Maternal and Infant Genetic Contributions to Spontaneous Very Preterm Birth Using a State-Based Biobank

Christopher Bean[1], Dhelia Williamson[1], Juan Yang[2], and Martin Kharrazi[3]

1Division of Reproductive Health, Centers for Disease Control and Prevention, Atlanta, GA; 2Sequoia Foundation, Richmond CA; and 3Genetic Disease Screening Program, California Department of Public Health, Richmond, CA

BACKGROUND

- Preterm birth (<37 weeks gestation) accounts for 12-13% of births in the United States and is the leading cause of infant mortality.
- The majority of this mortality is concentrated among very preterm infants (<32 weeks gestation).
- Striking disparities in the rate and consequences of very preterm birth exist across racial and ethnic groups.
- Non-Hispanic Blacks have more than twice the rate of very preterm births (3.8%) as either Non-Hispanic Whites (1.5%) or Mexican Hispanics (1.6%).

STUDY AIMS

- Examine the association between polymorphisms in candidate inflammatory, endocrine and vascular system genes, environmental factors and spontaneous very preterm birth among mother-infant pairs of three racial/ethnic groups.
- Identification of genetic markers and environmental factors to effectively predict spontaneous very preterm birth.

METHODS

- **Study Design:** Retrospective population-based case-control (*Table 1*).
- **Geographic area:** 3 southern California counties (Orange; San Diego; and Imperial).
- **Eligibility:**
  Mothers: Enrolled in the State's prenatal Expanded Alpha-Fetoprotein (XAFP) screening program November 1999 to December 2006; and banked maternal blood pellet specimen.
  Infants: Born January 2000 to April 2007; enrolled in State's newborn screening program; and banked newborn blood spot specimen.
- **Samples:** Torres, Utah State University: DNA extraction and amplification; Sequenom Genetic Services Lab: Genotyping; and Yolken, Johns Hopkins University: Immunologic analyses.

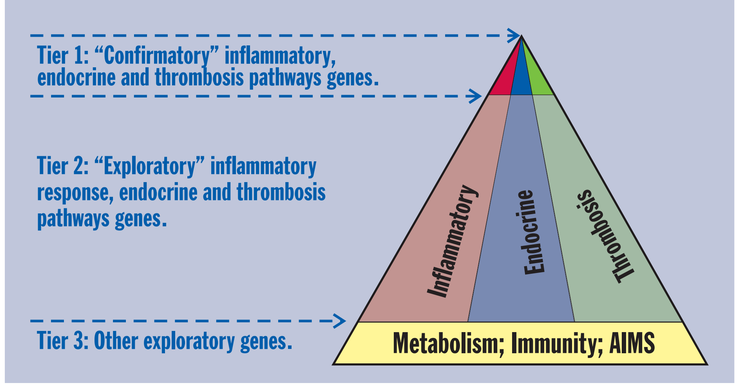
Table 1. Estimated Final Case and Control Sample Size

	Cases[1]	Controls[2]
	20 to 33 weeks	37 to 41 weeks
African American infants	162	324
African American mothers	162	324
	20 to 31 weeks	37 to 41 weeks
Mexican Hispanic infants	200	267
Mexican Hispanic mothers	200	267
	20 to 31 weeks	37 to 41 weeks
Non-Hispanic White infants	200	200
Non-Hispanic White mothers	200	200

1 Singleton, spontaneous live birth, birthweight <2500 grams, first pregnancy during study period, no registered birth defects; and 2 Singleton, spontaneous live birth, birthweight >=2500 grams, first pregnancy during study period, no registered birth defects, no fetal alcohol syndrome, hyaline membrane disease, respiratory distress syndrome, preeclampsia/eclampsia, Rh sensitization, hemoglobinopathy, incompetent cervix, premature labor, tocolysis, placenta previa or placental abruption reported on birth certificate.

- **Candidate Gene Selection:**
  - We are preparing to examine approximately 1,500 single nucleotide polymorphisms (SNPs) in about 90 candidate genes.
  - Candidate gene selection focuses primarily on the inflammatory, endocrine and thrombosis pathways, with genes from additional pathways included for analysis as possible (*Figure 1*).
  - Candidate genes will be prioritized into **3 Tiers** based on evidence from the available literature.
    - **Tier 1** will consist of "confirmatory" inflammatory, endocrine and thrombosis pathways genes implicated in previous association studies of preterm birth.
    - **Tier 2** will include "exploratory" inflammatory response, endocrine and thrombosis pathways genes that may play important regulatory roles but have not previously been linked to preterm birth.
    - **Tier 3** will contain genes in additional pathways, including those involved in innate immunity, metabolic and cardiovascular processes, that may contribute significantly to very preterm birth.

Figure 1. Candidate Gene Selection



Funders: This study is jointly funded by CDC and March of Dimes (#21-FY05-1248)

- **SNP Selection:**
  - SNPs will be selected by collecting and combining information from publicly available databases, including HapMap, with results published in the scientific literature.
  - Putative "functional" SNPs will be selected by searching Entrez PubMed for published evidence of polymorphisms that 1) alter expression of the candidate genes, or 2) are associated with relevant disease (*Figure 2A*).
  - Additionally, tagSNPs will be selected based on the r^2 linkage disequilibrium statistic to minimize the number of SNPs which must be typed to assay the genetic variation in the populations (*Figure 2B*).
    - For "confirmatory" Tier 1 genes, a stringent $r^2 > 0.8$ will be used.
    - Remaining "exploratory" genes will be assayed at $r^2 > 0.6$.

Figure 2. Candidate Gene SNP Selection

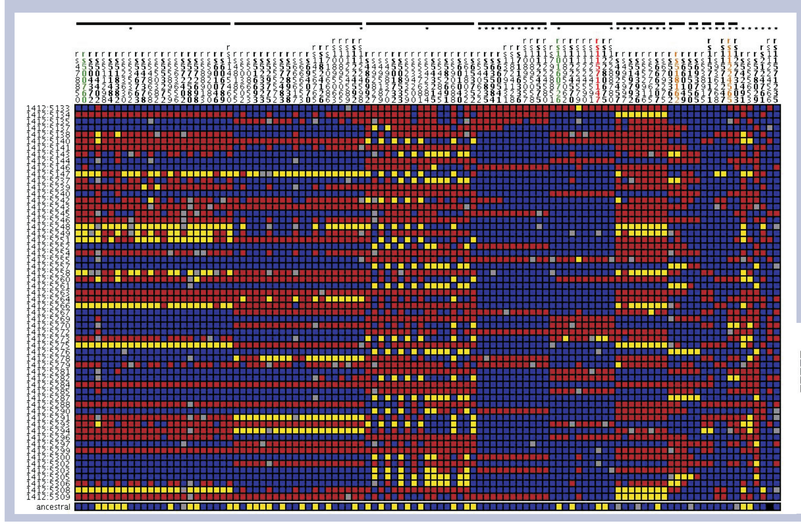


Figure 2A. "Functional" SNPs were selected based on published evidence of altered gene expression or association with preterm birth. Partial results from PubMed search of progesterone receptor (PGR) is shown.

Figure 2B. tagSNPs were selected with LD Select using publicly available genotype data. Example visual genotype for PGR is shown.

Gene = PGR; Dataset = HapMap YRI; $r^2 > 0.6$; Minor allele cutoff = 10%

tagSNP image generated with SeattleSNPs Genome Variation Server (http://gvs.gs.washington.edu/GVS/)

Genotyping:

- SNP assays will be multiplexed into pools of up to 40 SNPs using Sequenom Assay Design software.
- Assays will be validated and genotypes will be determined by single base primer extension and MALDI-TOF mass spectrometry using MassARRAY platform at Sequenom.
- Pilot genotyping studies using DNA samples from Mexican Hispanic mothers have been performed at the CDC (*Figure 3*).
- Sequenom's Genetic Services Laboratory will validate and run up to 1500 SNP assays on the provided maternal and infant DNA samples.

DISCUSSION:

Strengths of the design include:

- Specificity of outcome: the large population-based cohort of banked specimens will make sufficient sample size available to focus on the highest risk and potentially more homogenous group of spontaneous very preterm births (preterm labor and preterm premature rupture of membranes).
- Mother and child genetic contributions:
  - Mother-infant pairs will allow a comparative assessment of the contribution of both maternal and infant genotypes in a single study.
  - It will also assess potential maternal and/or infant gene-gene interactions (maternal-maternal, infant-infant, maternal-infant) and potential maternal and/or infant gene-environment interactions that contribute to the association.
- Multiple population groups: three ethnic/racial groups in a population based setting free from biased referral patterns will inform our understanding of preterm etiology.
- Multiple pathways: up to 1,500 polymorphisms in more than 90 candidate genes will allow simultaneous analysis of three overlapping biological pathways as they may contribute to the complex outcome of spontaneous very preterm birth.

Figure 3. MassARRAY Genotyping Results

Cluster plot summarizing 384-well plate results for rs10169916 (interleukin-1beta promoter region)

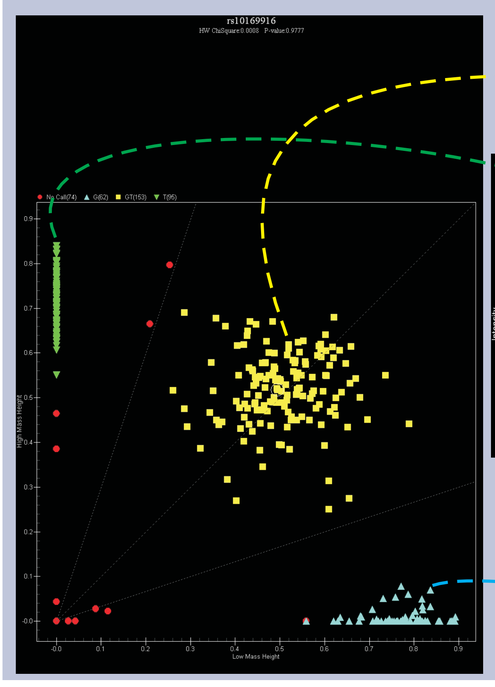


Figure 3A. Representative spectrum for TT homozygote

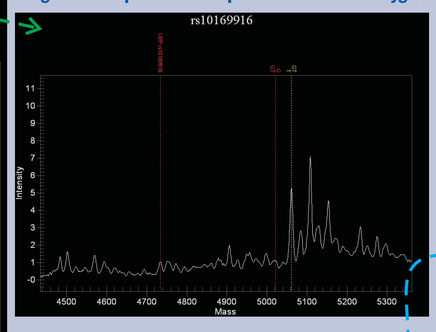


Figure 3B. Representative spectrum for GT heterozygote

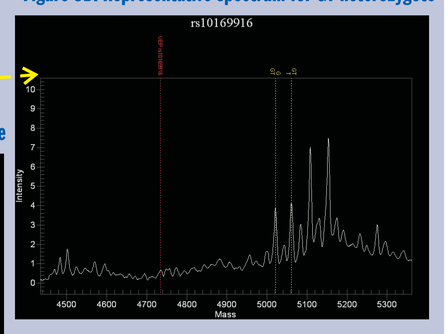


Figure 3C. Representative spectrum for GG homozygote