

## Validation of differential DNA methylation in newborns exposed to tobacco smoke during gestation using bisulfite pyrosequencing

Christine Crute<sup>1,2</sup>, Yihan Liao<sup>3</sup>, Esther Son<sup>4</sup>, Carole Grenier<sup>2</sup>, Zhiqing Huang<sup>2</sup>, Cathrine Hoyo<sup>5</sup> and Susan K. Murphy<sup>1,2§</sup>

<sup>1</sup>Integrated Toxicology and Environmental Health Program, Nicholas School of the Environment, Duke University, Durham, NC

<sup>2</sup>Division of Reproductive Sciences, Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC

<sup>3</sup>Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, NC

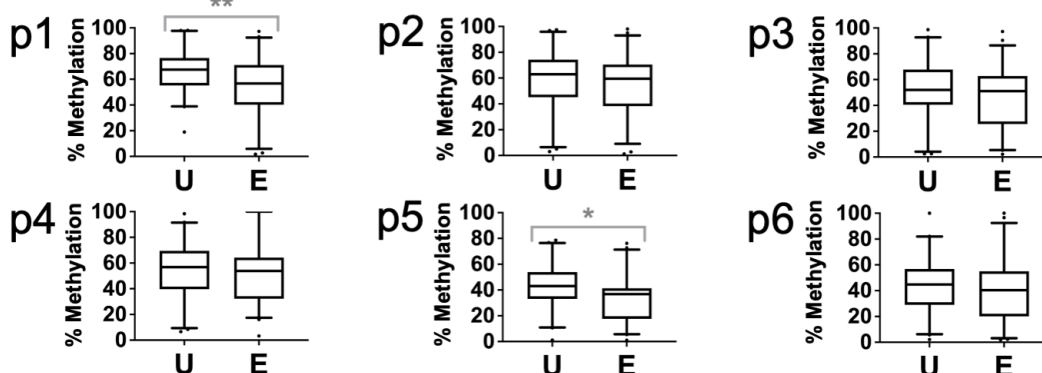
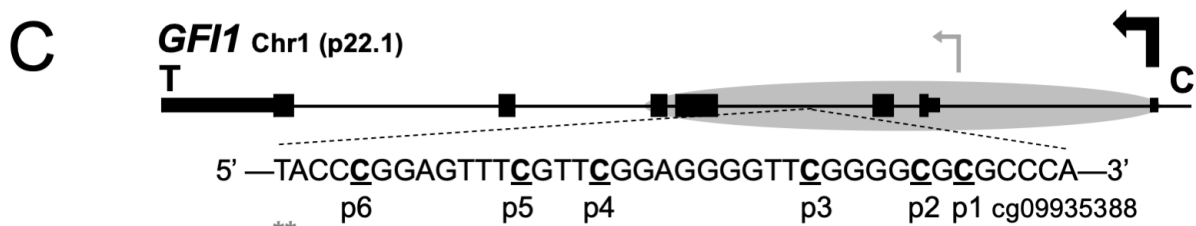
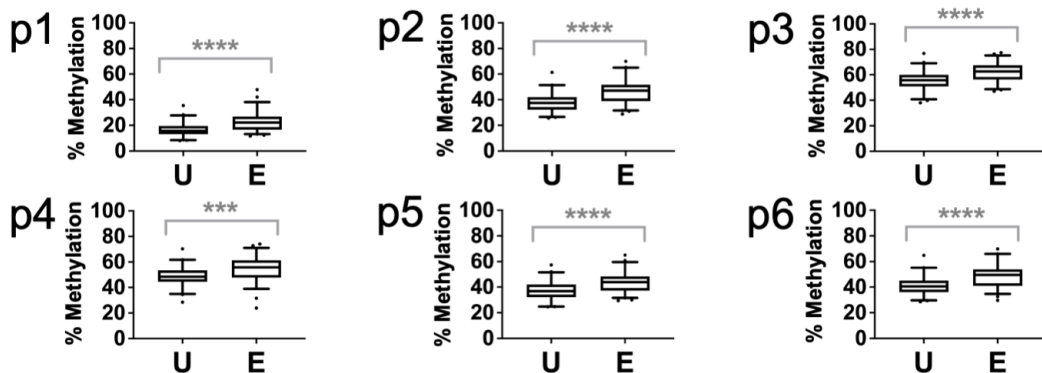
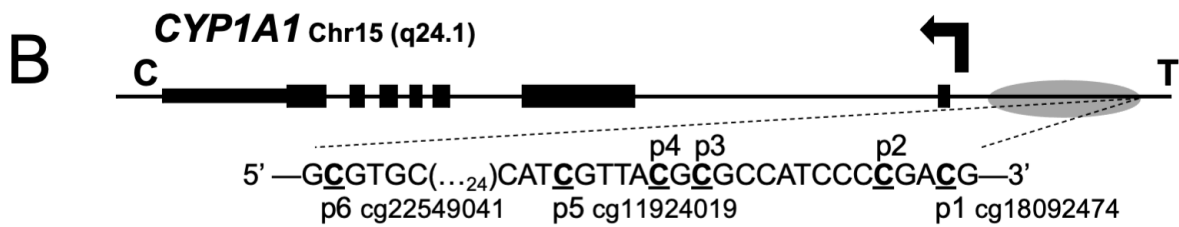
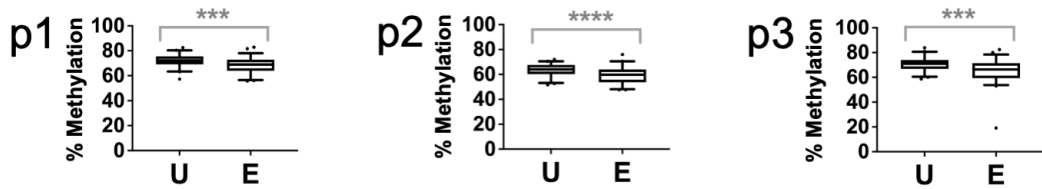
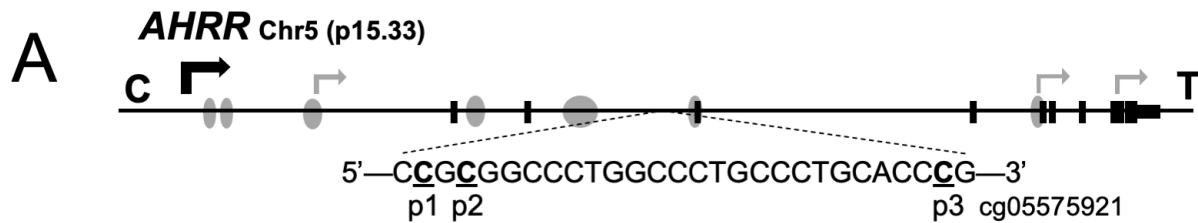
<sup>4</sup>University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>5</sup>Department of Biology, North Carolina State University, Raleigh, NC

<sup>§</sup>To whom correspondence should be addressed: susan.murphy@duke.edu

### Abstract

Maternal exposure to tobacco smoke during pregnancy has been associated with many negative child health outcomes. Tobacco smoke exposure alters DNA methylation in the developing embryo/fetus and may be a mechanism that increases risk of later life disease. Previous studies have identified CpG sites in umbilical cord blood that are associated with *in utero* tobacco smoke exposure. We sought to validate findings for CpG sites within several of the top hit genes, *AHRR*, *CYP1A1*, and *GFII*, using targeted quantitative bisulfite pyrosequencing. Comparing results from cord blood specimens of tobacco smoke-exposed to unexposed newborns, we confirmed significance at all previously identified CpG sites tested, including one in *AHRR* ( $p=0.007$ ), three in *CYP1A1* ( $p<0.0001$ ), and one in *GFII* ( $p=0.008$ ). These assays also captured novel differentially methylated CpGs located near the identified sites that were not included in the prior array-based studies ( $p$  value range, 0.02 to  $<0.0001$ ). These results validate the prior findings and provide a simplified and more economical approach to analysis of CpG sites for expanded use as biomarkers of *in utero* tobacco smoke exposure.



**Figure 1 : Figure 1.** Schematics and pyrosequencing results for unexposed (U) and exposed (E) individuals for the (A) *AHRR* gene (RefSeq: NM\_001242412, (B) *CYP1A1* gene (RefSeq: NM\_000499) and (C) *GFII* gene (RefSeq: NM\_001127216). Illumina BeadChip probe positions are designated. Black arrows, transcription start sites; gray arrows, alternative start sites. Gray ovals, CpG islands; black rectangles, exons; shorter rectangles, UTRs. Whiskers, 5-95 percentiles.  $p \leq 0.05^*$ ,  $p \leq 0.01^{**}$ ,  $p \leq 0.001^{***}$ ,  $p \leq 0.0001^{****}$ .

## Description

The Developmental Origins of Health and Disease (DOHaD) hypothesis is based on anecdotal and empirical findings that early life exposures are associated with later life susceptibility to disorders and diseases (Barker 1990). Our susceptibility to *in utero* toxicant exposures is now a major focal point in public health research. Gestational exposure to tobacco smoke is associated with low birth weight (Sbrana *et al.* 2011), airway hyperreactivity associated with asthma (Lee *et al.* 2015), and neurodevelopmental effects such as attention deficit hyperactivity disorder (Nomura *et al.* 2010; Sciberras *et al.* 2011). These outcomes have been linked to epigenetic alterations, including changes in DNA methylation (Suter *et al.* 2013; Markunas *et al.* 2014; Richmond *et al.* 2015).

A previous study utilizing the Infinium HumanMethylation450 (450k) BeadChip reported differential methylation related to *in utero* tobacco smoke exposure (Joubert *et al.* 2012) and identified 26 significant CpG probes. A number of these are in coding regions of the growth factor independent 1 transcription repressor (*GFII*) and the aryl-hydrocarbon receptor repressor (*AHRR*) genes, in addition to a region upstream of the cytochrome P450 gene, *CYP1A1*. Further studies have validated these results using the same or similar array-based methods (Richmond *et al.* 2015; Joubert *et al.* 2016; Rotroff *et al.* 2016; Rzehak *et al.* 2016). Here we sought to validate these findings using bisulfite pyrosequencing assays to measure DNA methylation in umbilical cord blood leukocytes from infants who were unexposed or exposed to tobacco smoke during pregnancy.

The sequence and positions of the specific CpGs analyzed are presented in **Figure 1A-1C**. The 146 bp PCR amplicon for *AHRR* contains three CpGs, one of which corresponds to probe cg05575921. The 226 bp *CYP1A1* amplicon contains six CpG sites, including CpGs corresponding to probes cg22549041, cg11924019, and cg18092474. The 61 bp *GFII* amplicon is 61 bp long and contains six CpG sites, one of which corresponds to probe cg09935388.

Pyrosequencing assays were tested in duplicate or triplicate with mixtures of Qiagen's Epitect fully methylated and unmethylated human DNAs containing 0%, 25%, 50%, 75% or 100% methylated DNA. Measured methylation was plotted against expected methylation levels. Goodness of fit assays for each CpG showed the assays measured incremental increases in DNA methylation as the amount of input methylated DNA increased across the range of possible values (*AHRR*:  $R^2=0.97$ ,  $p=0.0028$ ; *CYP1A1*:  $R^2>0.96$ ,  $p=0.0029$ ; *GFII*:  $R^2>0.92$ ,  $p=0.003$ ).

The assays were then used to measure methylation in human umbilical cord blood DNAs from the Newborn Epigenetics Study (NEST) cohort. Differential methylation was substantiated for the HumanMethylation450 probe CpGs for *AHRR* cg05575921 ( $p=0.0007$ ) (**Figure 1A**), *CYP1A1* cg11924019, cg22549041, and cg18092474 ( $p<0.0001$ ) (**Figure 1B**), and *GFII* cg09935388 ( $p=0.008$ ) (**Figure 1C**).

An advantage of targeted analyses by pyrosequencing is the ability to assess adjacent CpG sites that may not be represented on the 450k platform. For *AHRR*, sites p1 and p2 showed an average decrease in the exposed group of 3.9% ( $p=0.0009$ ) and 4.9% ( $p<0.0001$ ), respectively. For *CYP1A1*, sites p2, p3, and p4 showed an average increase in the exposed newborns of 8.4% ( $p<0.0001$ ), 6.7% ( $p<0.0001$ ), and 6.0% ( $p=0.0006$ ), respectively. The *GFII* assay identified only one additional CpG site as significantly differentially methylated. Site p5 exhibited a 9.1% ( $p<0.02$ ) decrease in the exposed group.

Using an independent quantitative method, we corroborated the findings from previously published studies of cord blood showing altered DNA methylation in association with *in utero* tobacco smoke exposure using array-based platforms (Joubert *et al.* 2012; Joehanes *et al.* 2016; Joubert *et al.* 2016; Tehranifar *et al.* 2018). We found significantly different levels of methylation between unexposed and exposed infant cord blood at identified sites for *AHRR*, *CYP1A1*, and *GFII*. We also showed significant differences in methylation at adjacent CpG sites that were not represented on the arrays used in the prior studies. These CpGs may collectively be useful as biomarkers for *in utero* tobacco smoke exposure using a relatively cost-efficient and scalable platform.

*AHRR* and *CYP1A1* are involved with the aryl-hydrocarbon receptor (AhR) pathway, which is known to facilitate changes in gene expression after toxicant exposure. Specifically, this pathway is known to mediate toxicity of polycyclic aromatic hydrocarbons (PAHs), several of which are found in cigarette smoke (Nguyen and Bradfield 2008). Although *GFII* has not been previously functionally associated with tobacco smoke exposure, it is involved with fundamental development processes such as hematopoiesis, inner ear and pulmonary neuroendocrine cell development, cellular proliferation, differentiation, and

apoptosis, and pre-mRNA splicing control as well as immune system response (Duan *et al.* 2005; Khandanpour *et al.* 2011; Moroy and Khandanpour 2011; Joubert *et al.* 2012; Rotroff *et al.* 2016).

A limitation of the study was that we used self-reported exposure rather than cotinine levels for some cases in which these data were missing. There is a tendency for participants to underestimate exposure, which is either intentional under-reporting due to societal expectations or a lack of awareness of secondhand exposure. The overall effect of under-reporting would diminish the differences between groups, as the group classified as unexposed may contain exposed individuals.

In conclusion, we have shown concordance between the previously reported 450k BeadChip findings and those from bisulfite pyrosequencing of an independent cohort. Our results provide the foundation for implementation of these relatively low-cost assays for defined analysis of these regions in other human cohort studies seeking to define exposure effects on these vulnerable regions of the genome.

## Methods

[Request a detailed protocol](#)

**Samples.** This research was approved by the Duke Institutional Review Board (protocol Pro00043033). Cord blood specimens were selected based on self-reported smoking status during pregnancy (exposed, mother or other person living in the household smoked tobacco during her pregnancy; unexposed, mother did not smoke and was not exposed to tobacco smoke in her residence during pregnancy). These specimens were derived from the Durham, NC-based Newborn Epigenetics Study (NEST), a mother-infant prospective cohort designed to determine the influence of early life exposures on the epigenome and health-related outcomes [see (Hoyo *et al.* 2014)].

Pregnant women were recruited from 2005 to 2011. Cord blood DNA was extracted from leukocytes separated by centrifugation from whole blood collected in vacutainer tubes containing EDTA. Samples were included in the analyses if we obtained clean PCR amplicons, the numbers of which were as follows: *AHRR*—42 unexposed, 58 exposed; *CYP1A1*—43 unexposed, 59 exposed; *GFII*—41 unexposed, 50 exposed. Of these, maternal blood collected at enrollment for ~88% underwent measurement for cotinine, a major metabolic breakdown product of nicotine. For 12 women with missing cotinine values, self-report of tobacco exposure was used for exposure status classification. There was 87% congruence between self-reported exposure and exposure status from cotinine levels for the 820 women with plasma cotinine measurements in the parent study (Schechter *et al.* 2018).

**DNA Methylation Analysis.** Genomic DNA was purified using Puregene Reagents (Qiagen) and 800 ng was modified with sodium bisulfite using the Zymo EZ DNA Methylation kit (Zymo Research). Pyrosequencing assays were designed to determine methylation levels at the previously identified significant CpG sites in *AHRR*, *CYP1A1*, and *GFII*<sup>14</sup>. All assay designs included adjacent CpGs not included on the array platform used in prior studies. Primers (Sigma) were designed using Qiagen Assay Design Software 1.0.6. *AHRR* PCR primers: F, 5'-TGG GGA TTG TTT ATT TTT GAG AG-3'; R, 5'-[biotin]AAA AAA CCC TAC CAA AAC CAC TC-3'. *CYP1A1* PCR primers: F, 5'-[biotin]ATG GGA GGT GAG GGG ATT-3'; R, 5'-CCC CAA TAC CAT TTA ACA TAA C-3'. *GFII* PCR primers: F, 5'-AGG GAG TTA AGT GGT TAG ATA A-3'; R, 5'-[biotin] TAT ACC CAC AAC ACT CCA ATT C -3'. PCR was performed in 8-well strip tubes with total reaction volumes of 10  $\mu$ L per well: 1  $\mu$ L (~20 ng) BS-treated DNA and 9  $\mu$ L master mix containing 5  $\mu$ L PyroMark PCR buffer, 0.6  $\mu$ L 25mM MgCl<sub>2</sub>, 1  $\mu$ L 10x Coral Load (Qiagen), 0.12  $\mu$ L of both primers at 10  $\mu$ M (Sigma), and 2.16  $\mu$ L molecular grade water. PCR conditions were: *AHRR*, 95°C for 15 minutes; then 55 cycles of 94°C for 30s, 58°C for 30s and 72°C for 30s; final extension 72°C for 10 minutes; *CYP1A1*, 95°C for 15 minutes; then 55 cycles of 94°C for 30s, 55°C for 30s and 72°C for 30s; final extension of 72°C for 10 minutes; *GFII*: 95°C for 15 minutes; then 5 cycles of 95°C for 30s, 66°C for 30s and 72°C for 30s; 5 cycles of 95°C for 30s, 63°C for 30s and 72°C for 30s; 5 cycles of 95°C for 30s, 60°C for 30s and 72°C for 30s; then 45 cycles for of 95°C for 30s, 57°C for 30s, and 72°C for 30s with final extension of 72°C for 10 minutes.

Sequencing primers were: *AHRR*, 5'-TTG TTT ATT TTT GAG AGG GTA-3'; *CYP1A1*, 5'- CCA AAA AAA AAA AAA TTA TAT T-3'; and *GFII*: 5'-TTA AGT GGT TAG ATA AGG AT-3'. Validation of pyrosequencing assays was performed using EpiTect DNAs (Qiagen). Pyrosequencing was performed on a Qiagen PyroMark Q96 MD instrument and percent methylation was determined using Qiagen PyroMark CpG software 1.0.11.

**Statistical analysis.** All statistical analysis was conducted using GraphPad Prism version 7.03. Data from the unexposed and exposed samples were tested for normal distribution using the D'Agostino and Pearson normality test, and statistical differences between the two groups were determined using the Welch's t-test for normally distributed data or Mann-Whitney test for data with a non-normal distribution.

**Acknowledgments:** We thank the Newborn Epigenetics Study participants without whom this research would not be possible.

## References

- Barker DJ. 1990. The fetal and infant origins of adult disease. *BMJ* 301: 1111. PMID: 2252919.
- Duan Z, Zarebski A, Montoya-Durango D, Grimes HL, Horwitz M. 2005. Gfi1 coordinates epigenetic repression of p21Cip/WAF1 by recruitment of histone lysine methyltransferase G9a and histone deacetylase 1. *Mol Cell Biol* 25: 10338-51. PMID: 16287849.
- Hoyo C, Daltveit AK, Iversen E, Benjamin-Neelon SE, Fuemmeler B, Schildkraut J, Murtha AP, Overcash F, Vidal AC, Wang F, Huang Z, Kurtzberg J, Seewaldt V, Forman M, Jirtle RL, Murphy SK. 2014. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics* 9: 1120-30. PMID: 24874916.
- Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, Guan W, Xu T, Elks CE, Aslibekyan S, Moreno-Macias H, Smith JA, Brody JA, Dhingra R, Yousefi P, Pankow JS, Kunze S, Shah SH, McRae AF, Lohman K, Sha J, Absher DM, Ferrucci L, Zhao W, Demerath EW, Bressler J, Grove ML, Huan T, Liu C, Mendelson MM, Yao C, Kiel DP, Peters A, Wang-Sattler R, Visscher PM, Wray NR, Starr JM, Ding J, Rodriguez CJ, Wareham NJ, Irvin MR, Zhi D, Barrdahl M, Vineis P, Ambatipudi S, Uitterlinden AG, Hofman A, Schwartz J, Colicino E, Hou L, Vokonas PS, Hernandez DG, Singleton AB, Bandinelli S, Turner ST, Ware EB, Smith AK, Klengel T, Binder EB, Psaty BM, Taylor KD, Gharib SA, Swenson BR, Liang L, DeMeo DL, O'Connor GT, Herceg Z, Ressler KJ, Conneely KN, Sotoodehnia N, Kardia SL, Melzer D, Baccarelli AA, van Meurs JB, Romieu I, Arnett DK, Ong KK, Liu Y, Waldenberger M, Deary IJ, Fornage M, Levy D, London SJ. 2016. Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet* 9: 436-447. PMID: 27651444.
- Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, Reese SE, Markunas CA, Richmond RC, Xu CJ, Küpers LK, Oh SS, Hoyo C, Gruziova O, Söderhäll C, Salas LA, Baiz N, Zhang H, Lepeule J, Ruiz C, Ligthart S, Wang T, Taylor JA, Duijts L, Sharp GC, Jankipersadsing SA, Nilsen RM, Vaez A, Fallin MD, Hu D, Litonjua AA, Fuemmeler BF, Huen K, Kere J, Kull I, Munthe-Kaas MC, Gehring U, Bustamante M, Saurel-Coubizolles MJ, Quraishi BM, Ren J, Tost J, Gonzalez JR, Peters MJ, Håberg SE, Xu Z, van Meurs JB, Gaunt TR, Kerkhof M, Corpeleijn E, Feinberg AP, Eng C, Baccarelli AA, Benjamin Neelon SE, Bradman A, Merid SK, Bergström A, Herceg Z, Hernandez-Vargas H, Brunekreef B, Pinart M, Heude B, Ewart S, Yao J, Lemonnier N, Franco OH, Wu MC, Hofman A, McArdle W, Van der Vlies P, Falahi F, Gillman MW, Barcellos LF, Kumar A, Wickman M, Guerra S, Charles MA, Holloway J, Auffray C, Tiemeier HW, Smith GD, Postma D, Hivert MF, Eskenazi B, Vrijheid M, Arshad H, Antó JM, Dehghan A, Karmaus W, Annesi-Maesano I, Sunyer J, Ghantous A, Pershagen G, Holland N, Murphy SK, DeMeo DL, Burchard EG, Ladd-Acosta C, Snieder H, Nystad W, Koppelman GH, Relton CL, Jaddoe VW, Wilcox A, Melén E, London SJ. 2016. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet* 98: 680-96. PMID: 27040690.
- Joubert BR, Håberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, Huang Z, Hoyo C, Middttun Ø, Cupul-Uicab LA, Ueland PM, Wu MC, Nystad W, Bell DA, Peddada SD, London SJ. 2012. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect* 120: 1425-31. PMID: 22851337.
- Khandanpour C, Kosan C, Gaudreau MC, Dührsen U, Hébert J, Zeng H, Möröy T. 2011. Growth factor independence 1 protects hematopoietic stem cells against apoptosis but also prevents the development of a myeloproliferative-like disease. *Stem Cells* 29: 376-85. PMID: 21732494.
- Lee JW, Jaffar Z, Pinkerton KE, Porter V, Postma B, Ferrini M, Holian A, Roberts K, Cho YH. 2015. Alterations in DNA methylation and airway hyperreactivity in response to in utero exposure to environmental tobacco smoke. *Inhal Toxicol* 27: 724-30. PMID: 26525176.
- Markunas CA, Xu Z, Harlid S, Wade PA, Lie RT, Taylor JA, Wilcox AJ. 2014. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ Health Perspect* 122: 1147-53. PMID: 24906187.
- Möröy T, Khandanpour C. 2011. Growth factor independence 1 (Gfi1) as a regulator of lymphocyte development and activation. *Semin Immunol* 23: 368-78. PMID: 21920773.
- Nguyen LP, Bradfield CA. 2008. The search for endogenous activators of the aryl hydrocarbon receptor. *Chem Res Toxicol* 21: 102-16. PMID: 18076143.
- Nomura Y, Marks DJ, Halperin JM. 2010. Prenatal exposure to maternal and paternal smoking on attention deficit hyperactivity disorders symptoms and diagnosis in offspring. *J Nerv Ment Dis* 198: 672-8. PMID: 20823730.
- Richmond RC, Simpkin AJ, Woodward G, Gaunt TR, Lyttleton O, McArdle WL, Ring SM, Smith AD, Timpson NJ, Tilling K, Davey Smith G, Relton CL. 2015. Prenatal exposure to maternal smoking and offspring DNA methylation across the

lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet* 24: 2201-17. PMID: 25552657.

Rotroff DM, Joubert BR, Marvel SW, Håberg SE, Wu MC, Nilsen RM, Ueland PM, Nystad W, London SJ, Motsinger-Reif A. 2016. Maternal smoking impacts key biological pathways in newborns through epigenetic modification in Utero. *BMC Genomics* 17: 976. PMID: 27887572.

Rzehak P, Saffery R, Reischl E, Covic M, Wahl S, Grote V, Xhonneux A, Langhendries JP, Ferre N, Closa-Monasterolo R, Verduci E, Riva E, Socha P, Gruszfeld D, Koletzko B, European Childhood Obesity Trial Study group.. 2016. Maternal Smoking during Pregnancy and DNA-Methylation in Children at Age 5.5 Years: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study. *PLoS One* 11: e0155554. PMID: 27171005.

Sbrana E, Suter MA, Abramovici AR, Hawkins HK, Moss JE, Patterson L, Shope C, Aagaard-Tillery K. 2011. Maternal tobacco use is associated with increased markers of oxidative stress in the placenta. *Am J Obstet Gynecol* 205: 246.e1-7. PMID: 21803321.

Schechter JC, Fuemmeler BF, Hoyo C, Murphy SK, Zhang JJ, Kollins SH. 2018. Impact of Smoking Ban on Passive Smoke Exposure in Pregnant Non-Smokers in the Southeastern United States. *Int J Environ Res Public Health* 15: 83. PMID: 29316617.

Sciberras E, Ukoumunne OC, Efron D. 2011. Predictors of parent-reported attention-deficit/hyperactivity disorder in children aged 6-7 years: a national longitudinal study. *J Abnorm Child Psychol* 39: 1025-34. PMID: 21468666.

Suter MA, Anders AM, Aagaard KM. 2013. Maternal smoking as a model for environmental epigenetic changes affecting birthweight and fetal programming. *Mol Hum Reprod* 19: 1-6. PMID: 23139402.

Tehranifar P, Wu HC, McDonald JA, Jasmine F, Santella RM, Gurvich I, Flom JD, Terry MB. 2018. Maternal cigarette smoking during pregnancy and offspring DNA methylation in midlife. *Epigenetics* 13: 129-134. PMID: 28494218.

**Funding:** This work was supported by NIEHS P01ES022831, USEPA RD-83543701, NCRR UL1TR001117 and NIEHS T32-ES021432.

**Author Contributions:** Christine Crute: Data curation, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review and editing. Yihan Liao: Investigation, Formal analysis, Visualization. Esther Son: Investigation, Formal analysis, Visualization. Carole Grenier: Investigation, Methodology. Zhiqing Huang: Formal analysis, Investigation, Methodology. Cathrine Hoyo: Funding acquisition, Resources. Susan K. Murphy: Conceptualization, Formal analysis, Funding acquisition, Methodology, Resources, Project administration, Supervision, Writing - review and editing.

**Reviewed By:** Wilfried Karmaus

**History:** Received August 23, 2021 **Revision received** December 5, 2021 **Accepted** December 29, 2021 **Published** January 7, 2022

**Copyright:** © 2022 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** Crute, C; Liao, Y; Son, E; Grenier, C; Huang, Z; Hoyo, C; Murphy, SK (2022). Validation of differential DNA methylation in newborns exposed to tobacco smoke during gestation using bisulfite pyrosequencing. *microPublication Biology*. <https://doi.org/10.17912/micropub.biology.000509>