A quantitative trait locus (QTL) on chromosome 6q influences birth weight in two independent family studies

Rector Arya1,4,*, Ellen Demerath5, Christopher P. Jenkinson2, Harald H.H. Göring4, Sobha Puppala4, Vidya Farook4, Sharon Fowler1, Jennifer Schneider4, Richard Granato1, Roy G. Resendez4, Thomas D. Dyer4, Shelley A. Cole4, Laura Almasy4, Anthony G. Comuzzie4, Roger M. Siervogel5, Benjamin Bradshaw3, Ralph A. DeFronzo5, Jean MacCluer4, Michael P. Stern1, Bradford Towne5, John Blangero4 and Ravindranath Duggirala4

1Division of Clinical Epidemiology, 2Division of Diabetes, Department of Medicine and 3School of Public Health, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, TX 78229-3900, USA, 4Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78245, USA and 5The Wright State University School of Medicine, Kettering, OH 45420, USA

Received January 25, 2006; Revised and Accepted March 22, 2006

Low birth weight is an important cause of infant mortality and morbidity worldwide. Birth weight has been shown to be inversely correlated with adult complex diseases such as obesity, type-2 diabetes and cardiovascular disease. However, little is known about the genetic factors influencing variation in birth weight and its association with diseases that occur in later life. We, therefore, have performed a genome-wide search to identify genes that influence birth weight in Mexican-Americans using the data from the San Antonio Family Birth Weight Study participants (n = 840). Heritability of birth weight was estimated as 72.0 ± 8.4% (P < 0.0001) after adjusting for the effects of sex and term. Multipoint linkage analysis yielded the strongest evidence for linkage of birth weight (LOD = 3.7) between the markers D6S1053 and D6S1031 on chromosome 6q. This finding has been replicated (LOD = 2.3) in an independent European-American population. Together, these findings provide substantial evidence (LODadj = 4.3) for a major locus influencing variation in birth weight. This region harbors positional candidate genes such as chorionic gonadotropin, alpha chain; collagen, type XIX, alpha-1; and protein-tyrosine phosphatase, type 4A, 1 that may play a role in fetal growth and development. In addition, potential evidence for linkage (LOD/C21 < 1.2) was found on chromosomes 1q, 2q, 3q, 4q, 9p, 19p and 19q with LODs ranging from 1.3 to 2.7. Thus, we have found strong evidence for a major gene on chromosome 6q that influences variation in birth weight in both Mexican- and European-Americans.

INTRODUCTION

Low birth weight (LBW), or small body size at birth, is one of the major public health problems in the USA, and an important cause of infant mortality and morbidity worldwide (1). The percentage of LBW infants was reported as 6.6 among non-Hispanic white mothers, 6.4 among Mexican-Americans (Hispanic) and 13.0 for African-Americans (2). On the basis of socio-demographic characteristics, Mexican-Americans would be expected to have an elevated risk of adverse health outcomes including high rates of LBW. Paradoxically, however, MAs have rates of LBW similar to those of non-Hispanic whites, and lower than those of African-Americans of similar socioeconomic status. Given the evidence that birth weight is positively associated with weight at later ages, LBW infants who gain weight very rapidly during childhood (i.e. who experience catch-up growth) may also be at increased risk of obesity later in life (2,3). There is evidence that LBW is strongly associated with obesity and type-2 diabetes (T2D) in childhood, two major public health problems, in

*To whom correspondence should be addressed. Tel: +1 2105675234; Fax: +1 2105676955; Email: arya@uthscsa.edu

© The Author 2006. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact: journals.permissions@oxfordjournals.org
non-European groups such as Mexican-Americans (2–7). LBW has been considered to be an important marker of various complex diseases in adulthood such as obesity, T2D and coronary heart disease (CHD) in adult life in a number of different populations, including western populations (8–15).

The association between birth weight and diseases in later life has been a much debated topic since Barker and colleagues proposed the ‘fetal programming’ hypothesis (8,10,11). According to this hypothesis, LBW or small body size is a repercussion of malnutrition in early pregnancy, and such fetal states might ‘programme’ tendencies to develop metabolic disease conditions in adult life (8,13). There is considerable evidence supporting this idea that fetal nutrition deficiencies during critical periods in early life may have adverse lifetime consequences (16,17). For example, LBW has been shown to predict the subsequent occurrence of insulin resistance and related conditions such as T2D, hypertension, dyslipidaemia and CHD (8). Thus, the term ‘thrifty phenotype’ has been introduced to refer to maternal and/or fetal nutritional deficiencies as primary environmental mechanisms that underlie the inverse correlations between birth weight and metabolic diseases in adulthood (18,19). However, some epidemiological studies have reported associations between birth weight (low and high) and T2D and IGT (20–23). Thus, the relationship between birth weight and the prevalence of T2D has been found to be U-shaped, indicating that both low and high birth weights are correlated with T2D (22). Interestingly, our recent family-based findings of positive genetic correlations between birth weight (low and high) and T2D and IGT (22).

RESULTS

As reported in Table 1, we used birth weight data obtained from birth certificates for 840 participants (449 from SAFDGS and 391 from SAHS for whom genotypic data were available. In the combined data, mean birth weight was 3.31 kg and the percentage of LBW was 5.6. The percentage of premature births was 7.6. Mean adult age at clinic examination was 28.2 years.

As the SAFDGS families were ascertained on a single T2D proband, all of our genetic analyses incorporated correction for possible ascertainment bias by computing the likelihood of a pedigree conditional on the birth weight of the proband. After accounting for the effects sex and term (premature: yes or no), the heritability of birth weight analyzed as a continuous trait, was 71.8 ± 8.4% which is highly significant (P < 0.0001). Following the estimation of heritability, we performed multipoint linkage analysis across 22 autosomes for birth weight with sex and term as covariates. As shown in Table 2 and Figures 1 and 2, we have found significant evidence for linkage (LOD = 3.7, nominal P-value = 0.0000175, empirical P-value = 0.00002) between markers D6S1053 and D6S1031 on chromosome 6. The empirical P-value that corresponds to our original observed LOD score (3.7) indicates that our nominal P-value may not have over-stated the evidence for linkage.

In addition, eight chromosomal regions representing seven chromosomes across the genome, exhibited modest evidence for linkage (LOD ≥ 1.2) with birth weight as shown in Table 2 and Figure 1. Genetic regions on chromosomes 2, 4 and 9 with suggestive evidence for linkage (LOD ≥ 1.9) appear to harbor genes influencing the variation in birth size is a repercussion of malnutrition in early pregnancy, and such fetal states might ‘programme’ tendencies to develop metabolic disease conditions in adult life (8,13). There is considerable evidence supporting this idea that fetal nutrition deficiencies during critical periods in early life may have adverse lifetime consequences (16,17). For example, LBW has been shown to predict the subsequent occurrence of insulin resistance and related conditions such as T2D, hypertension, dyslipidaemia and CHD (8). Thus, the term ‘thrifty phenotype’ has been introduced to refer to maternal and/or fetal nutritional deficiencies as primary environmental mechanisms that underlie the inverse correlations between birth weight and metabolic diseases in adulthood (18,19). However, some epidemiological studies have reported associations between birth weight (low and high) and T2D and IGT (20–23). Thus, the relationship between birth weight and the prevalence of T2D has been found to be U-shaped, indicating that both low and high birth weights are correlated with T2D (22). Interestingly, our recent family-based findings of positive genetic correlations between birth weight (low and high) and T2D and IGT (22).

RESULTS

As reported in Table 1, we used birth weight data obtained from birth certificates for 840 participants (449 from SAFDGS and 391 from SAHS for whom genotypic data were available. In the combined data, mean birth weight was 3.31 kg and the percentage of LBW was 5.6. The percentage of premature births was 7.6. Mean adult age at clinic examination was 28.2 years.

As the SAFDGS families were ascertained on a single T2D proband, all of our genetic analyses incorporated correction for possible ascertainment bias by computing the likelihood of a pedigree conditional on the birth weight of the proband. After accounting for the effects sex and term (premature: yes or no), the heritability of birth weight analyzed as a continuous trait, was 71.8 ± 8.4% which is highly significant (P < 0.0001). Following the estimation of heritability, we performed multipoint linkage analysis across 22 autosomes for birth weight with sex and term as covariates. As shown in Table 2 and Figures 1 and 2, we have found significant evidence for linkage (LOD = 3.7, nominal P-value = 0.0000175, empirical P-value = 0.00002) between markers D6S1053 and D6S1031 on chromosome 6. The empirical P-value that corresponds to our original observed LOD score (3.7) indicates that our nominal P-value may not have over-stated the evidence for linkage.

In addition, eight chromosomal regions representing seven chromosomes across the genome, exhibited modest evidence for linkage (LOD ≥ 1.2) with birth weight as shown in Table 2 and Figure 1. Genetic regions on chromosomes 2, 4 and 9 with suggestive evidence for linkage (LOD ≥ 1.9) appear to harbor genes influencing the variation in birth
weight (Fig. 1, Table 2). Thus, our genome-wide linkage analysis of birth weight implicated multiple regions on chromosomes 1, 2, 4, 6 and 9 with significant and potential/suggestive evidence of linkage with birth weight (Figs 1 and 2).

Importantly, our major linkage finding (LOD = 3.7) on 6q has been replicated in an independent European-American population study, the Fels Longitudinal Study, with a LOD of 2.3 between markers D6S460 and D6S1270 (Fig. 3). As shown in Figure 3, the distance between the study-specific maximum LOD peaks was about 12 cM. As can be seen from the same figure, the combined evidence for linkage (LOD$_{adj}$ = 4.3) occurred at a location between markers D6S1053 and D6S1031. The 1-LOD support interval approximately corresponds to the chromosomal region spanning ~16 cM that is flanked by markers D6S1960 and D6S1270.

As multiple linkage peaks with some evidence for linkage were observed in our genome-wide search, we extended variance components linkage analysis to oligogenic models to incorporate two or more loci simultaneously. As shown in Table 3, in a joint two-locus analysis, the two-locus model yielded a conditional LOD score of 2.9 for the second quantitative trait locus (QTL) ($P = 0.000127$) near D2S1776 marker.
on chromosome 2. The three-locus model yielded a conditional LOD score of 2.2 for the third QTL \( (P = 0.000798) \) near D1S518 marker on chromosome 1. The four-locus model yielded a conditional LOD score of 1.4 for the fourth QTL \( (P = 0.006303) \) between GATA187D09 and D9S925 markers on chromosome 9. Thus, genetic determinants of birth weight involve possibly four loci that are found on four different chromosomes.

**DISCUSSION**

Knowledge about the genetic basis of variation in birth weight and its association with diseases in adulthood is very limited. Although the available evidence is limited and indirect, it suggests a significant role for genetic factors in the phenotypic expression of birth weight. There have been suggestions that the variation in birth weight is multifactorially determined, and that the association between birth weight and adulthood diseases can be explained in terms of genetics and environment and their interactions (33,35,36). Birth weight, being highly heritable, showed significant positive genetic correlations with various components of the metabolic syndrome phenotypes, including fasting insulin (non-diabetics only), BMI, waist circumference and triglycerides, and a negative genetic correlation with HDL levels (24,25). However, little is known about the specific chromosomal regions harboring birth weight susceptibility genes. Also not known is whether the already identified susceptibility loci relating to various metabolic phenotypes in adults in our study and other studies have any common genetic effects on birth weight.

In this study, we utilized the wealth of information from our two family studies for which both phenotypic and genotypic data are currently available to understand the genetic basis of birth weight. We used previously collected marker data to scan the genome for specific genetic determinants of birth weight in 840 individuals using variance components linkage techniques. The significant heritability (72%) of birth weight estimated in this study indicates that variation in birth weight is highly heritable, which is consistent with the Fels heritability estimate \( (h^2 = 80\% ; \) Demerath et al., unpublished data). These heritability estimates are higher than some of those estimates \( (~50\%) \) reported elsewhere in the literature (25,37,38). However, it should be noted that heritability estimates are population-specific, and can be influenced by such factors as study population, design, ascertainment criteria and the covariates considered for the analysis. Such factors could have influenced the differences in the heritability estimates of birth weight across the populations.

To our knowledge, this is the first genome scan linkage analysis using birth weight as a continuous trait, to find loci with appreciable influences on birth weight. In particular, the present genome scan in Mexican-Americans provides strong evidence \( (LOD = 3.7) \) for the presence of a novel
major locus influencing variation in birth weight on chromosome 6q12-q14.1 near markers D6S1053–D6S1031. Importantly, our linkage finding in Mexican-Americans has been replicated in an independent non-Hispanic white population with a LOD of 2.3, given attention to the issues of localization (39,40). In fact, our combined data yielded substantial evidence for linkage (LOD adj = 4.3) of birth weight to the same chromosomal region. Markers D6S1960 and D6S1270 on chromosome 6p12.1 and 6q14.3, respectively, flank the 1-LOD support interval surrounding the genetic location, where we found the strongest linkage evidence from the combined data. This chromosomal region contains important positional candidate genes: protein-tyrosine phosphatase, type 4A, 1 (PTP4A1) also known as phosphatase of regenerating liver 1 (PRL-1, 6q12); collagen, type XIX, alpha-1 (COL19A1, 6q12-q14); and human choric gonadotropin (hCG) or chorionic gonadotropin, alpha chain (CGA, 6q12-q21). The protein product of hCG is a placental hormone produced by the trophoblast that acts like luteinizing hormone, stimulating the secretion of the pregnancy-sustaining steroid progesterone and also acts as a growth factor that facilitates endometrial receptivity (41,42). hCG plays a major role in the development and maintenance of pregnancy and hence it is known as the hormone of pregnancy and development. Col19a1 gene, with a transient embryonic expression, may be involved in muscle differentiation and function (43). Furthermore, alterations to human type XIX collagen may cause mild cartilage disorders such as osteoarthritides, mildly short stature and epiphyseal abnormalities (44,45) and modulate several functions including cell–matrix interactions and cell–cell communications (46). PTP4A1 (or PRL-1) is a unique protein tyrosine phosphatase, which is involved in normal cellular growth control, plays a role in liver regeneration and is positively associated with fetal and neoplastic hepatic growth (47,48).

The 6q12-14 region (D6S1053–D6S1031) has also been reported to be linked to several adult phenotypes such as height (49), blood pressure (50) and fasting plasma glucose (FPG) (51) with LODs ranging from 1.6 to 3.3 (Table 4). Specifically, this 6q region is concordant with a strong linkage signal for systolic blood pressure on chromosome 6 near marker D6S1031 in a large NHLBI family heart study (50). In another genome scan of blood pressure in families with familial combined hyperlipidaemia (FCHL), the same marker region has been implicated with a LOD of 2.5 for diastolic blood pressure (52). Thus, our birth weight linkage finding on chromosome 6q appears to overlap with the previously implicated regions influencing variation in height, hypertension and dyslipidaemia. It is worthwhile to note that, apart from these observations, a number of studies including our own implicated a broad overlapping genetic region on chromosome 6q16.1-q27 to contain susceptibility genes that influence obesity (both childhood and adult) and T2D or their related phenotypes, using the genome scan/linkage approach (53–55). Interestingly, oligogenic analysis yielded two additional linkage regions with improved LOD scores on chromosomes 2 and 1 that are similar to those obtained in the initial linkage analysis.

Among the other chromosomal regions of interest in our study, the 2q24 region showed evidence for linkage with birth weight with the second highest LOD, and contains the activin receptor-like kinase 7 (ALK7) gene, which belongs to the transforming growth factor-β family. ALK7 and its isoforms are expressed in human placenta and their expression is developmentally regulated (56). The 1q chromosomal region has been shown to be linked to HDL-C (57), metabolic syndrome phenotypes (58), obesity (59), glycemic traits (60), dyslipidaemia (61), T2D (62) and FCHL (63). This region harbors candidate genes such as apolipoprotein A2 (APOA2), pyruvate kinase deficiency of erythrocyte (PKLR), IL10 homeobox transcription factor 1 (LMO1) and interleukin-10 (IL-10), which could potentially influence birth weight and adult disorders such as T2D, metabolic syndrome and cardiovascular disease. The 9p22 chromosomal region has also been linked to HDL-C levels in Mexican-Americans (61), Finns and Dutch populations (64), and with adiponectin levels, T2D and age of onset of T2D (65–68). Although serum adiponectin levels are inversely associated with obesity, T2D and measures of insulin resistance, its association with birth weight is unclear. However, in a recent study, Cianfarani et al. (69) indicated that low adiponectin levels in LBW infants might predict the later development of visceral fat and insulin resistance.

The results of this study are consistent with findings from earlier studies that showed direct or indirect evidence that genetic factors may play a significant role in determining variation in birth weight. To our knowledge, there has been only one genome scan linkage study of birth weight in Pima Indians (n = 269): in this study, birth weight was analyzed as a discrete phenotype (i.e. LBW) (70). This study found evidence for linkage on chromosome 11p (LOD = 3.4 at 88 cM) and LBW with maternal imprinting (LOD = 4.1) on chromosome 11p (70). However, we failed to find any evidence for linkage of birth weight at this 11p region in our study. In addition, several genetic variants have been reported to be associated with alterations in birth size using animal models and humans. In animal models, the genes encoding insulin, the insulin-like growth factors (IGF-1 and IGF-II), their receptors and their regulatory proteins have been shown to have profound effects on fetal growth (71–73).

In humans, studies have been performed to examine the association between genetic variants in biologically plausible candidate genes, such as insulin, glucokinase and LBW (74,75). It has been shown that mutations in the glucokinase gene (7p15-p13) of the fetus have been associated with both LBW and subsequent diabetes (74). These mutations could cause impaired insulin secretion, hence their contributions to impaired fetal growth and to subsequent diabetes. In contrast, the class III allele of the variable number of tandem repeat locus in the promoter region of the insulin (INS) gene (11p15.5) has been found to be associated with both increased birth weight (76) and with many components of the metabolic syndrome (77). However, as we failed to identify evidence of linkage at these chromosomal regions in our present study, these regions may not have a major influence on birth weight in our population.

Some of the limitations of our findings include the following. We analyzed birth weight as a continuous trait in this study, and that LBW is a part of the continuum. It should, however, be noted that we did not have information on correlatives of birth weight such as mothers’ behavioral
Table 4. Correspondence between linkage findings of birth weight and adult phenotypes

<table>
<thead>
<tr>
<th>Population</th>
<th>Phenotype</th>
<th>Chromosomal band</th>
<th>Markers</th>
<th>Position (cM)</th>
<th>LOD</th>
<th>Candidate genes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>1q31.1-31.3</td>
<td>D1S518–D1S1660</td>
<td>202.2–212.4</td>
<td>1.71</td>
<td>APOA2, PKLR, LMX1, IL 10</td>
<td>Present study</td>
</tr>
<tr>
<td>Hispanic families</td>
<td>Birth weight</td>
<td>1q23–q31</td>
<td>D1S518</td>
<td>202.2</td>
<td>1.65NL</td>
<td></td>
<td>Langefeld et al. (58)</td>
</tr>
<tr>
<td>Caucasian pedigrees</td>
<td>Glycaemic traits</td>
<td>1q</td>
<td>D1S1678–D1S3462</td>
<td>218.5–247.2</td>
<td>1.8–2.33</td>
<td></td>
<td>Meigs et al. (60)</td>
</tr>
<tr>
<td>The DUK2 R UK population</td>
<td>T2D</td>
<td>1q24-25</td>
<td>D1S218</td>
<td>191.5</td>
<td>1.46</td>
<td></td>
<td>Wiltshire et al. (62)</td>
</tr>
<tr>
<td>Finnish population</td>
<td>FCHL</td>
<td>1q21-2q31</td>
<td>D1S1677</td>
<td>175.6</td>
<td>5.93</td>
<td></td>
<td>Pajukanta et al. (63)</td>
</tr>
<tr>
<td>Whites of Breda Study</td>
<td>Obesity (BMI)</td>
<td>1q31-q42</td>
<td>D1S1681–D1S179</td>
<td>202.2–252.1</td>
<td>1.5</td>
<td></td>
<td>Tilburg et al. 2003</td>
</tr>
<tr>
<td>African population</td>
<td>HDL-C</td>
<td>1q25</td>
<td>D1S1660</td>
<td>212.4</td>
<td>2.48</td>
<td></td>
<td>Adeyemo et al. (57)</td>
</tr>
<tr>
<td>Mexican-American</td>
<td>Lipid factor</td>
<td>1q</td>
<td>D1S451</td>
<td>182.4</td>
<td>1.9</td>
<td></td>
<td>Arya et al. (88)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>2q24.3</td>
<td>D2S1776</td>
<td>173.0</td>
<td>2.66</td>
<td></td>
<td>Present study</td>
</tr>
<tr>
<td>Italian families</td>
<td>BFIC</td>
<td>2q24</td>
<td>D2S139</td>
<td>186.2</td>
<td>1.69</td>
<td>ALK7</td>
<td>Malacarne et al. (89)</td>
</tr>
<tr>
<td>Hong Kong Chinese</td>
<td>Triglycerides</td>
<td>2q</td>
<td>D2S131</td>
<td>224.9</td>
<td>1.56</td>
<td>MUC4, MUC20, PPP1R2</td>
<td>Present study</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>3q29</td>
<td>D3S131</td>
<td>209.4</td>
<td>1.95NL</td>
<td></td>
<td>Ng et al. (51)</td>
</tr>
<tr>
<td>Framingham Heart Study</td>
<td>Atherogenic</td>
<td>3q29</td>
<td>D3S131</td>
<td>209.4</td>
<td>1.95NL</td>
<td></td>
<td>Present study</td>
</tr>
<tr>
<td>Population</td>
<td>dyslipidaemia</td>
<td>3q29</td>
<td>D3S1265–D3S3550</td>
<td>222.8/227.6</td>
<td>1.69/1.44</td>
<td></td>
<td>Bailey et al. (91)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>4q31.21</td>
<td>D4S1625</td>
<td>145.9</td>
<td>1.92</td>
<td>NPY1R, NPY2R, NPY3R, UCP1, CPE</td>
<td>Present study</td>
</tr>
<tr>
<td>The Heritage family study</td>
<td>Visceral fat</td>
<td>4q31.22</td>
<td>D4S2341</td>
<td>159.5</td>
<td>2.34</td>
<td></td>
<td>Rice et al. (92)</td>
</tr>
<tr>
<td>Rochester family heart study</td>
<td>BMI</td>
<td>4q31.1–q32</td>
<td>D4S1629</td>
<td>157.9</td>
<td>1.85ME</td>
<td></td>
<td>Gorlova et al. (93)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>6q12–14.1</td>
<td>D6S1053–D6S1031</td>
<td>80.5–88.6</td>
<td>3.72</td>
<td>HCG, PTP4A, COL19A</td>
<td>Present study</td>
</tr>
<tr>
<td>Dutch siblings</td>
<td>Height</td>
<td>6q</td>
<td>D6S1053–D6S1031</td>
<td>80.5–88.6</td>
<td>2.32</td>
<td></td>
<td>Willemens et al. (49)</td>
</tr>
<tr>
<td>White families</td>
<td>Blood pressure</td>
<td>6q</td>
<td>D6S1031</td>
<td>88.6</td>
<td>3.3</td>
<td></td>
<td>Hunt et al. (50)</td>
</tr>
<tr>
<td>Dutch families</td>
<td>Blood pressure</td>
<td>6q</td>
<td>D6S1031–D6S1056</td>
<td>88.6–102.8</td>
<td>2.5</td>
<td></td>
<td>Allayee et al. (52)</td>
</tr>
<tr>
<td>Hong Kong Chinese</td>
<td>FPG</td>
<td>6q</td>
<td>D6S2410–D6S1031</td>
<td>73.1–88.6</td>
<td>1.56</td>
<td></td>
<td>Ng et al. (51)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>9p22.2</td>
<td>GATA187D09–D9S925</td>
<td>32.2</td>
<td>2.28</td>
<td>VLDLR, INF</td>
<td>Present study</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>T2D, age of diabetes onset</td>
<td>9p24.1</td>
<td>D9S288–D9S925</td>
<td>9.8–32.2</td>
<td>2.06/2.38</td>
<td></td>
<td>Duggirala et al. (65)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>HDL-C</td>
<td>9p24</td>
<td>D9S925–D9S741</td>
<td>32.2–42.7</td>
<td>3.4</td>
<td></td>
<td>Arya et al. (61)</td>
</tr>
<tr>
<td>Dutch and Finnish families</td>
<td>HDL-C</td>
<td>9p23</td>
<td>D9S921</td>
<td>21.9</td>
<td>2.1</td>
<td></td>
<td>Pajukanta et al. (64)</td>
</tr>
<tr>
<td>Pima Indians</td>
<td>Adiponecin levels</td>
<td>9p</td>
<td>D9S168–D9S925</td>
<td>21.9–32.2</td>
<td>3.0/3.5</td>
<td></td>
<td>Lindsay et al. (67)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>19p13.2</td>
<td>D19S586</td>
<td>32.9</td>
<td>1.50</td>
<td>LDLR</td>
<td>Present study</td>
</tr>
<tr>
<td>Pima Indians</td>
<td>Total cholesterol</td>
<td>19p</td>
<td>D19S1034</td>
<td>20.8</td>
<td>3.89</td>
<td></td>
<td>Imperatore et al. (94)</td>
</tr>
<tr>
<td>Utah Caucasian families</td>
<td>TG-to-HDL ratio</td>
<td>19p</td>
<td>D19S247–D19S209</td>
<td>9.8–11.0</td>
<td>2.05/1.9</td>
<td></td>
<td>Elbein and Hasstedt (95)</td>
</tr>
<tr>
<td>Mexican-Americans</td>
<td>Birth weight</td>
<td>19q13.42</td>
<td>D19S589–D19S422</td>
<td>87.7–92.7</td>
<td>1.25</td>
<td>APOC2/ApoE/ApoC1/ApoC4/ATF3</td>
<td>Present study</td>
</tr>
<tr>
<td>Utah Caucasian families</td>
<td>Triglycerides</td>
<td>19q13.2</td>
<td>D19S178/ApoC2</td>
<td>68.1–69.5</td>
<td>3.16</td>
<td></td>
<td>Elbein and Hasstedt (95)</td>
</tr>
</tbody>
</table>

Candidate genes: APOA2, apolipoprotein A2; PKLR, pyruvate kinase deficiency of erythrocyte; LMX1, LIM homeobox transcription factor 1; IL-10, Interleukin 10; ALK7, activin receptor-like kinase 7; PTP4A1, protein-tyrosine phosphate type 4A; PRL-1, phosphatase of regenerating liver-1; MUC4, mucin 4; MUC20, mucin 20; PPP1R2, protein phosphatase inhibitor 2; NPYR, neuropeptide Y receptor 1, 2 and 5; VLDLR, very low-density lipoprotein receptor; LDLR, LDL receptor; INF, interferons; CPE, carboxypeptidase; UCP1, uncoupling protein 1; apolipoproteins—C1 (ApoC1), C2 (ApoC2), C4 (ApoC4) and E (ApoE); ATF5, activating transcription factor 5; BFIC, benign familial infantile convulsions.

aFor the purpose of comparison across the studies, we used Marshfield genetic distances. NPL, nonparametric multipoint linkage score; HLOD, heterogeneity LOD; ME, maternal expression.
attributes (e.g. smoking) or disease conditions (e.g. gestational diabetes and obesity) at the time of birth. Although it is interesting to note that most of the linked loci are shared by obesity/T2D genome scans, some of these findings could be chance occurrences, whereas some may harbor potential susceptibility genes that have pleiotropic influences on birth weight and adult phenotypes such as obesity and T2D. We, therefore, plan to conduct multivariate analyses to examine whether birth weight and adult metabolic diseases in our population have any common genetic influences (i.e. pleiotropy). We also plan to examine potential parent-of-origin effects on genetic variation in birth weight (i.e. imprinting), respectively.

In summary, in this genome-wide linkage study, perhaps for the first time, we have identified susceptibility loci that influence birth weight in human populations. Our genome-wide linkage analysis in Mexican-Americans yielded several genetic regions with significant and suggestive evidence for linkage to birth weight. In particular, we have identified a major locus (LOD = 3.7) for birth weight on chromosome 6 and this finding has been replicated in an independent study in European-Americans. In fact, data from these two studies yielded substantial evidence for linkage (LODadj = 4.3) of birth weight on chromosome 6. This locus appears to have substantial influence on the phenotypic variation of birth weight. Potential/suggestive evidence for linkage has also been found on chromosomes 1, 2, 4 and 9 indicating potential multi-locus influences on birth weight in Mexican-Americans. We plan to follow-up the present findings by fine-mapping our 6q region, and screening positional candidate genes in this region to examine if the sequence variants in such genes could explain our birth weight linkage findings in Mexican-Americans.

**MATERIALS AND METHODS**

**San Antonio Family Birth Weight Study**

The San Antonio Family Birth Weight Study (SAFBWS) sought to collect birth weights, gestational ages and related data for all individuals who had previously participated as adults in one of the three Mexican-American family studies: the San Antonio Family Diabetes/Gallbladder Study (SAFDGS), the San Antonio Family Heart Study (SAFHS) and the Veterans Administration Genetic Epidemiology Study (VAGES). However, to date, both phenotypic and genotypic data are currently available for SAFDGS and SAFHS. We included 449 subjects for whom birth weight data were available from 910 SAFDGS individuals. The SAFDGS families were ascertained on T2D probands, and the family members were distributed across 39 families that participated in our studies at baseline or recalls for whom genotypes for genome-wide markers at a 10 cM density, performed by the Center for Inherited Disease Research (CIDR), was available. In addition, we included 391 individuals with birth weight data from 893 individuals distributed across 24 pedigrees from the SAFHS study. These 24 families are part of the original 41 Mexican-American families ascertained at random without regard to disease status. A detailed description of SAFDGS and SAFHS including subjects, methods and phenotypes and genome scan data has been presented in previous publications (78–80). Data from these two studies are currently being examined for susceptibility genes of various metabolic diseases including T2D, obesity, hypertension and metabolic syndrome-related quantitative traits (e.g. insulin and lipids). Thus, for the present study, birth weight data for 840 individuals and genotypic information (a 10 cM map) were used from combined SAFDGS and SAFHS data sets as shown in Table 1.

**Birth weight phenotype**

Birth weight and related information, began late in 1949 to be routinely recorded on official state birth certificates in San Antonio and throughout Texas. It was recorded earlier than this in some parts of Texas, and in some other states. Our primary source of birth weight data for these analyses was an extensive set of previously microfilmed birth certificates from the San Antonio Metropolitan Health District, for the years 1949 and later. For participants born outside San Antonio but within Texas, actual birth weights were available from the Texas Department of Health (TDH) beginning in 1964. Participants for whom no birth weight data were available from either of these sources, a letter was sent to the individual, requesting birth weight documentation from the participant himself or herself. Through home visits to those participants who possessed such documentation, we were able to obtain data from photocopies of official birth certificates for participants born outside San Antonio or outside Texas, from hospital birth certificates, or from other materials, such as care instructions, provided by the hospital at the time of discharge.

We abstracted data from all available records on birth weight and gestational age, along with relevant co-variables, such as maternal and paternal age at the time of birth, and multiple-birth status. During different time periods, different methods were used to record gestational age. In earlier years, a dichotomous variable (whether the child was at ‘term’ or not) was used; later, weeks of gestation were recorded; and most recently, in a relatively small number of cases, the date of the last menstrual period of the mother has been recorded. We have therefore included gestational age as a dichotomized co-variable—whether the child was at term or not—in these analyses.

We have included in these analyses all individuals for whom both documented birth weight and gestational age data were available, on the one hand, and for whom genetic marker data were available from the full genome scans described earlier. Most individuals who were born before late 1949 could not be included in these analyses because of the lack of available data either from our microfilm set or from their own records. Similarly, most individuals who were born in Mexico, regardless of age, had to be excluded because they did not have copies of their birth certificates; in those few cases in which an individual shared with us a copy of his or her official birth certificate from Mexico, it did not include birth weight data. These analyses therefore primarily included US-born (and particularly San Antonio- and Texas-born) individuals aged 18–52 years at the time of their participation in the SAFDGS or the SAFHS. The Institutional Review Board of the University of Texas Health...
Science Center at San Antonio approved all procedures for the data collection process, and all individuals gave written informed consent to participate in the study.

STATISTICAL GENETIC ANALYSIS

Variance components linkage analysis

We used a pedigree-based multipoint variance components approach to test for linkage between marker loci and the birth weight phenotype using a maximum likelihood method (81,82). In this method, the expected genetic covariances between relatives are specified as a function of the IBD relationships at a marker locus, which is assumed to be closely linked to a locus influencing the quantitative trait in question. The covariance matrix for a pedigree can be written as shown in the following equation:

$$
\Omega = \sum_{j=1}^{n} \Pi_j \sigma_q^2 + 2\Phi \sigma_g^2 + I \sigma_e^2
$$

where $\Omega$ is the covariance matrix for a given pedigree; $\Pi_j$, a matrix with elements $(\pi_{ij})$ providing the expected proportion of alleles at the specific chromosomal location of the QTL that individuals $j$ and $l$ share IBD, which is estimated using genetic marker data; $\sigma_q^2$ the additive genetic variance due to the major locus; $\Phi$ the kinship matrix; $\sigma_g^2$ the variance due to residual additive genetic effects; $I$ an identity matrix; and $\sigma_e^2$ the variance due to random environmental effects.

The null hypothesis that the additive genetic variance due to QTL, $\sigma_q^2$, for a given phenotype equals zero (no linkage) can be tested using the likelihood ratio test. The likelihood of a restricted model in which $\sigma_q^2$ is constrained to 0 is compared with the likelihood of the general model in which genetic variance due to QTL ($\sigma_q^2$) is estimated. The natural logarithm ($\ln$) likelihood values of the general model and restricted model are then compared using the likelihood ratio test. Twice the difference between the $\ln$ likelihood values of these models yields a test statistic that is asymptotically distributed as $\chi^2_1$ and a point mass at 0.

Correction for ascertainment bias

As SAFDGS families were ascertained on a diabetic proband, all of our genetic analyses incorporated correction for ascertainment bias by computing the likelihood of a pedigree conditional on birth weight of the proband using SOLAR (84).

Multipoint mapping procedure

This method uses information on all available markers separated by known map distances and all possible biological relationships simultaneously in deciphering the genetic architecture of a given quantitative phenotype. For this study, we created a combined genetic map using the marker information available from the two studies (SAFDGS and SAFHS) and calculated multipoint identical by descent (MIBD) matrices using Markov Chain Monte Carlo (MCMC) methods implemented in Loki (85).

Oligogenic linkage analysis

We also conducted oligogenic linkage analysis which is a sequential strategy to identify multiple loci affecting the variation in birth weight. In this approach, the genome was scanned for linkage and the chromosomal location with the largest marginal LOD score was retained for further conditional analyses. Given the putative location of the first locus, we scanned the genome again and examined the resulting conditional LOD scores. Loci were added to the model sequentially until no locus met the selected criterion of a conditional LOD score of 0.59. In the two-loci model, the position of the first locus is fixed, and the location of the second locus is allowed to vary through the whole genome. The QTL variance ($h^2_q$) for the first locus is re-estimated along with the second locus. Similarly, in the three-loci model, the positions of the first two loci are fixed but the effect sizes are re-estimated along with the third locus. Finally, in the four-loci model, the positions of the first three loci are fixed, but the effect sizes are re-estimated along with the fourth locus. To obtain a conditional $P$-value, the likelihood values of the two competing models (e.g. one-locus versus two-locus models) are compared using the likelihood ratio test. Twice the difference between the $\ln$ likelihood values of these models yields a test statistic that is asymptotically distributed as a $\chi^2_1$ and a point mass at 0.

Empirical $P$-value estimation

We performed simulation analysis to verify our major finding on chromosome 6q by simulating a fully informative marker that was not linked to the QTL influencing variation in birth weight. For this simulated marker, IBD information was calculated and linkage analysis then conducted. Using simulations, the empirical distribution of the LOD scores under the assumption of multivariate normality was determined based on information obtained from 100 000 replicates using the SOLAR program (81).

Description of the replication study

The Fels Longitudinal Study (FLS), which was started in 1929, is the largest and longest ongoing family study of physical growth and development in the world (86). The Fels study consists of over 1200 serial participants, randomly ascertained from the greater Dayton, Ohio area from approximately 200 nuclear and multigenerational extended families. Measured birth weights are available for all of the Fels study participants who have been followed from birth. A trained staff member measured birth weights of infants in the hospital shortly after birth through 1968, after which time-measured birth weights were obtained from hospital and birth records.
For this replication analysis, we used birth weight data for 331 individuals (162 males and 169 females), distributed across 57 families for whom phenotypic and genotypic data were available. We generated a combined map incorporating the marker information from the two studies, SAFBWS and FLS, and calculated study-specific MIBDs for chromosome 6 using MCMC methods implemented in Loki (85). We then performed variance components linkage analysis.

**Combined evidence for linkage of birth weight on chromosome 6q**

After determining the evidence for linkage of birth weight in both SAFBWS and FLS independently, we combined the linkage information on chromosome 6 from both studies to obtain combined evidence for linkage of birth weight. However, thus obtained LOD score at a given location involves 2 df; in fact, the test statistic (L) is asymptotically distributed as a 1/4:1/2:1/4 mixture of $\chi^2_2$, $\chi^2_1$ and $\chi^2_0$. Given the evidence that birth weight is positively associated with weight at later ages (83). The combined LOD score with 2 df can be adjusted to 1 df, denoted as LOD$_{adj}$ score, by requiring it to provide the same $P$-value as is provided by the combined LOD score (87). Thus, the LOD$_{adj}$ score can be considered equivalent to the classical univariate LOD score.

**WEB RESOURCES**

Accession numbers and URLs for data presented herein are as follows: Center for Medical Genetics, Marshfield Medical Research Foundation http://research.marshfieldclinic.org/genetics (for genetic map information used for the purpose of comparison).


UCSC Bioinformatics Site, Human Genome Browser Gateway, May 2004 (hg17) Assembly, http://www.genome.ucsc.edu/cgi-bin/hgGateway (for physical map and cytogenetic band information regarding microsatellite markers and genes).

**ACKNOWLEDGEMENTS**

This study was supported by grants from the National Institute of Health (R01 HD 41111, DK53889, DK42273, P01 HL45522, DK47482, MH59490 and M01-RR-01346). We warmly thank the participants of SAFBWS subjects for their enthusiasm and co-operation. We thank Texas Department of Health for providing birth weight data for some individuals. We also thank the Center for Inherited Disease Research for providing a new genome scan using SAFDGS data.

**Conflict of Interest statement.** None declared.

**REFERENCES**


