, q-value < 1 x 10^{-6}) for bivalent chromHMM chromatin states as well (Figure 5C).

Early-life wildfire smoke exposure had a minimal effect on baseline genes expression levels.

To determine whether early-life exposure to wildfire smoke leads to detectable differences in gene expression later in life, we performed RNA-sequencing on 15 female rhesus macaques (6 born in 2008 and exposed to wildfire smoke, 9 born in 2009 and not directly exposed to the 2009 California wildfires). A principal component analysis (PCA) and hierarchical clustering of all detected transcripts were performed to visualize how samples clustered based on expression (Supplementary Figure 3D). The top two principal components in a principal component analysis (PCA) explained 62% of the variation in the dataset. Exposed and non-exposed samples did not cluster separately in either the PCA or the hierarchical clustering analysis, implying no

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

115 widespread transcriptomic difference between exposed and non-exposed individuals. After
116 multiple hypothesis correction (FDR < 0.05, fold change ≥ 1.2; Supplementary Table 6), there
117 was only one differentially expressed gene (*FLOT2*; Supplementary Table 6). None of the genes
118 annotated to DMRs were significantly differentially expressed.

119 To identify co-expressed genes whose expression correlated with wildfire smoke-
120 exposure status, we performed a weighted gene co-expression network analysis (WGCNA) (19).
121 We identified 16 co-expressed modules using WGCNA. None of the modules were significantly
122 associated with early-life exposure status ($p < 0.05$). The module that best correlated with
123 exposure status was the purple module ($p = 0.1$; consisting of 585 genes, including IFI44,
124 IFNA21, and IL24; Supplementary Figure 4, Supplementary Table 7). No genes in this module
125 were significantly differentially expressed at an individual level, 19 genes were associated with
126 significant DMRs, and two genes had significantly correlated methylation and expression. The
127 genes in this module were enriched (FDR < 0.05) for 21 IPA pathways, including *EIF2*
128 *signaling*, *mTOR signaling*, *Th17 activation pathway*, and *interferon signaling* (Supplementary
129 Table 8).
130
131 *Correlation of DNA methylation and gene expression differences resulting from wildfire smoke*
132 *exposure during infancy*
133 Out of the 2139 genes associated with DMRs, 2128 had enough corresponding expression data to
134 evaluate the correlation between expression and methylation. To identify genes where
135 differential methylation may be ultimately leading to differential expression, we calculated the
136 spearman rank correlation between methylation and expression levels for genes that were
137 associated with DMRs. In total, 172 genes were significantly correlated (spearman p-value <

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

0.05), with 76 genes showing a negative correlation and 96 showing a positive correlation between methylation and expression (Supplementary Table 9, two examples are shown in Figure 6). These 172 genes were enriched for 32 IPA pathway terms, including *leukocyte extravasation signaling, CCR5 signaling in macrophages,* and *MIF regulation of innate immunity* (Supplementary Table 10).

DISCUSSION

Utilizing rhesus macaques that experienced the harsh conditions of the 2008 California wildfire season in their first three months, we have elucidated some of the long-term effects of early-life exposure to wildfire smoke. Baseline methylation profiles generally clustered better by exposure status than expression profiles (Supplementary Figure 3). Many genes (2139) were associated with differentially methylated regions between exposed and control macaques (empirical $p < 0.05$), while only 1 gene (*FLOT2*) was differentially expressed between these groups after multiple hypothesis correction (FDR < 0.05). Out of the genes associated with differentially methylated regions, 172 had methylation levels that significantly correlated with expression levels across samples, indicating that the overall epigenetic regulatory landscape ultimately led to few significant differences in baseline expression. However, the changes in DNA methylation were significantly enriched at promoters and enhancers, and located at regions that transcription factors may bind, suggesting that they may have an impact on gene regulation.

FLOT2 (flotillin 2) encodes a caveolae-associated, integral membrane protein that belongs to the lipid raft family. Flotillins are implicated in variety of cellular functions, including regulation of G-protein coupled receptor signaling (20), endocytosis (21), cell-cell adhesion (22), uropod formation and migratory capacity of neutrophils and monocytes (23) and T cells (24).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FLOT2 also protected lung epithelial cells from Fas-signaling mediated apoptosis (25), and silica nanoparticles were found in Flotillin-1 and -2 marked vesicles in alveolar epithelial cell (26). However, its role in response to wildfire smoke exposure has not been reported. One potential explanation for the few gene expression changes despite more widespread methylation differences is that many of these DMRs are in regions associated with bivalent chromatin marks. The differential methylation at these regions may not affect gene expression because the bivalent chromatin marks generally keep expression repressed, but poised for rapid activation during early development (27) or in cancer (28). This would imply that some of the methylation differences were due to early-life events that were not reflected in baseline transcript levels later in life. Additionally, although baseline gene expression was relatively similar between exposed and control macaques, one hypothesis is that the altered regulatory landscape could lead to differences in expression upon additional immune (or other) challenges. This hypothesis is supported by a previous study on macaques from these same cohorts that found differences in IL-6 (significant in males) and IL-8 (significant in females) production in peripheral blood mononuclear cells (PBMC) from wildfire smoke-exposed macaques compared to controls after a challenge with media, LPS, or flagellin (11). Out of 84 genes tested, only two (RELB and REL) showed significant differences in expression following a media challenge (essentially a comparison of baseline expression), while five genes were differentially expressed following a challenge with either LPS or flagellin (11). RELB was the only gene that was differentially expressed in all three tests, but the direction of change in challenged cells (increased RELB in cells from exposed animals) was opposite of what was found at baseline (decreased RELB in cells from exposed animals) (11). In summary, there were very few differences in baseline expression in the previous study between exposed and control cells, and even when there was

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

differential expression, those patterns changed or became non-significant following an immune challenge. While the sample types (PBMCs vs. nasal epithelium) and ages of the macaques (adolescents vs. adults) differ between the prior study and the current study, they both support that early exposure to wildfire smoke did not lead to drastic differences in baseline expression profiles between samples. Another potential explanation for differences in the degree of differential expression and differential methylation is that we had fewer samples for our differential expression analysis, potentially limiting our ability to identify differential expression compared to our ability to identify differential methylation. If there were widespread differences in expression due to exposure status, however, we expect that wildfire smoke-exposed samples would have clustered together in the principal component analysis and in hierarchical clustering analyses, so we postulate that this is not the major reason for the lack of differential expression.

Long-term effects of wildfire smoke exposure on the methylome

Our data implies that there are long-term effects on the methylome due to wildfire smoke exposures during infancy. DMRs were enriched for many pathways linked to asthma, COPD, or other pulmonary diseases, including *IL-15 production* (29, 30), *CXCR4 signaling* (31, 32), *Actin cytoskeleton signaling* (33, 34), *VDR/RXR activation* (35, 36), *Th1 and Th2 activation pathway* (37, 38), and *Wnt/β-catenin signaling* (39, 40) (Supplementary Table 2). Cytokines derived from T helper type 2 (Th2) cells have long been thought to play a critical role in allergic asthma through regulation of immunoglobulin E (IgE) synthesis (38, 41), but other T helper subsets (such as Th1) are starting to gain recognition for their role in asthma as well. Increased levels of the Th1 cytokine IFN-γ have been shown to exacerbate existing asthmatic responses (42) and increase airway hyperresponsiveness (41) in transgenic mice. IFNGR2 (interferon gamma

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

receptor 2) was differentially methylated in our comparison (as were several other Th1 related genes, including IL6R, LOC694631/IFNA1/13-like, and NFATC1), perhaps indicating that the early life wildfire smoke exposure has altered Th1 responses and resulted in differential responses to bacterial and viral infection. Additionally, hypermethylation of IL6 and IFNA13 was associated with idiopathic pulmonary fibrosis (IPF) (43), while hypermethylation of IL6R was associated with COPD in prior studies (44). IL6R and IFNA13 were also hypermethylated in exposed macaques in our current study (Supplementary Table 11), indicating that changes in the Th1 pathway may contribute to the reduction in lung function noted in macaques exposed to wildfire smoke early in life (11). Th2-related genes that were differentially methylated in our dataset include IL4R and TIMD4, while there were several genes associated with DMRs that were related to both the Th1 and Th2 pathways (including CD4, IL10, IL12RB2, NFATC2, RUNX3, and SOCS3). Hypermethylation of NFATC2 (44), RUNX3 (44), and SOCS3 (45) has been associated with COPD (Supplementary Table 11). These three genes were also hypermethylated in wildfire smoke-exposed macaques versus controls.

Deletion of Fra1, the transcription factor with the most enriched motif in the DMRs (Table 2, Supplementary Table 5), in mice led to greater levels of progressive interstitial fibrosis (46). Fra1 is a bZIP transcription factor and bZIP transcription factor binding is generally inhibited by methylation (15, 16). Meanwhile, overexpression of Fra2 (another highly enriched bZIP TF motif in the DMRs) in mice lead to non-allergic asthma development (47). The other bZIP transcription factors whose motifs were among the top ten enriched motifs have also all been linked to pulmonary disease (ATF3 (48), JunB (49), BATF (50), and AP-1 (46, 51)). The role of bZIP transcription factors in pulmonary disease pathogenesis combined with the sensitivity of bZIP to changes in methylation imply that the differences in methylation noted

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

between wildfire smoke-exposed and non-exposed macaques could greatly impact how bZIP targets are regulated following a respiratory challenge. Overall, differences in methylation in Th1 and Th2-related genes (and the relation of those genes to asthma, IPF, and COPD pathogenesis; see Supplementary Table 11) may explain the long-term differences in lung function previously observed between wildfire smoke-exposed macaques and controls (11).

Interestingly, there is also recent research that suggests that exposure to air pollution can have negative neuropsychological effects in children (52, 53). The DMRs from our dataset were enriched in multiple IPA neurological pathways, including *axonal guidance signaling* (most significant pathway), *synaptogenesis signaling pathway* (third most significant pathway), and *neuropathic pain signaling in dorsal horn neurons* (Supplementary Table 2). Additionally, the top enriched biological process term in GOfuncR (54) was *neuron differentiation*, while the top enriched cellular component term was *synapse* (Supplementary Figure 5). The effect of wildfire smoke on neurological development is understudied, but studies have shown that particles less than 0.1 μm in diameter (which are produced by wildfires) can cross the blood-brain barrier (55). Additionally, exposure to these ultrafine particles has been associated with ADHD, autism, and declines in school performance and memory in children (53). Along with this evidence from prior studies, the differential methylation of regions near genes involved in neurological pathways indicates that early-life wildfire smoke exposure could have a long-lasting impact on nervous system function.

Genes with correlated changes in methylation and expression are enriched for pathways associated with respiratory diseases

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In addition to directly studying genes and enriched pathways associated with DMRs, we also wanted to identify genes that showed correlations between expression and methylation to get a better understanding of how differences in methylation modify mRNA expression. Though only one gene was differentially expressed between our groups following multiple hypothesis correction (*FLOT2*), there were many more genes associated with DMRs that had a significant correlation between methylation and expression (172 in total; Supplementary Table 9). *MAPK10* (Spearman's $\rho \sim 0.75$) and *WNT8B* (Spearman's $\rho \sim 0.82$) were two other genes that were associated with DMRs that showed a significant correlation between methylation and expression (Supplementary Table 9). Wnt signaling has been linked to *in utero* lung development and development/maturation during early life (alveologenesis) (56-58). Prior studies have shown that Wnt/β-catenin and the mitogen-activated protein kinase (MAPK) signaling pathway take part in the airway remodeling process in asthma (39). In a mouse model of asthma, blocking Wnt signaling reduced airway remodeling, while p38 MAPK expression was increased in asthmatic mice compared to controls (39). MAPK10 expression was slightly higher on average in wildfire smoke-exposed macaques than control macaques, and methylation was significantly positively correlated with expression (hypermethylated in exposed animals). WNT8B expression was slightly lower on average in exposed macaques, while methylation was significantly negatively correlated with methylation (hypomethylated in exposed animals). Neither of these genes were significantly differently expressed, however. Given the role of Wnt signaling and MAPK signaling in airway remodeling, it seems possible that changes in gene regulation could have contributed to the reduced lung function noted in wildfire smoke-exposed macaques (11).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Our study had several limitations. As previously touched upon, our current study included only female rhesus macaques, but a prior study with these macaques noted significant sex-specific differences in PBMCs challenged with LPS or flagellin. Male wildfire smoke-exposed macaques had significantly higher levels of IL-8 compared to controls, while female wildfire smoke-exposed macaques had significantly higher levels of IL-6 compared to controls (11). While IL-6 was not differentially expressed or methylated in our exposed macaques compared to controls, this does underscore that we may have missed some sex-specific differences in gene expression or methylation by sampling only female macaques for our current study. Indeed, studies have shown that there are sex-specific differences in expression between female and male asthmatics (59, 60), implying that the molecular underpinnings of asthma and other pulmonary issues may differ between the sexes. Additionally, our cohort of wildfire smoke-exposed macaques was roughly one year older than our cohort of control macaques. Studies have indicated that methylation patterns are associated with aging (epigenetic clocks) in humans (61, 62), so this is likely the case for rhesus macaques as well. Out of 2139 genes that were associated with DMRs in our dataset, 20 were differentially methylated in a pattern that was consistent with the models from the previously referenced studies on epigenetic clocks. Based on these results, most of the differential methylation we observed cannot be explained by known differences in how methylation correlates with age. Another potential alternative explanation is that the differences in methylation we observed were due to greater cumulative exposure to pollutants in the older macaques. Table 1 shows that the difference in cumulative exposures to high levels of $PM_{2.5}$ and ozone between the two groups were roughly equivalent to the differences observed in the first three months of life, implying that these early exposures were key drivers of the noted differences between the groups. However, cumulative exposures

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

297 below the current U.S. EPA standards were associated with increased mortality in a Medicare
298 population (63), and they may also have an impact on the epigenome. The epigenetic effects of
299 acute and chronic wildfire smoke exposure are worthy of further investigation. As previously
300 discussed, we had a smaller sample size for our expression dataset (n = 13 after removing two
301 outliers) than our methylation dataset (n = 22). This could explain why we saw fewer changes in
302 expression overall, however samples appeared to cluster more closely based on exposure status
303 for the methylation dataset than the expression dataset (Supplementary Figure 3). The p-values
304 from DMRichR were empirical p-values calculated from permutation tests (64, 65). Although
305 this puts our study at a higher risk of false-positive findings, these permutation p-values
306 calculated by DMRichR were used to determine DMR significance in multiple published studies
307 in combination with effect size (64-67). Given that our analysis of chromatin states relied on
308 human hg38 annotations, we compared our macaque rheMac10 annotations to the hg38
309 annotations to make sure they were similar enough. About 67% of the DMR gene annotations
310 were exact matches after lifting over the coordinates to hg38. While 30% of DMRs had a
311 different annotation, some of these differences were just due to differences in gene naming
312 convention between the species. For example, one DMR was annotated to LOC694631
313 (IFNA1/13-like) in rheMac10, while the lifted over DMR was annotated to IFNA13 in hg38. In a
314 broader pathway analysis, 90% of the IPA pathways enriched in DMRs using rheMac10
315 annotations were also enriched when we used hg38 annotations. We also focused our discussion
316 on genes that were consistent between the two annotations.

317 One area of interest for future studies would be the stability of these changes. The
318 exposure event took place in 2008, while samples were collected from the macaques in 2019.
319 Over that relatively long course of time (the average lifespan for macaques in captivity is ~27

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

years (68)), the methylation profiles still clustered based on exposure status (Figure 2, Supplementary Figure 3). This implies that there are long-term impacts of wildfire smoke exposure on methylation, and that at least some of these changes are highly stable. An early study on DNA methylation stability involved sampling individuals three days apart to check for differences in DNA methylation. This study on 12 gene promoters indicated that methylation stability was marker dependent and varied based on sequence composition (69). Meanwhile, a large-scale study on how storage conditions affect methylation stability indicated that storing DNA samples in temperatures as high as four degrees Celsius for up to 20 years had no significant impact on methylation (70).

CONCLUSIONS:

In summary, our study revealed differences in methylation and gene expression in nasal epithelial samples between macaques that were exposed to wildfire smoke during early life and macaques that were not exposed to wildfire smoke during early life. The wildfire smoke associated DMRs were enriched for a variety of immune processes, but there were few significant expression differences at baseline between exposed and non-exposed macaques. Given the differences in methylation, perhaps differences in expression between these two groups would become apparent following an immune/respiratory challenge, but future studies would be required to explore this hypothesis. Our study indicates that wildfire smoke exposure in early life can have long-term impacts on the epigenome.

METHODS:

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

342 *Animals*

343 Wildfire smoke-exposed rhesus macaque monkeys born between April 1 and June 8, 2008 were
344 housed in outdoor facilities at the CNPRC from birth to now (Table 1). Monkeys born between
345 April 1 and June 8, 2009 were used as controls. $PM_{2.5}$ and ozone were measured by a California
346 Air Resources Board air monitoring station (site no. 57,577) located 2.7 miles southeast of the
347 California National Primate Research Center on the University of California Davis campus
348 (Figure 1). Care and housing of animals complied with the provisions of the Institute of
349 Laboratory Animal Resources and conformed to practices established by the American
350 Association for Accreditation of Laboratory Animal Care. Procedures in this study were
351 approved by the UC Davis Institutional Animal Care and Use Committee.
352

353 *Sample Collection and DNA/RNA Extraction*

354 Nasal epithelium samples were collected from 22 female rhesus macaques (*Macaca mulatta*)
355 housed at the California National Primate Research Center. Exposures of these animals to
356 wildfire smoke were previously estimated (11). Eight of these macaques were born in 2008 and
357 exposed to wildfire smoke from birth to 3 months old, while the other 14 were born in 2009 with
358 low wildfire exposure from birth to 3 months old (Table 1, demographic comparison of two
359 groups). We collected these nasal epithelium samples in 2019. RNA and DNA were isolated
360 using the Allprep DNA and RNA kit (Qiagen) according to the manufacturer's instructions.
361

362 *Library preparation for whole genome bisulfite sequencing (WGBS)*

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Whole genome bisulfite sequencing (WGBS) libraries were prepared for all 22 samples. Library quality was checked prior to sequencing using an Agilent 2100 Bioanalyzer system; library concentration was measured using a Qubit DNA high sensitivity assay. Each library was comprised of sample from a single individual; these individually barcoded libraries were then pooled and sequenced on two lanes from a NovaSeq 6000 S4 flow cell at PE150 using Swift's Accel-NGS Methyl-Seq Kit at the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center. We sequenced approximately 475 million paired end reads per sample that passed initial filters. Reads were demultiplexed using the bcl2fastq Illumina software.

WGBS read alignment, differential methylation analysis, pathway analysis, and chromatin state analysis

The CpG_Me pipeline (71-74) was utilized to align the WGBS data. Reads were trimmed using Trim Galore (73) to address methylation biases at the 5' and 3' end of reads (10 bases were trimmed from the 3' end of both read 1 and read 2, and 10 and 20 bases were trimmed from the 5' end of reads 1 and 2 respectively). The reads were aligned to the *M. mulatta* genome using Bismark (72), which was also used to deduplicate the aligned reads and generate CpG count matrices. Read quality and mapping quality were assessed using MultiQC (74). Differentially methylated regions between exposed and non-exposed macaques were identified using DMRichR (64, 75, 76), which uses the dmrseq (75) and bsseq (76) algorithms. Animal weight was adjusted for as a covariate. We used the default paramters for DMRichR, including requiring at least 1x coverage for all samples for a CpG, requiring a minimum of 5 CpGs for a DMR, performing 10 permutations for DMR and block analyses, and setting the single CpG coefficient required to discover testable background regions to be at least 0.05. Using DMRichR, candidate

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regions are identified based on differences in mean methylation between groups, then region-level metrics that account for mean methylation, CpG correlation, and coverage are computed. These region-level metrics are then compared to a pooled null distribution generated via permutations to calculate an empirical p value for each candidate region (64, 65). Bsseq (76) was used to generate individual smoothed methylation values and heatmap visualizations. IPA (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) was used for pathway enrichment analysis. We also used GOfuncR (54) for GO enrichments based on DMR coordinates rather than gene names. HOMER (13) was used to identify enriched transcription factor binding motifs in the DMRs ($p < 0.05$) , while we utilized MethMotif (14) to characterize methylation frequency of transcription factors whose binding motifs were enriched in the DMRs. We used the UCSC liftover tool (77) to lift DMR coordinates from rheMac10 to hg38 because chromatin state information was not available for *M. mulatta*. Locus Overlap Analysis (LOLA) (78) was used to determine whether DMRs were enriched for chromHMM (79) states relative to the background regions. The spearman correlation coefficient between gene expression and methylation levels for genes associated with DMRs was used to determine whether significant methylation changes were associated with changes in gene expression ($p< 0.05$).

Library preparation for RNA-seq

RNAseq libraries were prepared for a total of 15 samples: 6 from wildfire smoke-exposed individuals and 9 from non-exposed individuals (Supplementary Table 12, comparison of these two groups). As some of the RNA samples were of low quantity, a special low-input RNA-seq pipeline were applied at the Genomics, Epigenomics and Sequencing Core of University of

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Cincinnati (80, 81). Briefly, polyA RNA was isolated using NEBNext Poly(A) mRNA Magnetic Isolation Module (New England BioLabs, Ipswich, MA) and enriched using SMARTer Apollo NGS library prep system (Takara Bio USA, Mountain View, CA). Libraries were prepared using NEBNext Ultra II Directional RNA Library Prep Kit (New England BioLabs), indexed, pooled and sequenced using Nextseq 550 sequencer (Illumina, San Diego, CA). Approximately 40 million reads passing filter per sample were generated under the sequencing setting of single read 1x85 bp. Reads were demultiplexed and adapters were trimmed using the bcl2fastq Illumina software.

RNA-seq read alignment, differential expression analysis, pathway analysis, and co-expression analysis

Read quality was checked using *FastQC* (82), then the reads were aligned to the *Macaca mulatta* genome (rheMac10, GenBank assembly accession: GCA_003339765.3) with *Bowtie2* (83). Transcripts were quantified using *RSEM* (84). The data from *RSEM* was congregated and converted into *DESeq2* (85) format using *tximport* (86). Sample clustering by expression (investigated via principal component analysis and hierarchical clustering) and detection of differentially expressed genes between wildfire smoke-exposed and non-exposed samples was done using *DESeq2* (85). Individual weight was included as a covariate in the differential expression analysis. Two samples (one wildfire smoke-exposed and one non-exposed sample) were excluded from all subsequent RNA-sequencing analyses because they were identified as outliers in the hierarchical clustering analysis (Supplementary Figure 6). The resulting log-fold change values were shrunken (following the recommendation from the *DESeq2* reference manual) using *apeglm* (87). Differentially expressed genes had FDR < 0.05 and an absolute shrunken fold change of at least 1.2. The Ingenuity Pathway Analysis (IPA) software (QIAGEN

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) was used for pathway analysis. Significantly enriched pathways in IPA had a p-value < 0.05.

Co-expressed modules of genes were found using WGCNA (19). The soft threshold (power) was set to 8 based on a plot of soft threshold vs scale free topology model fit. Modules that were too similar to one another (below a height of 0.5) were merged into one module. After merging, the final co-expression modules were tested for significant associations with wildfire smoke exposure and animal weight. Pathways enriched in genes in modules of interest were identified using IPA.

DECLARATIONS

Ethics approval and consent to participate

Procedures in this study were approved by the UC Davis Institutional Animal Care and Use Committee.

Consent for publication

Not applicable.

Availability of data and materials

WGBS and RNA-seq data will be deposited to GEO upon manuscript acceptance.

Competing interests

The authors declare they have no competing interests.

Funding

This study was supported by a pilot grant from Environmental Health Sciences Center (NIEHS-P30 ES006096) awarded to HJ. HJ was also supported by an Environmental Health Sciences Scholar Award (P30 ES006096), and NIH/NIAID R01AI141569-1A1. BL was supported

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

454 Canadian Institutes of Health Research (CIHR) postdoctoral fellowship (MFE-146824) and a
455 CIHR Banting postdoctoral fellowship (BPF-162684).

456 **Authors' contributions**

457 HJ conceived the study in discussion with LAM and JML. APB drafted the manuscript with the
458 help of HJ, BIL, JML and LAM. LC extracted DNA and RNA from nasal samples, prepared
459 WGBS libraries. APB performed WGBS analysis with the help of BIL. APB performed other
460 data analysis and visualization in discussion with HJ. All authors have approved the final
461 version of this manuscript.

462 **Acknowledgements**

463 Nasal sampling was performed by staff members at CNPRC Research Services. The sequencing
464 was carried out at the DNA Technologies and Expression Analysis Cores at the UC Davis
465 Genome Center, supported by NIH Shared Instrumentation Grant 1S10OD010786-01.

466

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

467 Table 1: Demographic characteristics of animal populations

	2008 Birth Year	2009 Birth year	P
Participants			
N	8	14	
Age at sample collection (yr)	11.2±0.2	10.3±0.1	P<0.001
Weight at sample collection (kg)	8.61±1.60	10.11±2.78	0.12
Genetic background			
Indian	7	13	1.0
Mixed Indian-Chinese	1	1	
Corral diversity	8	13	1.0
Maternal Background			
Age at parturition (yr)	5.5±1.7	5.8±2.5	0.8
Genetic background			
Indian	7	12	1.0
Mixed Indian-Chinese	1	2	
Corral diversity	8	14	1.0
Ambient pollutants months 0-3			
Days with $PM_{2.5}$ higher than 35ug/m^3	10±1.0	0±0	P<0.0001
Mean $PM_{2.5}$ concentration (μg/m^3)	14.6±16.3	9.1±3.7	P<0.0001
Median $PM_{2.5}$ concentration (μg/m^3)	9.5 (6.0-15.6)	8.8 (6.4-11.4)	P<0.0001
Hours over California 1-h ozone standard	13±2.9	0±0	P<0.0001
Mean ozone level (ppm)	0.032±0.020	0.029±0.017	P<0.001
Median ozone level (ppm)	0.030 (0.016-0.045)	0.027 (0.016-0.040)	P<0.001
Cumulative exposures through sampling date			

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Days with $PM_{2.5}$ higher than 35ug/m^3	40±0	29±0.4	P<0.0001
Mean $PM_{2.5}$ concentration (µg/m^3)	9.3±8.0	9.1±7.7	P<0.0001
Median $PM_{2.5}$ concentration (µg/m^3)	8.4 (5.5-13.4)	8.6 (6.3-11.4)	P<0.0001
Hours over California 1-h ozone standard	24±0	7±0	N/A (0 SD)
Mean ozone level (ppm)	0.026±0.015	0.026±0.015	P<0.001
Median ozone level (ppm)	0.025 (0.015-0.036)	0.025 (0.015-0.036)	P<0.001

Note: Age, weight and exposures are shown as mean ± (SD) or median (interquartile range), and compared using t test. Categorical variables (genetic background and corral diversity) are reported as group-specific numerical frequency and compared using Fisher's exact tests. 35ug/m^3 is the 24-hour $PM_{2.5}$ National Ambient Air Quality Standard. 0.09ppm is the California 1-h ozone standard.

468

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure legends

Figure 1: Average daily $PM_{2.5}$ from April 2008 through October 2019 at the California Air Resources Board air monitoring station (site no. 57,577) located 2.7 miles southeast of the California National Primate Research Center on the University of California Davis campus. The dotted line at 35ug/m^3 represent the 24-hour $PM_{2.5}$ National Ambient Air Quality Standard. Note the arrow pointing to the early-life exposure event in macaques born in 2008. All other exposure events were shared between the two groups.

Figure 2: Heatmap showing sample clustering based on methylation. The heatmap includes only differentially methylated regions (DMRs). The heatmap was normalized on a per row basis for visualization, therefore the values on the scales are relative rather than absolute.

Figure 3: Examples of differentially methylated regions (DMRs) between rhesus macaques exposed in the first three months of life to wildfire smoke and those that were not. A) IL4R. B) RXRG. C) TLR5. D) ITGB6. Each dot represents the methylation percentage of one individual at one CpG site, while each line represents the smoothed average methylation level moving across the region. The red shaded boxes denote the specific DMR locations. Tracks for CpG islands (if present) or genes are included underneath each plot. For the gene tracks, a solid box indicates an exon, while the arrows indicate the direction of transcription.

Figure 4: Enriched pathway analyses for differentially methylated regions (DMRs). Only the top ten (out of 186) enriched Ingenuity Pathway Analysis (IPA) canonical pathways are shown.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5: Enrichment in chromHMM (79) states in A) all differentially methylated regions (DMRs), B) DMRs that were hypomethylated in wildfire smoke-exposed macaques, and C) DMRs that were hypermethylated in wildfire smoke-exposed macaques. The rows in the plot represent different datasets from different cell types from the NIH Roadmap Epigenomics Consortium (88). Epithelial and IMR90 are highlighted in the plots, as these are the closest to the nasal epithelial samples in our current study.

Figure 6. Correlation plots between expression and methylation for A) MAPK10and B) CD44. Each individual point represents one sample. Expression and methylation were significantly correlated (spearman p-value < 0.05) for both genes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Knowlton K. Where there's fire, there's smoke: wildfire smoke affects communities distant from deadly flames. 2013.
2. Dockery DW. Health effects of particulate air pollution. Ann Epidemiol. 2009;19(4):257-63.
3. Brunekreef B, Holgate ST. Air pollution and health. Lancet. 2002;360(9341):1233-42.
4. Li YJ, Takizawa H, Kawada T. Role of oxidative stresses induced by diesel exhaust particles in airway inflammation, allergy and asthma: their potential as a target of chemoprevention. Inflamm Allergy Drug Targets. 2010;9(4):300-5.
5. Zhou Z, Liu Y, Duan F, Qin M, Wu F, Sheng W, et al. Transcriptomic Analyses of the Biological Effects of Airborne PM2.5 Exposure on Human Bronchial Epithelial Cells. PloS one. 2015;10(9):e0138267-e.
6. Gehring U, Gruzieva O, Agius RM, Beelen R, Custovic A, Cyrys J, et al. Air pollution exposure and lung function in children: the ESCAPE project. Environ Health Perspect. 2013;121(11-12):1357-64.
7. Urman R, McConnell R, Islam T, Avol EL, Lurmann FW, Vora H, et al. Associations of children's lung function with ambient air pollution: joint effects of regional and near-roadway pollutants. Thorax. 2014;69(6):540-7.
8. Spears D, Dey S, Chowdhury S, Scovronick N, Vyas S, Apte J. The association of early-life exposure to ambient PM2.5 and later-childhood height-for-age in India: an observational study. Environmental Health. 2019;18(1):62.
9. Rosa MJ, Hair GM, Just AC, Kloog I, Svensson K, Pizano-Zárate ML, et al. Identifying critical windows of prenatal particulate matter (PM(2.5)) exposure and early childhood blood pressure. Environ Res. 2020;182:109073.
10. To T, Zhu J, Stieb D, Gray N, Fong I, Pinault L, et al. Early life exposure to air pollution and incidence of childhood asthma, allergic rhinitis and eczema. European Respiratory Journal. 2020;55(2):1900913.
11. Black C, Gerriets JE, Fontaine JH, Harper RW, Kenyon NJ, Tablin F, et al. Early Life Wildfire Smoke Exposure Is Associated with Immune Dysregulation and Lung Function Decrements in Adolescence. Am J Respir Cell Mol Biol. 2017;56(5):657-66.
12. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics. 2012;13(7):484-92.
13. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.
14. Xuan Lin QX, Sian S, An O, Thieffry D, Jha S, Benoukraf T. MethMotif: an integrative cell specific database of transcription factor binding motifs coupled with DNA methylation profiles. Nucleic Acids Research. 2018;47(D1):D145-D54.
15. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science. 2017;356(6337).
16. Héberlé É, Bardet AF. Sensitivity of transcription factors to DNA methylation. Essays in biochemistry. 2019;63(6):727-41.
17. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nature methods. 2012;9(3):215-6.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

18. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006;125(2):315-26.
19. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):559.
20. Sugawara Y, Nishii H, Takahashi T, Yamauchi J, Mizuno N, Tago K, et al. The lipid raft proteins flotillins/reggies interact with Galphaq and are involved in Gq-mediated p38 mitogen-activated protein kinase activation through tyrosine kinase. Cell Signal. 2007;19(6):1301-8.
21. Otto GP, Nichols BJ. The roles of flotillin microdomains--endocytosis and beyond. Journal of cell science. 2011;124(Pt 23):3933-40.
22. Bodin S, Planchon D, Rios Morris E, Comunale F, Gauthier-Rouviere C. Flotillins in intercellular adhesion - from cellular physiology to human diseases. Journal of cell science. 2014;127(Pt 24):5139-47.
23. Ludwig A, Otto GP, Riento K, Hams E, Fallon PG, Nichols BJ. Flotillin microdomains interact with the cortical cytoskeleton to control uropod formation and neutrophil recruitment. The Journal of cell biology. 2010;191(4):771-81.
24. Giri B, Dixit VD, Ghosh MC, Collins GD, Khan IU, Madara K, et al. CXCL12-induced partitioning of flotillin-1 with lipid rafts plays a role in CXCR4 function. European journal of immunology. 2007;37(8):2104-16.
25. Wei S, Moon HG, Zheng Y, Liang X, An CH, Jin Y. Flotillin-2 modulates fas signaling mediated apoptosis after hyperoxia in lung epithelial cells. PloS one. 2013;8(10):e77519.
26. Kasper J, Hermanns MI, Bantz C, Utech S, Koshkina O, Maskos M, et al. Flotillin-involved uptake of silica nanoparticles and responses of an alveolar-capillary barrier in vitro. Eur J Pharm Biopharm. 2013;84(2):275-87.
27. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007;448(7153):553-60.
28. Bernhart SH, Kretzmer H, Holdt LM, Juhling F, Ammerpohl O, Bergmann AK, et al. Changes of bivalent chromatin coincide with increased expression of developmental genes in cancer. Sci Rep. 2016;6:37393.
29. Jonakowski M, Zioło J, Koćwin M, Przemęcka M, Mokros Ł, Panek M, et al. Role of IL-15 in the modulation of TGF-β1-mediated inflammation in asthma. Exp Ther Med. 2017;14(5):4533-40.
30. Liu Z, Fan W, Chen J, Liang Z, Guan L. The role of Interleukin 15 in protein degradation in skeletal muscles in rats of chronic obstructive pulmonary disease. Int J Clin Exp Med. 2015;8(2):1976-84.
31. Chen H, Xu X, Teng J, Cheng S, Bunjhoo H, Cao Y, et al. CXCR4 inhibitor attenuates allergen-induced lung inflammation by down-regulating MMP-9 and ERK1/2. Int J Clin Exp Pathol. 2015;8(6):6700-7.
32. Henrot P, Prevel R, Berger P, Dupin I. Chemokines in COPD: From Implication to Therapeutic Use. International journal of molecular sciences. 2019;20(11):2785.
33. Tang DD, Gerlach BD. The roles and regulation of the actin cytoskeleton, intermediate filaments and microtubules in smooth muscle cell migration. Respiratory Research. 2017;18(1):54.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

591 34. D'Anna C, Cigna D, Di Sano C, Di Vincenzo S, Dino P, Ferraro M, et al. Exposure to
592 cigarette smoke extract and lipopolysaccharide modifies cytoskeleton organization in bronchial
593 epithelial cells. Experimental Lung Research. 2017;43(9-10):347-58.
594 35. Iqbal SF, Freishtat RJ. Mechanism of action of vitamin D in the asthmatic lung. J Investig
595 Med. 2011;59(8):1200-2.
596 36. Hu G, Dong T, Wang S, Jing H, Chen J. Vitamin D(3)-vitamin D receptor axis
597 suppresses pulmonary emphysema by maintaining alveolar macrophage homeostasis and
598 function. EBioMedicine. 2019;45:563-77.
599 37. Martín-Orozco E, Norte-Muñoz M, Martínez-García J. Regulatory T Cells in Allergy and
600 Asthma. Frontiers in Pediatrics. 2017;5(117).
601 38. Barnes PJ. Th2 cytokines and asthma: an introduction. Respiratory research.
602 2001;2(2):64-5.
603 39. Jia X-X, Zhu T-T, Huang Y, Zeng X-X, Zhang H, Zhang W-X. Wnt/β-catenin signaling
604 pathway regulates asthma airway remodeling by influencing the expression of c-Myc and cyclin
605 D1 via the p38 MAPK-dependent pathway. Experimental and therapeutic medicine.
606 2019;18(5):3431-8.
607 40. Guo L, Wang T, Wu Y, Yuan Z, Dong J, Li Xo, et al. WNT/β-catenin signaling regulates
608 cigarette smoke-induced airway inflammation via the PPARδ/p38 pathway. Laboratory
609 Investigation. 2016;96(2):218-29.
610 41. Kim Y-M, Kim Y-S, Jeon SG, Kim Y-K. Immunopathogenesis of allergic asthma: more
611 than the th2 hypothesis. Allergy Asthma Immunol Res. 2013;5(4):189-96.
612 42. Durrant DM, Metzger DW. Emerging roles of T helper subsets in the pathogenesis of
613 asthma. Immunol Invest. 2010;39(4-5):526-49.
614 43. Sanders YY, Ambalavanan N, Halloran B, Zhang X, Liu H, Crossman DK, et al. Altered
615 DNA methylation profile in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med.
616 2012;186(6):525-35.
617 44. Morrow JD, Cho MH, Hersh CP, Pinto-Plata V, Celli B, Marchetti N, et al. DNA
618 methylation profiling in human lung tissue identifies genes associated with COPD. Epigenetics.
619 2016;11(10):730-9.
620 45. Vucic EA, Chari R, Thu KL, Wilson IM, Cotton AM, Kennett JY, et al. DNA
621 methylation is globally disrupted and associated with expression changes in chronic obstructive
622 pulmonary disease small airways. Am J Respir Cell Mol Biol. 2014;50(5):912-22.
623 46. Rajasekaran S, Vaz M, Reddy SP. Fra-1/AP-1 Transcription Factor Negatively Regulates
624 Pulmonary Fibrosis In Vivo. PLOS ONE. 2012;7(7):e41611.
625 47. Gungl A, Biasin V, Wilhelm J, Olschewski A, Kwapiszewska G, Marsh LM. Fra2
626 Overexpression in Mice Leads to Non-allergic Asthma Development in an IL-13 Dependent
627 Manner. Front Immunol. 2018;9(2018).
628 48. Gilchrist M, Henderson WR, Jr., Clark AE, Simmons RM, Ye X, Smith KD, et al.
629 Activating transcription factor 3 is a negative regulator of allergic pulmonary inflammation. J
630 Exp Med. 2008;205(10):2349-57.
631 49. Chen S, Yun F, Yao Y, Cao M, Zhang Y, Wang J, et al. USP38 critically promotes
632 asthmatic pathogenesis by stabilizing JunB protein. Journal of Experimental Medicine.
633 2018;215(11):2850-67.
634 50. Übel C, Sopel N, Graser A, Hildner K, Reinhardt C, Zimmermann T, et al. The activating
635 protein 1 transcription factor basic leucine zipper transcription factor, ATF-like (BATF),

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

636 regulates lymphocyte- and mast cell-driven immune responses in the setting of allergic asthma. J
637 Allergy Clin Immunol. 2014;133(1):198-206.e1-9.
638 51. Caramori G, Casolari P, Adcock I. Role of transcription factors in the pathogenesis of
639 asthma and COPD. Cell Commun Adhes. 2013;20(1-2):21-40.
640 52. Suades-González E, Gascon M, Guxens M, Sunyer J. Air Pollution and
641 Neuropsychological Development: A Review of the Latest Evidence. Endocrinology.
642 2015;156(10):3473-82.
643 53. Holm SM, Miller MD, Balmes JR. Health effects of wildfire smoke in children and
644 public health tools: a narrative review. Journal of Exposure Science & Environmental
645 Epidemiology. 2020.
646 54. Grote S. GOfuncR: Gene ontology enrichment using FUNC. R package version 1.10.0
647 ed2020.
648 55. Schraufnagel DE. The health effects of ultrafine particles. Exp Mol Med.
649 2020;52(3):311-7.
650 56. Frank DB, Peng T, Zepp JA, Snitow M, Vincent TL, Penkala IJ, et al. Emergence of a
651 Wave of Wnt Signaling that Regulates Lung Alveologenesis by Controlling Epithelial Self-
652 Renewal and Differentiation. Cell Reports. 2016;17(9):2312-25.
653 57. Hussain M, Xu C, Lu M, Wu X, Tang L, Wu X. Wnt/β-catenin signaling links embryonic
654 lung development and asthmatic airway remodeling. Biochimica et Biophysica Acta (BBA) -
655 Molecular Basis of Disease. 2017;1863(12):3226-42.
656 58. De Langhe SP, Reynolds SD. Wnt signaling in lung organogenesis. Organogenesis.
657 2008;4(2):100-8.
658 59. Gautam Y, Afanador Y, Abebe T, López JE, Mersha TB. Genome-wide analysis revealed
659 sex-specific gene expression in asthmatics. Hum Mol Genet. 2019;28(15):2600-14.
660 60. Mersha TB, Martin LJ, Biagini Myers JM, Kovacic MB, He H, Lindsey M, et al.
661 Genomic architecture of asthma differs by sex. Genomics. 2015;106(1):15-22.
662 61. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide
663 Methylation Profiles Reveal Quantitative Views of Human Aging Rates. Molecular Cell.
664 2013;49(2):359-67.
665 62. Horvath S. DNA methylation age of human tissues and cell types. Genome biology.
666 2013;14(10):R115.
667 63. Shi L, Zanobetti A, Kloog I, Coull BA, Koutrakis P, Melly SJ, et al. Low-Concentration
668 PM2.5 and Mortality: Estimating Acute and Chronic Effects in a Population-Based Study.
669 Environ Health Perspect. 2016;124(1):46-52.
670 64. Laufer BI, Hwang H, Vogel Ciernia A, Mordaunt CE, LaSalle JM. Whole genome
671 bisulfite sequencing of Down syndrome brain reveals regional DNA hypermethylation and novel
672 disorder insights. Epigenetics. 2019;14(7):672-84.
673 65. Mordaunt CE, Jianu JM, Laufer BI, Zhu Y, Hwang H, Dunaway KW, et al. Cord blood
674 DNA methylome in newborns later diagnosed with autism spectrum disorder reflects early
675 dysregulation of neurodevelopmental and X-linked genes. Genome Med. 2020;12(1):88.
676 66. Murat El Houdigui S, Adam-Guillermin C, Armant O. Ionising Radiation Induces
677 Promoter DNA Hypomethylation and Perturbs Transcriptional Activity of Genes Involved in
678 Morphogenesis during Gastrulation in Zebrafish. International Journal of Molecular Sciences.
679 2020;21(11):4014.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

67. Laufer BI, Gomez JA, Jianu JM, LaSalle JM. Stable DNMT3L overexpression in SH-SY5Y neurons recreates a facet of the genome-wide Down syndrome DNA methylation signature. Epigenetics & Chromatin. 2021;14(1):13.
68. Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM, et al. Caloric restriction delays disease onset and mortality in rhesus monkeys. Science. 2009;325(5937):201-4.
69. Byun H-M, Nordio F, Coull BA, Tarantini L, Hou L, Bonzini M, et al. Temporal Stability of Epigenetic Markers: Sequence Characteristics and Predictors of Short-Term DNA Methylation Variations. PLOS ONE. 2012;7(6):e39220.
70. Li Y, Pan X, Roberts ML, Liu P, Kotchen TA, Jr AWC, et al. Stability of global methylation profiles of whole blood and extracted DNA under different storage durations and conditions. Epigenomics. 2018;10(6):797-811.
71. Laufer BI, Hwang H, Jianu JM, Mordaunt CE, Korf IF, Hertz-Picciotto I, et al. Low-pass whole genome bisulfite sequencing of neonatal dried blood spots identifies a role for RUNX1 in Down syndrome DNA methylation profiles. Human Molecular Genetics. 2020;29(21):3465-76.
72. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics (Oxford, England). 2011;27(11):1571-2.
73. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. 2011;17(1):3.
74. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-8.
75. Korthauer K, Chakraborty S, Benjamini Y, Irizarry RA. Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. Biostatistics. 2018;20(3):367-83.
76. Hansen KD, Langmead B, Irizarry RA. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. Genome Biology. 2012;13(10):R83.
77. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. Brief Bioinform. 2013;14(2):144-61.
78. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. Bioinformatics. 2015;32(4):587-9.
79. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc. 2017;12(12):2478-92.
80. Walsh KB, Zhang X, Zhu X, Wohleb E, Woo D, Lu L, et al. Intracerebral Hemorrhage Induces Inflammatory Gene Expression in Peripheral Blood: Global Transcriptional Profiling in Intracerebral Hemorrhage Patients. DNA Cell Biol. 2019;38(7):660-9.
81. Rapp SJ, Dershem V, Zhang X, Schutte SC, Chariker ME. Varying Negative Pressure Wound Therapy Acute Effects on Human Split-Thickness Autografts. J Burn Care Res. 2020;41(1):104-12.
82. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010.
83. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012;9(4):357-9.
84. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12(1):323.
85. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15(12):550.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

725 86. Soneson C, Love M, Robinson M. Differential analyses for RNA-seq: transcript-level
726 estimates improve gene-level inferences [version 1; peer review: 2 approved]. F1000Research.
727 2015;4(1521).
728 87. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data:
729 removing the noise and preserving large differences. Bioinformatics. 2018;35(12):2084-92.
730 88. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al.
731 Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-30.

732

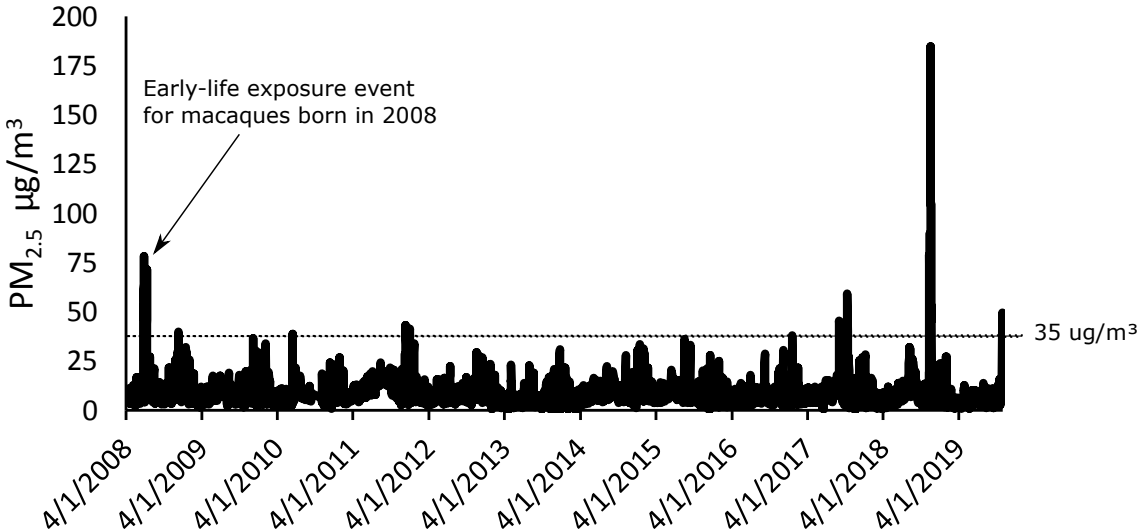

Average Daily $PM_{2.5}$ Over Time, Davis, CA
200
175
150
125
100
75
50
25
0
$PM_{2.5}$ µg/m³
Early-life exposure event
for macaques born in 2008
35 ug/m³
4/1/2008
4/1/2009
4/1/2010
4/1/2011
4/1/2012
4/1/2013
4/1/2014
4/1/2015
4/1/2016
4/1/2017
4/1/2018
4/1/2019

Figure 1


Clustering by methylation, differentially methylated regions only
Exposure
4
0
−4
Exposure
Exposed
Control

Figure 2

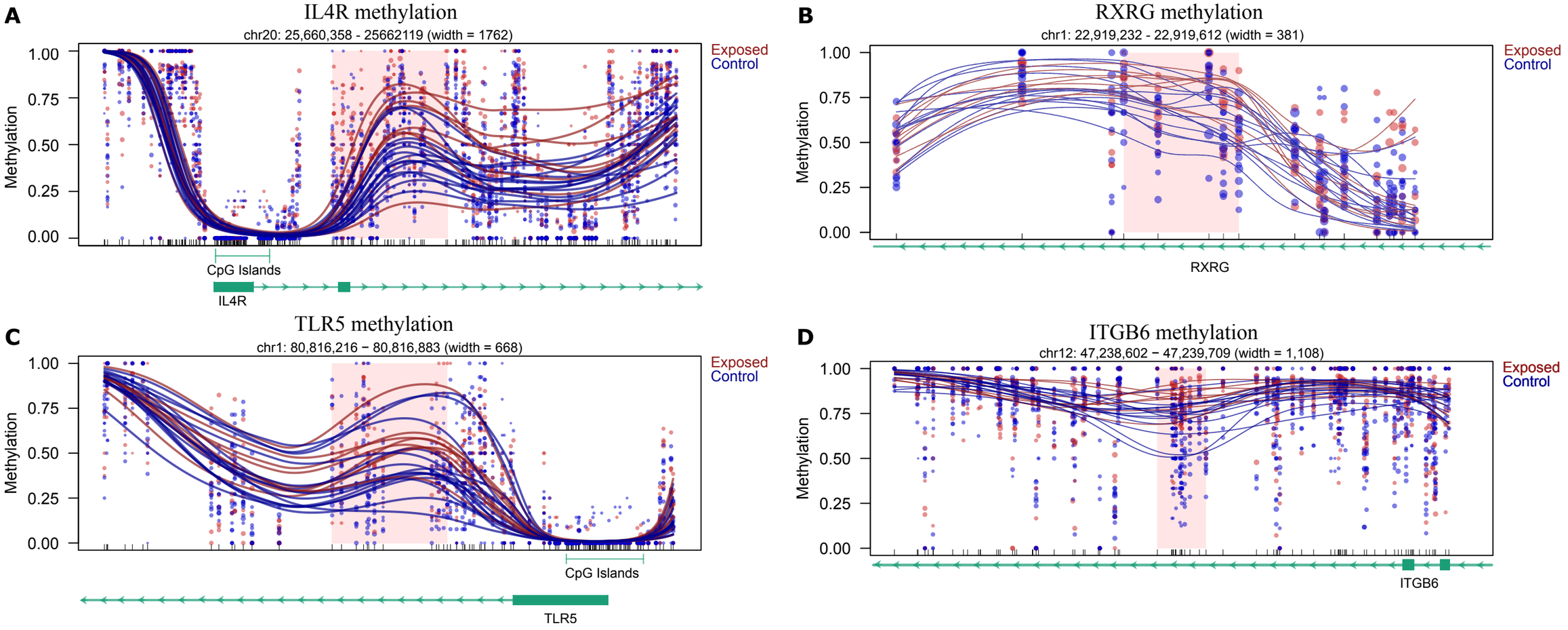

A
IL4R methylation
chr20: 25,660,358 - 25662119 (width = 1762)
Methylation
1.00
0.75
0.50
0.25
0.00
Exposed
Control
CpG Islands
IL4R
B
RXRG methylation
chr1: 22,919,232 - 22,919,612 (width = 381)
Methylation
1.00
0.75
0.50
0.25
0.00
Exposed
Control
RXRG
C
TLR5 methylation
chr1: 80,816,216 − 80,816,883 (width = 668)
Methylation
1.00
0.75
0.50
0.25
0.00
Exposed
Control
CpG Islands
TLR5
D
ITGB6 methylation
chr12: 47,238,602 − 47,239,709 (width = 1,108)
Methylation
1.00
0.75
0.50
0.25
0.00
Exposed
Control
ITGB6

Figure 3

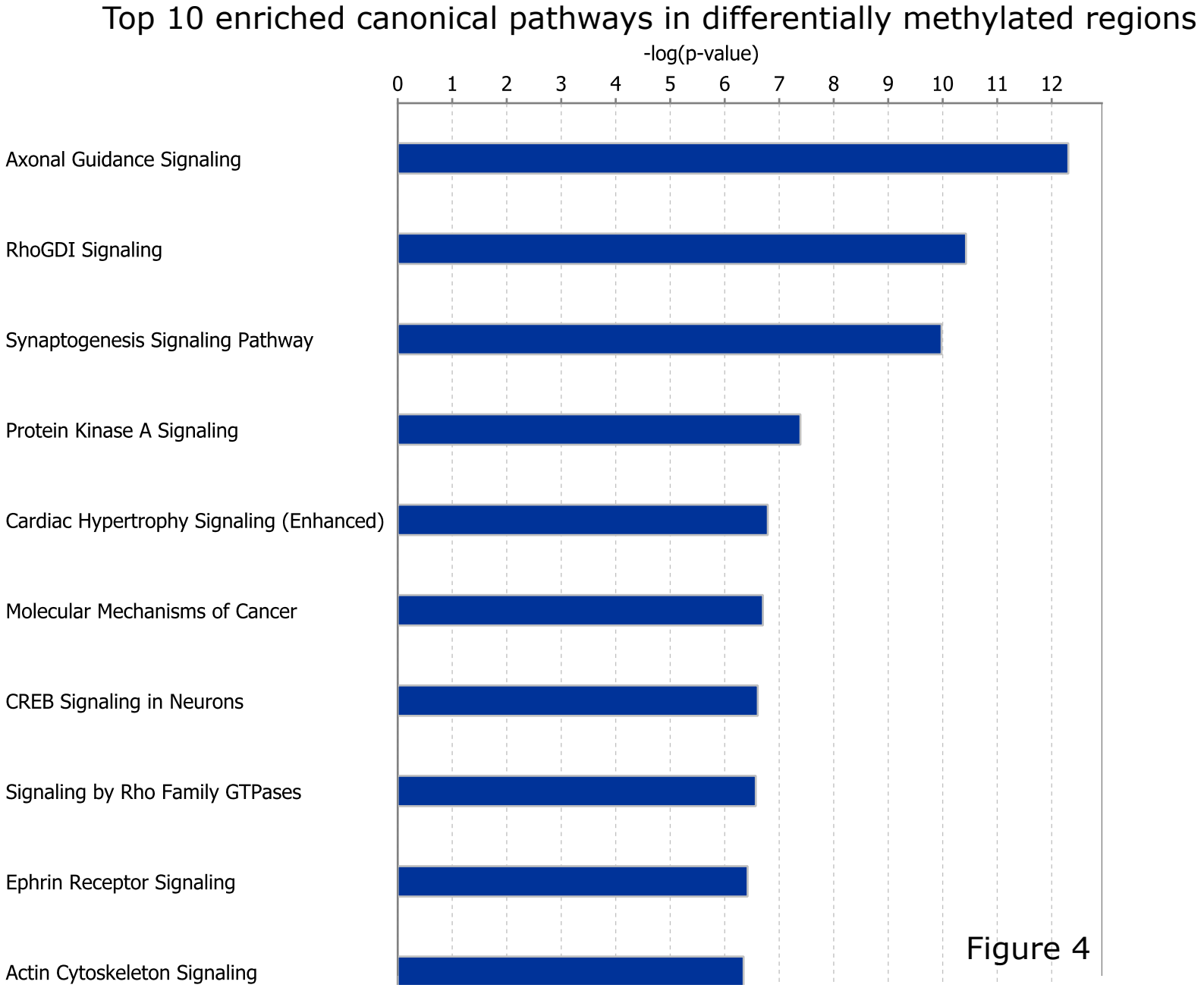

Top 10 enriched canonical pathways in differentially methylated regions
-log(p-value)
0
1
2
3
4
5
6
7
8
9
10
11
12
Axonal Guidance Signaling
RhoGDI Signaling
Synaptogenesis Signaling Pathway
Protein Kinase A Signaling
Cardiac Hypertrophy Signaling (Enhanced)
Molecular Mechanisms of Cancer
CREB Signaling in Neurons
Signaling by Rho Family GTPases
Ephrin Receptor Signaling
Actin Cytoskeleton Signaling

Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

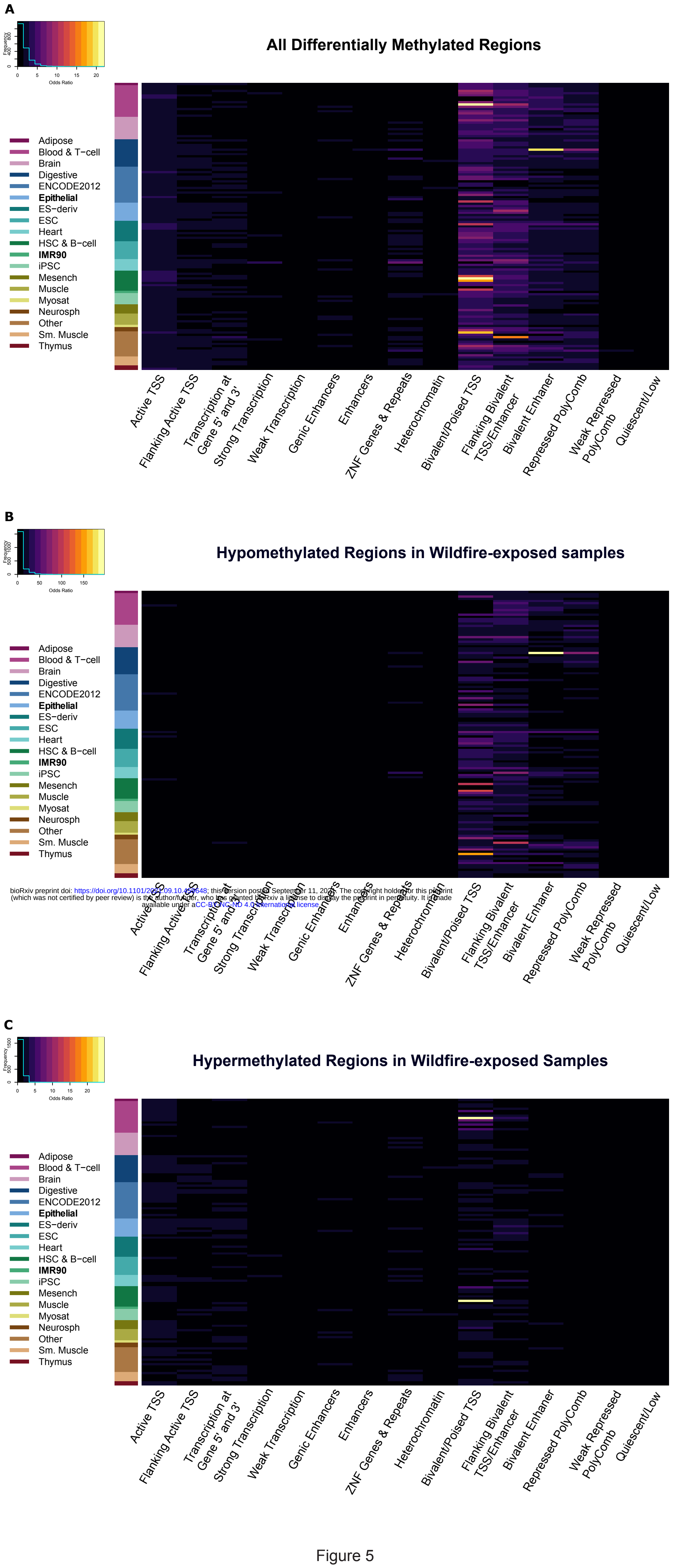


Figure 5

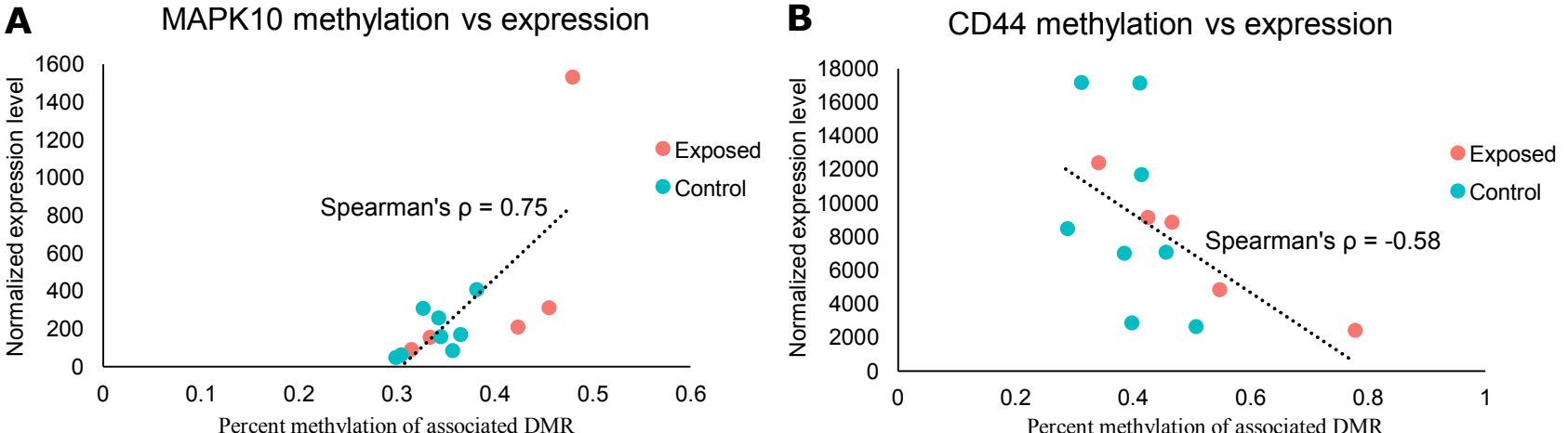

A
MAPK10 methylation vs expression
Normalized expression level
1600
1400
1200
1000
800
600
400
200
0
0
0.1
0.2
0.3
0.4
0.5
0.6
Percent methylation of associated DMR
Spearman's ρ = 0.75
Exposed
Control
B
CD44 methylation vs expression
Normalized expression level
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
0
0.2
0.4
0.6
0.8
1
Percent methylation of associated DMR
Spearman's ρ = -0.58
Exposed
Control

Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Motif	unmethylated	partially methylated	methylated
Fra1/FOSL1	0.68	0.32	0.00
Atf3	0.82	0.18	0.01
JunB	0.88	0.12	0.00
BATF	0.69	0.28	0.03
Fra2/FOSL2	0.90	0.10	0.00
AP-1/Jun	0.81	0.19	0.00
p63/TP53	0.46	0.35	0.19
NF1-halfsite(CTF)/NFIA	0.64	0.31	0.05

Table 2

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Table 1. Differentially methylated regions between rhesus macaques exposed to wildfire smoke in early life and rhesus macaques with no early life exposure to wildfire smoke. Note: betaCoefficient is presented with respect to exposed macaques (*i.e.* a positive betaCoefficient implies hypermethylation in exposed macaques compared to control macaques)

Supplementary Table 2. Canonical pathways from Ingenuity Pathway Analysis that were enriched in all differentially methylated regions.

Supplementary Table 3. Canonical pathways from Ingenuity Pathway Analysis that were enriched in differentially methylated regions hypermethylated in macaques exposed to wildfire smoke in early life.

Supplementary Table 4. Canonical pathways from Ingenuity Pathway Analysis that were enriched in differentially methylated regions hypomethylated in macaques exposed to wildfire smoke in early life.

Supplementary Table 5. Transcription factor binding site motifs that were significantly enriched in differentially methylated regions (from HOMER (1)).

Supplementary Table 6. The differentially expressed gene between rhesus macaques exposed to wildfire smoke in early life and rhesus macaques with no early life exposure to wildfire smoke. Note: log2FoldChange is presented with respect to exposed macaques (*i.e.* a positive

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

log2FoldChange implies greater expression in exposed macaques compared to control macaques).

Supplementary Table 7. Genes in the purple module (the module most significantly associated with exposure) from the weighted gene coexpression network analysis (WGCNA (2)).

Supplementary Table 8. Canonical pathways from Ingenuity Pathway Analysis that were enriched in genes in the purple module from the WGCNA (2) analysis.

Supplementary Table 9. Genes that showed significant correlation ($p \leq 0.05$) between methylation and expression across all samples.

Supplementary Table 10. Canonical pathways from Ingenuity Pathway Analysis that were enriched in genes that had significantly correlated methylation and expression.

Supplementary Table 11. Comparison between differentially methylated genes from the current study and other studies on respiratory diseases.

Supplementary Table 12. Extended information on the samples in our current study.

Supplementary Figure 1. Enrichment of different CpG features associated with all differentially methylated regions, regions hypermethylated in wildfire-exposed macaques, and regions

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hypomethylated in wildfire-exposed macaques. Asterisks indicate a significant deficit or enrichment of the feature in a given set ($p \leq 0.05$).

Supplementary Figure 2. Enrichment of different genic features associated with all differentially methylated regions, regions hypermethylated in wildfire-exposed macaques, and regions hypomethylated in wildfire-exposed macaques. Asterisks indicate a significant deficit or enrichment of the feature in a given set ($p \leq 0.05$).

Supplementary Figure 3. Heatmaps showing sample clustering by A) methylation and B) gene expression, and principal component analysis showing sample clustering by C) methylation and D) gene expression.

Supplementary Figure 4. Module-trait relationship between clusters identified in WGCNA and either exposure or animal weight. The top number in each box is the correlation value (ranging from -1 to 1), while the bottom number in parentheses is the p-value for this correlation.

Supplementary Figure 5. Top enriched biological process, cellular component, and molecular function gene ontology terms identified by GOfuncR (3) associated with differentially methylated regions between wildfire-exposed macaques and control macaques.

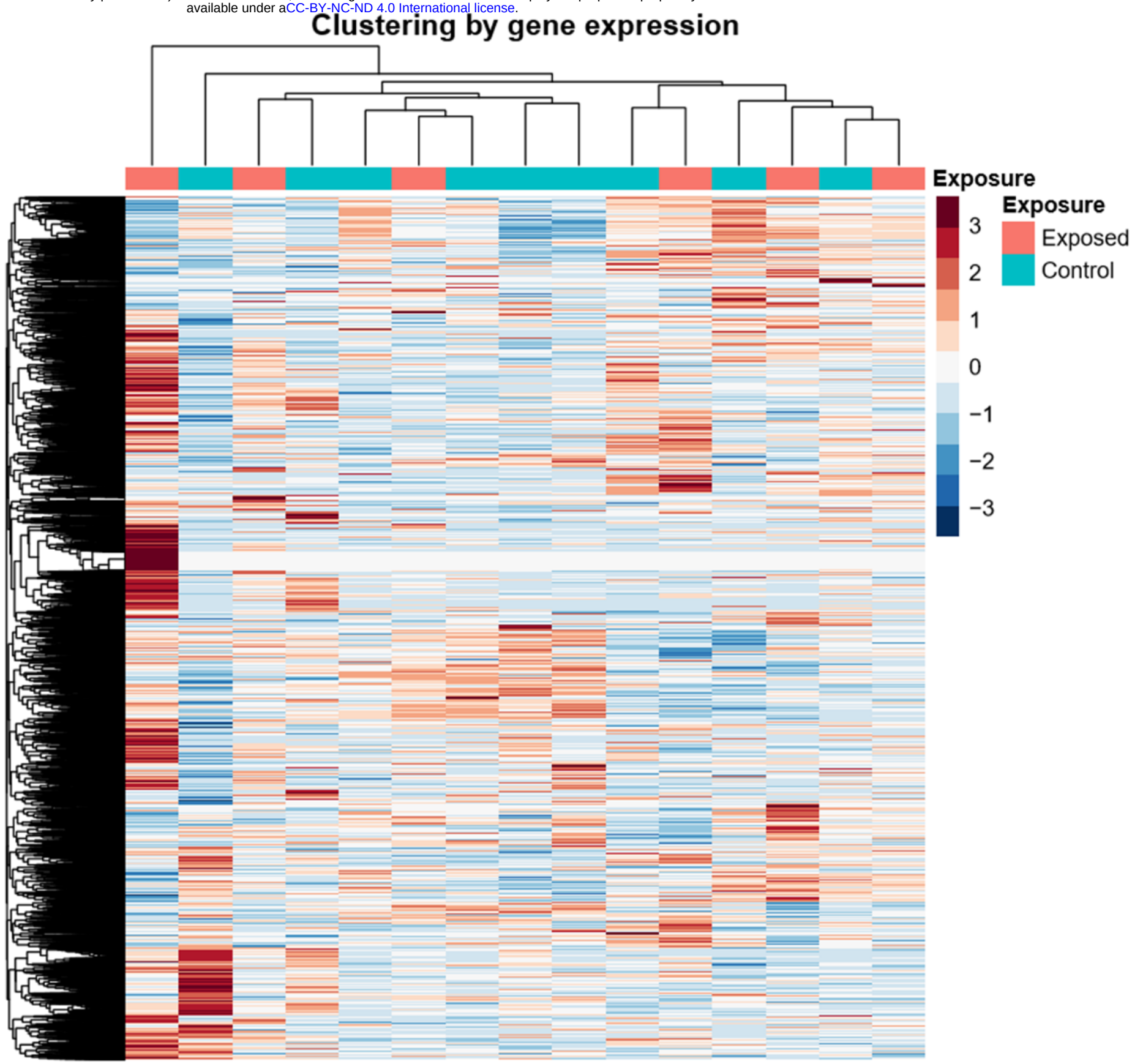
Supplementary Figure 6. Heatmap showing all samples clustering by gene expression. Gene expression data from the two leftmost samples were removed from the study as outliers.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

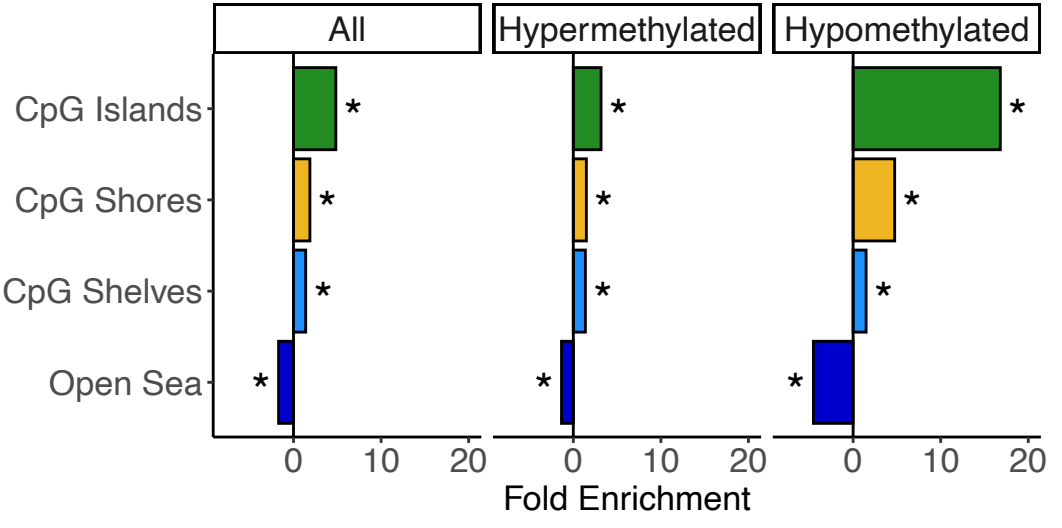
References

1. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.
2. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):559.
3. Grote S. GOfuncR: Gene ontology enrichment using FUNC. R package version 1.10.0 ed2020.

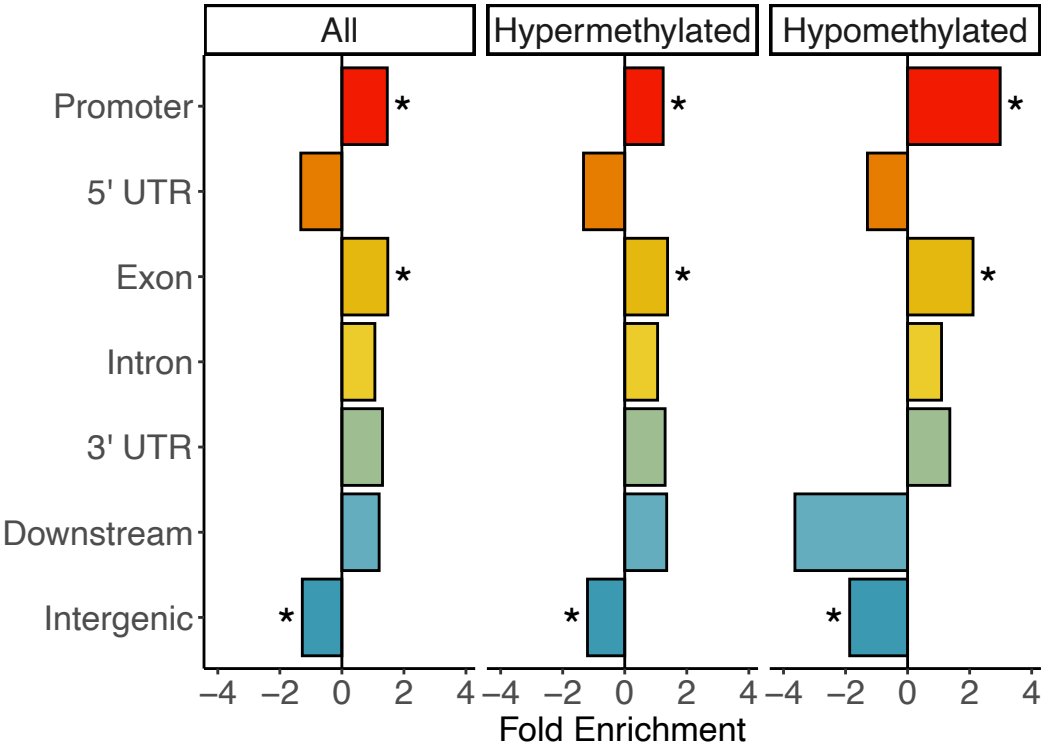
bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplemntary Figure 1

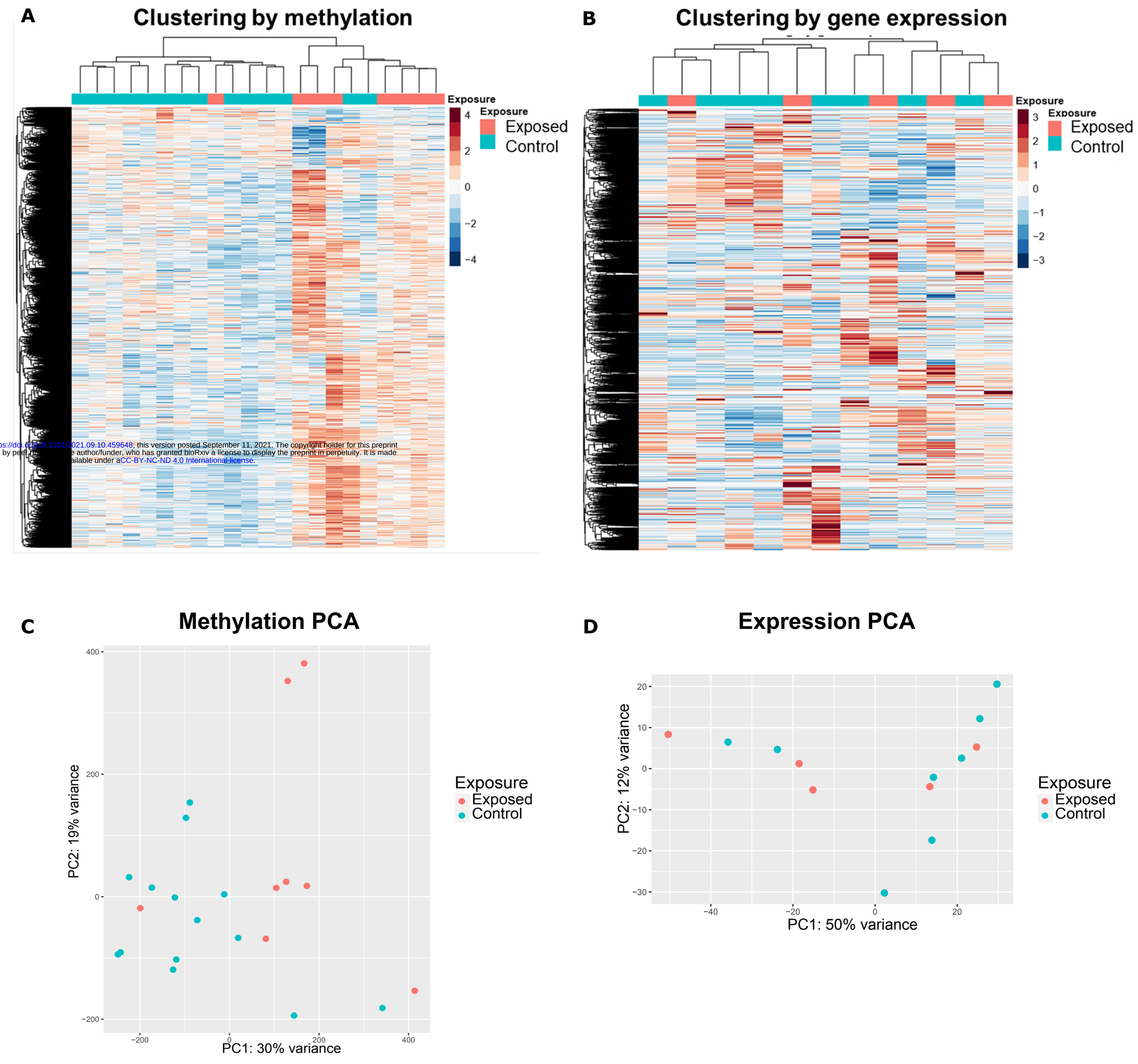


Supplemntary Figure 2

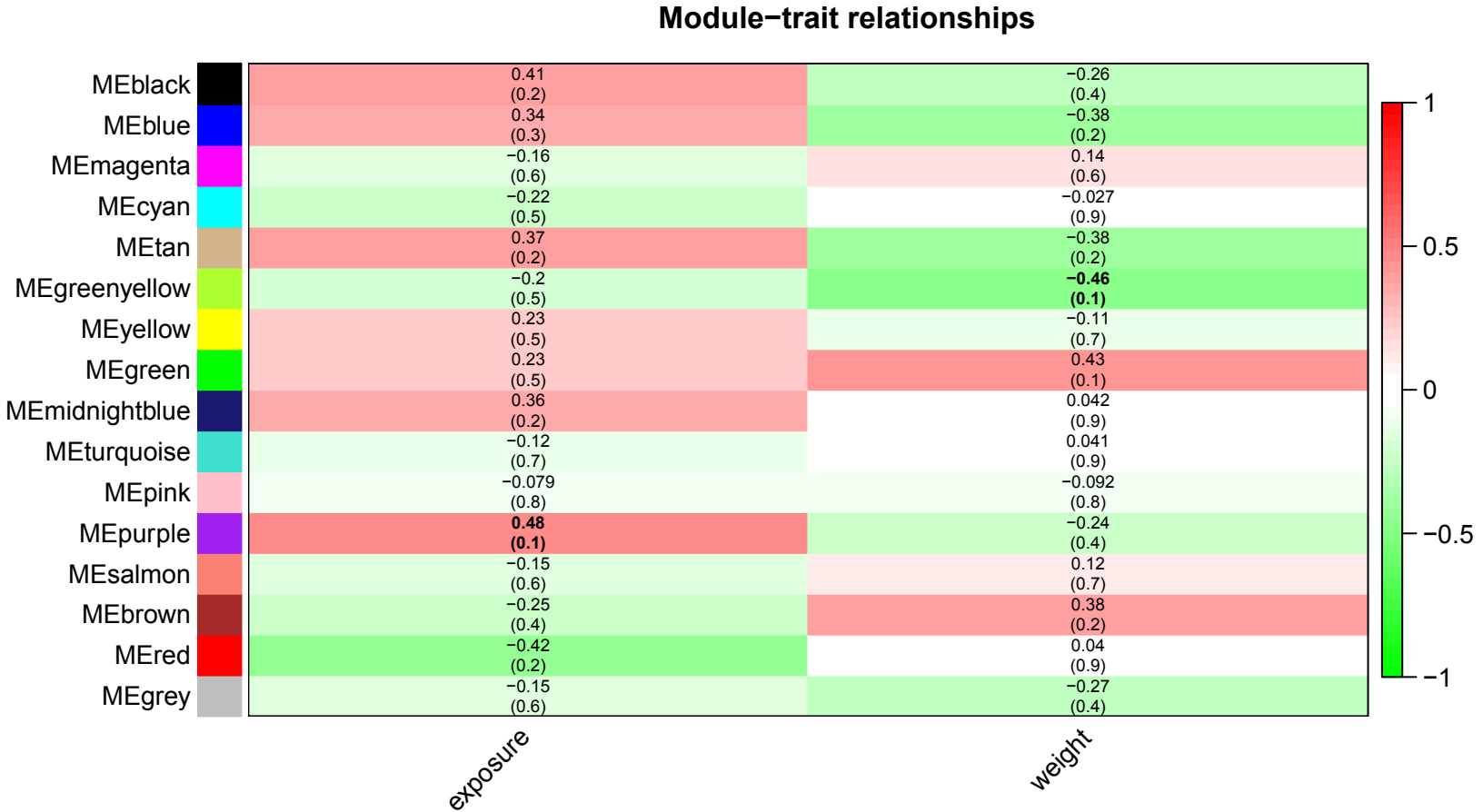


Supplemntary Figure 3

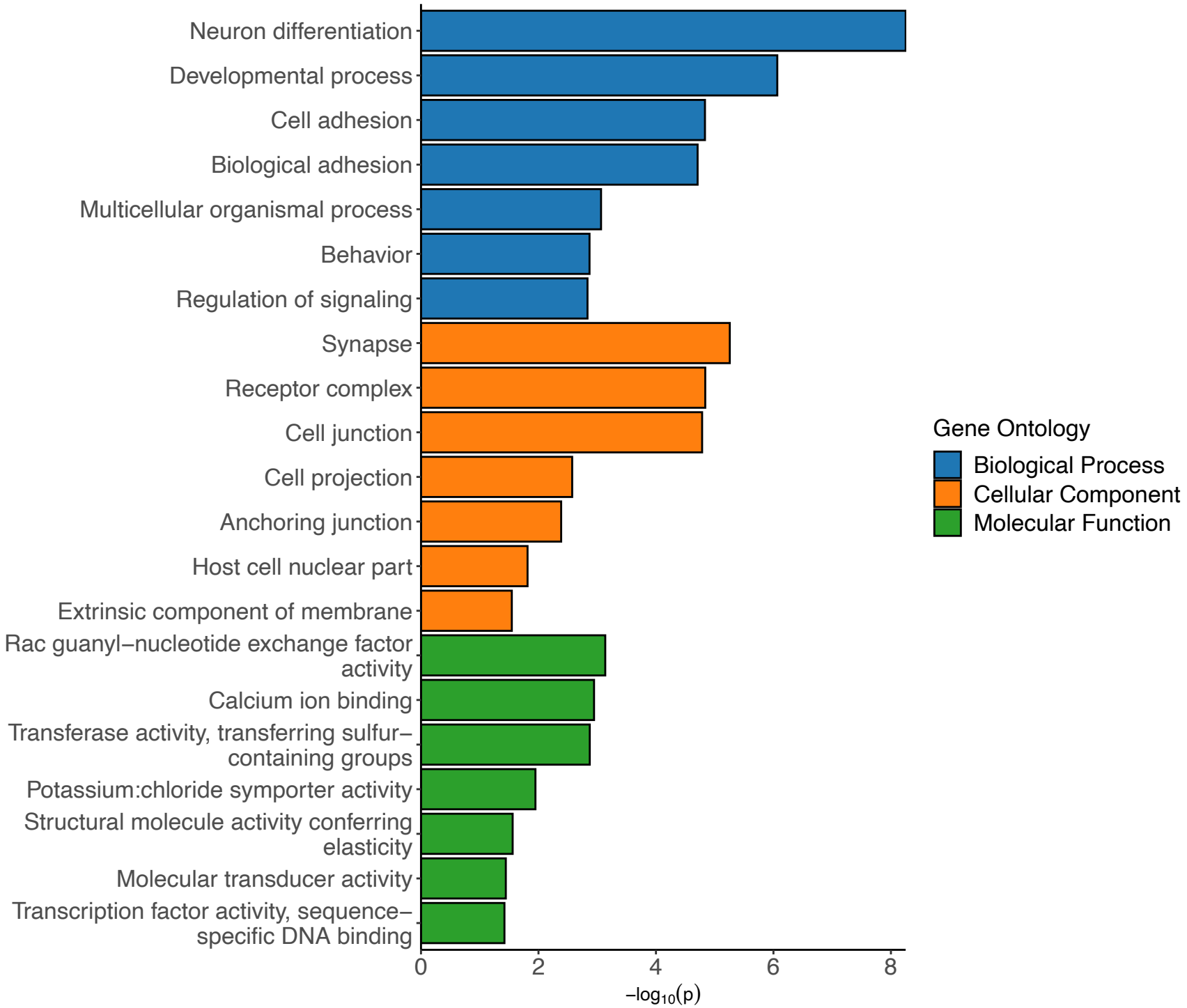
bioRxiv preprint doi: https://doi.org/10.1101/2021.09.10.459648; this version posted September 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplemntary Figure 4



Supplemntary Figure 5



Supplemntary Figure 6