

Evidence of Gene–Environment Interaction for Two Genes on Chromosome 4 and Environmental Tobacco Smoke in Controlling the Risk of Nonsyndromic Cleft Palate

Tao Wu^{1,2*}, Holger Schwender³, Ingo Ruczinski², Jeffrey C. Murray⁴, Mary L. Marazita⁵, Ronald G. Munger⁶, Jacqueline B. Hetmanski², Margaret M. Parker², Ping Wang¹, Tanda Murray², Margaret Taub², Shuai Li¹, Richard J. Redett⁷, M. Daniele Fallin², Kung Yee Liang^{2,8}, Yah Huei Wu-Chou⁹, Samuel S. Chong¹⁰, Vincent Yeow¹¹, Xiaoqian Ye^{12,13}, Hong Wang¹, Shangzhi Huang¹⁴, Ethylin W. Jabs^{7,13}, Bing Shi¹⁵, Allen J. Wilcox¹⁶, Sun Ha Jee¹⁷, Alan F. Scott⁷, Terri H. Beaty²

1 Peking University Health Science Center, Beijing, China, **2** Johns Hopkins University, School of Public Health, Baltimore, Maryland, United States of America, **3** Mathematical Institute, Heinrich Heine University Duesseldorf, Duesseldorf, Germany, **4** University of Iowa, Children's Hospital, Iowa City, Iowa, United States of America, **5** Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **6** Utah State University, Logan, Utah, United States of America, **7** Johns Hopkins University, School of Medicine, Baltimore, Maryland, United States of America, **8** National Yang-Ming University, Taipei, Taiwan, **9** Chang Gung Memorial Hospital, Taoyuan, Taiwan, **10** National University of Singapore, Singapore, Singapore, **11** KK Women's & Children's Hospital, Singapore, Singapore, **12** Wuhan University, School of Stomatology, Wuhan, China, **13** Mount Sinai Medical Center, New York, New York, United States of America, **14** Peking Union Medical College, Beijing, China, **15** State Key Laboratory of Oral Disease, West China College of Stomatology, Sichuan University, Chengdu, China, **16** NIEHS/NIH, Epidemiology Branch, Durham, North Carolina, United States of America, **17** Yonsei University, School of Public Health, Seoul, Korea

Abstract

Nonsyndromic cleft palate (CP) is one of the most common human birth defects and both genetic and environmental risk factors contribute to its etiology. We conducted a genome-wide association study (GWAS) using 550 CP case-parent trios ascertained in an international consortium. Stratified analysis among trios with different ancestries was performed to test for GxE interactions with common maternal exposures using conditional logistic regression models. While no single nucleotide polymorphism (SNP) achieved genome-wide significance when considered alone, markers in *SLC2A9* and the neighboring *WDR1* on chromosome 4p16.1 gave suggestive evidence of gene-environment interaction with environmental tobacco smoke (ETS) among 259 Asian trios when the models included a term for GxE interaction. Multiple SNPs in these two genes were associated with increased risk of nonsyndromic CP if the mother was exposed to ETS during the peri-conceptual period (3 months prior to conception through the first trimester). When maternal ETS was considered, fifteen of 135 SNPs mapping to *SLC2A9* and 9 of 59 SNPs in *WDR1* gave *P* values approaching genome-wide significance ($10^{-6} < P < 10^{-4}$) in a test for GxETS interaction. SNPs rs3733585 and rs12508991 in *SLC2A9* yielded $P = 2.26 \times 10^{-7}$ in a test for GxETS interaction. SNPs rs6820756 and rs7699512 in *WDR1* also yielded $P = 1.79 \times 10^{-7}$ and $P = 1.98 \times 10^{-7}$ in a 1 *df* test for GxE interaction. Although further replication studies are critical to confirming these findings, these results illustrate how genetic associations for nonsyndromic CP can be missed if potential GxE interaction is not taken into account, and this study suggest *SLC2A9* and *WDR1* should be considered as candidate genes for CP.

Citation: Wu T, Schwender H, Ruczinski I, Murray JC, Marazita ML, et al. (2014) Evidence of Gene–Environment Interaction for Two Genes on Chromosome 4 and Environmental Tobacco Smoke in Controlling the Risk of Nonsyndromic Cleft Palate. *PLoS ONE* 9(2): e88088. doi:10.1371/journal.pone.0088088

Editor: Nicholas M. Pajewski, Wake Forest University Health Sciences, United States of America

Received: July 1, 2013; **Accepted:** January 6, 2014; **Published:** February 6, 2014

Copyright: © 2014 Wu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Funding to support data collection, genotyping and analysis came from several sources, some to individual investigators (R01-DE-014581, R01-DE-016148, P50-DE-016215, R21-DE-016930, R01-DE-09886, R37-DE-08559, U01-DE-020057, R37-DE-0-8559, U01-DE-020057, R01-DE-012472, R01-DE-0106877) and some to the cleft consortium itself (U01-DE-018993). This project was part of the Gene, Environment Association Studies (GENEVA) Consortium funded by the National Human Genome Research Institute (NHGRI) to enhance communication and collaboration among researchers conducting genome-wide studies of complex diseases. The authors' group benefited greatly from the work and efforts of the entire consortium, especially the Coordinating Center (directed by B. Weir and C. Laurie of the University of Washington; U01- HG004446) in data cleaning and preparation for submission to the Database for Genotypes and Phenotypes (dbGaP). The authors also acknowledge the leadership of T. Manolio of NHGRI and E. Harris of National Institute of Dental and Craniofacial Research (NIDCR). Genotyping services were provided by the Center for Inherited Disease Research (CIDR). The International Cleft Consortium including genotyping and analysis was supported by the National Institute for Dental and Craniofacial Research through U01-DE-018993; "International Consortium to Identify Genes & Interactions Controlling Oral Clefts", 2007–2009; TH Beaty, PI. TW was supported by the International Collaborative Genetics Research Training Program (ICGRTP), NIH D43 TW06176 and 81102178 from National Natural Science of China. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: twu@bjmu.edu.cn

Introduction

Nonsyndromic cleft palate (CP) is a common birth defects and has a complex and heterogeneous etiology, involving both genetic and environmental risk factors [1]. The prevalence of CP is about 1/2500 live births, much lower than the 1/1000 live births prevalence for nonsyndromic cleft lip with or without cleft palate (CL/P). About half of all CP cases have another congenital anomaly or represent a recognized malformation syndrome, with the remaining half representing isolated nonsyndromic CP cases [2].

Genetic risk factors play an important role in the etiology of CP. A recent twin study in Denmark showed heritability of CP is as high as 90%, and the proband-wise concordance rate for CP among monozygotic twins was much higher compared to dizygotic twins: 33% vs. 7% [3]. Both family studies and population based studies have identified multiple candidate genes associated with increased risk of CP, including *FOXE1*, *ALX3*, *MKX*, *PDGFC*, and *SUMO1* [4–6]. However, evidence of association between reported candidate genes and CP remains inconsistent. Compared to the candidate gene approach, genome wide association studies (GWAS) have the advantage of providing better coverage of the human genome and are unbiased from a genetic perspective. Although several GWAS have identified strong signals at several chromosomal regions in multiple populations for CL/P [7–11], the variants controlling risk of CP have proven more difficult to find.

A few studies of CP have investigated potential GxE interaction for candidate genes and maternal exposure to cigarette smoking [12–14]. Environmental tobacco smoke (ETS) has also been reported to interact with certain SNPs to influence the risk of nonsyndromic CL/P and CP [15–20]. However, the evidence of GxE interaction has been rather inconclusive [21]. Possible reasons for the difficulty in documenting potential GxE interactions include: limited power due to modest sample size, different study designs and lack of available replication data. Integrating GxE interaction analysis into GWAS design is a powerful strategy for identifying more genetic factors influencing risk of complex disease, which could be overlooked when such interaction is ignored. A recent GWAS using the case-parent trio design found markers in several genes (*MLLT3*, *SMC2*, *TBK1*, *ZNF236*, and *BAALC*) showed statistically significant interaction with common maternal exposures, although no single SNP achieved genome-wide significance when such GxE interaction was ignored [22]. Beaty et al. (2011) combined CP trios from 12 different recruitment sites in their analysis, which involved considerably different rates of exposure to certain maternal exposures [22]. While the case-parent trio design has the advantage of being robust to confounding due to population stratification (compared to case-control designs), therefore allowing multi-site studies to amass large sample sizes, this advantage may not hold when considering GxE interaction especially if the exposure rates vary across sites.

In this study, we performed stratified analysis of the CP case-parent trios from the International Cleft Consortium [22] among trios with different ancestries to test for GxE interactions with common maternal exposures, including maternal cigarette smoking, alcohol consumption, ETS and multivitamin supplementation. Here we classified the trios used by Beaty et al. (2011) into groups of Asian and European ancestry and explored the potential GxE interactions.

Subjects and Methods

Case-parent Trios

Research protocols were reviewed and approved by institutional review boards (IRB) at each institution, including IRBs at The Johns Hopkins School of Public Health, University of Iowa, University of Pittsburgh, Utah State University, and all foreign collaborators. The review process for the consortium was approved by Johns Hopkins' IRB. Written informed consent was obtained from parents. Case-parent trios were drawn from an international consortium which conducted a GWAS using a case-parent trio design to search for genes controlling risk of nonsyndromic, isolated oral clefts [10]. Most cases were ascertained through surgical treatment centers at a surgical or post-surgical visit. Racial/ethnic background of participants was originally based on self-report and most of the 550 CP trios were of European or Asian ancestry, but this was confirmed by genotyping. Table 1 lists characteristics of the CP probands noting gender and recruitment site, stratified by European or Asian ancestry. To minimize potential misclassification of nonsyndromic CP, all probands were examined for other congenital anomalies or major developmental delays by either a clinical geneticist or experienced health care provider to rule out syndromic forms of oral clefts. As expected, there were slightly more female CP cases (56.1%) compared to males. None of the parents of these CP cases were themselves affected.

Table 1. Gender of isolated, nonsyndromic cleft palate (CP) cases in the International Cleft Consortium by recruitment site.

Asian ancestry			
Site	Males	Females	Total
Singapore	20	30	50
Taiwan	29	50	79
Shangdong Prov,China	16	22	38
Hubei Prov., China	19	26	45
Sichuan Prov., China	18	22	40
Other*	4	3	7
Subtotal	106 (40.9%)	153 (59.1%)	259
European ancestry			
Site	Males	Females	Total
Denmark	8	5	13
Norway	52	58	110
Iowa	18	22	40
Maryland	14	21	35
Pittsburg	6	8	14
Utah	29	27	56
Singapore	1	3	4
Subtotal	128(47.1%)	144(52.9%)	272
Other ancestries**			
	10(52.0%)	9(48.0%)	19
Total	244(43.9%)	306(56.1%)	550

*other sites include Maryland, Utah, and Korea.

**other ancestries include African American, Hispanic, Malay and others.

doi:10.1371/journal.pone.0088088.t001

Genotyping

The Center for Inherited Disease Research (CIDR) genotyped DNA samples using Illumina's 610 Quad platform and 99.1% passed CIDR quality control (QC) [10]. Genotypes on 589,945 SNPs (99.56% of those attempted) were released and then underwent further QC analysis to set up 4 types of QC flags for each SNP: 1) unacceptably high rates (>5%) of missing genotype calls, 2) low minor allele frequency (MAF<0.01), 3) unacceptably high rates of Mendelian errors (>5%) between parents and child, and 4) significant deviation ($p<10^{-5}$) from Hardy Weinberg equilibrium (HWE) among parents within recruitment site or across European and Asian populations separately. This QC process flagged 14.6% of all SNPs (mostly for low MAF), leaving ~498 K SNPs available for analysis.

Exposure Assessment

Maternal exposure information, including cigarette smoking, ETS, multivitamin supplementation, and alcohol consumption was collected through direct interview of mothers. Only Asian sites collected complete information on ETS. Environmental exposures were defined as being exposed from three months prior to pregnancy through the first trimester. The question measuring ETS status during certain periods asked “did someone smoke in your home, workplace or any other place near you?”. Maternal exposures were assessed as simple yes/no responses. See Table II in Beaty et al. (2011) for details of the exposure rates for all CP trios. The proportion of infants exposed to maternal cigarette smoking and alcohol consumption was very low among Asian mothers (around 4%), so only maternal ETS and multivitamin supplementation could be analyzed in this group. The proportion of exposure to ETS and multivitamin supplementation among Asians were 40.3% and 20.2%, respectively.

Statistical Analysis

MAFs were computed using parents only. Pairwise linkage disequilibrium (LD) was measured as r^2 for all SNPs using the Haploview program, and was used to identify linkage disequilibrium (LD) blocks [23]. In this study, we used a closed form genotypic transmission disequilibrium test (gTDT) developed by Schwender et al. (2012) to test for genetic association of each SNP [24]. To perform this gTDT, a “pseudo-control” dataset was created based on the observed genotype of the case and all alternative possible genotypes given the parental mating type. The gTDT has a number of advantages compared to allelic TDT [25]. While assuming different models of inheritance, the gTDT can be used to estimate the relative risks (RRs) of each genotype and a term for GxE interaction can also be incorporated. Schwender et al. (2012) developed a method with a closed form solution providing parameter estimates for genome-wide markers efficiently [24] and is implemented in the R package Trio (v 1.5.0).

All autosomal markers were examined using the conditional logistic regression model assuming an additive model of inheritance. The log-odds of being the observed case in the i -th trio is modeled as: $\text{logit}[P(\text{case}_i)] = \beta_G(G_i) + \beta_{G \times E}(G_i \times E_i)$, where $G = 0, 1,$ or 2 stands for the number of risk alleles in the case: “pseudo-control” set (representing a 1:3 matching), and where $E = 0$ or 1 reflects unexposed or exposed mothers, respectively. A 2 degree of freedom (df) likelihood ratio test (LRT) for joint effects of G and $G \times E$ interaction was first performed, followed by a 1 df LRT for $G \times E$ interaction alone. The 2 df test examines the inherited effect of the SNP after taking into account effects of $G \times E$ interaction, while the 1 df test focuses exclusively on $G \times E$ interaction. We used $\text{RR}(CP|G \text{ no } E) = \exp(\beta_G)$ to represent the estimated RRs of being a case with one copy of the risk allele in the absence of maternal

environment exposures, while $\text{RR}(CP|G \text{ and } E) = \exp(\beta_G + \beta_{G \times E})$ reflects the RR of being a case carrying one copy of the risk allele in the presence of maternal exposure.

Results

A conventional search for marginal gene (G) effects in the total sample of 550 CP trios, as well as in the stratified analysis of trios of Asian and European ancestry, showed no markers achieved significance at a genome-wide level ($P \leq 10^{-7}$, data not shown).

A genome-wide screen for $G \times E$ interaction was carried out using Trio (1.5.0), where conditional logistic regression models were used to estimate effects of $G \times E$ interaction alone (LRT with 1 df), as well as the combined effects of gene (G) and gene-environment ($G \times E$) interaction (LRT with 2 df). This screening process yielded no significant signals among European trios (see Figure S1 in File S1 for $G \times E$ interaction results on maternal smoking, alcohol consumption and multivitamin supplementation among European trios), but revealed several markers with suggestive evidence of $G \times E$ interaction ($10^{-6} < P < 10^{-4}$) among 259 Asian trios clustered on chromosome 4p16, especially in the 1 df test for $G \times E$ interaction. Figure 1 presents a conventional Manhattan plot for all autosomal SNPs where $-\log_{10}(P)$ from the 1 df LRT for $G \times E$ interaction was plotted (See Figure S2 in File S1 for a Q-Q plot of $G \times E$ interaction among Asian trios). Therefore, we mainly present results for $G \times E$ interaction among Asian trios here.

To further investigate this evidence, Figure 2 presents a “double Manhattan plot” to summarize joint evidence for G and $G \times E$ interaction effects on chromosome 4p (over the region 8988690 kb~10636912 kb). Table S1 in File S1 showed the physical location and MAFs of SNPs in this region (19 SNPs with $MAF < 0.01$ were dropped in this region). The bottom half of this plot shows the $\log_{10}(P)$ for the conventional family-based test of SNP effects *ignoring exposure* (where more significant results fall farther below the mid-line). In the top half of Figure 2, $-\log_{10}(P)$ are shown for each autosomal SNP from both the 2 df test of G and $G \times E$ interaction together (red dots) and the 1 df test for $G \times E$ interaction alone (blue dots). Dashed lines connect P -values from the marginal test ignoring exposures (below the mid-line) to those models considering $G \times E$ interaction (above the mid-line). As seen in Figure 2, more than 20 markers gave P values approaching genome-wide significance level in tests for $G \times E$ interaction, including 15 SNPs in *SLC2A9* and 9 in *WDR1* on chr. 4p16.1 (Figure S3 in File S1 shows LD plots for these two genes).

Although none of 135 SNPs mapping to *SLC2A9* approached genome-wide significance level when maternal exposure to ETS was ignored (lower half of Figure 2), a cluster of 61 SNPs identified a region spanning 125 kb yielded P values approaching genome-wide significance levels when interaction with maternal ETS was considered. In this region, fifteen SNPs showed suggestive evidence of $G \times E$ interaction in the 1 df test (Table 2). SNPs rs3733585 and rs12508991 suggested $G \times E$ interaction in the 1 df test ($P = 2.26 \times 10^{-7}$).

Regression coefficients from the conditional logistic regression model provide an estimate of exposure specific RRs under this additive model. When both G and $G \times E$ terms were included in the conditional logistic regression model, RRs were also calculated for both exposed and unexposed heterozygous carriers of the apparent risk allele. Figure 3 shows estimated $\text{RR}(CP|G \text{ no } E)$ and $\text{RR}(CP|G \text{ and } E)$ for 15 SNPs in *SLC2A9* along with P values from the LRT for both the 2 df and 1 df test. Here, the apparent “risk allele” became the target allele (which was the minor allele for rs10022499, rs10016075, rs2240723, rs3733585, rs733175, but

CP Asian ETS G×E Manhattan Plot

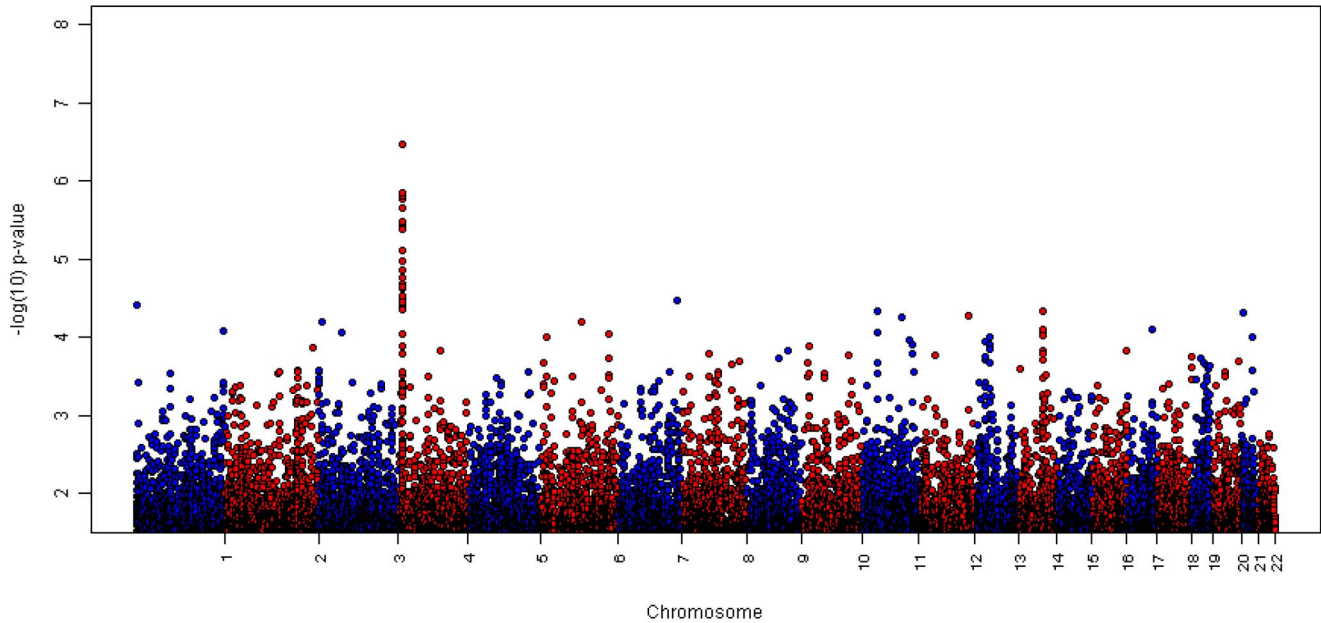


Figure 1. CP Asian ETS G×E Manhattan Plot. Manhattan plot with P values from likelihood ratio tests with 1 degree of freedom testing for G×ETS interaction among 259 Asian CP trios (492,698 SNPs were left in Asian trios after quality control). doi:10.1371/journal.pone.0088088.g001

the major allele for rs4447863, rs998676, rs6849717, rs11723970, rs17187075, rs12499857, rs10939650, rs4622999, rs7657096, rs12508991—see Table 2). Estimated RR(CP|G and E) and their 95% CI for a heterozygous child whose mother was exposed to ETS were distinctly higher (open circles) compared to a similar heterozygous child of unexposed mothers (solid circles). For the two most significant SNPs (rs3733585 and rs12508991) being a heterozygous child of an exposed mother was associated with a

2.58-fold increase in risk (RR = 2.58; 95% CI: 1.61–4.14), but not among children of unexposed mothers (RR = 0.60; 95% CI: 0.43–0.83). The 1 *df* LRT for G×ETS interaction in this conditional logistic regression model approached genome-wide significance ($P = 2.26 \times 10^{-7}$).

WDR1 on chr. 4p16.1 is located next to *SLC2A9* and encompasses 59 SNPs. Like *SLC2A9*, none of these SNPs achieved genome-wide significance levels alone, however, a block of 9 SNPs

Gene X Environment Tobacco Smoke Interaction among Asian CP Group

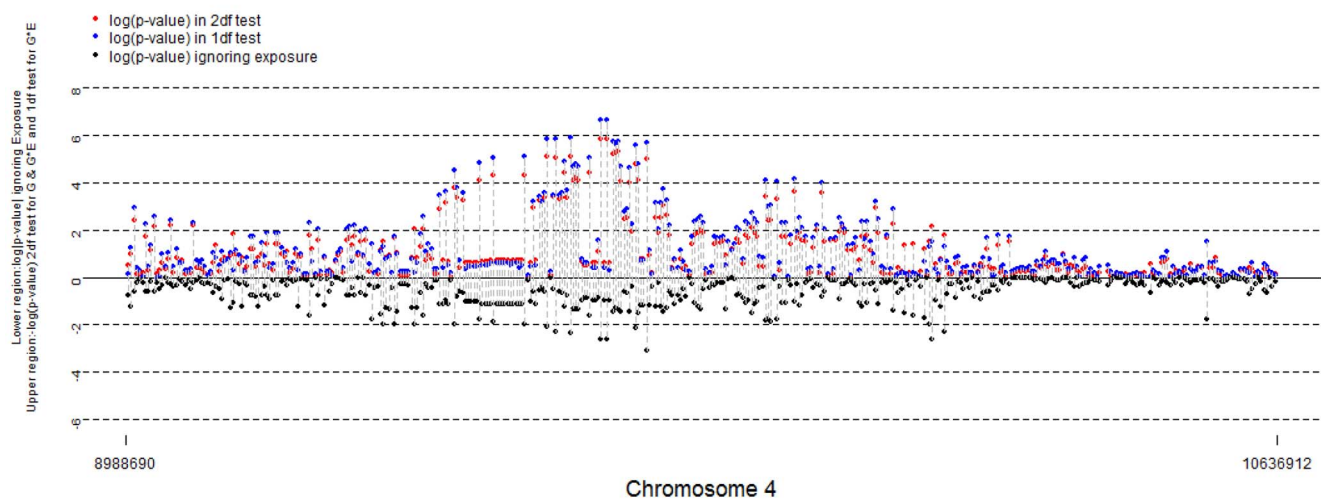


Figure 2. Gene × Environmental Tobacco Smoke Interaction among Asian CP Group. Double Manhattan plots for SNP effects ignoring maternal exposures (black dots in the lower half) and considering G and G×E interaction for environmental tobacco smoke on selected region on chromosome 4p among 259 Asian trios. Blue dots represent $-\log_{10}(P)$ from the 1 *df* test of G×E interaction alone; red dots represent $-\log_{10}(P)$ from the 2 *df* test of G and G×E interaction. Dashed lines connect SNP showing this level of significance in one test considering G×E interaction with their corresponding P -value when interaction was ignored. doi:10.1371/journal.pone.0088088.g002

Table 2. Estimated RR(case|G no E) and RR(case|G and E) from conditional logistic regression using cases and 3 pseudo-controls in 259 Asian CP case-parent trios for 15 SNPs in *SLC2A9* considering GxE interaction between each SNP and maternal exposure to environmental tobacco smoke.

SNP	Physical location	TA (freq)	RR(case G no E)	RR(case G and E)	LRT 2 <i>df</i> P values	LRT 1 <i>df</i> P values
rs4447863	9548067	C(0.593)	0.64(0.46,0.90)	2.07(1.32,3.25)	1.61×10^{-4}	2.94×10^{-5}
rs998676	9557662	G(0.587)	0.64(0.46,0.90)	2.07(1.32,3.25)	7.51×10^{-5}	1.36×10^{-5}
rs6849717	9567817	C(0.590)	0.66(0.47,0.91)	2.26(1.44,3.55)	5.07×10^{-5}	9.01×10^{-6}
rs11723970	9589560	T (0.588)	0.64(0.46,0.90)	2.21(1.42,3.46)	4.67×10^{-5}	8.14×10^{-6}
rs17187075	9599426	G(0.596)	0.64(0.46,0.89)	2.48(1.56,3.95)	8.22×10^{-6}	1.38×10^{-6}
rs12499857	9604474	G(0.600)	0.62(0.44,0.86)	2.46(1.53,3.95)	8.57×10^{-6}	1.36×10^{-6}
rs10939650	9607538	T(0.500)	0.71(0.50,1.01)	2.46(1.56,3.88)	3.91×10^{-5}	1.22×10^{-5}
rs4622999	9612493	C(0.600)	0.61(0.44,0.86)	2.46(1.53,3.95)	8.08×10^{-6}	1.28×10^{-6}
rs7657096	9613098	A(0.528)	0.70(0.49,0.99)	2.33(1.49,3.66)	7.74×10^{-5}	2.03×10^{-5}
rs10022499	9615635	C(0.492)	0.71(0.50,1.00)	2.37(1.51,3.72)	5.87×10^{-5}	1.60×10^{-5}
rs10016075	9615761	G(0.492)	0.71(0.50,1.00)	2.33(1.49,3.66)	8.24×10^{-5}	2.15×10^{-5}
rs2240723	9630249	A(0.465)	0.67(0.47,0.95)	2.32(1.49,3.62)	3.93×10^{-5}	8.81×10^{-6}
rs3733585	9645437	C(0.412)	0.60(0.43,0.83)	2.58(1.61,4.14)	1.52×10^{-6}	2.26×10^{-7}
rs12508991	9650202	C(0.588)	0.60(0.43,0.83)	2.58(1.61,4.14)	1.52×10^{-6}	2.26×10^{-7}
rs733175	9659239	C(0.488)	0.69(0.49,0.98)	2.71(1.70,4.33)	6.12×10^{-6}	1.89×10^{-6}

TA: target allele and its frequency among parents of Asian ancestry.
doi:10.1371/journal.pone.0088088.t002

(spanning 213 kb) showed suggestive GxETS interaction in the 1 *df* test for GxE interaction (Table 3). SNPs rs6820756 and rs7699512 yielded $P=1.79 \times 10^{-6}$ and $P=1.98 \times 10^{-6}$ in the 1 *df* test for GxE interaction, and two adjacent SNPs also approached genome-wide significance (rs6834555, $P=2.28 \times 10^{-6}$; rs6834555, $P=2.73 \times 10^{-6}$). Figure 4 shows estimated RR(CP|G no E) and RR(CP|G and E), plus their 95%CI, for these 9 SNPs under an additive model. The risk of having nonsyndromic CP was 1.97–2.75 times higher when the fetus carried the risk allele and the mother was exposed to ETS compared to carriers whose mothers were not exposed.

Examining the imputed genotypes generated by the GENEVA Coordinating Center [26] using 1000 Genomes reference populations after pre-phasing haplotypes using IMPUTE2 [27] yielded additional evidence of GxETS interaction. Analysis of imputed SNPs in the region of these two genes yielded genome-wide significance for several markers (see Figure S4 in File S1).

Similar analysis for potential GxE interaction with maternal multivitamin supplementation in these same Asian CP trios showed no significant GxE interaction. Because the exposure rate for maternal multivitamin supplementation was lower in this sample of Asian CP trios (~20%), however, this sample had less statistical power to detect GxE interaction unless the causal allele

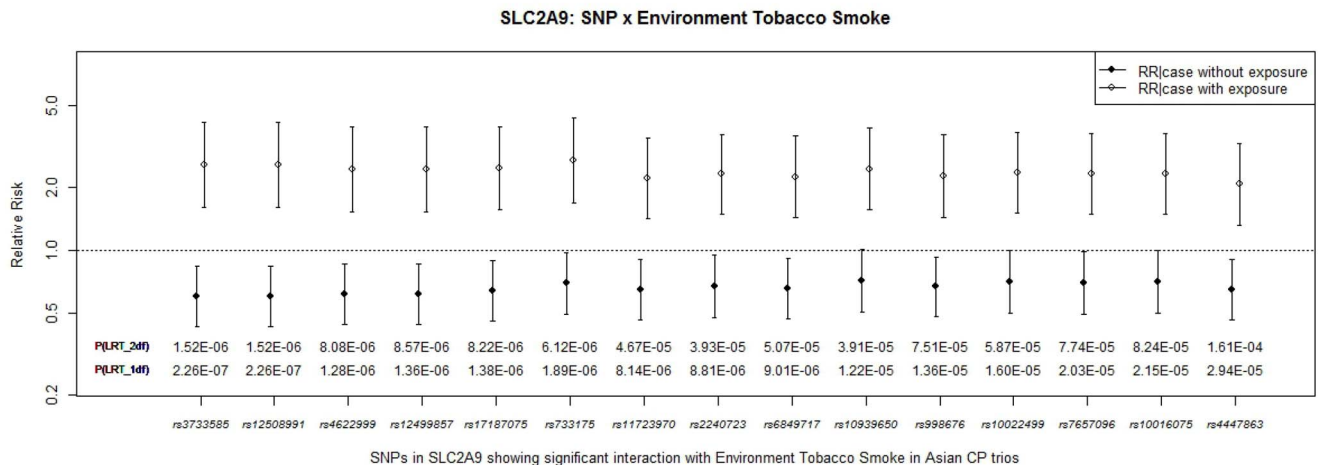


Figure 3. SLC2A9: SNP × Environmental Tobacco Smoke. Estimated RR(CP|G no E) and RR(CP|G and E) from conditional logistic regression model considering SNP effects and their interaction with maternal exposure to ETS on 259 CP case-parent trios of Asian ancestry for fifteen SNPs in *SLC2A9*. P-values from the 2 *df* and 1 *df* LRT for GxE interaction are shown along the X axis.
doi:10.1371/journal.pone.0088088.g003

Table 3. Estimated RR(case|G no E) and RR(case|G and E) from conditional logistic regression using cases and 3 pseudo-controls in 259 CP case-parent trios for 9 SNPs in *WDR1* considering maternal exposure to environmental tobacco smoke.

SNP	Physical location	TA (freq)	RR(case G no E)	RR(case G and E)	LRT 2 df p-value	LRT 1 df p-value
rs6834555	9671424	G(0.482)	0.71(0.50,1.00)	2.75(1.72,4.39)	5.61×10^{-6}	2.28×10^{-6}
rs6820756	9671947	A(0.482)	0.70(0.50,0.99)	2.75(1.72,4.39)	4.93×10^{-6}	1.79×10^{-6}
rs2241469	9689560	A(0.652)	0.68(0.48,0.97)	2.36(1.48,3.77)	8.73×10^{-5}	1.97×10^{-5}
rs2241482	9708912	G(0.654)	0.69(0.49,0.98)	2.37(1.47,3.83)	1.06×10^{-4}	2.31×10^{-5}
rs717615	9713768	C(0.527)	0.63(0.45,0.88)	2.42(1.50,3.89)	1.67×10^{-5}	2.73×10^{-6}
rs4697922	9719703	C(0.657)	0.68(0.48,0.97)	2.43(1.50,3.96)	7.70×10^{-5}	1.58×10^{-5}
rs7699512	9734906	T(0.519)	0.56(0.40,0.79)	2.11(1.34,3.30)	9.90×10^{-6}	1.98×10^{-6}
rs10489072	9882342	G(0.499)	0.65(0.46,0.92)	1.97(1.27,3.05)	4.08×10^{-4}	7.81×10^{-5}
rs6833142	9885080	G(0.494)	0.66(0.47,0.93)	1.97(1.27,3.05)	4.65×10^{-4}	8.98×10^{-5}

TA: target allele and its frequency among parents of Asian ancestry.
doi:10.1371/journal.pone.0088088.t003

were highly polymorphic (MAF>0.15) and the true interaction effects were at least as large as those seen in test of GxETS interaction ($RR_{GE}>2.5$).

Discussion

While the initial GWAS of 550 CP case-parent trios stratified by European and Asian ancestry did not yield any markers achieving genome-wide significance (i.e. when GxE interaction was ignored), multiple markers in two adjacent genes on chr. 4p16.1 (*SLC2A9* and *WDR1*) showed *P*-values approaching genome-wide significance when GxETS interaction was incorporated into the analysis of Asian trios. Our results suggested *SLC2A9* and/or *WDR1* located at position 9 Mb on chromosome 4p16.1 may influence risk of nonsyndromic CP through interaction with maternal exposure to ETS, though independent replication studies are still needed to confirm these findings. Our study did not yield any compelling evidence of GxE interactions approaching genome-wide significance among trios of European ancestry.

Identifying GxE interaction will lead to better understanding of the etiology of common birth defects and potential biological mecha-

nisms, as well as create opportunities for designing effective prevention strategies. Several studies have shown maternal smoking is not only an independent risk factor for CP [28,29], but may interact with genetic variants to influence risk [12–14]. GxSmoking interaction has been suggested for markers in the chr. 4p16 region. A previous case-control study and case-parent trio studies showed evidence of GxSmoking for markers near *MSX1* on chr. 4p16 among CP trios or combined CL/P and CP trios [30,31]. This 4p16 region has been suggested to be associated with increased risk of nonsyndromic oral clefts, including CL/P and CP in a previous analysis [32]. Ingersoll et al. (2010) used 381 case-parent trios from four populations including Asian samples from Singapore, Korea and Taiwan [32]. Their analysis focused on the 2 Mb region around *MSX1* and showed SNP effects in *STK32B*, the *EVC–EVC2–CRMP1* region, and the *STX18–MSX1* region were significantly associated with risk to CP, especially among Asian trios. A Dutch study showed smoking by both parents may interact with SNPs in *MSX1* to increase the risk of nonsyndromic oral clefts [16]. *SLC2A9* and *WDR1*, the most significant genes seen here, are located about 3 Mb downstream of *MSX1*. ETS has been shown to interact with candidate genes to influence risk of nonsyndromic oral clefts in different populations.

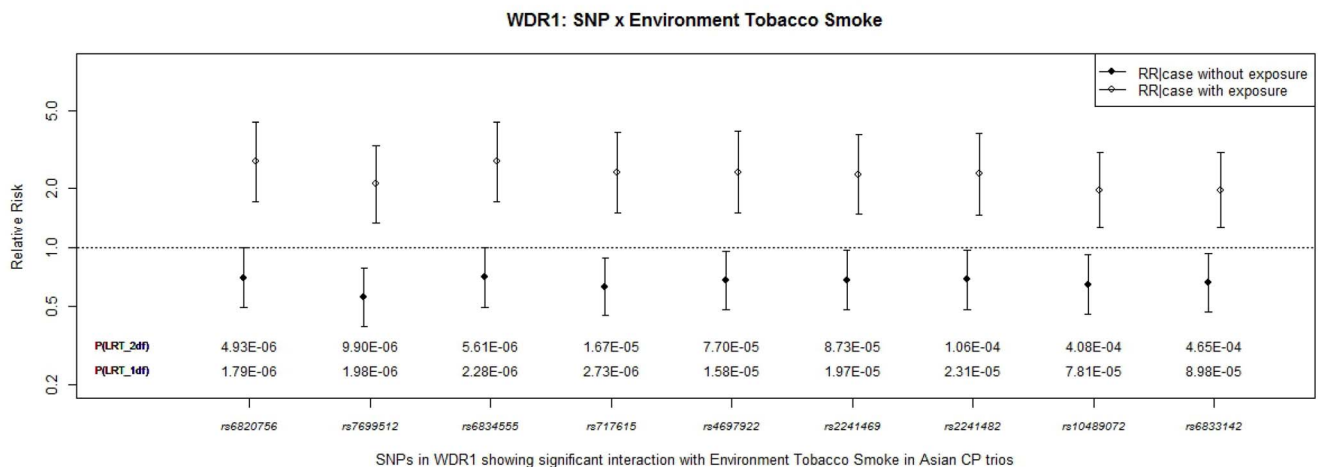


Figure 4. WDR1: SNP × Environmental Tobacco Smoke. Estimated RR(CP|G no E) and RR(CP|G and E) from conditional logistic regression model considering SNP effects and their interaction with maternal exposure to ETS on 259 CP case-parent trios of Asian ancestry for nine SNPs in *WDR1*. *P*-values from the 2 *df* and 1 *df* LRT for GxE interaction are shown along the X axis.
doi:10.1371/journal.pone.0088088.g004

Previous studies have shown ETS may interact with *IRF6*, *RUNX2*, and *BMP4* among Chinese CL/P case-parent trios [17,19,33]. Another French study yielded suggestive evidence for interaction between *CYP1A1* and ETS among nonsyndromic oral cleft trios [15]. Li et al. (2011) also found maternal ETS interacted with one SNP in *microRNA-140* gene to increase the risk of nonsyndromic CP using case-control design in a Chinese population [18].

The *WDR1* gene (WD repeat domain 1, also called actin-interacting protein 1) is downstream from *SLC2A9* and it is highly conserved in eukaryotes and promotes cofilin-mediated actin filament disassembly [34]. Kato et al. (2008) noted *WDR1* has an important role in unidirectional cell migration by promoting cofilin activity. Protein aggregates of actin and cofilin in the brains of twins with dystonia and CL/P were described by Gearing, et al. (2010) [35]. While neither of these studies is proof of any link between *WDR1* and nonsyndromic oral clefts, they suggest a possible biological mechanism involving disruption of cell migration during development of the palate. The *SLC2A9* gene (solute carrier family 2, member 9) is located on chromosome 4p16.1, and encodes a member of the *SLC2A* facilitative glucose transporter gene family, which is critical for maintaining glucose homeostasis. Multiple association studies across several populations showed consistent findings that this gene is associated with uric acid concentration and risk of gout [36–43], with a higher effect size among females compared to males. In our study, markers in these genes were in high LD. Therefore, the significant findings of GxETS interaction in *SLC2A9* may reflect its close physical proximity to *WDR1*. We also performed the GxETS analysis using imputed genotype data in this chromosomal region among Asian trios, and the imputed genotypes yielded greater significance (including several achieving genome-wide significance) in the region (Figure S4 in File S1). In addition, we tested for parent-of-origin effects among exposed and unexposed trios using the parent-of-origin likelihood ratio test [44], but found no significant signals. Our study suggested genes in this region may play a role in the etiology of CP not only through gene effects but also may through potential GxE interactions, at least among Asian populations. Although it is unclear how either of these genes affects cleft development, our results suggest these two genes (especially *WDR1*) should be considered as candidate genes for nonsyndromic CP.

We acknowledge the suggestive GxETS interactions on chr. 4p16.1 region seen in the present study require further confirmation in independent samples. However, adequate sample size will be a challenge. Our results argue maternal ETS appears to increase risk of nonsyndromic CP in Asian cases carrying certain genotypes in *SLC2A9* and *WDR1*. Exposure rate of ETS among Asian mothers in this sample was as high as 40%, reflecting the high prevalence of smoking among Asian males (about 60%) [45]. If this observation can be confirmed, such a GxE interaction creates opportunities for an effective intervention to reduce the risk. Our suggestive evidence of interaction between ETS and two genes on Chromosome 4 would be strengthened if we could test for GxSmoking interaction in this same population. However, such analysis would be severely underpowered due to low rates of personal smoking among Asian women. Further analyses will be required to understand how maternal exposure to ETS could interact with genes to affect fetal development. In addition, the analysis testing for the interaction between maternal genes and environmental exposures could also be informative. While the case-parent trio design is robust to population stratification [46,47], and stratification into Asian/European ancestries minimizes potential confounding due to differences in exposure rates, this study illustrates the importance of considering possible GxE interaction in the etiology of CP. Still statistical interaction does not guarantee biological interaction, and the functional gene may be located some distance from the statistical signals for GxE interaction seen here.

Supporting Information

File S1 Supporting information. Figure S1.1: Manhattan plot with *P* values from likelihood ratio tests with 1 degree of freedom testing for GxSmoking interaction among 272 European CP trios. Figure S1.2 Manhattan plot with *P* values from likelihood ratio tests with 1 degree of freedom testing for GxAlcohol consumption interaction among 272 European CP trios. Figure S1.3 Manhattan plot with *P* values from likelihood ratio tests with 1 degree of freedom testing for GxMultivitamin supplementation consumption interaction among 272 European CP trios. Figure S2: Q-Q plot with *P* values from likelihood ratio tests with 1 degree of freedom testing for GxETS interaction among 259 Asian CP trios (492,698 SNPs were left in Asian trios after quality control). The gray shaded region indicates 95% confidence band for order statistics. The numbers on the top axis indicate the respective locations for (ordered) expected $-\log_{10}$ p-values. (e.g., the number 1 (10) indicates the expected value, on the $-\log_{10}$ scale, for the minimum (i.e. the tenth smallest) p-value). Figure S3.1 LD plots for *SLC2A9* among 259 Asian CP trios. Black squares represent $r^2 = 1$; gray squares represent $0 < r^2 < 1$; white squares represent $r^2 = 0$. Figure S3.2 LD plots for *WDR1* among 259 Asian CP trios. Black squares represent $r^2 = 1$; gray squares represent $0 < r^2 < 1$; white squares represent $r^2 = 0$. Figure S4: *P* values from likelihood ratio test with 1 degree of freedom testing for GxETS interaction after including the imputed SNPs among Asian CP trios. Circles represent imputed genotypes using 1000 Genomes as a reference population and squares represent observed SNPs. Table S1. (DOCX)

Acknowledgments

We sincerely thank all of the families at each recruitment site for participating in this study, and we gratefully acknowledge the invaluable assistance of clinical, field and laboratory staff who contributed to making this work possible. Funding to support data collection, genotyping and analysis came from several sources, some to individual investigators (R01-DE-014581, R01-DE-016148, P50-DE-016215, R21-DE-016930, R01-DE-09886, R37-DE-08559, U01-DE-020057, R37-DE-0-8559, U01-DE-020057, R01-DE-012472, R01-DE-0106877) and some to the cleft consortium itself (U01-DE-018993). This project was part of the Gene, Environment Association Studies (GENEVA) Consortium funded by the National Human Genome Research Institute (NHGRI) to enhance communication and collaboration among researchers conducting genome-wide studies of complex diseases. Our group benefited greatly from the efforts of the entire consortium, especially the Coordinating Center (directed by B. Weir and C. Laurie of the University of Washington; U01-HG004446) in data cleaning and preparation for submission to the Database for Genotypes and Phenotypes (dbGaP). We also acknowledge the leadership of T. Manolio of NHGRI and E. Harris of National Institute of Dental and Craniofacial Research (NIDCR). Genotyping services were provided by the Center for Inherited Disease Research (CIDR). The International Cleft Consortium including genotyping and analysis was supported by the National Institute for Dental and Craniofacial Research through U01-DE-018993; “International Consortium to Identify Genes & Interactions Controlling Oral Clefts”, 2007–2009; TH Beaty, PI. T Wu was supported by the International Collaborative Genetics Research Training Program (ICGRTP, NIH D43 TW06176) and “The genome wide association study of nonsyndromic oral clefts in Chinese populations using case parent trios design” (81102178, 2012–2014) from National Natural Science Foundation of China.

Author Contributions

Conceived and designed the experiments: THB JCM MLM RGM. Analyzed the data: TW IR JBH MMP PW TM MT SL KYL. Contributed reagents/materials/analysis tools: HS IR. Wrote the paper: TW IR AFS MDF EWJ THB. Conducted field investigation and collected samples:

THB JCM MLM RGM JBH RJR YHWC SSC VY XY HW SH EWJ BS AJW SHJ.

References

- Dixon MJ, Marazita ML, Beaty TH, Murray JC (2011) Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev Genet* 12: 167–178.
- Tolarová MM, Cervenka J (1998) Classification and birth prevalence of orofacial clefts. *Am J Med Genet* 75: 126–137.
- Grosen D, Bille C, Petersen I, Skytthe A, Hjelmborg Jv, et al. (2011) Risk of oral clefts in twins. *Epidemiology* 22: 313–319.
- Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, et al. (2009) FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. *Hum Mol Genet* 18: 4879–4896.
- Jugessur A, Shi M, Gjessing HK, Lie RT, Wilcox AJ, et al. (2009) Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia. *PLoS One* 4: e5385. doi:10.1371/journal.pone.0005385.
- Carter TC, Molloy AM, Pangilinan F, Troendle JF, Kirke PN, et al. (2010) Testing reported associations of genetic risk factors for oral clefts in a large Irish study population. *Birth Defects Res A Clin Mol Teratol* 88: 84–93.
- Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, et al. (2009) Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet* 41: 473–477.
- Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, et al. (2009) A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *J Pediatr* 155: 909–913.
- Mangold E, Ludwig KU, Birnbaum S, Baluado C, Ferriani M, et al. (2010) Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nat Genet* 42: 24–26.
- Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, et al. (2010) A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet* 42: 525–529.
- Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, et al. (2012) Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nat Genet* 44: 968–971.
- Hwang SJ, Beaty TH, Panny SR, Street NA, Joseph JM, et al. (1995) Association study of transforming growth factor alpha (TGF alpha) TaqI polymorphism and oral clefts: indication of gene-environment interaction in a population-based sample of infants with birth defects. *Am J Epidemiol* 141: 629–636.
- Maestri NE, Beaty TH, Hetmanski J, Smith EA, McIntosh I, et al. (1997) Application of transmission disequilibrium tests to nonsyndromic oral clefts: including candidate genes and environmental exposures in the models. *Am J Med Genet* 73: 337–344.
- Shi M, Christensen K, Weinberg CR, Romitti P, Bathum L, et al. (2007) Orofacial cleft risk is increased with maternal smoking and specific detoxification-gene variants. *Am J Hum Genet* 80: 76–90.
- Chevrier C, Bahuau M, Perret C, Iovannisci DM, Nelva A, et al. (2008) Genetic susceptibilities in the association between maternal exposure to tobacco smoke and the risk of nonsyndromic oral cleft. *Am J Med Genet A* 146A(18): 2396–406.
- van den Boogaard MJ, de Costa D, Krapels IP, Liu F, van Duijn C, et al. (2008) The MSX1 allele 4 homozygous child exposed to smoking at periconception is most sensitive in developing nonsyndromic orofacial clefts. *Hum Genet* 124(5): 525–34.
- Jiyan L, Zeqiang G, Yongjuan C, Kaihong D, Bing D, et al. (2010) Analysis of interactions between genetic variants of BMP4 and environmental factors with nonsyndromic cleft lip with or without cleft palate susceptibility. *Int J Oral Maxillofac Surg* 39(1): 50–6.
- Li L, Zhu GQ, Meng T, Shi JY, Wu J, et al. (2011) Biological and epidemiological evidence of interaction of infant genotypes at rs7205289 and maternal passive smoking in cleft palate. *Am J Med Genet A* 155A(12): 2940–8.
- Wu T, Fallin MD, Shi M, Ruczinski I, Liang KY, et al. (2012) Evidence of gene-environment interaction for the RUNX2 gene and environmental tobacco smoke in controlling the risk of cleft lip with/without cleft palate. *Birth Defects Res A Clin Mol Teratol* 94(2): 76–83.
- Jia ZL, Li Y, Chen CH, Li S, Wang Y, et al. (2010) Association among polymorphisms at MYH9, environmental factors, and nonsyndromic orofacial clefts in western China. *DNA Cell Biol* 29(1): 25–32.
- Zeiger JS, Beaty TH, Liang KY (2005) Oral clefts, maternal smoking, and TGFA: a meta-analysis of gene-environment interaction. *Cleft Palate Craniofac J* 42: 58–63.
- Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, et al. (2011) Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. *Genet Epidemiol* 35: 469–478.
- Barret JC, Fry B, Maller J, Daly MJ (2005) Haploview analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265.
- Schwender H, Taub MA, Beaty TH, Marazita ML, Ruczinski I (2012) Rapid testing of SNPs and gene-environment interactions in case-parent trio data based on exact analytic parameter estimation. *Biometrics* 68: 766–773.
- Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52: 506–516.
- Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ et al. (2010) Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* 34: 591–602.
- Howie B, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 5: e1000529. doi: 10.1371/journal.pgen.1000529.
- Little J, Cardy A, Munger RG (2004) Tobacco smoking and oral clefts: a meta-analysis. *Bull World Health Organ* 82: 213–218.
- Honein MA, Rasmussen SA, Reefhuis J, Romitti PA, Lammer EJ, et al. (2007) Maternal smoking and environmental tobacco smoke exposure and the risk of orofacial clefts. *Epidemiology* 18: 226–233.
- Romitti PA, Lidral AC, Munger RG, Daack-Hirsch S, Burns TL, et al. (1999) Candidate genes for nonsyndromic cleft lip and palate and maternal cigarette smoking and alcohol consumption: evaluation of genotype-environment interactions from a population-based case-control study of orofacial clefts. *Teratology* 59: 39–50.
- Beaty TH, Hetmanski JB, Zeiger JS, Fan YT, Liang KY, et al. (2002) Testing candidate genes for non-syndromic oral clefts using a case-parent trio design. *Genet Epidemiol* 22: 1–11.
- Ingersoll RG, Hetmanski J, Park JW, Fallin MD, McIntosh I, et al. (2010) Association between genes on chromosome 4p16 and non-syndromic oral clefts in four populations. *Eur J Hum Genet* 18: 726–732.
- Wu T, Liang KY, Hetmanski JB, Ruczinski I, Fallin MD, et al. (2010) Evidence of gene-environment interaction for the IRF6 gene and maternal multivitamin supplementation in controlling the risk of cleft lip with/without cleft palate. *Hum Genet* 128: 401–410.
- Kato A, Kurita S, Hayashi A, Kaji N, Ohashi K, et al. (2008) Critical roles of actin-interacting protein 1 in cytokinesis and chemotactic migration of mammalian cells. *Biochem J* 414: 261–270.
- Gearing M, Juncos JL, Procaccio V, Gutekunst CA, Marino-Rodriguez EM, et al. (2002) Aggregation of actin and cofilin in identical twins with juvenile-onset dystonia. *Ann Neurol* 52: 465–476.
- Döring A, Gieger C, Mehta D, Gohlke H, Prokisch H, et al. (2008) SLC2A9 influences uric acid concentrations with pronounced sex-specific effects. *Nat Genet* 40: 430–436.
- Vitart V, Rudan I, Hayward C, Gray NK, Floyd J, et al. (2008) SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. *Nat Genet* 40: 437–442.
- McArdle PF, Parsa A, Chang YP, Weir MR, O'Connell JR, et al. (2008) Association of a common nonsynonymous variant in GLUT9 with serum uric acid levels in old order amish. *Arthritis Rheum* 58: 2874–2881.
- Dehghan A, Köttgen A, Yang Q, Hwang SJ, Kao WL, et al. (2008) Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *Lancet* 372: 1953–1961.
- Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, et al. (2009) Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet* 5: e1000504. doi: 10.1371/journal.pgen.1000504.
- Yang Q, Köttgen A, Dehghan A, Smith AV, Glazer NL, et al. (2010) Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circ Cardiovasc Genet* 3: 523–530.
- Charles BA, Shriner D, Doumatey A, Chen G, Zhou J, et al. (2011) A genome-wide association study of serum uric acid in African Americans. *BMC Med Genomics* 4: 17. Available: <http://www.biomedcentral.com/1755-8794/4/17>.
- Li C, Chu N, Wang B, Wang J, Luan J, et al. (2012) Polymorphisms in the presumptive promoter region of the SLC2A9 gene are associated with gout in a Chinese male population. *PLoS One* 7: e24561. doi:10.1371/journal.pone.0024561.
- Weinberg CR. (1999) Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am J Hum Genet* 65: 229–235.
- Gu D, Wu X, Reynolds K, Duan X, Xin X, et al. (2004) Cigarette smoking and exposure to environmental tobacco smoke in China: the international collaborative study of cardiovascular disease in Asia. *Am J Public Health* 94: 1972–1976.
- Cordell HJ, Barratt BJ, Clayton DG (2004) Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genetic Epidemiology* 26: 167–185.
- Starr JR, Hsu L, Schwartz SM (2005) Assessing maternal genetic associations: a comparison of the log-linear approach to case-parent triad data and a case-control approach. *Epidemiology* 16: 294–303.