Title: Do genetic variants modify the effect of smoking on risk of preeclampsia in pregnancy?

Authors: Anna E. Bauer, PhD[a,1], Christy L. Avery, PhD[a, b], Min Shi, PhD[c], Clarice R. Weinberg, PhD[c], Andrew F. Olshan, PhD[a], Quaker E. Harmon, MD, PhD[d], Jingchun Luo, PhD[e], Jenny Yang, PhD[f], Tracy Manuck, MD, MSCI[g], Michael C. Wu, PhD[h], Kari Klungsøyr, PhD[i,j], Lill Trogstad, MD, PhD[i], Per Magnus, MD, PhD[k], Stephanie M. Engel, PhD[a]

[a] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, CB# 7435, Chapel Hill, NC, 27599-7435, United States

[b] Carolina Population Center, University of North Carolina at Chapel Hill, 123 West Franklin St, Chapel Hill, NC, 27516, United States

[c] Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, P.O. Box 12233, Mail Drop A3-03, Durham, NC, 27709, United States

[d] Epidemiology Branch, National Institute of Environmental Health Sciences, P.O. Box 12233, Mail Drop A3-05, Durham, NC, 27709, United States

[e] Mammalian Genotyping Core, University of North Carolina at Chapel Hill, Carolina Crossing C, 2234 Nelson Highway, Chapel Hill, NC, 27517, United States

[f] Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, CB# 7420, Chapel Hill, NC, 27599-7420, United States

[g] Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill, 3009 Old Clinic Building, CB# 7570, Chapel Hill, NC, 27599-7570, United States

[h] Biostatistics and Biomathematics Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-8500, Seattle, WA 98109, United States

[i] Division for Mental and Physical Health, Norwegian Institute of Public Health, P.O. Box 222 Skøyen, 0213 Oslo, Norway

[j] Department of Global Public Health and Primary Care, University of Bergen, P.O. Box 7804, N-5020, Bergen, Norway

[k] Centre for Fertility and Health, Norwegian Institute of Public Health, P.O. Box 222 Skøyen, N-0213 Oslo, Norway.

[1]Present Address: Department of Psychiatry, Perinatal Psychiatry Program**,** School of Medicine, University of North Carolina at Chapel Hill, CB# 7160**,** Chapel Hill, NC 27599-7160

The authors report no conflict of interest.

MS, CRW, and QEH are employees of US federal agency, National Insitute of Environmental Health Sciences.

Corresponding Author:

Anna E. Bauer, PhD, MPH

234 Medical Wing C

University of North Carolina at Chapel Hill

Chapel Hill, NC 27599-7160

O: (919) 962-6892

C: (919) 945-4063

anna_bauer@med.unc.edu

Abstract word count: 264

Main text word count: 2,994

ABSTRACT

Objective: Maternal smoking is associated with as much as a 50% reduced risk of preeclampsia, despite increasing risk of other poor pregnancy outcomes that often co-occur with preeclampsia, such as preterm birth and fetal growth restriction. Researchers have long sought to understand whether the this perplexing association is biologically based, or a result of non-causal mechanisms. We examined whether smoking-response genes modify the smoking-preeclampsia association to investigate potential biological explanations.

Study Design: We conducted a nested case-control study within the Norwegian Mother, Father and Child Birth Cohort (1999-2008) of 2,596 mother-child dyads. We used family-based log-linear Poisson regression to examine modification of the maternal smoking-preeclampsia relationship by maternal and fetal single nucleotide polymorphisms involved in cellular processes related to components of cigarette smoke (n=1,915 with minor allele frequency ≥10%). We further investigated the influence of smoking cessation during pregnancy.

Results: Three polymorphisms showed overall ($p<0.001$) multiplicative interaction between smoking and maternal genotype. For rs3765692 (*TP73*) and rs10770343 (*PIK3C2G*), protection associated with smoking was reduced with two maternal copies of the risk allele and was stronger in continuers than quitters (interaction *P*=0.02 for both loci, based on testing 3-level smoking by 3-level genotype). For rs2278361 (*APAF1*) the inverse smoking-preeclampsia association was eliminated by the presence of a single risk allele, and again the trend was stronger in continuers than in quitters (interaction *P*=0.01).

Conclusion: Evidence for gene-smoking interaction was limited, but differences by smoking cessation warrant further investigation. We demonstrate the potential utility of expanded dyad

methods and gene-environment interaction analyses for outcomes with complex relationships between maternal and fetal genotypes and exposures.

Keywords: carbon monoxide; detoxification; gene-environment interaction; genetics; MoBa; mother-child dyad; nitric oxide; Norwegian Mother, Father and Child Cohort Study; preeclampsia; pregnancy; smoking

Abbreviations: *APAF1*: Apoptotic Peptidase Activating Factor 1 (gene); CO: Carbon monoxide**;** HO-1: Heme oxygenase 1; MoBa: Norwegian Mother, Father and Child Cohort Study**;** PIK3C2G: Phosphatidylinositol-4-phosphate 3-kinase C2 domain-containing gamma polypeptide (gene)**;** NO: Nitric oxide**;** sFlt-1: Soluble fms-like tyrosine kinase 1; SNP: Single nucleotide polymorphism**;** TP73: Tumor protein P73 (gene)

Key Points:

- Maternal and fetal genotype may differentially influence preeclampsia
- Smoking-related genes did not strongly modify smoking-preeclampsia association
- Smoking cessation reduced strength of gene by smoking interactions

INTRODUCTION

Preeclampsia is a common pregnancy complication, affecting approximately 2-7% of pregnant women, typically characterized by new-onset gestational hypertension and proteinuria after 20 weeks gestation.[1] Preeclampsia is associated with serious maternal and fetal morbidity and mortality and there are limited options for treatments.[1] The etiology of preeclampsia is unknown, but poor placental development is hypothesized to be involved, which may be influenced by both maternal and fetal factors.[2]

Many risk factors for preeclampsia have been identified across studies.[1] One of the strongest and most consistent associations, yet most poorly understood, is the inverse relationship between maternal smoking and preeclampsia. Maternal smoking is associated with as much as a 50% reduced risk of preeclampsia, despite increasing risk of other adverse pregnancy outcomes that often co-occur with preeclampsia.[3,4] The reason for this association remains unknown; both biological causes[5] and methodological[6] reasons (e.g., survival bias) have been suggested.

However, there is good reason to believe that at least some part of the overall smoking-preeclampsia association is biologically mediated. Some of the effect appears to depend on combustion, as Swedish users of snus (an oral tobacco product) have increased risk.[7] Further, women who stop smoking later in pregnancy have an attenuated reduction in risk, which is not compatible with a survival bias scenario.[5] One biological hypothesis for the reduced risk of preeclampsia among smokers is through the powerful, vasodilatory action of carbon monoxide (CO) and nitric oxide (NO),[8,9] which are produced by combustion of cigarettes[10] and endogenously in the body.[11–13] NO and CO are thought to be required for fetal trophoblast

differentiation and invasion into the maternal spiral arteries,[9,12] and trophoblast invasion and subsequent spiral artery remodeling are required for normal placental development.[2]

Reduced NO bioavailability during invasion may result in high vascular resistance and hypoxia,[14,15] contributing to hypertensive disorders of pregnancy.[16] NO production and serum metabolites of NO are lower among women with preeclampsia than women with normal pregnancy[13] and inadequate trophoblast invasion occurs more frequently in pregnancies affected by preeclampsia.[17] Similarly, high early-pregnancy levels of the antiangiogenic factor soluble fms-like tyrosine kinase 1 (sFlt-1) are strongly associated with preeclampia,[18] and CO directly inhibits sFlt-1.[19] CO is produced by degradation of heme catalyzed by heme oxygenase-1 (HO-1), and HO-1 expression is lower in preeclampsia placentae than those from normal pregnancies.[11] Experimental induction of HO-1 in rodents reduces hypertension and corrects the angiogenic imbalance characteristic of preeclampsia[20] and in human studies, women with preeclampsia have decreased concentrations of CO in exhaled breath.[19,21]

We recently investigated the relationship of genetic variation in NO and CO signaling pathways with preeclampsia risk, but not variability by maternal smoking.[22] Smoking detoxification pathways may differentially influence preeclampsia in the presence of maternal smoking, suggesting a potential role for gene-smoking interactions.[23] We selected a comprehensive list of potentially relevant genes in pathways involved in response to cigarette smoke components based on biological plausibility, and prior evidence for interactions with smoking in other conditions.[24–27] Some of these genes (*TNFα, CYP1A1, GSTT1)* have been investigated for their interaction with smoking for other perinatal outcomes,[28,29] but not for

preeclampsia; a study of *MTHFR* and folate is the only other investigation of gene by environment interactions for preeclampsia.[29]

The investigation of gene-smoking interactions in the etiology of preeclampsia is challenging for multiple reasons. First, preeclampsia is a disease that is likely influenced by both maternal and fetal genetic features;[30,31] maternal and fetal genetics are correlated which must be accounted for in the analysis. Second, cigarette smoke is a complex and time-varying exposure, consisting of over 5,000 chemicals of potential interest,[32] which interact with multiple pathways influenced by genetic contributions. A thorough investigation of this hypothesis, therefore, requires the consideration of both detoxification and signaling pathways, while accounting for the relationship between maternal and fetal genetic contributions.

We aimed to investigate the occurrence of multiplicative interactions between maternal smoking and genetic variants in pathways involved in the biological response to cigarette smoke components on risk of preeclampsia, considering both maternal and fetal genes. Observing differential associations between smoking and preeclampsia by genotype may help to explain the enigmatic inverse association of smoking with preeclampsia, and potentially identify mechanistic targets for future research and ultimate prevention.

MATERIALS AND METHODS

Study Population

We performed a nested case-control study within the Norwegian Mother, Father and Child Cohort Study (MoBa), a large prospective birth cohort of pregnant women and their offspring recruited throughout Norway from 1999 to 2008 (N=112,908 pregnancies), which has

been previously described.[33] Participants completed two prenatal questionnaires about their health and behaviors. Maternal blood was collected at the first ultrasound appointment and cord (child) blood was collected at birth. DNA was extracted from both and stored at the MoBa Biobank.[34]

Women provided informed consent prior to participation in MoBa. Data collection for MoBa was approved by the Norwegian Data Inspectorate and the Norwegian Committee for Medical and Health Research Ethics. The study was also approved by the Institutional Review Board at UNC Chapel Hill.

Outcome Assessment

Preeclampsia information from MoBa was obtained through linkage with the Medical Birth Registry of Norway[35] and verified by antenatal records through an independent validation study.[36] Preeclampsia was defined as *de novo* hypertension (systolic blood pressure ≥ 140 mm Hg or diastolic blood pressure of ≥ 90 mm Hg) with proteinuria (urine protein ≥ 0.3 g/24-hour or 1+ on urine dipstick) using criteria of the American College of Obstetrics and Gynecologists (ACOG) current at the time of the validation study.[37] Preeclampsia is diagnosed after gestational week 20. For the present study, we included all women from the validation study with a singleton pregnancy who conceived spontaneously, were verified as cases or controls, returned both (early and late) pregnancy questionnaires, and had no history of chronic hypertension. In total, 2,682 preeclampsia case samples (1,564 maternal and 1,118 fetal blood samples) and 1,967 non-preeclampsia control samples (999 maternal and 968 fetal samples) met inclusion criteria and were genotyped.

Maternal Smoking

We created a dichotomous smoking variable to indicate self-report of any smoking in gestational weeks 11 through 20. We selected this etiological window to capture smoking during the time in which physiologic processes facilitate maternal blood flow perfusion and trophoblasts become invasive to complete spiral artery remodeling. Women were surveyed in early (13-17 weeks) and late (~30 weeks) pregnancy about smoking habits, including smoking prior to pregnancy, current smoking, quantity of cigarettes smoked, and gestational week of quitting if they stopped smoking during pregnancy. A woman was considered a non-smoker if she indicated she never smoked or did not currently smoke, and we had no other evidence of smoking after 10 weeks of gestation.

Single Nucleotide Polymorphism Selection

For this study, 124 genes involved in signaling, metabolism, or dexotification of cigarette smoke or its components of were identified from the following 8 canonical pathways (Supplementary Table S1): 1) endothelial nitric oxide synthase signaling, 2) heme degradation, 3) hypoxia-inducible factor 1-alpha, 4) xenobiotic metabolism, 5) aryl hydrocarbon receptor signaling, 6) glutathione-mediated detoxification, 7) nicotine degradation II, and 8) nicotine degradation III. A total of 1,915 single nucleotide polymorphisms (SNPs; Minor Alelle Frequency $\geq$ 10%) were selected and analyzed for this study, using a 10kb upstream and downstream margin around the transcription start and end sites of each gene.

Genotyping and Quality Control

SNPs were genotyped by the UNC Mammalian Genotyping Core using the HumanCoreExome+ array from Illumina (Illumina, Inc., San Diego, CA). Samples and SNPs in controls were examined using PLINK 1.07 (http://pngu.mgh.harvard.edu/purcell/plink) for quality control. SNPs were excluded for missing rate > 5%, deviation from Hardy-Weinberg Equilibrium ($P < 1x10^{-3}$) in controls, or minor allele frequency <10%. Known genotype and DNA replicates were included on each plate and exhibited high genotyping quality. All subject-specific call rates were acceptable (minimum 97.2%). Sex-specific markers were inspected and potential inbreeding was examined and parent-child relationships were confirmed by identity by descent. Quantile-quantile plots for interaction *P* values and calculation of genomic control lambda (lambda mom=1.01, lambda child=1.03) indicated no systematic test statistic inflation, bias due to unidentified relationships, or cryptic admixture. To assess admixture, we plotted principal components of genetic variation for our sample with 1000 Genomes reference populations; individuals or dyads in which the mother was more than >3 standard deviations from the mean were excluded. The final analysis sample (n=4,514 total samples) included dyads with both mother and child genotype data as well as incomplete dyads with only mother or child genotype data (n=2,596 preeclampsia case samples [1,063 mother/child pairs, 450 mother only, 20 child only], n=1,918 control samples [925 mother/child pairs, 45 mother only, 23 child only]). Complete exposure and outcome classification and genetic quality control methods are described in the Supplementary Methods.

Statistical Analysis

The current analysis is based on version 7 of the MoBa quality-assured data files. Our primary goal was to evaluate whether genetic variants in pathways involved in response to cigarette smoke components modify the maternal smoking-preeclampsia relationship. We extended the case-mother control-mother log-linear modeling approach proposed by Shi et al.,[38] which simultaneously models effects due to maternal and fetal genotypes, as well as maternal smoking and maternal smoking-genotype interaction (both mother and child). This method uses Poisson regression to model expected counts of each possible genetic mating type combination (the set of genotypes in the parents) with child genotype under the assumption of Mendelian inheritance (full model described in Supplementary Methods). We modeled the interaction linearly to improve statistical power, however, present a more flexible model with genotype indicator variables in the Supplementary Methods and Supplementary Table S2.

LEM software[39] was used to fit these models. The expectation-maximization algorithm was used to incorporate dyads with missing genotypes. The proportion of mothers with missing smoking data was low (<5%), so dyads missing smoking information were excluded from analysis. Likelihood ratio tests comparing reduced models with terms for main effects for mother and child with full models containing either a maternal or child interaction term were used to determine interaction *P*-values.

Point estimates and 95% confidence intervals were calculated for the relative risks of preeclampsia in relation to maternal smoking, stratified in turn by both maternal and child genotype. To account for multiple testing, we calculated Bonferroni corrected *P*-values, with a threshold of $P < 2.6 \times 10^{-5}$ for an experimentwise error rate of 0.05 and 1,915 statistical tests.

Because interaction tests have low power to reject homogeneity, and we also report as noteworthy associations yielding uncorrected *P*-interaction <0.001.

Exposure-related population stratification bias within family-based studies can occur when risk allele frequencies, exposure prevalences, and disease prevalences differ in subgroups,[40] so we also performed a sensitivity analysis extending the model to include smoking by mating-type-stratification interaction parameters. We also examined gene-smoking interactions separately for women who smoked during the entire 11-20 week window, and those who quit at some point during that window. Further sensitivity analyses stratified our data by maternal age (≤ 25 and >25 years) and parity to investigate the potential for residual confounding by these factors.

RESULTS

The final analysis was based on 4,288 individual samples for 1,888 complete mother-child dyads (1003 cases), 471 dyads with only maternal genotype data (429 cases), and 41 dyads with only child genotype data (19 cases) (n=2,400 pregnancies) (Table 1). Of the 2,526 pregnancies for which we had genotype information, 126 (5.0%) were missing information on smoking status during our selected etiologic window (gestational weeks 11 to 20). The distribution of covariates did not differ between all pregnancies and those missing smoking data (data not shown), therefore we performed a complete case analysis, requiring complete data for smoking. Of the 2,400 pregnancies for which we had smoking information, 214 (8.9%) smoked during weeks 11 to 20. Fewer women with preeclampsia smoked during weeks 11 to 20 than those without (8.2% vs 10.0%). Overall, smoking intensity was light in this population, with

few women (<1%) smoking more than half a pack a day. A higher proportion of smoking women who later developed preeclampsia quit smoking during the 11 to 20 week window than those without preeclampsia (33% vs 25%).

No maternal or child genotype-smoking interactions met a Bonferroni-corrected threshold ($P < 2.6 \times 10^{-5}$). However, three maternal SNPs yielded interaction $P < 0.001$ (Table 2). No SNPs met this threshold for the child genotype interactions. Of the maternal genes for which we found potential interactions, the main effects for each of the SNPs were null.

We explored how the smoking-preeclampsia association differed by genotype (Table 2). For both rs3765692 in the tumor protein P73 gene (*TP73*) and rs10770343 in the phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 gamma gene (*PIK3C2G*), the reduced relative risk of preeclampsia among smokers was estimated as eliminated with two copies of the risk allele (rs3765692 in *TP73:* RR=0.28, 95% CI: 0.16, 0.51 (*CC*) and RR=1.00, 95% CI: 0.74, 1.56 (*TT*); rs10770343 in *PIK3C2G*: RR=0.37, 95% CI: 0.26, 0.53 (*CC*) and RR=1.07, 95% CI: 0.78, 1.48 (*AA*)). For rs2278361 in the apoptotic peptidase activating factor 1 gene (*APAF1*), the inverse association between smoking and preeclampsia was eliminated by the presence of a single copy of the risk allele and increased with two copies of the risk allele (RR = 0.62, 95% CI: 0.44, 0.86 (*TT*); RR = 1.04, 95% CI: 0.76, 1.41 (*TC*); RR = 1.74, 95% CI: 1.41, 2.15 (*CC*)). RRs for the association between smoking and preeclampsia are presented in Table 2 as stratified by genotype for ease of interpretation; however, the dose-response relationship is imposed by modelling a linear interaction term and should be interpreted with caution. Results of the flexible model also demonstrate a dose-response trend for all except rs10770343, where estimates were similar for heterozygotes and homozygous carriers (Supplementary Table S2).

Results of the sensitivity analysis for exposure-related population stratification are similar; the same set of SNPs had the lowest *P*-values for smoking-SNP interactions in mothers, and risk ratios were similar in magnitude and direction for those SNPs (Supplementary Table S3a). Two additional smoking-SNP interactions in the fetus emerged as potentially noteworthy, but estimates were imprecise due to sparse data (Supplementary Table S3b).

There were few heavy smokers in our population, limiting our ability to address smoking dose. However, we were able to explore the role of smoking cessation, which revealed some differences among groups. Table 3 compares risk ratios for the association of smoking and preeclampsia by genotype for groups in which smokers consisted of only those who quit during weeks 11-20 and only those who did not quit during weeks 11-20. In each case, the effect estimate direction was consistent with the overall association among all smokers seen in the primary analysis; however it was attenuated in those who quit smoking and stronger in those who continued.

COMMENT

Principal Findings

The inverse relationship between smoking and preeclampsia is well established but not well understood.[5] We hypothesized that this association may be in part biological, and investigated the plausibility of a biologic mechanism by assessing gene-smoking interactions in pathways known to be activated by smoking. A biological mechanism is not implausible. Human studies demonstrate associations between NO and CO levels and hypertension in pregnancy.[13,19,21] Despite the strong relationship between smoking and reduced risk of

preeclampsia, we found only limited evidence that smoking-related gene variants modified this association.

Clinical and Research Implications

While none met a Bonferroni corrected threshold, the patterns for our most noteworthy SNPs, rs3765692 (*TP73*), rs1077343 (*PIK3C2G*), and rs2278361 (*APAF1*) suggest the potential for a biological mechanism related to exposure effects. The P3IK pathway (which includes *PIK3C2G*) may affect placental maternal-fetal resource allocation; mouse fetuses and placentas heterozygous for *Pik3ca* were lighter, vascularization was impaired, and circulating leptin, insulin, and plasma tryglyceride levels were reduced.[41] Aryl hydrocarbon receptor activation represses *TP73* and *APAF1*, resulting in pro-proliferative, anti-apoptotic effects,[42] and consequently both are plausibly involved in the pathogenesis of preeclampsia.

Our sensitivity analysis comparing women who report quitting smoking with those who did not also suggests the potential for a dose-related effect. Pregnancy is a time in which many women quit or try to reduce smoking. When we looked just within our selected biologically relevant window, we found that 63 of 214 (29%) reported quitting at some point during that period. Although the directions of association were the same comparing quitters to non-quitters, the inverse associations with smoking were stronger when the mother continued to smoke throughout pregnancy (Table 3).

Strengths and Limitations

Our approach has several strengths. Given the suspected pathophysiology of preeclampsia, both maternal and child genotypes influence susceptibility. Our method accounted for the parent-child relationship by simultaneously modeling effects of mother and child genotypes, each adjusting for the other. Although some studies have measured child genotype,[30,43,44] none have modeled both mother and child genotype while addressing gene by environment interactions. The case-mother control-mother design also improves our statistical power as compared to a traditional case-control study. By applying the expectation maximization algorithm to account for missing data we could include families where only the mother or the child was genotyped. Finally, gene-by-environment studies may be vulnerable to exposure-related population stratification, which has been shown to bias results if both smoking exposure and mating type frequencies differ across exposure categories.[40] We were able to address this concern by including additional family-based exposure parameters, and found little evidence for bias (Supplementary Table S3). And finally, a significant strength of our study is the careful validation of preeclampsia status by antenatal medical and hospital discharge records.[36]

Although we assessed exposure-related population stratification and found none, the model structure is limited by the assumption of genetic homogeneity. Additionally, self-reported smoking is subject to measurement error by under-reporting. A validation study of self-reported smoking and plasma cotinine indicated that reported smoking on the MoBa questionnaire as a marker of tobacco exposure has a sensitivity of 82% and specificity of 99%.[45] To the extent that the nonsmoker category could have actually had some smokers in it this ascertainment error could have biased our results and reduced our power to detect differences.

Our most significant limitation is the large sample size required to confidently identify gene-based effect modification, despite a multiple-testing penalty. Even with our large study, we were underpowered for minor allele frequencies < 0.3. To maximize our power, we modeled interactions on the log-additive scale. We also limited our analysis to SNPs with both a plausible biological relationship or known interactions with smoking, and minor allele frequencies greater than 10%.

Conclusions

In conclusion, we found little evidence of multiplicative gene-smoking interactions with preeclampsia. However, for the three variants that appeared to be effect modifiers, differences by smoking cessation supported a biological interpretation. The case-mother control-mother design as well as environmental interaction analysis may be useful methods to include in the toolkit for the genetic study of pregnancy outcomes, particularly preeclampsia; however very large studies or consortia will be needed to ensure adequate power to find multiplicative gene-by-exposure interactions.

Acknowledgements

This research was supported in part by NICHD R01HD058008 and the Intramural Research Program of the NIH, NIEHS Z01-ES103086. AEB was supported in part by NICHD T32 HD052468. The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health, contract N01-ES-75558 with the NIH/NIEHS, NIH/NINDS (grant no.1 U01 NS 047537-01), and the Norwegian Research Council/FUGE (grant no. 151918/S10). This work was partly

supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 262700. We gratefully thank the participating families in Norway who take part in this ongoing cohort study.

REFERENCES

1. Sibai B, Dekker G, Kupferminc M. Pre-eclampsia. *Lancet*. 2005;365:785-799.
2. Roberts J, Hubel C. The two stage model of preeclampsia: variations on the theme. *Placenta*. 2009;30(Suppl A):S32-S37. doi:10.1016/j.placenta.2008.11.009.
3. Conde-Agudelo A, Althabe F, Belizán JM, Kafury-Goeta AC. Cigarette smoking during pregnancy and risk of preeclampsia: A systematic review. *American Journal of Obstetrics and Gynecology*. 1999;181:1026-1035. doi:10.1016/S0002-9378(99)70341-8
4. Engel SM, Scher E, Wallenstein S, et al. Maternal active and passive smoking and hypertensive disorders of pregnancy: risk with trimester-specific exposures. *Epidemiology*. 2013;24(3):379-386. doi:10.1097/EDE.0b013e3182873a73
5. England L, Zhang J. Smoking and risk of preeclampsia: a systematic review. *Front Biosci*. 2007;12:2471-2483.
6. Kinlaw AC, Buckley JP, Engel SM, Poole C, Brookhart MA, Keil AP. Left Truncation Bias as an Explanation for the Protective Effect of Smoking on Preeclampsia: Potential, but how plausible? *Epidemiology*. 2017;28(3):428-434. doi:10.4315/0362-028X.JFP-13-395.
7. Wikström A-K, Stephansson O, Cnattingius S. Tobacco use during pregnancy and preeclampsia risk: effects of cigarette smoking and snuff. *Hypertension*. 2010;55(5):1254-1259. doi:10.1161/HYPERTENSIONAHA.109.147082
8. Ryter SW, Otterbein LE, Morse D, Choi AMK. Heme oxygenase/carbon monoxide signaling pathways: regulation and functional significance. *Mol Cell Biochem*. 2002;234-235(1-2):249-263.
9. Bainbridge SA, Smith GN. HO in pregnancy. *Free Radic Biol Med*. 2005;38(8):979-988. doi:10.1016/j.freeradbiomed.2004.11.002
10. Bacsik Z, McGregor J, Mink J. FTIR analysis of gaseous compounds in the mainstream smoke of regular and light cigarettes. *Food Chem Toxicol*. 2007;45(2):266-271. doi:10.1016/j.fct.2006.08.018
11. Morse D, Sethi J. Carbon monoxide and human disease. *Antioxid Redox Signal*. 2002;4(2):331-338. doi:10.1089/152308602753666389
12. Krause BJ, Hanson MA, Casanello P. Role of nitric oxide in placental vascular development and function. *Placenta*. 2011;32(11):797-805. doi:10.1016/j.placenta.2011.06.025
13. Choi JW, Im MW, Pai SH. Nitric oxide production increases during normal pregnancy and decreases in preeclampsia. *Ann Clin Lab Sci*. 2002;32(3):257-263. http://www.ncbi.nlm.nih.gov/pubmed/12175088.
14. Casas JP, Cavalleri GL, Bautista LE, Smeeth L, Humphries SE, Hingorani AD. Endothelial Nitric Oxide Synthase Gene Polymorphisms and Cardiovascular Disease: A HuGE Review. *Am J Epidemiol*. 2006;164(10):921-935. doi:10.1093/aje/kwj302

15. Seligman SP, Buyon JP, Clancy RM, Young BK, Abramson SB. The role of nitric oxide in the pathogenesis of preeclampsia. *Am J Obstet Gynecol*. 1994;171(4):944-948. doi:0002-9378(94)90011-6 [pii]

16. Sandrim VC, Palei ACT, Metzger IF, Gomes V, Cavalli RC, Tanus-Santos JE. Nitric oxide formation is inversely related to serum levels of antiangiogenic factors soluble fms-like tyrosine kinase-1 and soluble endogline in preeclampsia. *Hypertension*. 2008;52(2):402-407. doi:10.1161/HYPERTENSIONAHA.108.115006

17. Lyall F, Robson SC, Bulmer JN. Spiral artery remodeling and trophoblast invasion in preeclampsia and fetal growth restriction relationship to clinical outcome. *Hypertension*. 2013;62(6):1046-1054. doi:10.1161/HYPERTENSIONAHA.113.01892

18. Levine RJ, Maynard SE, Qian C, et al. Circulating angiogenic factors and the risk of preeclampsia. *N Engl J Med*. 2004;350(7):672-683. doi:10.1056/NEJMoa031884

19. Cudmore M, Ahmad S, Al-Ani B, et al. Negative regulation of soluble Flt-1 and soluble endoglin release by heme oxygenase-1. *Circulation*. 2007;115(13):1789-1797. doi:10.1161/CIRCULATIONAHA.106.660134

20. George EM, Arany M, Cockrell K, Storm MV, Stec DE, Granger JP. Induction of heme oxygenase-1 attenuates sFlt-1-induced hypertension in pregnant rats. *AJP Regul Integr Comp Physiol*. 2011;301(5):R1495-R1500. doi:10.1152/ajpregu.00325.2011

21. Kreiser D, Baum M, Seidman DS, et al. End tidal carbon monoxide levels are lower in women with gestational hypertension and pre-eclampsia. *J Perinatol*. 2004;24(4):213-217. doi:10.1038/sj.jp.7211062

22. Bauer AE, Avery CL, Shi M, et al. A family based study of carbon monoxide and nitric oxide signalling genes and preeclampsia. *Paediatr Perinat Epidemiol*. 2018;32(1):1-12. doi:10.1111/ppe.12400

23. Raijmakers MTM, Peters WHM, Steegers EAP, Poston L. Amino Thiols, Detoxification and Oxidative Stress in Pre-Eclampsia and Other Disorders of Pregnancy. *Curr Pharm Des*. 2005;11:711-734.

24. Taioli E. Gene – environment interaction in tobacco-related cancers. *Carcinogenesis*. 2008;29(8):1467-1474. doi:10.1093/carcin/bgn062

25. Hoyt JC, Robbins RA, Habib M, et al. Cigarette smoke decreases inducible nitric oxide synthase in lung epithelial cells. *Exp Lung Res*. 2003;29(1):17-28. doi:10.1080/01902140390116490

26. Zhan P, Wang Q, Qian Q, Wei S, Yu L. CYP1A1 MspI and exon7 gene polymorphisms and lung cancer risk : An updated meta-analysis and review. *J Exp Clin cancer Res*. 2011;30:99. doi:10.1186/1756-9966-30-99

27. Zhang Z, Hao K, Shi R, et al. Human Genome Epidemiology (HuGE) Review Glutathione S-Transferase M1 (GSTM1) and Glutathione S-Transferase T1 (GSTT1) Null Polymorphisms, Smoking, and their Interaction in Oral Cancer : A HuGE Review and Meta-Analysis. *Am J Epidemiol*. 2011;173(8):847-857. doi:10.1093/aje/kwq480

28. Tsai HJ, Liu X, Mestan K, et al. Maternal cigarette smoking, metabolic gene polymorphisms, and preterm delivery: New insights on G × E interactions and pathogenic pathways. *Hum Genet*. 2008;123(4):359-369. doi:10.1007/s00439-008-0485-9

29. Cummings AM, Kavlock RJ. Gene-environment interactions: a review of effects on reproduction and development. *Crit Rev Toxicol*. 2004;34(6):461-485. doi:10.1080/10408440490519786

30. McGinnis R, Steinthorsdottir V, Williams NO, et al. Variants in the fetal genome near FLT1 are associated with risk of preeclampsia. *Nat Genet*. 2017;49(8). doi:10.1038/ng.3895

31. Skjaerven R, Vatten LJ, Wilcox AJ, Rønning T, Irgens LM, Lie RT. Recurrence of pre-eclampsia across generations: exploring fetal and maternal genetic components in a population based cohort. *BMJ*. 2005;331(September):877. doi:10.1136/bmj.38555.462685.8F

32. Talhout R, Schulz T, Florek E, van Benthem J, Wester P, Opperhuizen A. Hazardous compounds in tobacco smoke. *Int J Environ Res Public Health*. 2011;8(2):613-628. doi:10.3390/ijerph8020613

33. Magnus P, Birke C, Vejrup K, et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol*. 2016;45(2):382-388. doi:10.1093/ije/dyw029

34. Paltiel L, Haugan A, Skjerden T, et al. The biobank of the Norwegian Mother and Child Cohort Study -- present status. *Nor J Epidemiol*. 2014;24(1-2):29-35.

35. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. *Acta Obstet Gynecol Scand*. 2000;79(6):435-439.

36. Klungsøyr K, Harmon QE, Skard LB, et al. Validity of Pre-Eclampsia Registration in the Medical Birth Registry of Norway for Women Participating in the Norwegian Mother and Child Cohort Study, 1999-2010. *Paediatric and Perinatal Epidemiology*. 2014;28(5):362-371.

37. American College of Obstetricians and Gynecologists. ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. American College of Obstetricians and Gynecologists. *Int J Gynaecol Obstet*. 2002;77(1):67-75.

38. Shi M, Umbach DM, Vermeulen SH, Weinberg CR. Making the most of case-mother/control-mother studies. *Am J Epidemiol*. 2008;168(5):541-547. doi:10.1093/aje/kwn149

39. Vermunt JK. LEM 1.0 : A general program for the analysis of categorical data. 1997. *Tilburg: Tilburg University.*

40. Shi M, Umbach DM, Weinberg CR. Family-based Gene-by-Environment Interaction Studies: Revelations and Remedies. *Epidemiology*. 2011;22(3):400-407. doi:110.1097/EDE.0b013e318212fec6

41. Gerhardt K. The Placenta: A Bridge Between Mother and Baby. *Biol Reprod*. 2016;95(6):114-114. doi:10.1095/biolreprod.116.145789

42. Marlowe JL, Fan Y, Chang X, et al. The Aryl Hydrocarbon Receptor Binds to E2F1 and Inhibits E2F1-induced Apoptosis. *Mol Biol Cell*. 2008;19:3263-3271. doi:10.1091/mbc.E08

43. Goddard KAB, Tromp G, Romero R, et al. Candidate-gene association study of mothers with pre-eclampsia, and their infants, analyzing 775 SNPs in 190 genes. *Hum Hered*. 2007;63(1):1-16. doi:10.1159/000097926

44. Hill LD, Hilliard DD, York TP, et al. Fetal ERAP2 variation is associated with preeclampsia in African Americans in a case-control study. *BMC Med Genet*. 2011;12(1):64. doi:10.1186/1471-2350-12-64

45. Kvalvik LG, Nilsen RM, Skjærven R, et al. Self-reported smoking status and plasma cotinine concentrations among pregnant women in the Norwegian Mother and Child Cohort Study. *Pediatr Res*. 2012;72(1):101-107. doi:10.1038/pr.2012.36

Table 1. Characteristics of Participants by Case and Control Status (N=2,400 Pregnancies)

	Preeclampsia Cases (N = 1,451)		Controls (N = 949)	
	No.	**%**	**No.**	**%**
Maternal Age[a]				
≤ 20 years	45	3.1	14	1.5
21 – 30 years	834	57.5	462	48.7
31 – 40 years	551	38.0	468	49.3
≥ 41 years	20	1.4	5	0.5
Maternal Education[a]				
< High School	127	9.3	69	7.8
High School Graduate	431	31.6	255	28.7
University Degree	804	59.0	565	63.6
Body Mass Index (kg/m^2)[a]				
Underweight (<18.5)	23	1.6	32	3.5
Normal weight (18.5-24.9)	686	48.8	612	66.6
Overweight (25.0-29.9)	437	31.1	181	19.7
Obese (30.0+)	259	18.4	94	10.2
Any Maternal Smoking (11-20 weeks)	119	8.2	95	10.0
Smokers who quit during 11-20 wks[b]	39	32.8	24	25.3
Smoking Intensity[a, b]				
0-2 cigarettes/day	34	36.5	37	46.3
3-5 cigarettes/day	30	31.3	19	23.8
6-10 cigarettes/day	28	29.2	19	23.8
>10 cigarettes/day	3	6.3	5	3.1
Nulliparous[a]	954	65.8	392	41.3
Preterm (< 37 weeks)[a]	305	21.2	28	3.0
Small for gestational age (SGA) (< 10th percentile)[a, c]	323	22.5	74	7.8
Severe preeclampsia	285	19.7	--	--
Birthweight (mean grams (SD))[a]	3173.4 (820.3)		3676.74 (512.4)	

[a]Missing observations for each covariate: maternal age (1), maternal education (149), body mass index (76), smoking intensity (39), parity (1), preterm birth (11), small for gestational age (16), birthweight (1).

[b] Percentage calculated as proportion of women who smoked during 11-20 week window.

[c] Population percentiles derived from Norwegian distribution, eSnurra Norway.

Table 2. Relative Risks for the Association of Smoking and Preeclampsia for All Maternal and Child Smoking-SNP Interactions With Interaction $p < 0.001$, Stratified by Genotype.

						Mother							Child						
						Number of copies of risk allele							Number of copies of risk allele						
						0 copies		1 copy		2 copies			0 copies		1 copy		2 copies		
Marker	**Chr**	**Position**	**MAF**	**Gene**	**Alleles[1]**	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P-Int***	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P*-Int**
rs3765692	1	3584771	0.22	*TP73*	*T/C*	0.28	0.16, 0.51	0.53	0.34, 0.79	1.00	0.74, 1.56	5.9×10^{-4}	0.48	0.28, 0.82	0.66	0.45, 0.97	0.91	0.66, 0.97	0.09
rs10770343	12	18414253	0.31	*PIK3C2G*	*A/C*	0.37	0.26, 0.53	0.63	0.45, 0.88	1.07	0.78, 1.48	5.9×10^{-4}	0.43	0.32, 0.61	0.67	0.48, 0.93	1.02	0.73, 1.43	0.02
rs2278361	12	99043207	0.21	*APAF1*	*C/T*	0.62	0.44, 0.86	1.04	0.76, 1.41	1.74	1.41, 2.15	7.5×10^{4}	0.66	0.47, 0.92	0.99	0.72, 1.37	1.49	1.15, 1.94	0.02

[1]Risk allele/other allele

Table 3. Relative Risks for the Association of Smoking and Preeclampsia for all Maternal Smoking-SNP Interactions with Interaction $p < 0.001$ in the Original Analysis (of All Smokers, n=214 smokers), Stratified by Genotype. Smokers were included as: 1) only smokers who quit smoking during 11-20 weeks gestation (n=63 smokers) and 2) only smokers who did not quit during 11-20 weeks gestation (n=151 smokers).

							Number of copies of risk allele						
							0 copies		1 copy		2 copies		
Marker	**Chr**	**Position**	**MAF**	**Gene**	**Alleles[1]**	**Smoking Group[2]**	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P*-Interaction**
rs3765692	1	3584771	0.22	*TP73*	*T/C*								
						Quit smoking	0.48	0.19, 1.24	0.78	0.40, 1.52	1.24	0.72, 2.16	0.12
						Did not quit smoking	0.25	0.11, 0.54	0.48	0.29, 0.79	0.92	0.65, 1.32	5.7×10^{-3}
rs10770343	12	18414253	0.31	*PIK3C2G*	*A/C*								
						Quit smoking	0.57	0.32, 1.04	0.87	0.49, 1.56	1.32	0.74, 2.35	0.10
						Did not quit smoking	0.28	0.15, 0.51	0.54	0.34, 0.84	1.04	0.71, 1.52	2.8×10^{-3}
rs2278361	12	99043207	0.21	*APAF1*	*C/T*								
						Quit smoking	0.92	0.51, 1.64	1.20	0.68, 2.14	1.57	0.90, 2.75	0.33
						Did not quit smoking	0.49	0.33, 0.75	0.97	0.68, 1.39	1.92	1.81, 2.02	2.6×10^{-4}

[1]Risk allele/other allele

[2]The reference group was non-smokers. The 'quit smoking' group excludes smokers who did not quit from the index and reference groups; The 'did not quit' group excludes smokers who quit from the index and reference groups.

Online Supplementary Material

Contents

Supplementary Methods

We performed a nested case-control study within the Norwegian Mother and Child Cohort Study (MoBa),[1] a large prospective birth cohort of pregnant women and their offspring recruited throughout Norway from 1999 to 2008 (N=112,908 pregnancies).[1] All consenting pregnant women living in Norway who gave birth at a hospital or maternity unit with more than 100 births annually and who could speak Norwegian were eligible. Pregnant women were recruited by mail prior to their routine ultrasound appointment at 17 to 20 weeks of gestation. Of all women invited to participate, 41% enrolled in the study.[1] Participants completed two prenatal questionnaires about their health and environment. Survey completion rate was 91% for the early pregnancy questionnaire and 83% for the late pregnancy questionnaire.[1] Maternal blood was collected at the first ultrasound appointment and cord (child) blood was collected at birth. Maternal blood was received from 89% of participants and child blood from 81% of children in the cohort.[1] DNA was extracted at the time of collection before being stored at the MoBa Biobank.

Women provided informed consent prior to participation in MoBa and the study was approved by the Institutional Review Board of the Norwegian Institute of Public Health. The current study was approved by the Institutional Review Boards of the Norwegian Institute for Public Health and the University of North Carolina at Chapel Hill.

Outcome Assessment

Preeclampsia was defined by the American College of Obstetrics and Gynecologists (ACOG),[2] which specifies that both the following be present:

1) Systolic blood pressure ≥ 140 mm Hg or diastolic blood pressure of ≥ 90 mm Hg occurring after 20 weeks gestation in a woman whose blood pressure has been previously normal, and
2) Proteinuria, with excretion of ≥ 0.3 g of protein in a 24-hour urine specimen or as measured by 1+ on urine dipstick.

Preeclampsia information from MoBa was obtained through linkage with the Medical Birth Registry of Norway (MBRN).[3] A 2013 revision of the ACOG definition also included clinical symptoms, however, we used the above criteria, current at the time of the validation study. We also considered as "true cases" any case with an ICD-10 code of severe PE (O14.1) or HELLP syndrome (O14.2) and delivery < 37 weeks, or ICD-10 code of eclampsia (O15), which are routinely validated by the MBRN.

All registered preeclampsia cases (n=3500) and a random sample of 2000 pregnancies registered as being unaffected by preeclampsia were selected from the MoBa cohort to be verified by antenatal records through an independent validation study[4] Of the 3500 registered preeclampsia cases and 1840 registered to be unaffected by preeclampsia for which records were received, 2936 pregnancies identified as preeclampsia cases based on the Medical Birth Registry of Norway (MBRN) were verified to have been affected by preeclampsia, and 1745 pregnancies identified as unaffected by preeclampsia were found to be negative for preeclampsia. For the current study, we included all women from the validation study with a singleton pregnancy who conceived spontaneously, were verified cases or controls, returned both the early and late pregnancy questionnaires, and had no history of chronic hypertension.

In total, 2682 preeclampsia case samples (1564 maternal and 1118 fetal blood samples) and 1967 non-preeclampsia control samples (999 maternal and 968 fetal samples) met inclusion criteria and were genotyped.

Maternal Smoking Assessment

We created a dichotomous smoking variable to indicate self-report of any smoking in gestational weeks 11 through 20. We selected this time period to specifically capture smoking during the window in which maternal blood flow perfusion into the intervillous space occurs and trophoblasts move to an invasive state completing spiral artery remodeling.[5] Women were surveyed twice during pregnancy about smoking habits. In the early survey (13-17 weeks), women were asked whether they had smoked prior to pregnancy, whether they currently smoked, and if so, how many cigarettes per day or week. If they were not current smokers, they were also asked if they had stopped smoking after becoming pregnant, and if so, at what gestational age. In the late questionnaire (~30 weeks), women were asked whether they currently smoked and if so, how much. They were also asked if they had quit smoking during pregnancy, and if so, at what gestational age they stopped.

For our primary smoking variable during the window of 11 to 20 weeks gestation, we used smoking information from the early pregnancy questionnaire unless it was missing or ambiguous, in which case we supplemented with information from the late pregnancy questionnaire. A woman was considered to be a non-smoker if she indicated she had never smoked and that she did not currently smoke, and we had no other evidence of smoking after 10 weeks of gestation. A woman was considered a smoker in weeks 11 to 20 if she indicated

that she currently smoked or quit after 10 weeks gestation on either survey. To determine the week in which a woman quit smoking, we obtained the latest quit week reported on either the early or late pregnancy questionnaire. If a woman was missing smoking information on the early questionnaire and indicated being a daily smoker on the late questionnaire, we assumed she had also been smoking in weeks 11 to 20. To determine smoking quantity, we categorized smoking quantity as number of cigarettes reported daily or weekly. In a sensitivity analysis, we separately considered women who smoked during the entire 11-20 week window, and women who quit during the 11-20 week window.

Tag SNP Selection for the Parent Study

Genotyping for the current study was completed as part of a larger parent study of genetics of preeclampsia. Genotyping was done using the HumanCoreExome+ array from Illumina (Illumina, Inc., San Diego, CA). Additional custom selected SNPs for the parent study included SNPs on the Illumina Cardio-Metabolic chip not already included in the HumanCoreExome+ manifest, with particular emphasis on three regions of interest: 1) regions associated with systolic blood pressure, diastolic blood pressure, and hypertension; 2) regions associated with myocardial infarction, chronic heart disease, and chronic kidney disease; and 3) regions associated with body mass index, lipids, and C-reactive protein. Additional candidate genes were selected based on the following pathways and sources: 1) existing associations with preeclampsia and/or commonly hypothesized genes, 2) inflammation, 3) angiogenesis, 4) apoptosis, 5) smoking detoxification, 6) carbon monoxide signaling, 7) smoking addiction, and 8) novel pathways including Vitamin D and in vitro studies. For each gene, the Illumina database

was queried for all polymorphism design scores within the genes of interest, allowing for 20kb upstream and 10kb downstream margins. A scoring algorithm for each SNP was created, taking into account Illumina design score, Illumina error codes, DNA coding changes, and presence in a possible 5' promoter site. The composite SNP database was then analyzed using TagZilla (http://tagzilla.nci.nih.gov) to identify haplotype tagging SNPs with an R^2 criteria of 80%. In total, 525,125 variants were genotyped for the larger parent study of genetics and preeclampsia.

SNP Selection

For current study, 124 genes involved in cigarette smoke component signaling, metabolism, and detoxification were identified from 8 canonical pathways (Table S2), which included: 1) endothelial nitric oxide synthase (eNOS) signaling pathway, 2) heme degradation, 3) hypoxia-inducible factor 1-alpha (HIF1A), 4) xenobiotic metabolism, 5) aryl hydrocarbon receptor signaling, 6) glutathione-mediated detoxification, 7) nicotine degradation II, and 8) nicotine degradation III. A total of 1,915 SNPs (MAF $\geq$ 10%) were selected and analyzed for this study, using a 10kb upstream and downstream margin around the transcription start and end sites of each gene.

Genotyping and Quality Control

SNPs were genotyped by the UNC Mammalian Genotyping Core using the HumanCoreExome+ array from Illumina (Illumina, Inc., San Diego, CA). Samples and SNPs were examined using PLINK 1.07 (http://pngu.mgh.harvard.edu/purcell/plink) for quality control.

SNPs were excluded if the missing rate exceeded 5%, there was substantial deviation from Hardy-Weinberg Equilibrium ($p < 1x10^{-3}$) or minor allele frequency was <10%. Known genotype (n = 84) and DNA replicates (n = 51) were included on each plate and generally exhibited high genotyping quality; however, one pair of discordant duplicate samples was excluded. All other quality control samples on the implicated plates were found to be perfectly concordant. All subject-specific call rates were acceptable (minimum 97.2%). Sex-specific markers were inspected and 3 samples with sex discrepancies were excluded. Parent-child relatedness and inbreeding within the cohort was confirmed by identity by descent. Thirteen mother-child pairs were dropped because relatedness could not be confirmed (expected pi-hat = 0.5, observed pi-hat<0.128). Additionally, 16 pairs of siblings or cousins among the mothers were flagged as related (pi-hat>0.125). For each such pair of related mothers, we preferentially included the parent-child pair with the most complete genetic data, or in the case of equivalence, randomly sampled between them.

Quantile-quantile plots and calculation of genomic control lambda[6] (lambda mom=1.01, lambda child=1.03) indicated no systematic test statistic inflation, unidentified relationships, or cryptic admixture. To assess admixture, we plotted principal components of genetic variation for our sample with 1000 Genomes reference populations (CEU: Utah Residents with Northern and Western Ancestry; CHB: Han Chinese in Beijing, China; PUR: Puerto Ricans from Puerto Rico; MXL: Mexican Ancestry from Los Angeles, USA; CLM: Colombians from Medellin, Colombia; YRI: Yoruban in Ibadan, Nigeria).[7] Outliers for each of the first three principal components >3 standard deviations from the mean were excluded. The final analysis sample (n=4514 total samples) included dyads with both mother and child genotype data as well as

incomplete dyads with only mother or child genotype data (n=2596 preeclampsia case samples [1063 mother/child pairs, 450 mother only, 20 child only], n=1918 control samples [925 mother/child pairs, 45 mother only, 23 child only]).

Statistical Analyses

We extended the case-mother control-mother log-linear modeling approach proposed by Shi et al.,[8] which simultaneously models effects due to maternal and fetal genotypes, as well as maternal smoking and maternal smoking-genotype interaction (both mother and child). This method uses Poisson regression to model expected counts of each possible genetic mating type combination (the set of genotypes in the parents) with child genotype under the assumption of Mendelian inheritance as follows:

$$\ln[E(N_{mcde})] = \theta_{mc} + \delta d + \gamma I_e + \sigma d I_e + \alpha_1 d I_{m=1} + \alpha_2 d I_{m=2} + \beta_1 d I_{c=1} + \beta_2 d I_{c=2} + \omega d I_e \times G$$

Where $E(N_{mcde})$ is the expected value of the counts of families with the subscript-specified maternal genotype, child genotype, case or control status, and smoking status; m or c = 0, 1, or 2 for the number of copies of the variant allele carried by the mother or child, respectively; $d = 1$ for a case and $d = 0$ for a control; $I_{(e=1)}$ is an indicator for maternal smoking; and G=m when assessing maternal interaction or G=c when assessing child interaction. The θ_{mc} parameters are constrained as in Table 1 of Shi, et al.,[8] to impose parent-child relatedness for controls but otherwise allow flexibility of the control-mother distribution (avoiding the need to assume

Hardy-Weinberg for the source population) and ensure that the parental genotype distribution is only constrained by the family relationships.

The primary GxE analysis imposes a linear constraint on the smoking by genotype interaction term. Here, we present the flexible model that includes the smoking by maternal interaction:

$$\ln[E(N_{mcde})] = \theta_{mc} + \delta d + \gamma I_e + \sigma d I_e + \alpha_1 d I_{m=1} + \alpha_2 d I_{m=2} + \beta_1 d I_{c=1} + \beta_2 d I_{c=2} + \omega_1 d I_e \times I_{m=1} + \omega_2 d I_e \times I_{m=2}$$

Where $E(N_{mcde})$ is the expected value of the counts of families with each of maternal genotypes, child genotypes, case or control status, and smoking status; m or c = 0, 1, or 2 for the number of copies of the variant allele carried by the mother or child, respectively; $d = 1$ for a case and $d = 0$ for a control; $I_{(e=1)}$ is an indicator for maternal smoking. The θ_{mc} parameters are constrained as in Table 1 of Shi, et al. [8], to impose parent-child relatedness for controls but otherwise allow flexibility of the control-mother distribution (avoiding the need to assume Hardy-Weinberg for the source population) and ensure that the parental genotype distribution is only constrained by the family relationships. Likewise, the model including the smoking by child genotype interaction is similar but contains indicator variables for the child genotypes instead of the maternal genotypes in the interaction terms. The results for these flexible models are presented in Table S2.

Exposure-related population stratification bias within family-based studies can occur when risk allele frequencies, exposure prevalences, and disease prevalences differ in subgroups

[9], so we also performed a sensitivity analysis extending the model to include smoking by mating-type-stratification interaction parameters. This model is as follows:

$$\ln[E(N_{mcde})] = \theta_{mc} + \theta'_{mc} I_e + \delta d + \gamma I_e + \sigma d I_e + \alpha_1 d I_{m=1} + \alpha_2 d I_{m=2} + \beta_1 d I_{c=1} + \beta_2 d I_{c=2} + \omega d I_e \times G$$

in which $\theta'_{mc} I_e$ is now a correction to the mating type stratification parameters to account for possible allele-frequency differences in the subpopulation where the mother smoked. In this model, both θ_{mc} and $\theta_{mc} + \theta'_{mc}$ must satisfy the constraints of Table 1 in Shi et al.[8] This model assumes no population stratification, as it presumes the same set of stratification parameters for the case parents and control parents.

References

1. Magnus P, Birke C, Vejrup K, et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology*. 2016;45(2):382-388. doi:10.1093/ije/dyw029
2. American College of Obstetricians and Gynecologists. ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. American College of Obstetricians and Gynecologists. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*. 2002;77(1):67-75.
3. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. *Acta obstetricia et gynecologica Scandinavica*.

2000;79(6):435-439. http://www.ncbi.nlm.nih.gov/pubmed/10857866.

4. Klungsøyr K, Harmon QE, Skard LB, et al. Validity of Pre-Eclampsia Registration in the Medical Birth Registry of Norway for Women Participating in the Norwegian Mother and Child Cohort Study, 1999-2010. *Paediatric and Perinatal Epidemiology*. http://www.ncbi.nlm.nih.gov/pubmed/25040774. Published September 2014. Accessed December 2, 2014.

5. Wang Y, Zhao S. *Vascular Biology of the Placenta*. Vol 2. (Granger DN, Granger JP, eds.). Lousiana State University: Morgan and Claypool Life Sciences; 2010. doi:10.4199/C00016ED1V01Y201008ISP009

6. Devlin B, Roeder K. Genomic Control for Association Studies. *Biometrics*. 1999;55(4):997-1004.

7. The 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393

8. Shi M, Umbach DM, Vermeulen SH, Weinberg CR. Making the most of case-mother/control-mother studies. *American Journal of Epidemiology*. 2008;168(5):541-547. doi:10.1093/aje/kwn149

9. Shi M, Umbach DM, Weinberg CR. Family-based Gene-by-Environment Interaction Studies: Revelations and Remedies. *Epidemiology*. 2011;22(3):400-407. doi:110.1097/EDE.0b013e318212fec6

Supplementary Tables

Table S1. Canonical pathways related to response to cigarette smoke components selected for this analysis.

Pathway	Description	Genes*
Endothelial nitric oxide synthase signaling	Describes the synthesis of nitric oxide from L-arginine. Endothelial nitric oxide synthase plays a crucial role in the state of blood vessel vasodilation and blood pressure regulation.	*CASP9, ADCY3, CASP8, CASP3, CCNA2, GUCY1A3, GUCY1B3, ESR1, CAV1, ADCY8, GUCY1A2, CHRNA3, CHRNA5, EPAS1, STAT1, PDE1C, PDGFA, PON1, ENG, TLR4, GUCY2C, MPO, PRKCA, PDGFB, PDGFC, NOS3, PIK3C2G, AKT1, FIGF, MAPK10, MAPK9, MAPK13, MAPK14, MAPK8, MAPK3, MAPK1, MAPK11, MAPK12, NOS1, NOS2*
Heme degradation	Describes the breakdown of hemoglobin into carbon monoxide, biliverdin, iron, and bilirubin.	*HBB, HBA1, HBA2, HMOX2, HMOX1, CA9, NGB, HP, MB*
Hypoxia-inducible factor 1-alpha signaling	Describes regulation of oxygen homeostasis and response to hypoxia. Activates transcription of nitric oxide synthase. Also involved in the xenobiotic response via aryl hydrocarbon receptor nuclear translocator.	*PDGFC, NOS3, PIK3C2G, AKT1, FIGF, MAPK10, MAPK9, MAPK13, MAPK14, MAPK8, MAPK3, MAPK1, MAPK11, MAPK12, NOS1, NOS2, ARNT, NCOA1, EDN1, EPO, MAPK15, MMP12, MMP3, MDM2, APEX1, HIF1A, MAPK6, MMP2, MAPK7, MAPK4, EGLN2, MMP9*
Xenobiotic metabolism	Describes the three groups of enzymes that metabolize, eliminate, and detoxify harmful substances. Phase I: introduces polar moiety, Phase II: conjugates toxins to small hydrophilic molecules, Phase III: transporters that export toxins.	*IL1A, MAPK1, HS2ST1, CYP3A7, MAPK13, IL6, NFKB1, MAPK11, PTGES3, ARNT, HMOX1, GSTT1, MAOB, CYP1A2, ALDH1A1, MAPK3, NOS2, AHR, GSTK1, PRKCA, ATM, GSTA3, ABCB1, GSTM1, CYP1A1, GSTM3, GSTA4, NQO1, MAPK8, PIK3C2G, MAPK9, NFKB2, MAPK12, GSTO1, CYP1B1, AIP, MAPK14, SULT1A1, NCOA1, IL1B, CYP2B6, MAPK7, GSTO2, TNF, UGT1A9, GSTP1, MAOA*
Aryl hydrocarbon receptor signaling	Describes mediation of halogenated and polycyclic aromatic hydrocarbons by the aryl hydrocarbon receptor. Activates	*CDKN2A, MAPK1, TP73, IL6, CCND1, ARNT, MYC, RB1, ALDH1A1, CYP1A2, CCND3,*

	xenobiotic metabolizing enzymes and other growth factors and proteins involved in cell cycle progression and apoptosis.	*TGFB1, MAPK3, AHR, FASLG, GSTK1, ATM, TP53, CCNE2, GSTM3, CDK6, NFKB2, CCND2, E2F1, TGFB3, ESR1, TNF, GSTP1, CDK2, IL1A, CDK4, NFKB1, PTGES3, FAS. CCNA2, GSTT1, TGFB2, CHEK2, GSTA3, GSTM1, CYP1A1, GSTA4, NQO1, APAF1, MAPK8, MDM2, BAX, CYP1B1, GSTO1, AIP, CCNE1, CDKN1A, IL1B, ATR, CDKN1B GSTO2*
Glutathione-mediated detoxification	Describes detoxification in which the first step is catalyzed by glutathione transferases.	*GSTA3, GSTM1, GSTT1, GSTM3, GSTA4, GSTO2, GSTP1, GSTO1, GSTK1*
Nicotine degradation II and III	Describes degradation of nicotine primarily through the metabolic action of cytochrome P450.	*CYP2D6, ADH7, CYP1A1, CYP1A2, CYP2E1, CYP2A7, CYP3A7, CYP2A6, CYP2B6, UGT1A9, CYP1B1*

*Some genes are involved in multiple pathways

Table S2. Relative risks for the association of smoking and preeclampsia for all maternal and child smoking-SNP Interactions, stratified by genotype for mother and child genotype interactions for SNPs with p<0.001 presented in the primary analysis. Multiplicative smoking by genotype interactions were modeled flexibly using indicator variables for maternal or child genotype.

						Mother							Child						
						Number of copies of variant allele							Number of copies of variant allele						
						0 copies		1 copy		2 copies			0 copies		1 copy		2 copies		
Marker	**Chr**	**Position**	**MAF**	**Gene**	**Alleles[1]**	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P-Int***	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P*-Int**
rs3765692	1	3584771	0.22	*TP73*	*T/C*	0.99	0.73, 1.35	0.58	0.31, 1.06	0.14	0.02, 0.99	6.1×10^{-3}	0.93	0.68, 1.28	0.65	0.32, 1.31	0.72	0.28, 1.82	0.20
rs10770343	12	18414253	0.31	*PIK3C2G*	*A/C*	1.13	0.82, 1.55	0.53	0.28, 0.99	0.57	0.28, 1.16	0.01	1.03	0.74, 1.46	0.64	0.32, 1.31	0.49	0.20, 1.20	0.41
rs2278361	12	99043207	0.21	*APAF1*	*T/C*	0.63	0.45, 0.89	0.98	0.51, 1.86	2.01	1.04, 3.89	0.02	0.66	0.47, 0.94	0.97	0.47, 1.99	1.59	0.69, 3.63	0.27

[1]Major allele/minor allele; The minor allele is considered the variant allele.

Table S3a. Relative risks for the association of smoking and preeclampsia for maternal and child smoking-SNP interactions stratified by genotype, in the sensitivity analysis model allowing smoking exposure to vary by mating type. Results are presented for SNPs that had interaction $p < 0.001$ in the original model presented in the main manuscript.

						Mother							Child						
						Number of copies of variant allele							Number of copies of variant allele						
						0 copies		1 copy		2 copies			0 copies		1 copy		2 copies		
Marker	**Chr**	**Position**	**MAF**	**Gene**	**Alleles[1]**	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P-Int***	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P*-Int**
rs3765692	1	3584771	0.22	*TP73*	*T/C*	1.16	0.81, 1.66	0.45	0.29, 0.71	0.18	0.11, 0.30	5.5×10^{-4}	1.02	0.71, 1.46	0.58	0.39, 0.88	0.34	0.21, 0.53	0.03
rs10770343	12	18414253	0.31	*PIK3C2G*	*A/C*	0.99	0.71, 1.39	0.67	0.43, 1.05	0.46	0.27, 0.78	0.07	0.99	0.70, 1.42	0.69	0.45, 1.06	0.48	0.30, 0.79	0.12
rs2278361	12	99043207	0.21	*APAF1*	*T/C*	0.63	0.44, 0.91	1.01	0.69, 1.47	1.62	1.10, 2.37	0.04	0.68	0.47, 0.99	0.93	0.64, 1.37	1.29	0.88, 1.89	0.20

[1]Major allele/minor allele; The minor allele is considered the variant allele.

Table S3b. Additional child SNPs with interaction $p < 0.001$ in the sensitivity analysis model allowing smoking exposure to vary by mating type. As in Table S2a, table presents relative risks for the association of smoking and preeclampsia for all maternal and child smoking-SNP interactions stratified by genotype, in the model allowing smoking exposure to vary by mating type.

						Mother							Child						
						Number of copies of variant allele							Number of copies of variant allele						
						0 copies		1 copy		2 copies			0 copies		1 copy		2 copies		
Marker	**Chr**	**Position**	**MAF**	**Gene**	**Alleles[1]**	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P-Int***	**RR**	**95% CI**	**RR**	**95% CI**	**RR**	**95% CI**	***P*-Int**
rs995647	2	24810255	0.13	*NCOA1*	*A/G*	0.66	0.48, 0.91	1.54	0.85, 2.80	3.61	1.65, 7.88	0.10	0.58	0.42, 0.81	1.93	1.14, 3.27	6.39	3.27, 12.47	3.0×10^{-5}
rs7929753	11	106795127	0.25	*GUCY1A2*	*C/T*	1.00	0.68, 1.46	0.68	0.48, 0.97	0.47	0.34, 0.64	0.01	1.38	0.94, 2.03	0.50	0.34, 0.73	0.18	0.12, 0.26	1.2×10^{-4}