

# Maternal cigarette smoking, metabolic gene polymorphisms, and preterm delivery: new insights on G × E interactions and pathogenic pathways

Hui-Ju Tsai · Xin Liu · Karen Mestan · Yunxian Yu · Shanchun Zhang · Yaping Fang · Colleen Pearson · Katherin Ortiz · Barry Zuckerman · Howard Bauchner · Sandra Cerda · Phillip G. Stubblefield · Xiping Xu · Xiaobin Wang

Received: 15 November 2007 / Accepted: 24 February 2008 / Published online: 5 March 2008  
© Springer-Verlag 2008

**Abstract** Preterm delivery (PTD, <37 weeks of gestation) is a significant clinical and public health problem. Previously, we reported that maternal smoking and metabolic gene polymorphisms of *CYP1A1* MspI and *GSTT1* synergistically increase the risk of low birth weight. This study investigates the relationship between maternal smoking and metabolic gene polymorphisms of *CYP1A1* MspI and *GSTT1* with preterm delivery (PTD) as a whole and preterm subgroups. This case–control study included 1,749 multi-ethnic mothers (571 with PTD and 1,178 controls) enrolled at Boston Medical Center. After adjusting covariates, regression analyses were performed to identify individual and joint associations of maternal smoking, two functional variants of *CYP1A1* and *GSTT1* with PTD. We

observed a moderate effect of maternal smoking on PTD (OR = 1.6; 95% CI: 1.1–2.2). We found that compared to non-smoking mothers with low-risk genotypes, there was a significant joint association of maternal smoking, *CYP1A1* (*Aa/aa*) and *GSTT1* (absent) genotypes with gestational age ( $\beta = -3.37$ ; SE = 0.86;  $P = 9 \times 10^{-5}$ ) and with PTD (OR = 5.8; 95% CI: 2.0–21.1), respectively. Such joint association was particularly strong in certain preterm subgroups, including spontaneous PTD (OR = 8.3; 95% CI: 2.7–30.6), PTD < 32 weeks (OR = 11.1; 95% CI: 2.9–47.7), and PTD accompanied by histologic chorioamnionitis (OR = 15.6; 95% CI: 4.1–76.7). Similar patterns were observed across ethnic groups. Taken together, maternal smoking significantly increased the risk of PTD among women with high-risk *CYP1A1* and *GSTT1* genotypes. Such joint associations were strongest among PTD accompanied by histologic chorioamnionitis.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00439-008-0485-9) contains supplementary material, which is available to authorized users.

H.-J. Tsai (✉) · X. Liu · Y. Yu · S. Zhang · Y. Fang · X. Wang  
The Mary Ann and J. Milburn Smith Child Health  
Research Program, Children’s Memorial Hospital and Children’s  
Memorial Research Center, Department of Pediatrics,  
Feinberg School of Medicine, Northwestern University,  
Chicago, IL, USA  
e-mail: htsai@childrensmemorial.org; h-tsai@northwestern.edu

K. Mestan  
Division of Neonatology, Children’s Memorial Hospital,  
Department of Pediatrics, Feinberg School of Medicine,  
Northwestern University, Chicago, IL, USA

S. Zhang  
Institute for Biomedicine, Anhui Medical University,  
Hefei, People’s Republic of China

C. Pearson · K. Ortiz · B. Zuckerman · H. Bauchner  
Department of Pediatrics,  
Boston University School of Medicine and Boston  
Medical Center, Boston, MA, USA

S. Cerda  
Department of Pathology and Laboratory Medicine,  
Boston University School of Medicine and Boston  
Medical Center, Boston, MA, USA

P. G. Stubblefield  
Department of Obstetrics and Gynecology,  
Boston University School of Medicine and Boston  
Medical Center, Boston, MA, USA

X. Xu  
Center for Population Genetics,  
University of Illinois at Chicago School of Public Health,  
Chicago, IL, USA

## Introduction

Preterm delivery (PTD, <37 weeks of gestation) is a significant clinical and public health problem. The rate of PTD remains high in the U.S. at about 12.5% (Anonymous 2006a) and has actually increased in the past two decades (Anonymous 2006b). The highest rates of PTD occur among African-Americans (17.8%). The health consequences and economic cost to society of PTD are enormous. There is compelling evidence that PTD is a complex disease influenced by multiple genetic and environmental factors and gene-environment (G×E) interactions (Anonymous 2006a; Crider et al. 2005; Macones et al. 2004). Until recently, research on the causative factors of PTD has primarily focused on demographic, socio-behavioral and environmental risk factors. There are limited studies on the roles of genetic susceptibility and G×E interactions in the development of PTD.

Maternal cigarette smoking is an important preventable environmental factor known to increase the risk of PTD, with odds ratios ranging from 1.1 to 2.4 (Cnattingius 2004; Pollack et al. 2000; Shah and Bracken 2000; Wang et al. 1997). However, given the same amount and duration of cigarette smoking, not all women who smoke during pregnancy deliver preterm. One possible explanation is the individual variation in genetic susceptibility. Two functional genes that have received attention are Cytochrome P-450 1A1 (*CYP1A1* MspI) and glutathione S-transferases Theta 1 (*GSTT1*). *CYP1A1* and *GSTT1* encoded enzymes are involved in metabolism and detoxification of polycyclic aromatic hydrocarbon (PAH), an important carcinogen in tobacco smoke (Bartsch 1996; Bartsch et al. 1995; Kriek et al. 1998). The adverse health effects of cigarette smoke may depend on the combined effects of phase I and phase II metabolic detoxification (Hayashi et al. 1992; Katoh et al. 1995; Nakachi et al. 1993). Others and we have described the role of cigarette smoking, including the *CYP1A1* or *GSTT1* gene variants in relation to adverse birth outcomes (Chen et al. 2005; Cnattingius 2004; Nukui et al. 2004; Wang et al. 2001). Specifically, we previously reported that maternal smoking and metabolic gene polymorphisms of *CYP1A1* MspI and *GSTT1* can synergistically increase the risk of low birth weight, gestational age and birth weight ratio (Wang et al. 2002). The *CYP1A1* MspI polymorphic allele is associated with alterations in regulation and transcription half-life of the enzyme (Landi et al. 1994), whereas deletion of *GSTT1* polymorphism results in no functional enzymatic activity (Pemble et al. 1994; Seidegard et al. 1988). However, the synergistic effect of these two genes in conjunction with maternal cigarette smoking during pregnancy on the risk of PTD has not been well-studied in a large population.

The aim of this study was to investigate the relationship between *CYP1A1* and *GSTT1* genotypes and maternal smok-

ing with PTD as a whole and preterm subgroups defined by mode of delivery (spontaneous vs. induced PTD); degree of prematurity (<32 vs. 34.0–36.86 weeks); and major pregnancy complications (intrauterine infection/inflammation vs. hypertensive disorders). We hypothesize that the smoking-PTD association can be significantly modified by *CYP1A1* and *GSTT1* gene polymorphisms. Since PTD is a heterogeneous entity, we further hypothesize that the association will be stronger among PTD subgroups due to underlying pathogenic pathway by which smoking affects PTD.

## Materials and methods

### Study population and data collection

We included the first 1,749 mothers enrolled in an ongoing case-control study of preterm birth at the Boston Medical Center (BMC) since 1998 (Wang et al. 2002). Case mothers were those with singleton live births occurring at less than 37 weeks gestation, and controls were defined as mothers delivering at greater than or equal to 37 weeks of gestation, and with birthweight appropriate for gestational age as defined by the National Center for Health Statistics/CDC guidelines (birthweight between 2,500 and 4,000 g) (Hamill et al. 1979). Pregnancies resulting in multiple births and newborns with major birth defects were excluded. Detailed description was published previously (Wang et al. 2002). Briefly, we comprehensively collected the epidemiologic data, the clinical data, and the maternal venous blood samples. Of note, the data collected in this study contained all the variables and phenotypic information suggested by the International Preterm Birth Collaborative (PREBIC), Genetics Working Group (<http://www.prebic.org>) for the genetic epidemiological studies of preterm birth (Pennell et al. 2007). In addition, placentas were sent for histopathology based on routine indications including preterm birth (Gupta et al. 2007). The institutional review boards of the BMC, the Massachusetts Department of Public Health, the Children's Memorial Hospital in Chicago, and the Harvard School of Public Health approved the study protocol. All the participants gave written informed consent.

### Definition of phenotypes

#### *Preterm delivery (PTD) and gestational age*

The primary phenotype was evaluated as both a binary (PTD) and a continuous (gestational age) variable. PTD was defined as birth occurring before 37 weeks of gestation. Gestational age was assessed using an algorithm based on last menstrual period and the result of early ultrasound (<20 weeks gestation). The last menstrual period estimate

was used only if confirmed by ultrasound within 7 days or if no ultrasound estimate was obtained; otherwise, the ultrasound estimate was used. This approach has been used in previous studies (Kramer et al. 1998; Wang et al. 2002).

### PTD subgroups

We further divided all the preterm cases into the following subgroups:

1. *By mode of delivery.* We categorized PTD cases as *spontaneous PTD* (vaginally or by cesarean section) that occurred secondary to documented active preterm labor (uterine contractions with cervical effacement and dilation at less than 37 weeks), or preterm premature rupture of membranes (PPROM) (<37 weeks without uterine contractions) or by both uterine contractions and PPRM occurring simultaneously; or as *medically induced PTD* defined as delivery (vaginally or by cesarean section) that was not preceded by the presence of uterine contractions and/or rupture of membranes.
2. *By degree of prematurity.* In this study, we used a cut-point of <32 weeks to define very preterm delivery and 34.0–36.86 weeks to define late preterm delivery, which has been used by other groups (Engle et al. 2007; Mathews et al. 2004).
3. *By major pregnancy complications.* Due to sample size constraints, we focused on two relatively common pregnancy complications: (1) PTD with intrauterine infection/inflammation using placental histologic chorioamnionitis as a proxy. Detailed description on placental collection, pathological methods, definition of maternal and fetal inflammatory responses and quality control was published previously (Gupta et al. 2007), and (2) PTD complicated by maternal hypertensive disorders. This group consisted of PTD cases with an accompanying diagnosis of maternal preeclampsia, eclampsia, gestational hypertension, or HELLP syndrome as defined in our previous publication (Wang et al. 2006), with or without a history of chronic hypertension.

### Maternal cigarette smoking

The information on maternal smoking was obtained from the postpartum questionnaire for four time periods: 3 months before the index pregnancy and the first, second, and third trimesters of the index pregnancy. Mothers' smoking data were clustered into three groups: (1) those who did not smoke throughout the index pregnancy (never smokers); (2) those who smoked during early pregnancy but quit smoking during the first trimester (quitters); and (3) those who smoked continuously during the index pregnancy (persistent smokers).

### Genotyping

#### DNA extraction and PCR condition

DNA was extracted from samples of venous whole blood in accordance with standard protocol (Sambrook et al. 1989). *CYP1A1* polymorphism was amplified using the primer pair: 5'-TAGGAGTCTTGTCTCATGCCT-3' and 5'-CAGTGAAGAGGTGTAGCCGCT-3'. PCR products were digested by MspI. The detailed method for PCR amplification can be found elsewhere (Wang et al. 2002; Xu et al. 1996). *GSTT1* variant was amplified using the primer pair: 5'-TTCCTTACTGGTCCTCACATCTC-3' and 5'-TCACCGGATCATGGCCAGCA-3'. The detailed PCR amplification for *GSTT1* can be obtained elsewhere (Palli et al. 2005; Wang et al. 2002). Due to sample size constraints, the *CYP1A1* polymorphism was coded as AA or Aa/aa, whereas *GSTT1* polymorphism was coded as present or absent.

#### Quality checking

The genotype completion rates for *CYP1A1* and *GSTT1* polymorphisms were 98 and 96%, respectively. The consistency rates of genotyping calls made by two technicians independently were 99.9% for both *CYP1A1* and *GSTT1*. We repeated genotyping for *CYP1A1* in 54 subjects and for *GSTT1* in 43 subjects, respectively. The genotyping consistency rates were 100% for both *CYP1A1* and *GSTT1*.

#### AIMs genotyping

Population stratification is a critical concern for case–control genetic studies in admixed populations such as African Americans. One common approach to deal with this potential confounding effect is to adjust individual admixture in the analysis. To estimate individual admixture proportions, we genotyped 59 ancestry informative markers (AIMs) with averaged  $\delta$  (difference of allele frequencies between African and European ancestral populations) equal to 0.59 in 812 black mothers (326 preterm cases and 486 controls). Detailed information regarding polymorphic site, flanking sequence and other relevant information for all the 59 AIMs is available elsewhere (Yang et al. 2005).

#### Statistical analyses

Prior to data analysis, we computed genotype frequencies of *CYP1A1* and *GSTT1* and examined whether *CYP1A1* genotype frequency was under Hardy–Weinberg equilibrium (HWE) by using exact tests of HWE implemented in the SNP–HWE program (Wigginton et al. 2005).

We applied a Bayesian approach implemented in the structural program to estimate individual admixture proportions using 59 genotyping AIMs data (Falush et al. 2003; Pritchard et al. 2000). Specifically, the admixture model assuming that each individual inherits some proportion of their ancestry from each ancestral population was selected for estimating individual ancestral proportions. The average and corresponding standard deviation (SD) of African proportions were 0.90 (SD = 0.11) in PTD cases, and were 0.89 (SD = 0.12) in controls.

We applied logistic regression and general linear regression to examine the individual effects of metabolic gene polymorphisms and maternal smoking status, and the joint effects of smoking, *CYP1A1* and *GSTT1* on preterm phenotypes, as binary and continuous outcomes, respectively. For *CYP1A1* polymorphism, we combined subjects with Aa or aa genotypes in data analysis, since this reflects the functional activity of *CYP1A1* (Landi et al. 1994). Subsequently, we examined the interaction effect of maternal smoking, *CYP1A1* and *GSTT1* with PTD by adding all the product terms (both 2-way and 3-way terms) in the model. To assess heterogeneity across ethnic and preterm subgroups, we applied similar analyses described above for specific ethnic and PTD subgroups.

All analyses include the following covariates: maternal age (<20, 20–24, 25–29, 30–34 and  $\geq 35$  years), ethnicity (white, black, Hispanic, and other), education ( $\leq$ middle school, =high school, and >high school), parity (0, 1, and  $\geq 2$ ), marital status (married, others), maternal pre-pregnant body mass index (BMI) (<20, 20–24, 25–29, and  $\geq 30$ ), alcohol use (yes/no), passive smoke exposure (yes/no), illicit drug use (yes/no) and infant sex. Of note, false positive results have been a critical concern when a number of association tests are performed. To address this issue, we applied false discovery rate (FDR) to correct for multiple testing (Benjamini and Yekutieli 2001). As mentioned earlier, population stratification is another concern in genetic association studies. In addition to performing stratified analyses by ethnic groups (black mothers vs. Hispanic/white mothers), we also adjusted for individual ancestry estimated from 59 ancestry informative markers (AIMs) in black mothers. Data analyses were performed using statistical packages R 2.3.1, Bioconductor 1.8 and Intercooled STATA 8.0 (College Station, TX, USA).

## Results

### Demographic and clinical characteristics

Our analysis included a total of 1,749 mothers: 571 with PTD, and 1,178 with full term delivery. The distributions of age, ethnicity, education, parity, marital status, alcohol

intake, infant sex and genotype frequencies of *CYP1A1* and *GSTT1* polymorphisms were similar in both groups (Table 1). Maternal pre-pregnant BMI was slightly higher among the PTD than the term delivery group. The rates of maternal smoking during pregnancy, passive smoke exposure and maternal illicit drug use were higher among the PTD. In addition, we examined whether *CYP1A1* variant was under HWE across different ethnic groups. Only the Hispanic group was slightly deviated from HWE ( $P = 0.01$ ).

There were 71% of PTD cases that were spontaneous deliveries, and 24% of PTD cases that were delivered before 32 gestational weeks (Table 2). Twenty-seven percent of PTD cases had evidence of placental histologic chorioamnionitis, and 24% of PTD cases were complicated by maternal hypertensive disorders. The PTD subgroup with spontaneous delivery had a significantly higher proportion of placental histologic chorioamnionitis when compared to the PTD subgroup with medically induced delivery, 33 versus 13% (OR = 3.2, 95% CI: 2.0–5.3). This was also true for the very PTD subgroup when compared to late PTD subgroup, 51 versus 18% (OR = 4.7, 95% CI: 3.0–7.4).

### Individual and joint associations

When examined individually, maternal smoking was associated with a moderately increased risk of PTD (OR = 1.6; 95% CI: 1.1–2.2) and lower gestational age ( $\beta = -0.70$ ; SE = 0.29;  $P = 0.02$ ); neither *CYP1A1* nor *GSTT1* polymorphisms were associated with PTD or gestational age (Table 3). We excluded “quitters” in the analysis due to small sample size.

When examined jointly, using never-smoking mothers with low-risk genotypes of *CYP1A1* (AA) and *GSTT1* (present) as the reference group, we found a significantly higher risk of PTD (OR = 5.8; 95% CI: 2.0–21.1) and lower gestational age ( $\beta = -3.37$ ; SE = 0.86;  $P = 9 \times 10^{-5}$ ) among smoking mothers with high-risk genotypes of *CYP1A1* (Aa/aa) and *GSTT1* (absent) (Table 3). The interaction effect of maternal smoking, *CYP1A1* and *GSTT1* genotypes, was marginally significant on gestational age ( $\beta = -3.30$ ; SE = 1.22; multiple testing corrected  $P = 0.06$ ). To assess if the associations varied among different ethnic groups, we further examined joint association in black mothers only, and in white and Hispanic subjects combined. We found very similar results (Table 1 in supplementary material). All the significant joint associations remained statistically significant after correcting for multiple comparison. To examine whether the associations and interactions we observed may be modified by past history of PTD, we stratified study subjects into two groups: null parity versus parity  $\geq 1$ . The trends of associations and interactions were comparable between these two groups.

**Table 1** Demographic, clinical and genetic characteristics of BMC cohort, by preterm status

	Preterm delivery ( <i>N</i> = 571)	Term delivery ( <i>N</i> = 1,178)	<i>P</i>
<i>Demographic characteristics</i>			
Age (%)			
<20	11.2	11.2	0.15
20–24	24.3	27.3	
25–29	24.5	25.2	
30–34	21.4	22.2	
≥35	18.6	14.0	
Ethnicity (%)			
African Americans	57.4	52.6	0.17
Caucasians	13.8	13.1	
Hispanic Americans	23.6	25.8	
Education (%)			
≤Middle school	29.6	34.5	0.38
=High school	38.0	31.9	
>High school	32.4	33.6	
Parity (%)			
0	40.5	37.0	0.54
1	26.6	30.9	
≥2	32.9	32.1	
Marital status (%)			
Married	32.1	37.0	0.04
Others	67.9	63.0	
Smoking during pregnancy (%)			
Never	75.3	84.7	<0.01
Quitter	4.8	3.4	
Persistent	19.9	11.9	
Passive smoking during pregnancy (%)			
No	77.5	70.6	<0.01
Yes	22.5	29.4	
Illicit drug use (%)			
No	82.8	91.0	<0.01
Yes	17.2	9.0	
Alcohol use (%)			
No	95.3	94.2	0.37
Yes	4.7	5.8	

**Table 1** continued

	Preterm delivery ( <i>N</i> = 571)	Term delivery ( <i>N</i> = 1,178)	<i>P</i>
<i>Maternal pre-pregnant BMI (%)</i>			
<20	15.2	12.1	0.03
20–24	35.6	40.7	
25–29	28.9	30.7	
30+	20.3	16.6	
<i>Infant sex (%)</i>			
Male	48.7	50.3	0.52
Female	51.3	49.7	
<i>Clinical characteristics</i>			
Gestational age (week)	33.6 (3.4)	39.7 (1.1)	<0.01
Birth weight (g)	2,144.1 (748.3)	3,375.1 (460.2)	<0.01
<i>Genotype counts and frequency [n (%)]</i>			
<i>CYP11A1</i>			
AA	323 (56.6)	674 (57.2)	0.70
Aa	206 (36.1)	395 (33.5)	
aa	42 (7.3)	109 (9.3)	
<i>GSTT1</i>			
Present	440 (77.1)	895 (76.0)	0.60
Absent	131 (22.9)	283 (24.0)	

Furthermore, the results were not substantially different between the models with and without adjustment for individual ancestry in black mothers (data not shown).

#### Heterogeneity in joint association among preterm subgroups

To determine whether the above joint associations varied among preterm subgroups, we repeated the above analyses among the following: spontaneous versus medically induced PTD, very PTD versus late PTD, and PTD accompanied by histologic chorioamnionitis versus hypertensive disorders (Table 4). In the presence of maternal smoking, significant joint association with spontaneous PTD (OR = 8.3; 95% CI: 2.7–30.6), very PTD (OR = 11.1; 95% CI: 2.9–47.7), and PTD accompanied by histologic chorioamnionitis (OR = 15.6; 95% CI: 4.1–76.7) was found among the mothers with high-risk genotypes of *CYP11A1* and *GSTT1* variants. Interaction of maternal smoking, *CYP11A1* and *GSTT1* genotypes was marginally significant

**Table 2** Distribution and inter-relationship of preterm subgroups

	PTD subjects, N (%)	PTD accompanied by histologic chorioamnionitis		Odds ratio (95% CI)
		Yes (N = 156)	No (N = 415)	
<b>Spontaneous PTD</b>				
Yes <sup>a</sup>	405 (70.9)	134 (33.1%)	271 (66.9%)	3.2 (2.0–5.3)
No <sup>b</sup>	166 (29.1)	22 (13.3%)	144 (86.7%)	
<b>Severity of PTD<sup>c</sup></b>				
<32 weeks	137 (24.0)	70 (51.1%)	67 (48.9%)	4.7 (3.0–7.4)
34–36.86 weeks	362 (63.4)	66 (18.2%)	296 (81.8%)	
<b>Hypertensive disorders<sup>d</sup></b>				
Yes	135 (23.8)	22 (16.3%)	113 (83.7%)	0.4 (0.3–0.7)
No	436 (76.2)	134 (30.7%)	302 (69.3%)	

<sup>a</sup> Spontaneous PTD (yes): documented active preterm labor (uterine contractions with cervical effacement and dilation at less than 37 weeks), or preterm premature rupture of membranes (PPROM) (<37 weeks without uterine contractions) or by both uterine contractions and PPRM occurring simultaneously

<sup>b</sup> Spontaneous PTD (no): delivery that was not preceded by the presence of uterine contractions and/or rupture of membranes

<sup>c</sup> Severity of PTD: a cut-point of <32 weeks to define very preterm delivery and 34.0–36.86 weeks to define late preterm delivery

<sup>d</sup> PTD complicated by maternal hypertensive disorders: PTD cases with an accompanying diagnosis of maternal preeclampsia, eclampsia, gestational hypertension, or HELLP syndrome, with or without a history of chronic hypertension

**Table 3** Individual and joint association of maternal cigarette smoking during pregnancy, *CYP1A1* and *GSTT1* gene polymorphisms with PTD and gestational age

Smoking	<i>CYP1A1</i>	<i>GSTT1</i>	PTD (<37 weeks) <sup>a</sup>					Gestational age (weeks) <sup>b</sup>		
			Total subjects	% <sup>c</sup>	OR (95% CI)	<i>P</i> <sup>d</sup>	Adjusted <i>P</i> <sup>d</sup>	$\beta$ (SE)	<i>P</i>	Adjusted <i>P</i>
Never	–	–	1,417	30.1	1.6 (1.1–2.2)	0.01	0.06	–0.70 (0.29)	0.02	0.10
Persistent	–	–	252	44.8						
–	AA	–	968	32.7	1.1 (0.9–1.4)	0.33	0.62	–0.14 (0.18)	0.42	0.72
–	Aa/aa	–	728	33.4						
–	–	Present	1,312	33.1	0.9 (0.7–1.2)	0.46	0.77	0.21 (0.21)	0.31	0.60
–	–	Absent	407	31.7						
Never	AA	Present	583	32.6	Ref			Ref		
Never	AA	Absent	177	26.0	0.7 (0.5–1.0)	0.08	0.23	0.65 (0.31)	0.03	0.12
Never	Aa/aa	Present	448	29.0	0.9 (0.7–1.2)	0.42	0.72	0.16 (0.23)	0.47	0.77
Never	Aa/aa	Absent	138	34.1	1.1 (0.7–1.7)	0.58	0.83	–0.12 (0.34)	0.72	0.90
Persistent	AA	Present	113	42.5	1.3 (0.8–2.1)	0.22	0.48	–0.61 (0.40)	0.13	0.34
Persistent	AA	Absent	43	37.2	0.9 (0.5–1.8)	0.86	0.96	0.16 (0.58)	0.79	0.95
Persistent	Aa/aa	Present	69	46.4	1.5 (0.9–2.6)	0.15	0.35	–0.11 (0.48)	0.82	0.95
Persistent	Aa/aa	Absent	19	73.7	5.8 (2.0–21.1)	$3 \times 10^{-3}$ *	0.03	–3.37 (0.86)	$9 \times 10^{-5}$ *	$9 \times 10^{-3}$
Interaction					3.1 (0.7–14.6)	0.15	0.35	–3.30 (1.22)	$7 \times 10^{-3}$	0.06

\* *P* values with asterisk sign are the ones retaining statistical significance after multiple-testing correction

<sup>a</sup> Logistic regression was performed with adjusting covariates: maternal age, ethnicity, education, parity, marital status, maternal pre-pregnant BMI, alcohol use, passive smoke exposure, illicit drug use and infant sex

<sup>b</sup> Linear regression was performed with adjusting the same covariates described above

<sup>c</sup> Percentage of PTD subjects

<sup>d</sup> The first *P* value is raw *P* value without multiple testing correction, the adjusted *P* value is after multiple testing correction;  $P \leq 4 \times 10^{-3}$  is the significance level after correcting for multiple testing

with PTD accompanied by histologic chorioamnionitis (OR = 18.6; 95% CI: 2.1–162.7, multiple testing corrected *P* = 0.07). Likewise, we observed interaction of maternal

smoking, *CYP1A1* and *GSTT1* with spontaneous PTD (OR = 4.4; 95% CI: 0.9–21.8) and with very PTD (OR = 10.8; 95% CI: 2.5–51.4), but the interaction was not

**Table 4** Joint association of maternal cigarette smoking during pregnancy, *CYP1A1* and *GSTT1* gene polymorphisms with preterm subgroups

Smoking	<i>CYP1A1</i>	<i>GSTT1</i>	Total subjects	%	OR (95% CI)	<i>P</i>	Adjusted <i>P</i>
<i>Spontaneous PTD</i>							
Never	AA	Present	506	22.3 <sup>a</sup>	Ref		
Never	AA	Absent	166	21.1 <sup>a</sup>	0.9 (0.6–1.4)	0.71	0.90
Never	Aa/aa	Present	412	22.8 <sup>a</sup>	1.1 (0.8–1.5)	0.62	0.86
Never	Aa/aa	Absent	125	27.2 <sup>a</sup>	1.4 (0.9–2.2)	0.15	0.35
Persistent	AA	Present	104	37.5 <sup>a</sup>	1.7 (1.0–2.8)	0.04	0.15
Persistent	AA	Absent	40	32.5 <sup>a</sup>	1.2 (0.6–2.5)	0.65	0.88
Persistent	Aa/aa	Present	63	41.3 <sup>a</sup>	2.0 (1.1–3.5)	0.02	0.10
Persistent	Aa/aa	Absent	18	72.2 <sup>a</sup>	8.3 (2.7–30.6)	4 × 10 <sup>-4</sup> *	0.01
Interaction					4.4 (0.9–21.8)	0.07	0.21
<i>Very PTD (&lt;32 weeks)</i>							
Never	AA	Present	438	10.3 <sup>b</sup>	Ref		
Never	AA	Absent	138	5.1 <sup>b</sup>	0.4 (0.2–0.9)	0.05	0.17
Never	Aa/aa	Present	347	8.4 <sup>b</sup>	0.8 (0.5–1.4)	0.52	0.81
Never	Aa/aa	Absent	106	11.7 <sup>b</sup>	1.3 (0.6–2.5)	0.48	0.77
Persistent	AA	Present	79	17.7 <sup>b</sup>	1.6 (0.7–3.3)	0.27	0.55
Persistent	AA	Absent	30	10.0 <sup>b</sup>	0.7 (0.2–2.4)	0.65	0.88
Persistent	Aa/aa	Present	43	14.0 <sup>b</sup>	1.1 (0.4–3.0)	0.82	0.95
Persistent	Aa/aa	Absent	12	58.3 <sup>b</sup>	11.1 (2.9–47.7)	6 × 10 <sup>-4</sup> *	0.01
Interaction					10.8 (2.5–51.4)	2 × 10 <sup>-3</sup> *	0.03
<i>PTD accompanied by histologic chorioamnionitis</i>							
Never	AA	Present	413	10.9 <sup>c</sup>	Ref		
Never	AA	Absent	129	10.1 <sup>c</sup>	0.9 (0.5–1.8)	0.87	0.96
Never	Aa/aa	Present	328	10.7 <sup>c</sup>	1.0 (0.6–1.6)	0.99	0.99
Never	Aa/aa	Absent	97	10.3 <sup>c</sup>	0.9 (0.4–1.9)	0.81	0.95
Persistent	AA	Present	77	23.4 <sup>c</sup>	2.2 (1.1–4.3)	0.03	0.12
Persistent	AA	Absent	35	22.9 <sup>c</sup>	1.9 (0.7–4.7)	0.16	0.36
Persistent	Aa/aa	Present	40	12.5 <sup>c</sup>	1.1 (0.3–2.8)	0.92	0.97
Persistent	Aa/aa	Absent	12	66.7 <sup>c</sup>	15.6 (4.1–76.7)	2 × 10 <sup>-4</sup> *	0.01
Interaction					18.6 (2.1–162.7)	8 × 10 <sup>-3</sup>	0.06

Logistic regression was performed with adjusting covariates: maternal age, ethnicity, education, parity, marital status, maternal pre-pregnant BMI, alcohol use, passive smoke exposure, illicit drug use and infant sex

\* *P* values with asterisk sign are the ones retaining statistical significance after multiple-testing correction

<sup>a</sup> Percentage of subjects with spontaneous PTD

<sup>b</sup> Percentage of subjects with very PTD

<sup>c</sup> Percentage of subjects with PTD and histologic chorioamnionitis

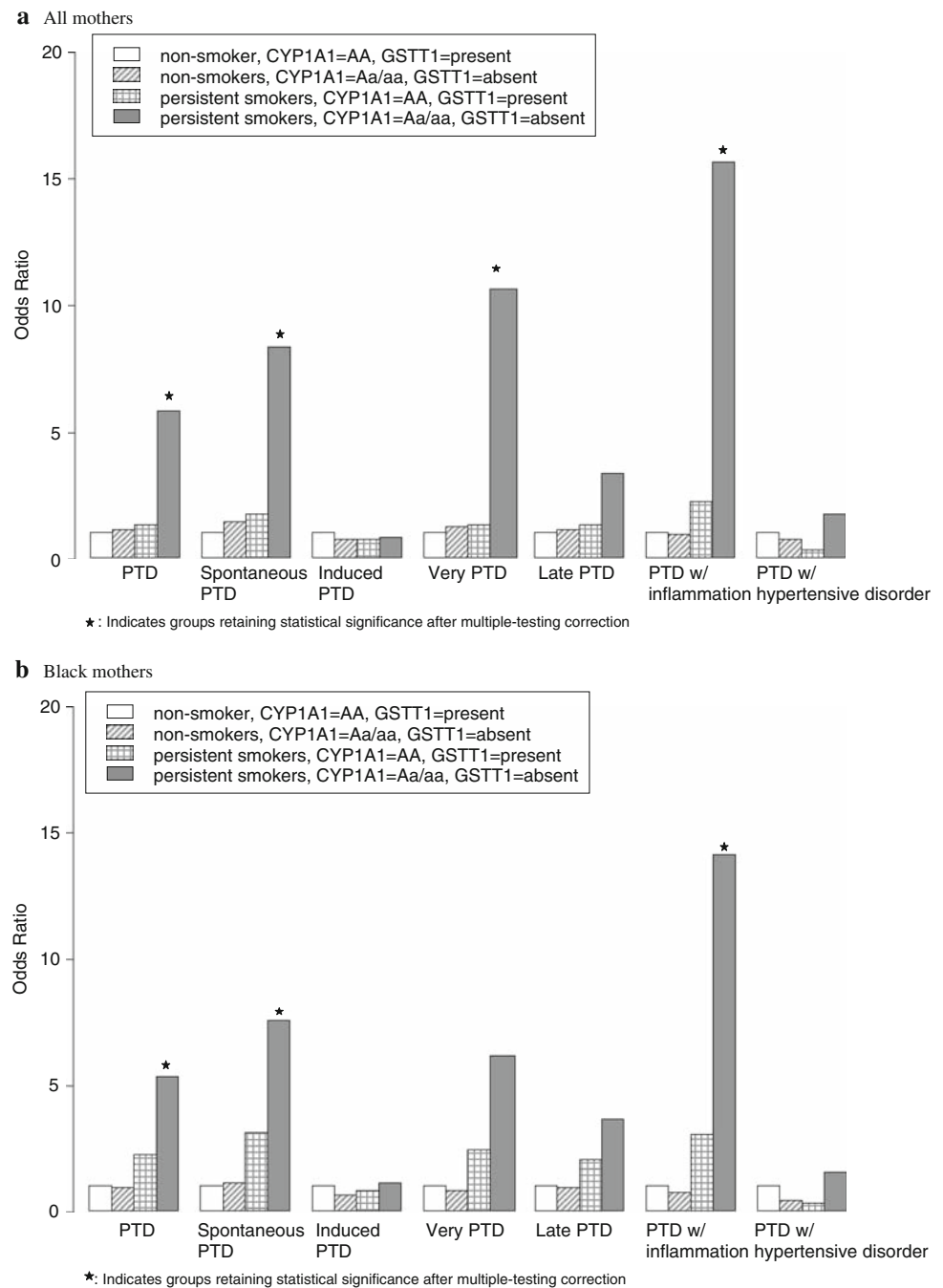
statistically significant (Table 4). In contrast, such joint associations or interactions were not found for medically induced PTD, late PTD, and PTD with hypertensive disorders (Fig. 1). Furthermore, we repeated above analyses in black mothers only and found the similar results (Fig. 1, Table 2 in supplementary material).

## Discussion

The findings of this study contribute new information to ongoing preterm gene-environmental research. First, we confirmed our previous finding in a sample half the size of this analysis of the joint association of maternal smoking and *CYP1A1* and *GSTT1* polymorphisms with gestational

age. Second, this is one of the largest study populations for investigation of preterm birth. The larger sample size allowed us to further elucidate this relationship among preterm subgroups. Interestingly, the joint effect of smoking and genetic susceptibility was significant in spontaneous PTD, very PTD, and most significant in PTD accompanied by histologic chorioamnionitis, but not in medically indicated PTD, late PTD, and PTD complicated by hypertensive disorders. We further demonstrated a marked difference in the percentages of histologic chorioamnionitis among significantly and non-significantly associated preterm subgroups, suggesting that intrauterine infection/inflammation may be a potential pathogenic pathway by which maternal cigarette smoking affects PTD. Indeed, an odds ratio (OR) of 15.6 for preterm subgroup with histo-

**Fig. 1** Joint association of maternal smoking during pregnancy, *CYP1A1* and *GSTT1* polymorphisms with PTD and preterm subgroups



logic chorioamnionitis among smoking mothers with high-risk genotypes is among the highest ever reported for PTD.

Of note, our finding that links gene–smoking interaction with histologic chorioamnionitis is biologically plausible. Although the adverse effect of maternal smoking on PTD is well established, the underlying biological mechanisms remain unclear. Previous studies have suggested several plausible explanations. First, tobacco smoke impairs placental vasculature function and subsequent transplacental transport of oxygen and nutrients (Bush et al. 2000; Muller et al. 2002; Shiverick and Salafia 1999). Second, toxic

chemicals within tobacco smoke could disturb fetal and placental cellular regulation via elevated DNA adducts and DNA damage. On the basis of our results, it is likely that smoking mothers with high-risk genotypes may have higher levels of PAH-DNA adducts and DNA strand breakage due to the increased activity of enzymes that metabolize cigarette toxins (e.g., *CYP1A1* Aa and aa) and lower or absent activity of enzymes that detoxify these compounds (e.g. *GSTT1* null genotypes). Moreover, such gene–smoking interactions may exert their synergistic effects on PTD through maternal and fetal inflammatory responses. Reac-



tive oxygen species (ROS) are important metabolites of tobacco smoke (Palackal et al. 2002) and may cause inflammatory responses by the activation of transcription factors, including nuclear factor- $\kappa$ B (NF- $\kappa$ B) and hypoxia-inducible factor-1 $\alpha$  (HIF-1 $\alpha$ ) (Haddad 2002; Irani 2000). In addition, the effects of tobacco smoke on the host response by triggering inflammatory responses have been reported previously in periodontal diseases (Ryder 2007). The effect of tobacco on periodontal diseases may be due to increasing expression or activities of inflammatory-related genes. Previous studies have also suggested that maternal smoking during pregnancy has an impact on fetal immune function measured by cord blood Immunoglobulin E (Magnusson 1986) and cytokine (Macaubas et al. 2003) concentrations, cord blood mononuclear cell cytokine responses (Noakes et al. 2003), cord blood lymphoproliferative response (Willwerth et al. 2006), and neonatal toll-like-receptor mediated immune responses (Noakes et al. 2006). Interestingly, when we examined the cord blood concentrations of 5 pro-inflammatory cytokines between 22 never-smoking mothers with low-risk genotypes of *CYP1A1* (AA) and *GSTT1* (present) and 11 smoking mothers with high-risk genotypes of *CYP1A1* (Aa/aa) and *GSTT1* (absent), 2 of the 5 were significantly higher in smoking mothers with high-risk genotypes of *CYP1A1* and *GSTT1* than the ones in never-smoking mothers with low-risk genotypes of *CYP1A1* and *GSTT1* (Table 3 in supplementary material). Together, previous studies and our finding support the possibility that intrauterine infection/inflammation is a pathogenic pathway by which maternal cigarette smoking affects PTD.

There are several strengths in the present study. This is one of the largest gene-environment studies of PTD, allowing us to study gene-smoking interaction across ethnic groups and among distinct preterm subgroups. By studying more homogeneous PTD subgroups, we were also able to explore possible pathogenic pathways by which maternal cigarette smoking may affect PTD. Multiple comparison is a common concern in genetic studies when certain number of association tests is examined. In this report, all the significant joint associations remained statistically significant after correcting for multiple comparison. Population stratification is another concern. We further adjusted for individual ancestry in black mothers. The results were not substantially changed when compared to the models without such adjustment.

A number of limitations should be considered. First, information on maternal smoking may be subject to recall bias. However, previous studies have demonstrated fair agreement between self-reported smoking amounts with serum and urinary levels of cotinine, a biochemical marker of cigarette smoke (George et al. 2006; Klebanoff et al. 2001; Peacock et al. 1998). In addition, we compared 277

subjects' smoking status between self-report in the postpartum questionnaire and medical record documentation. Of note, the data in medical record documentation were collected prospectively during each trimester and may be subject to less recall bias. There was over 99% agreement between self-reported smoking and medical record documentation. Moreover, any recall bias should be independent of maternal genotypes. Second, while placental histologic chorioamnionitis is objective, it is only a proxy for intrauterine infection. However, our previous study demonstrated that histologic chorioamnionitis had much stronger association with degree of prematurity than clinical chorioamnionitis (Gupta et al. 2007). Third, due to sample size constraints, we did not evaluate the joint effect in white and Hispanic groups separately, and future investigation of this subgroup is warranted. Fourth, due to limited sample size, we only examined two relatively common pregnancy complications: PTD with intrauterine infection/inflammation and PTD complicated by maternal hypertensive disorders. It is likely that we may miss such joint effect on some other important PTD subgroups in the current study. Fifth, although *CYP1A1* MspI and *GSTT1* deletion are well-studied functional variants, other SNPs in these two genes may be also of interest for investigation. Last, we did not include the genotypes from newborns, it is important to further investigate the joint maternal and fetal effect on PTD.

In summary, we demonstrate coherent evidence that smoking women carrying high-risk genotypes of *CYP1A1* (Aa&aa) and *GSTT1* (absent) are at significantly increased risk of PTD, especially in certain PTD subgroups, including PTD with spontaneous deliveries, PTD < 32 weeks, and PTD accompanied by histologic chorioamnionitis. These data suggest that intrauterine infection/inflammation may be a potential pathogenic pathway by which maternal cigarette smoking affects PTD and should provide a useful framework for future investigations. Our data also raised the possibility that we can identify women at high risk of PTD by taking into account both environmental exposure and gene polymorphisms. Identification of high-risk women for targeted intervention is very important from a research and public health perspective.

**Acknowledgments** We are grateful to two reviewers for the very helpful comments. The study was supported in part by grants from the National Institute of Child Health and Human Development (R01 HD41702), National Institute of Environmental Health Sciences (R01ES11682, R21ES11666), March of Dimes Birth Defects Foundation (20-FY98-0701, 20-FY02-56 and #21-FY07-605) and NICHD (K24 HD 042489). We thank the nursing staff of Labor and Delivery at Boston Medical Center for their continuous support and assistance to the study and Lingling Fu for data management, and Ann Ramsay for administrative support. We would like to particularly thank the outstanding expert consultants of the BMC Preterm Study team: Drs. Paul Wise, Jerome Klein, John M. Kasznica, and Milton Kotelchuck. Finally, we thank all of the participating mothers and their families.

## Appendix: Web resources

The URLs for data presented herein are as follows: SNP-HWE program, Center for Statistical Genetics, University of Michigan, <http://www.sph.umich.edu/csg/abecasis/Exact/index.html>; R website, The R Project for Statistical Computing, <http://www.r-project.org>; Bioconductor website, <http://www.bioconductor.org/>.

## References

- Anonymous (2006a) Preterm birth: cause, consequences, and prevention. In: press Tna (ed). Institute of medicine of the national academies (Committee on understanding premature birth and assuring healthy outcomes board on health sciences policy). Washington DC
- Anonymous (2006b) Preterm birth: crisis and opportunity. *Lancet* 368:339
- Bartsch H (1996) DNA adducts in human carcinogenesis: etiological relevance and structure–activity relationship. *Mutat Res* 340:67–79
- Bartsch H, Rojas M, Alexandrov K, Camus AM, Castegnaro M, Malaveille C, Anttila S, Hirvonen K, Husgafvel-Pursiainen K, Hietanen E et al (1995) Metabolic polymorphism affecting DNA binding and excretion of carcinogens in humans. *Pharmacogenetics* 5 (Spec No):S84–S90
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188
- Bush PG, Mayhew TM, Abramovich DR, Aggett PJ, Burke MD, Page KR (2000) Maternal cigarette smoking and oxygen diffusion across the placenta. *Placenta* 21:824–833
- Chen D, Hu Y, Yang F, Li Z, Wu B, Fang Z, Li J, Wang L (2005) Cytochrome P450 gene polymorphisms and risk of low birth weight. *Genet Epidemiol* 28:368–375
- Cnattingius S (2004) The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. *Nicotine Tob Res* 6(Suppl 2):S125–S140
- Crider KS, Whitehead N, Buus RM (2005) Genetic variation associated with preterm birth: a HuGE review. *Genet Med* 7:593–604
- Engle WA, Tomashek KM, Wallman C (2007) “Late-preterm” infants: a population at risk. *Pediatrics* 120:1390–1401
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587
- George L, Granath F, Johansson AL, Cnattingius S (2006) Self-reported nicotine exposure and plasma levels of cotinine in early and late pregnancy. *Acta Obstet Gynecol Scand* 85:1331–1337
- Gupta M, Mestan KK, Martin CR, Pearson C, Ortiz K, Fu L, Stubblefield P, Cerda S, Kasznica JM, Wang X (2007) Impact of clinical and histologic correlates of maternal and fetal inflammatory response on gestational age in preterm births. *J Matern Fetal Neonatal Med* 20:39–46
- Haddad JJ (2002) Antioxidant and prooxidant mechanisms in the regulation of redox(y)-sensitive transcription factors. *Cell Signal* 14:879–897
- Hamill PV, Drizd TA, Johnson CL, Reed RB, Roche AF, Moore WM (1979) Physical growth: National Center for Health Statistics percentiles. *Am J Clin Nutr* 32:607–629
- Hayashi S, Watanabe J, Kawajiri K (1992) High susceptibility to lung cancer analyzed in terms of combined genotypes of P450IA1 and Mu-class glutathione S-transferase genes. *Jpn J Cancer Res* 83:866–870
- Irani K (2000) Oxidant signaling in vascular cell growth, death, and survival: a review of the roles of reactive oxygen species in smooth muscle and endothelial cell mitogenic and apoptotic signaling. *Circ Res* 87:179–183
- Katoh T, Inatomi H, Nagaoka A, Sugita A (1995) Cytochrome P450IA1 gene polymorphism and homozygous deletion of the glutathione S-transferase M1 gene in urothelial cancer patients. *Carcinogenesis* 16:655–657
- Klebanoff MA, Levine RJ, Morris CD, Hauth JC, Sibai BM, Ben Curet L, Catalano P, Wilkins DG (2001) Accuracy of self-reported cigarette smoking among pregnant women in the 1990s. *Paediatr Perinat Epidemiol* 15:140–143
- Kramer MS, Platt R, Yang H, Joseph KS, Wen SW, Morin L, Usher RH (1998) Secular trends in preterm birth: a hospital-based cohort study. *JAMA* 280:1849–1854
- Kriek E, Rojas M, Alexandrov K, Bartsch H (1998) Polycyclic aromatic hydrocarbon-DNA adducts in humans: relevance as biomarkers for exposure and cancer risk. *Mutat Res* 400:215–231
- Landi MT, Bertazzi PA, Shields PG, Clark G, Lucier GW, Garte SJ, Cosma G, Caporaso NE (1994) Association between CYP1A1 genotype, mRNA expression and enzymatic activity in humans. *Pharmacogenetics* 4:242–246
- Macaubas C, de Klerk NH, Holt BJ, Wee C, Kendall G, Firth M, Sly PD, Holt PG (2003) Association between antenatal cytokine production and the development of atopy and asthma at age 6 years. *Lancet* 362:1192–1197
- Macones GA, Parry S, Elkousy M, Clothier B, Ural SH, Strauss JF III (2004) A polymorphism in the promoter region of TNF and bacterial vaginosis: preliminary evidence of gene-environment interaction in the etiology of spontaneous preterm birth. *Am J Obstet Gynecol* 190:1504–1508 discussion 3A
- Magnusson CG (1986) Maternal smoking influences cord serum IgE and IgD levels and increases the risk for subsequent infant allergy. *J Allergy Clin Immunol* 78:898–904
- Mathews TJ, Menacker F, MacDorman MF (2004) Infant mortality statistics from the 2002 period: linked birth/infant death data set. *Natl Vital Stat Rep* 53:1–29
- Muller J, Petkovic M, Schiller J, Arnold K, Reichl S, Arnhold J (2002) Effects of lysophospholipids on the generation of reactive oxygen species by fMLP- and PMA-stimulated human neutrophils. *Luminescence* 17:141–149
- Nakachi K, Imai K, Hayashi S, Kawajiri K (1993) Polymorphisms of the CYP1A1 and glutathione S-transferase genes associated with susceptibility to lung cancer in relation to cigarette dose in a Japanese population. *Cancer Res* 53:2994–2999
- Noakes PS, Holt PG, Prescott SL (2003) Maternal smoking in pregnancy alters neonatal cytokine responses. *Allergy* 58:1053–1058
- Noakes PS, Hale J, Thomas R, Lane C, Devadason SG, Prescott SL (2006) Maternal smoking is associated with impaired neonatal toll-like-receptor-mediated immune responses. *Eur Respir J* 28:721–729
- Nukui T, Day RD, Sims CS, Ness RB, Romkes M (2004) Maternal/newborn GSTT1 null genotype contributes to risk of preterm, low birthweight infants. *Pharmacogenetics* 14:569–576
- Palackal NT, Lee SH, Harvey RG, Blair IA, Penning TM (2002) Activation of polycyclic aromatic hydrocarbon trans-dihydrodiol proximate carcinogens by human aldo-keto reductase (AKR1C) enzymes and their functional overexpression in human lung carcinoma (A549) cells. *J Biol Chem* 277:24799–24808
- Palli D, Saieva C, Gemma S, Masala G, Gomez-Miguel MJ, Luzzi I, D’Errico M, Matullo G, Ozzola G, Manetti R, Nesi G, Sera F, Zanna I, Dogliotti E, Testai E (2005) GSTT1 and GSTM1 gene polymorphisms and gastric cancer in a high-risk Italian population. *Int J Cancer* 115:284–289
- Peacock JL, Cook DG, Carey IM, Jarvis MJ, Bryant AE, Anderson HR, Bland JM (1998) Maternal cotinine level during pregnancy and birthweight for gestational age. *Int J Epidemiol* 27:647–656

- Pemble S, Schroeder KR, Spencer SR, Meyer DJ, Hallier E, Bolt HM, Ketterer B, Taylor JB (1994) Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *Biochem J* 300(Pt 1):271–276
- Pennell CE, Jacobsson B, Williams SM, Buus RM, Muglia LJ, Dolan SM, Morken NH, Ozcelik H, Lye SJ, Relton C (2007) Genetic epidemiologic studies of preterm birth: guidelines for research. *Am J Obstet Gynecol* 196:107–118
- Pollack H, Lantz PM, Frohna JG (2000) Maternal smoking and adverse birth outcomes among singletons and twins. *Am J Public Health* 90:395–400
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Ryder MI (2007) The influence of smoking on host responses in periodontal infections. *Periodontol* 2000 43:267–277
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor
- Seidegard J, Vorachek WR, Pero RW, Pearson WR (1988) Hereditary differences in the expression of the human glutathione transferase active on trans-stilbene oxide are due to a gene deletion. *Proc Natl Acad Sci USA* 85:7293–7297
- Shah NR, Bracken MB (2000) A systematic review and meta-analysis of prospective studies on the association between maternal cigarette smoking and preterm delivery. *Am J Obstet Gynecol* 182:465–472
- Shiverick KT, Salafia C (1999) Cigarette smoking and pregnancy. I: Ovarian, uterine and placental effects. *Placenta* 20:265–272
- Wang X, Tager IB, Van Vunakis H, Speizer FE, Hanrahan JP (1997) Maternal smoking during pregnancy, urine cotinine concentrations, and birth outcomes: a prospective cohort study. *Int J Epidemiol* 26:978–988
- Wang X, Zuckerman B, Kaufman G, Wise P, Hill M, Niu T, Ryan L, Wu D, Xu X (2001) Molecular epidemiology of preterm delivery: methodology and challenges. *Paediatr Perinat Epidemiol* 15(Suppl 2):63–77
- Wang X, Zuckerman B, Pearson C, Kaufman G, Chen C, Wang G, Niu T, Wise PH, Bauchner H, Xu X (2002) Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. *JAMA* 287:195–202
- Wang L, Feng Y, Zhang Y, Zhou H, Jiang S, Niu T, Wei LJ, Xu X, Xu X, Wang X (2006) Prolylcarboxypeptidase gene, chronic hypertension, and risk of preeclampsia. *Am J Obstet Gynecol* 195:162–171
- Wigginton JE, Cutler DJ, Abecasis GR (2005) A note on exact tests of Hardy–Weinberg equilibrium. *Am J Hum Genet* 76:887–893
- Willwerth BM, Schaub B, Tantisira KG, Gold DR, Palmer LJ, Litonjua AA, Perkins DL, Schroeter C, Gibbons FK, Gillman MW, Weiss ST, Finn PW (2006) Prenatal, perinatal, and heritable influences on cord blood immune responses. *Ann Allergy Asthma Immunol* 96:445–453
- Xu X, Kelsey KT, Wiencke JK, Wain JC, Christiani DC (1996) Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 5:687–692
- Yang N, Li H, Criswell LA, Gregersen PK, Alarcon-Riquelme ME, Kittles R, Shigeta R, Silva G, Patel PI, Belmont JW, Seldin MF (2005) Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Hum Genet* 118: 382–392