Replication of a Genome-Wide Association Study of Birth Weight in Preterm Neonates

Kelli K. Ryckman, PhD[1], Bjarke Feenstra, PhD[2], John R. Shaffer, PhD[3], Elise N. A. Bream, BA[1], Frank Geller, MSc[2], Eleanor Feingold, PhD[3], Daniel E. Weeks, PhD[3], Enrique Gadow, MD, PhD[4], Viviana Cosentino, MD, PhD[4], Cesar Saleme, MD[5], Hyagriv N. Simhan, MD[6], David Merrill, MD, PhD[7], Chin-To Fong, MD[8], Tamara Busch, BS[1], Susan K. Berends, ARNP, NNP[1], Belen Comas, BS[4], Jorge L. Camelo, PhD[4], Heather Boyd, PhD[2], Cathy C. Laurie, PhD[9], David Crosslin, PhD[9], Qi Zhang, PhD[9], Kimberly F. Doheny, PhD[10], Elizabeth Pugh, PhD[10], Mads Melbye, MD, PhD[2], Mary L. Marazita, PhD[3], John M. Dagle, MD, PhD[1], and Jeffrey C. Murray, MD[1]

Objective To examine associations between rs9883204 in *ADCY5* and rs900400 near *LEKR1* and *CCNL1* with birth weight in a preterm population. Both markers were associated with birth weight in a term population in a recent genome-wide association study of Freathy et al.

Study design A meta-analysis of mother and infant samples was performed for associations of rs900400 and rs9883204 with birth weight in 393 families from the US, 265 families from Argentina, and 735 mother–infant pairs from Denmark. *Z*-scores adjusted for infant sex and gestational age were generated for each population separately and regressed on allele counts. Association evidence was combined across sites by inverse-variance weighted meta-analysis.

Results Each additional C allele of rs900400 (*LEKR1/CCNL1*) in infants was marginally associated with a 0.069 SD lower birth weight (95% CI, −0.159 to 0.022; $P = .068$). This result was slightly more pronounced after adjusting for smoking ($P = .036$). No significant associations were identified with rs9883204 or in maternal samples.

Conclusions These results indicate the potential importance of this marker on birth weight regardless of gestational age. *(J Pediatr 2012;160:19-24)*.

See editorial, p 3

Birth weight is a complex trait, the extremes of which are associated with high rates of perinatal morbidity and mortality.[1-3] Infants with low (<2500 g) and very low (<1500 g) birth weight are at an increased risk for developing such neonatal complications as bronchopulmonary dysplasia, intraventricular hemorrhage, and necrotizing enterocolitis.[3] In addition, there is evidence that low birth weight is associated with disease in adulthood, including type 2 diabetes, cardiovascular disease, and hypertension.[4-9] Many factors influence birth weight, the most significant of which is gestational age. Other environmental contributors include maternal age, race/ethnicity, body mass index, parity, education, and smoking.[10,11]

Twin and family studies have demonstrated that genetic factors also play a role in birth weight.[12-15] The heritability of birth weight increases with decreasing gestational age and is ~38% at 25 weeks gestation, compared with 15% at 42 weeks.[16] Therefore, more of the variation in birth weight is due to genetic factors at an earlier gestational age. However, to date, most genetic association studies of birth weight have examined individuals born at term. Identifying genetic associations with birth weight in preterm infants will provide insight into the biology underlying low birth weight and may lead to the development of predictive and preventive treatment for infants at risk for being born at low birth weight for gestational age.

Previous studies found numerous genes associated with birth weight, most of which are also associated with type 2 diabetes, including transcription factor 7-like 2 (*TCF7L2*), peroxisome proliferator-activated receptor-γ (*PPARG*), and insulin-like growth factor binding protein 3 (*IGFBP3*).[17-19] Most of these associations have not been replicated in independent populations, however.[20,21] Recently, a genome-wide association study (GWAS) of birth weight was performed in 10 623 European infants of term gestation (>37 weeks) from several

GWAS	Genome-wide association study
HWE	Hardy-Weinberg equilibrium
QTDT	Quantitative transmission disequilibrium test
SNP	Single nucleotide polymorphism

From the [1]Department of Pediatrics, University of Iowa, Iowa City, IA; [2]Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark; [3]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; [4]Centro de Educación Médica E Investigaciones Clínicas, Buenos Aires, Capital Federal, Argentina; [5]Instituto de Maternidad y Ginecología Nuestra Señora de las Mercedes, San Miguel de Tucumán, Argentina; [6]Obstetrics, Gynecology, and Reproductive Sciences, University of Pittsburgh School of Medicine, Magee-Women's Research Institute, Pittsburgh, PA; [7]Wake Forest University Baptist Medical Center, Wake Forest University School of Medicine, Winston-Salem, NC; [8]Strong Children's Research Center, University of Rochester School of Medicine, Rochester, NY; [9]Department of Biostatistics, University of Washington, Seattle, WA; and [10]Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD

Funding information is available at www.jpeds.com (**Appendix**). The authors declare no conflicts of interest.

 10.1016/j.jpeds.2011.07.038

large pregnancy and birth cohorts.[22] This study found 2 fetal genetic associations with birth weight, with one intergenic marker between the leucine, glutamate, and lysine rich 1 (*LEKR1*) and cyclin L1 (*CCNL1*) genes and the other marker within the adenylate cyclase 5 (*ADCY5*) gene. These loci contribute to 0.1% and 0.3% of the variation in birth weight, respectively. Both replicated in a similar population of 27 591 individuals. Specifically, the C alleles of rs900400 (*LEKR1/CCNL1*) and rs9883204 (*ADCY5*) were associated with lower birth weight ($P = 2 \times 10^{-35}$ and $P = 7 \times 10^{-15}$, respectively). We examined these associations with birth weight in a diverse collection of Caucasian preterm (23-36 weeks gestation) populations, including samples from 4 sites in the US, 2 sites in Argentina, and 1 site in Denmark. We tested these associations using meta-analysis as well as a family-based approach.

Methods

The US and Argentina study samples were collected at 4 sites in the US (University of Iowa, Iowa City, Iowa; Magee-Women's Hospital, Pittsburgh, Pennsylvania; University of Rochester Medical Center, Rochester, New York; and Wake Forest University, Wake Forest, North Carolina) and 2 centers in Argentina (Instituto de Maternidad y Ginecología Nuestra Señora de las Mercedes in Tucumán and Hospital Provincial de Rosario in Rosario). Signed informed consent (IRB 199911068, 200411759, and 200506792) for samples collected from Iowa, Wake Forest, Rochester, Pittsburgh, and Argentina were obtained from all families before enrollment. DNA was extracted from cord blood or buccal swabs collected from the infant and from venous blood, saliva samples, or buccal swabs collected from the parents and other relatives. Demographic information and phenotype data were collected through an interview with the mother and/or medical chart review.

The present study is a secondary analysis of data that have been described previously in relation to genetic factors associated with spontaneous preterm birth.[23,24] Because the underlying genetic etiologies may differ between indicated and spontaneous preterm birth, individuals with any of the following outcomes were excluded from the study: multiple gestations, use of assisted reproductive technology, fetal demise, elective termination of pregnancy, congenital anomalies, uterine structure anomalies, maternal autoimmune disorder, placental abruptions caused by hypertensive or traumatic clinical techniques, placenta previa, and delivery resulting from obstetric intervention due to maternal or fetal indications. This information was missing or limited for the samples from Argentina and for 172 of the samples from Iowa. Gestational age was defined by best obstetrical estimate using menstrual dating and obstetrical ultrasound corroboration. A total of 658 families, including 393 families from the US (241 from Iowa City, 95 from Wake Forest, 38 from Pittsburgh, and 19 from Rochester) and 265 families from Argentina (245 from Tucumán and 20 from Rosario), with complete birth weight and gestational age information were included in the analysis.

The Danish study samples came from a GWAS of prematurity and its complications. In this GWAS, 1000 mother–infant case pairs (20-36 weeks gestation) and 1000 mother–infant control pairs (40 weeks gestation) were selected from the Danish National Birth Cohort, a population-based cohort of 101 042 pregnancies recruited between 1996 and 2002.[25] These samples were selected to study the primary outcome of spontaneous preterm birth; therefore, individuals were considered for inclusion in this study only if there was no evidence of placental abnormalities, preeclampsia/eclampsia, congenital abnormalities, or stillbirth. Gestational age was determined by a consensus algorithm taking up to 6 different gestational age variables (self-reported and from health registers) into account. The study protocol was approved by the Danish Scientific Ethical Committee and the Danish Data Protection Agency.

Blood samples from mothers were obtained at recruitment (~week 8 of pregnancy), and infant samples were obtained at birth from umbilical cord blood. All samples were stored in a biobank at Statens Serum Institut (Copenhagen, Denmark) as buffy coats and/or on filter paper, and DNA was extracted for the samples used in the GWAS. All participating women in the Danish National Birth Cohort underwent thorough phenotype characterization based on information from 4 computer-assisted telephone interviews conducted during pregnancy (2 interviews) and after delivery (2 interviews). Additional phenotype information was obtained from Danish health registers. Preterm singleton births occurring after spontaneous labor were included in these analyses. A total of 735 mother–infant pairs with genotype data for mother and/or infant were included in the analysis.

Genotyping

DNA samples from Iowa, Pittsburgh, Wake Forest, Rochester, and Argentina were genotyped for the single nucleotide polymorphism (SNP) markers *ADCY5* rs9883204 and *LEKR1/CCNL1* rs900400 using Taqman chemistry (Applied Biosystems, Foster City, California) under standard conditions. Allele determination was done in the endpoint analysis mode on an Applied Biosystems 7900 HT Sequence Detection System machine using SDS 2.3 software (Applied Biosystems). Genotypes were entered into the Progeny Lab database (Progeny Software, Delray Beach, Florida), which was used to create datasets for analysis.

GWAS genotyping of the Danish study sample was performed using the Illumina Human 660W-Quadv1_A chip (Illumina Inc, San Diego, California) as part of the Gene Environment Association Studies consortium. Because this array does not include rs900400 and rs9883204, allele dosages for these SNPs were imputed using the MACH software package[26,27] (Data Management Software Inc, Cincinnati, Ohio) and the HapMap CEU sample as the reference panel. Imputation was done separately for mothers and infants; in both cases, imputation quality was excellent for both SNPs (estimated $R^2 > 0.95$).

Statistical Analysis

We chose an approach similar to the GWAS of Freathy et al[22] and generated *z*-scores for the individual study populations