Localization of a major susceptibility locus influencing preterm birth


1Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX 78245-0549, USA 2Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA 3Department of Pediatrics, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA 4Department of Obstetrics and Gynecology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA 5Centre for Genetic Epidemiology and Biostatistics, The University of Western Australia, Perth, WA, Australia 6School of Public Health, University of Texas Health Science Center at Houston, San Antonio, TX, USA

*Correspondence address. Department of Genetics, Texas Biomedical Research Institute, PO Box 760549, San Antonio, TX 78245-0549, USA. Tel: +1-210-258-9452; Fax: +1-210-258-9132; E-mail: geethac@txbiomedgenetics.org

Submitted on December 14, 2012; resubmitted on April 29, 2013; accepted on May 4, 2013

ABSTRACT: Preterm birth (PTB) is a complex trait, but little is known regarding its major genetic determinants. The objective of this study is to localize genes that influence susceptibility to PTB in Mexican Americans (MAs), a minority population in the USA, using predominantly microfilmed birth certificate-based data obtained from the San Antonio Family Birth Weight Study. Only 1302 singleton births from 288 families with information on PTB and significant covariates were considered for genetic analysis. PTB is defined as a childbirth that occurs at <37 completed weeks of gestation, and the prevalence of PTB in this sample was 6.4%. An ~10 cM genetic map was used to conduct a genome-wide linkage analysis using the program SOLAR. The heritability of PTB was high (h² = 0.75 ± 0.20) and significant (P = 4.5 × 10⁻⁵), after adjusting for the significant effects of birthweight and birth order. We found significant evidence for linkage of PTB (LOD = 3.6; nominal P = 2.3 × 10⁻³⁵; empirical P = 1.0 × 10⁻⁵) on chromosome 18q between markers D18S1364 and D18S541. Several other chromosomal regions (2q, 9p, 16q and 20q) were also potentially linked with PTB. A strong positional candidate gene in the 18q linked region is SERPINB2, a member of the plasminogen activator system that is associated with various reproductive processes. In conclusion, to our knowledge, perhaps for the first time in MAs or US populations, we have localized a major susceptibility locus for PTB on chromosome 18q21.33-q23.

Key words: preterm birth / family studies / susceptibility locus / variance-components linkage analysis / Mexican Americans

Introduction

Preterm birth (PTB) or prematurity, defined as a childbirth that occurs at <37 completed weeks of gestation, is a major global health issue, and the leading cause of neonatal morbidity and mortality around the world (Goldenberg et al., 2008; Beck et al., 2010; Lawn et al., 2010). Despite the advancing knowledge on etiology of PTB and public health efforts to prevent it, globally, 15 million babies are born preterm each year; 1.1 million babies die due to PTB complications; and 60% of PTBs occur in 10 countries including the USA (MOD et al., 2012). Although the burden of PTB has declined slightly in the USA since 2007, its occurrence in 2010 was reported to be ~12.0% of births or about half a million babies in the USA (Martin et al., 2011, 2012). There are remarkable ethnic disparities in its prevalence in the US populations: 10.8% in non-Hispanic European Americans, 17.1% in non-Hispanic African Americans and 11.8% in Hispanics (Martin et al., 2012). In Texas, the highest PTB rate of 17.4% was found in non-Hispanic African Americans, followed by Hispanics [13.3%] and non-Hispanic European Americans [11.6%] (Martin et al., 2012). Such epidemiological patterns of PTB are associated with substantial socioeconomic costs (Petrou and Khan, 2012). For example, the annual societal economic burden associated with PTB in the USA was at least $26.2 billion in 2005 (Institute of Medicine, 2007).

The etiology of PTB is multifactorial; and it is influenced by genetic and environmental factors and their interactions. Numerous maternal and/or fetal characteristics are associated with PTB (Tucker and McGuire, 2004; Institute of Medicine, 2007; Goldenberg et al., 2008; Menon, 2008; Carter et al., 2011; Arbour et al., 2012; Petrou and Khan, 2012; Ruiz et al., 2012a). Some such PTB risk factors include: previous PTB, multiple pregnancies, uterine/cervical problems (e.g. uterine infections),...
high blood pressure, diabetes, pre-eclampsia, adverse behaviors (e.g. maternal smoking), ethnicity, socioeconomic factors, mother’s age, maternal malnutrition, mother’s psychological/social stress, preconceptional health status of childbearing-aged women and problems associated with the fetus such as infections, developmental problems and birth defects. Some common complications found in premature babies at birth are: low birthweight, underdeveloped organs, respiratory distress syndrome, apnea, intraventricular hemorrhage, patent ductus arteriosus, retinopathy of prematurity, jaundice, anemia and infections. Above all, the biggest concern is long-term outcomes such as mental retardation, autism symptoms including social, learning, behavioral and speech problems, cerebral palsy, chronic lung disease, vision and hearing loss (Barker, 2002; Aylward, 2003; Ward et al., 2004; Hovi et al., 2007; Limperopoulos et al., 2008; Schendel and Bhasin, 2008; MOD et al., 2012).

Several lines of evidence support a genetic basis of PTB including its familial aggregation, population differences in its prevalence and genetic association studies of PTB with variants in candidate genes (Hao et al., 2012). To our knowledge, so far, one PTB linkage study has been performed, using data from a Finnish population (Haataja et al., 2011). This study found evidence for a major locus on chromosome 15q26.3 influencing PTB; subsequently, variants in a positional candidate gene IGF1R were shown to be associated with PTB. Recently, using an evolutionary genomic approach, Plunkett et al. (2011) tested several accelerated genes and found evidence for significant association of variants in FSHR with PTB.

Despite such advances in our understanding of the etiology of prematurity, knowledge on specific, major genetic determinants of PTB is extremely limited. Therefore, we performed a genome-wide linkage screen to localize PTB susceptibility loci in the Mexican American population, the major subgroup of the largest and the fastest growing Hispanic minority population in the USA, which is associated with a high fertility rate and a high occurrence of PTB (Martin et al., 2012; Ruiz et al., 2012a). To carry out this study, we used data collected as part of the San Antonio Family Birth Weight Study (SAFBWS), a complex-pedigree based study composed of predominantly low-income Mexican American families in San Antonio, Texas.

Materials and Methods

San Antonio Family Birth Weight Study

A major objective of SAFBWS is to investigate genetic and environmental influences and their interactions on birthweight and to examine its correlation with different metabolic syndrome-related traits in adulthood. The present study focuses on the genetics of PTB, a correlate of birthweight. The SAFBWS individuals were original participants in three ongoing large Mexican American family studies in San Antonio, Texas: the San Antonio Family Heart Study (SAFHS; ~1400 participants from 42 large families), the San Antonio Family Diabetes/Gallbladder Study (SAFDGS; ~900 participants from 40 large families) and the Veterans Administration Genetic Epidemiology Study (VAGES; ~1000 participants from about 300 families, mostly nuclear in structure) (Mitchell et al., 1996; Hunt et al., 2005; Arya et al., 2006; Puppala et al., 2006; Coletta et al., 2009). The phenotypic data for birthweight, PTB and other related information were collected predominantly from microfilmed birth certificates for those individuals who were born in San Antonio in 1949 or thereafter, and the Texas Department of State Health Services. For those individuals who were either born before 1949 in San Antonio or born outside San Antonio, the data were obtained through home visits and other sources including family Bibles, and medical records from hospitals and health departments (Arya et al., 2006). The PTB information was available for 1439 individuals. In addition to birthweight, other available information included sex of the infant, mother’s age at childbirth, birth order and type of birth (singleton versus multiple). For genetic analysis, the phenotype PTB is considered to represent the affected infant, and PTB is defined as a childbirth that occurs at <37 completed weeks of gestation. However, for 62/1439 (4.3%) babies only term or no term information was available, and they were considered for this study.

As described previously, the genotyping for genome-wide microsatellite markers (~10 cM density for SAFDGS and VAGES) was conducted at the Center for Inherited Disease Research (CIDR) and similar activities for SAFHS were performed at Texas Biomedical Research Institute (Arya et al., 2006; Puppala et al., 2006; Coletta et al., 2009). The marker genotyping errors and the pedigree discrepancies were resolved using analytical techniques as implemented in programs Simwalk2 and PREST. SAFBWS has been approved by the Institutional Review Board (IRB) at the University of Texas Health Science Center at San Antonio (UTHSCSA), San Antonio, Texas.

Analytical procedures

Variance components linkage analysis

The variance-component approach uses information from all possible biological relationships simultaneously in an attempt to disentangle the genetic architecture of a quantitative trait. In a simple model, variances or covariances between relatives as a function of the genetic relationships from the pedigrees can be specified, and the proportion of phenotypic variance that is attributed to (additive) genetic effects (i.e. heritability: \( h^2 \)) can be estimated from the components of variance (Hopper and Mathews, 1982; Falconer, 1989; Amos, 1994; Almasy and Blangero, 1998). Although this variance components method is appropriate to quantitative traits, it can be extended to the dichotomous or discrete (disease = yes or no) traits such as PTB, using a liability or threshold model (Duggirala et al., 1997, 1999). It is assumed that an individual belongs to a specific disease category if the liability or underlying genetically determined risk exceeds a certain threshold, \( T \), on a normally distributed liability curve. The liability is assumed to have an underlying multivariate normal distribution with equal unit variances both in the general population and in relatives of affected individuals. The correlation in liability between pairs of individuals is estimated using the affected status of unrelated individuals and various categories of relatives (Duggirala et al., 1997, 1999). Because the calculation of the likelihood for this multifactorial model requires high dimensional integration, we evaluated it approximately, using the Mendel–Elston algorithm (Mendel and Elston, 1974).

The variance components such as heritability attributed to the susceptibility locus \( h^2_g \) and heritability attributed to the residual additive genetic effects \( h^2 \), individual-specific random environmental factors \( e^2 = 1 - [h^2 + h^2_g + h^2_r] \) and covariate effects (e.g. birthweight) for PTB were estimated in likelihood terms, and hypothesis tests were performed using likelihood ratio tests (Self and Liang, 1987; Duggirala et al., 1999). The hypothesis of no linkage (i.e. \( h^2_g = 0 \)) was tested by comparing the likelihood of this restricted model with that of a model in which \( h^2_g \) was estimated. Twice the difference in the \( \ln \) likelihoods of these models yields a test statistic that is asymptotically distributed as \( \chi^2 \) mixture of \( \chi^2 \) and a point mass at 0, denoted by \( \chi^2_0 \), where the degree of freedom is equal to the difference in the number of parameters estimated between the two competing models. To obtain LOD scores, the \( \ln \) likelihood values were converted into values of \( \log_{10} \). As strong evidence in
support of linkage, genetic locations across the genome with a LOD score ≥ 3.0 were considered (Lander and Kruglyak, 1995); and for discussion purposes, other regions with a LOD score ≥ 1.2 (i.e. nominal $P \leq 0.01$ or LOD score ≥ 1.18) were considered as evidence for potential linkage.

The Marshfield genetic maps (map distance in Kosambi cM) were used; and the multipoint identical-by-descent (IBD) matrices given a number of genetic markers (map distance in Haldane cM) were calculated using Markov chain Monte Carlo methods implemented in the program Loki, as described previously (Puppala et al., 2006). However, for the purpose of discussion, the locations of our linkage findings have been placed on the Marshfield genetic map (i.e. Kosambi cM). To verify our PTB linkage we performed simulation analysis to determine the empirical $P$-value, using information obtained from 100,000 replicates. The analytical procedures described above are incorporated in the program SOLAR (Almasy and Blangero, 1998).

**Results**

The characteristics of the study participants are reported in Table I. The PTB data combined from three studies were available for 1439 individuals (629 from SAFHS, 448 from SAFDGS and 362 from VAGES). The prevalence of PTB in the combined sample was 6.7% (SAFHS = 6.5%; SAFDGS = 8.3%; VAGES = 5.2%). Approximately 60% of births examined were females, and the mean mother’s age at childbirth was 26.6 years. The mean birthweight was 3.3 kg and the prevalence rates of low birthweight (< 2.5 kg) and macrosomia (birthweight > 4.0 kg) were 5.9 and 7.8%, respectively. The singleton births were the majority (97.7%), and the mean birth order was 4.1. The average adult age at clinic examination of the study participants was 31.3 years.

For the genetic analysis, however, we considered only singleton births. Prior to performing genetic analysis, we evaluated whether the variables—study, sex, birthweight, mother’s age at childbirth and birth order—were significant ($P < 0.05$) covariates of PTB, in turn finding birthweight ($P = 9.9 \times 10^{-13}$) and birth order ($P = 0.0151$) as significant covariates. The data for PTB, birthweight and birth order were available for 1302 study participants, and for 59/1302 babies only term or no term information was available. The prevalence of PTB in this sample was 6.4% (SAFHS = 6.1%; SAFDGS = 7.7%; VAGES = 5.4%). All genetic analyses of PTB included birthweight and birth order as covariates. These individuals were distributed across 288 families; however, 125 were found to be represented by single individuals with PTB data. These 125 single, unrelated individuals were considered for the analysis because they contribute to the evaluation of covariate effects. As shown in Table II, the remaining 1177 participants from 163 families generated 7525 relative pairs: 40 PTB-PTB pairs; 931 PTB-term birth (TB) pairs and 6554 TB-TB pairs. The heritability of PTB was determined to be high ($h^2 ± SE: 0.75 ± 0.20$) and statistically significant ($P = 4.5 \times 10^{-5}$), after accounting for the significant covariate effects of birthweight and birth order. The covariates explained ~9% of total phenotypic variation in PTB.

As reported in Table III and Figs 1 and 2, after accounting for the significant covariate influences, the strongest evidence for linkage of PTB (LOD = 3.6; nominal $P = 2.3 \times 10^{-5}$; empirical $P = 1.0 \times 10^{-5}$) occurred at a genetic location on chromosome 18q between markers D18S1364 and D18S541 (18q22.1-q22.3, 99–107 cM). The 1 LOD-support interval surrounding the peak extends between markers D18S1270 (18q21.33, 96 cM) and D18S1371 (18q23, 116 cM). The second strongest evidence for PTB linkage (LOD = 2.7; nominal $P = 2.1 \times 10^{-4}$; empirical $P = 2.0 \times 10^{-4}$) was found on chromosome 16q between markers D16S242 and D16S539 (16q23.3-q24.1, 111–125 cM). The 1 LOD-support interval around the linkage peak extends between markers D16S3096 (16q23.1, 99 cM) and D16S2621 (16q24.2, 130 cM). In addition, as shown in Table III and Fig. 1, the genome-wide linkage screen identified three genetic locations on chromosomes 2q, 9p and 20q, respectively, to be potentially linked (LOD ≥ 1.2) with PTB.

We performed an additional analysis. We assessed genetic correlation, a measure of the shared genetic basis of the two phenotypes (i.e. pleiotropy), between PTB and birthweight and found a negative correlation (i.e. $-0.38$) between them, as expected. Since this study aimed to localize PTB susceptibility loci, all our genetic analyses included birthweight as a covariate. We reanalyzed the data without birthweight as a covariate, in turn finding a remarkable change in the LOD score value (from 3.6 to 2.2) on chromosome 18q. However, the LOD score attenuated minimally on chromosome 16q, where it decreased from 2.7 to 2.2. These linkage patterns suggest that the PTB susceptibility locus on chromosome 18q could be relevant to risk associated with birthweight also.

**Discussion**

PTB is a major public health issue, which has a multifactorial origin. The available evidence suggests complex genetic involvement in PTB (Porter et al., 1998).
et al., 1997; Treloar et al., 2000; Dolan, 2010). Although a number of susceptibility loci for complex traits/diseases have been localized using genome-wide linkage and association analytical approaches, knowledge on specific, major genetic factors that underlie variation in PTB is extremely limited. Epidemiological studies suggest that premature births cluster in families, and the available evidence from two twin studies suggests that PTB heritability estimates range from 17 to 36% (Clausson et al., 2000; Treloar et al., 2000; Ward et al., 2005; Dolan et al., 2010). Using an extended twin design approach, heritability of birth timing in women was estimated to be 34% (Kistka et al., 2008). Given that heritability estimates are population sample specific, our data reveal that PTB is highly heritable (75%) in the Mexican American population. Most importantly, perhaps for the first time in the US populations, we found significant evidence for a novel, major locus on chromosome 18q.

### Table II  Numbers and types of relative pairs by PTB status used for this study.

<table>
<thead>
<tr>
<th>Type of relative pair</th>
<th>Relationship coefficient</th>
<th>Relative pairs by PTB status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PTB-PTB</td>
</tr>
<tr>
<td>Parent-offspring</td>
<td>0.5000</td>
<td>3</td>
</tr>
<tr>
<td>Sibs</td>
<td>0.5000</td>
<td>10</td>
</tr>
<tr>
<td>Avuncular</td>
<td>0.2500</td>
<td>7</td>
</tr>
<tr>
<td>Half-sibs</td>
<td>0.2500</td>
<td>–</td>
</tr>
<tr>
<td>1st/2nd cousins</td>
<td>0.1562</td>
<td>–</td>
</tr>
<tr>
<td>Grand avuncular</td>
<td>0.1250</td>
<td>2</td>
</tr>
<tr>
<td>Half avuncular</td>
<td>0.1250</td>
<td>12</td>
</tr>
<tr>
<td>1st cousins</td>
<td>0.1250</td>
<td>12</td>
</tr>
<tr>
<td>Half grand avuncular</td>
<td>0.0625</td>
<td>–</td>
</tr>
<tr>
<td>1st cousins, once removed</td>
<td>0.0625</td>
<td>–</td>
</tr>
<tr>
<td>Half 1st cousins</td>
<td>0.0625</td>
<td>1</td>
</tr>
<tr>
<td>1st cousins, twice removed</td>
<td>0.0312</td>
<td>–</td>
</tr>
<tr>
<td>Half 1st cousins, once removed</td>
<td>0.03125</td>
<td>–</td>
</tr>
<tr>
<td>2nd cousins</td>
<td>0.0312</td>
<td>3</td>
</tr>
<tr>
<td>2nd cousins, 1 removed</td>
<td>0.0156</td>
<td>2</td>
</tr>
<tr>
<td>Half 2nd cousins</td>
<td>0.0156</td>
<td>–</td>
</tr>
<tr>
<td>3rd cousins</td>
<td>0.0078</td>
<td>–</td>
</tr>
<tr>
<td>Double 1st cousins</td>
<td>0.2500</td>
<td>–</td>
</tr>
<tr>
<td>Double 1st cousins, 1 removed</td>
<td>0.1250</td>
<td>–</td>
</tr>
<tr>
<td>Double 2nd cousins</td>
<td>0.0625</td>
<td>–</td>
</tr>
<tr>
<td>Half 1st cousins and cousins</td>
<td>0.09375</td>
<td>–</td>
</tr>
<tr>
<td>1st cousins, 1 removed and 2nd cousins, 1 removed</td>
<td>0.0781</td>
<td>–</td>
</tr>
<tr>
<td>2nd cousins and 3rd cousins</td>
<td>0.0391</td>
<td>–</td>
</tr>
<tr>
<td>1st cousins and Half 1st cousins</td>
<td>0.1875</td>
<td>–</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>931</td>
</tr>
</tbody>
</table>

a The relationship coefficient refers to twice the coefficient of kinship of two individuals.
PTB, preterm birth; TB, term birth.

### Table III  Chromosomal regions potentially linked to PTB with multipoint LOD scores ≥ 1.2.

<table>
<thead>
<tr>
<th>Marker region</th>
<th>Distance from p-ter (cM)a</th>
<th>Chromosomal location</th>
<th>Maximum LOD scoreb</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2S1363</td>
<td>227</td>
<td>2q36.3</td>
<td>1.6</td>
</tr>
<tr>
<td>D9S1779</td>
<td>0</td>
<td>9p24.3</td>
<td>1.3</td>
</tr>
<tr>
<td>D16S422-D16S339</td>
<td>111–125</td>
<td>16q23.3-q24.1</td>
<td>2.7</td>
</tr>
<tr>
<td>D18S1364-D18S541</td>
<td>99–107</td>
<td>18q22.1-q22.3</td>
<td>3.6</td>
</tr>
<tr>
<td>D20S480-D20S1085</td>
<td>80–82</td>
<td>20q13.2–q13.2</td>
<td>1.4</td>
</tr>
</tbody>
</table>

a Marshfield map (Kosambi cM).
bPTB adjusted for the significant covariate effects of birthweight and birth order; please see text for details.

Chittoor et al. 690
Chittoor et al. 690
influencing susceptibility to PTB in Mexican Americans, a population at high risk for PTB. In addition, our study also provided strong evidence for a novel locus on chromosome 16q suggestively linked to PTB.

Numerous studies have examined associations between PTB and genetic variants in candidate genes involved in a variety of pathways related to prematurity, but replications of such findings are rather limited. Some of the associated genes seem to be involved in response to infection and inflammation such as TNF, IL1RN, IL, IL6R and TIMP; genes involved in the rupture of amniotic membranes: MMP-1, MMP-8 and MMP-9, progesterone receptor gene, genes involved in the cholesterol metabolism and vitamin C transporter gene (Ferrand et al., 2002; Romero et al., 2002, 2010a, b; Macones et al., 2004; Wang et al., 2004; Engel et al., 2005a, b; Erichsen et al., 2006; Menon et al., 2006; Ehn et al., 2007; Steffen et al., 2007; Chaves et al., 2008; Holtegaard et al., 2008; Velez et al., 2008, 2009; Dolan et al., 2010; Ruiz et al., 2012b).

In contrast, to date, only a single study has performed a genome-wide linkage screen for spontaneous preterm birth (SPTB) and subsequent association analysis to identify the novel PTB susceptibility gene IGF1R on chromosome 15q26.3, using data from seven large northern Finnish families with recurrent SPTB for linkage analysis and a follow-up Finnish population-based association analysis (Haataja et al., 2011). In addition, several other genetic regions on chromosomes 2, 4, 8, 10, 12 and 13 were reported to harbor possible candidate genes for SPTB. However, none of the chromosomal regions with relevance to SPTB in the sub-isolated Finnish population including the major locus on chromosome 15q26.3 appears to be related to the linkage findings in our study, emphasizing the potential population genetic differences.

The 1 LOD-support interval surrounding the linkage signal on chromosome 18q between markers D18S1270 (18q21.33, 96 cM, 61.4 Mb) and D18S1371 (18q23, 116 cM, 73.1 Mb) encompasses an ~20 cM or 11.7 Mb chromosomal region. Given that the observed linkage signal may correspond to potential PTB functional variants representing one or more genes, a search of the current genome databases revealed about 52 genes being mapped to this critical region. An important positional candidate gene for PTB in this region is serpin peptidase inhibitor, clade B (ovalbumin), member 2 (SERPINB2 [MIM 173390]; also called plasminogen activator inhibitor 2 [PAI-2]; 18q21.3; 61.6 Mb), a member of the plasminogen activator (PA) system, including plasminogen activators (i.e. tissue type [tPA] and urokinase type [uPA]) and their SERPIN inhibitors (i.e. SERPINAS/PCI, SERPINB2/PAI-2, SERPINE1/PAI-1 and SERPINE2/PN-1) (Lee et al., 2011). The PA
system is associated with various reproductive processes including embryogenesis and implantation, and pregnancy-related issues such as placental development and functioning, hemostasis and labor-associated rupture of fetal membranes (Tsatas et al., 1998; Bogic et al., 1999; Coolman et al., 2006, 2012; Lee et al., 2011). Using data from Australian children with later-diagnosed cerebral palsy, some evidence for association between certain variants in PAI-2 gene and PTB was reported (Gibson et al., 2007). These data together with our findings of PTB linkage appear to highlight SERPNB2 as a strong positional candidate for PTB risk on chromosome 18q. Interestingly, the potential PTB linkage found at marker D2S1363 (2q36.3, 227.0 Mb) in our data is proximal to the gene SERPNB2/PN-1, another member of the PA system, on chromosome 2q36.1 (224.8 Mb; MIM 177010).

Some other positional candidate genes in the 18q chromosomal region with potential relevance to PTB include the following: CD226 is a member of the immunoglobulin (Ig) superfamily, and the variants in CD226 antigen (CD226) gene (Wieczorek et al., 2009; Qiu et al., 2012; Song et al., 2011; the cadherin 7 gene (CDH7) (Prasad and Paulson, 2011; Redies et al., 2012); the rotatin gene (RTTN) (Kheradmand Kia et al., 2012); and, the precerebellin 2 (CBLN2) (Matsuda and Yuzuki, 2011; Clarke et al., 2012).

The 1 LOD-support interval around the linkage peak on chromosome 16q extends between markers D16S3096 (16q23.1, 99 cM, 79.0 Mb) and D16S2621 (16q24.2, 130 cM, 88.0 Mb), spanning an ~31 cM or 8.9 Mb long chromosomal region that contains 86 genes. In comparison with the critical 18q chromosomal region, the 16q 1-LOD support interval is gene rich. A potential positional candidate gene for PTB is β-carotene 15, 15-prime-monooxygenase 1 (BCMO1, 16q23.2, 81.3 Mb, MIM 605748). BCMO1 is a key enzyme for β-carotene conversion into retinal, the active form of vitamin A, and plays a critical role in embryonic retinoid metabolism (McConaha et al., 2011; Lietz et al., 2012). As reviewed by Gutierrez-Mazariegos et al. (2011), maternal vitamin A status is critical for implantation and subsequent normal development of the fetus and neonate, and abnormal levels of vitamin A during gestation could lead to teratogenesis. Some other potential candidate genes for PTB on 16q chromosomal region include the following: 17 β-hydroxysteroid dehydrogenase 2 (HSD17B2) (Andersson et al., 2008; Rantakari et al., 2008; Gao et al., 2012); dehydrated hereditary stomatocytosis (DHS) (Grootenboer et al., 2000; Basu et al., 2003; Steiner and Gallagher, 2007); cysteine-rich secretory protein, icl domain-containing 2 (CRISPLD2) (Chiquet et al., 2011); phospholipase Cγ2 (PLCG2) (Zhou et al., 2012) and, forkhead box f1 (FOXF1) (Stankiewicz et al., 2009). It is worth noting that there is evidence based on linkage and association studies for the existence of loci in the 16q23-q24 chromosomal region (containing such genes as CRISPLD2, FOXF1 and CDH13) that influence variation in human height (Gudbjartsson et al., 2008; Axenovich et al., 2009). As described by Yang et al. (2012), the blood pressure or hypertension susceptibility loci have also been reported in this chromosomal region near CDH13 and WWOX genes.

Some limitations of our study are described as follows. Given the multifactorial basis of PTB’s phenotypic expression, several studies reported a major role for the maternal factors (Boyd et al., 2009; Svensson et al., 2009), while there is also some evidence for the fetal genome to be informative for genetic studies compared with the maternal genome regarding susceptibility to PTB (York et al., 2010; Haataja et al., 2011). Given the limitations of the data availability to consider mothers as affected individuals with PTB in our study for a genetic analysis, the reported linkage findings correspond to the newborn PTB affection status. Indeed, as remarked by Dolan and Christiaens (2013), the study of pregnancy adds additional complexity, wherein two phenotypes (i.e. mother and newborn) and three genomes (i.e. maternal, paternal, and child’s) are to be considered. Since the PTB and other related information were obtained predominantly from microfilmed birth certificates for study participants born in or after 1949, the use of completed weeks of gestation as reported on birth certificates potentially based on last menstrual period (LMP) to define PTB is another limitation of our study, although it appears to be the best assignment procedure at that time. Also, the available recorded data lack information on other correlates of PTB such as mother’s behavioral attributes (e.g. mother’s smoking or dietary profile) and disease conditions (e.g. diabetes and hypertension) at the time of birth. However, the available PTB, birthweight and genotypic data from well-established large family studies provided a unique opportunity to localize PTB susceptibility loci in the Mexican American population, which is currently associated with high occurrence of PTBs. In consideration of the issue of the Hispanic/Acculturation Paradox (Ruiz et al., 2008), it should be noted that the reported PTB prevalence of 6.4% appears to be reflective of its appropriateness to the period of data collection before 1980, which spanned ~40 years (~1940–1980).

In conclusion, perhaps for the first time in a Mexican American population, we have localized a major susceptibility locus for PTB on chromosome 18q21.33-q22. In addition, strong evidence for linkage of PTB was also found on chromosome 16q23.1-q24.2, and potential evidence for linkage occurred on chromosomes 2q, 9p and 20q. Generalization of our main observations on chromosomes 18q and 16q would require validation of these linkage findings in ethnically diverse populations. A search of the existing genomic databases has revealed several potential positional candidate genes for PTB in these twochromosomal regions. Our plans are to conduct fine mapping in these linked chromosomal regions, and subsequently to perform high throughput sequencing of the same chromosomal regions to identify the potential functional variants responsible for our original PTB linkage signals on chromosomes 18q and 16q. In addition, following the preliminary evidence for potential common genetic or pleiotropic influences on PTB and birthweight in our data, we plan to conduct a genome-wide bivariate linkage analysis to detect major loci that simultaneously influence both PTB and birthweight. Lastly, the current findings may ultimately contribute to the development of effective strategies for the diagnosis, prevention and treatment of PTB.

Acknowledgements

We warmly thank the participants of SAFBWS (i.e. SAFHS, SAFDGS and VAGES) for their enthusiasm and co-operation. We sincerely thank Dr Richard A. Granato, Dr Mary Bartley, Mr Harold V. Gaskill IV and Mr Ben Barloco for their help regarding data collection and verification. We thank the Texas Department of Health for providing birthweight and its related information for some individuals. We also thank the Center for Inherited Disease Research (CIDR) for providing new genome scans using SAFDGS and VAGES data. None of the authors have any conflicts of interest with respect to the study.

Authors’ roles

G.C. was responsible for writing the paper and analyzed and interpreted the data. V.S.F., S.P., R.A., T.D.D. and J.S. helped with the statistical

Funding
This study was supported by grants from the National Institute of Health [HD041111, DK053889, DK042273, P01 HL05522, DK047482, MH059490 and M01-RR-01346]. In addition, this work was supported by a Veterans Administration Epidemiologic grant. The AT&T Genomics Computing Center supercomputing facilities used for this work were supported in part by a gift from the AT&T Foundation and with support from the National Center for Research Resources Grant Number S10 RR029392. This investigation was conducted in facilities constructed with support from Research Facilities Improvement Program grants C06 RR13556 and C06 RR017515 from the National Center for Research Resources of the National Institutes of Health.

Conflict of interest
None declared.

Web resources
Accession numbers and URLs for data presented in the manuscript are as follows:
Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics (for genetic map information).
Online Mendelian Inheritance in Man (OMIM) http://www.ncbi.nlm.nih.gov/Omim (for the positional candidate genes mentioned in the manuscript).

References


