nature publishing group

# Polymorphisms in *NR5A2*, gene encoding liver receptor homolog-1 are associated with preterm birth

Dinushan C. Kaluarachchi[1], Allison M. Momany[1], Tamara D. Busch[1], Lucas G. Gimenez[2], Cesar Saleme[3], Viviana Cosentino[2], Kaare Christensen[4], John M. Dagle[1], Kelli K. Ryckman[5], and Jeffrey C. Murray[1]

**BACKGROUND:** Preterm birth (PTB) is a major cause of neonatal mortality and morbidity. There is strong evidence of genetic susceptibility. Objective of this study was to identify genetic variants contributing to PTB.

**METHODS:** Genotyping was performed for 24 single nucleotide polymorphisms (SNPs) in 4 candidate genes (*NR5A2*, *FSHR*, *FOXP3*, and *SERPINH1*). Genotyping was completed on 728 maternal triads (mother and maternal grandparents of a preterm infant). Data were analyzed with Family Based Association Test.

**RESULTS:** For all maternal triads rs2737667 of *NR5A2* showed significant association at $P = 0.02$. When stratifying by gestational age three SNPs in *NR5A2* had *P* values <0.05 in the <32-wk gestational age group (rs12131233, $P = 0.007$; rs2737667, $P = 0.04$; rs2816949, $P = 0.02$). When preterm premature rupture of membranes cases were excluded rs2737667 of *NR5A2* showed the strongest association with a *P* value <0.0002. This association remained significant after correction for multiple testing.

**CONCLUSION:** This study suggests a potential association between intronic SNPs in the *NR5A2* gene and PTB. *NR5A2* gene encodes for the liver receptor homolog-1 protein, which plays a critical role in regulation of cholesterol metabolism, steroidogenesis, and progesterone synthesis. These findings suggest that *NR5A2* may be important in the pathophysiology of PTB and exploring noncoding regulators of *NR5A2* is warranted.

Preterm birth (PTB), defined as a birth before 37-wk gestation, is a major perinatal health problem. Approximately 10% of all births worldwide are preterm (1). There is wide variation by geography and socioeconomic status. Africa, Asia, and North America have the highest rates of PTB. PTB rate in United States was 11.4% in 2013 (2), significantly higher than other developed nations. PTB is directly responsible for an estimated one million neonatal deaths annually (3). It is also an important contributor to long-term morbidities such as cerebral palsy, developmental delay, visual and hearing impairments, and chronic lung disease (4,5).

The etiology of PTB is multifactorial and there is strong evidence for genetic susceptibility (6,7). PTBs tend to recur in mothers (8). Familial trends and racial disparities of prematurity also suggest that genetics influence this trait (9–11). A number of case–control and family-based studies have evaluated candidate genes chosen from some of the pathways associated with PTB (12–17).

In this study, we hypothesize that single nucleotide polymorphisms (SNPs) in four candidate genes (*NR5A2*, *FSHR*, *FOXP3*, and *SERPINH1)* contribute to genetic predisposition to PTB. We evaluated the association between PTB and SNPs in the four candidate genes.

The four candidate genes we studied act on fundamental mechanisms associated with the establishment and maintenance of pregnancy. *NR5A2* gene encodes Nuclear Receptor Subfamily 5, Group A, Member 2, also known as liver receptor homolog-1 (LRH1). LRH1 has been shown to play an important role in establishing and sustaining pregnancy in animal models (18). Polymorphisms in *FSHR* encoding follicle-stimulating hormone receptor have been found to be associated with PTB in a Finnish as well as an African-American population (19,20). Dysregulation of *FSHR* may contribute to early uterine contractility. *SERPINH1* encodes heat shock protein 47, which serves as a chaperone stabilizing the collagen triple helix. Polymorphisms in *SERPINH1* increase risk of preterm premature rupture of membranes (PPROM) in African Americans (21). *FOXP3* encodes a member of the FOX protein family that appears to function as a master regulator in the development and function of regulatory T cells an important pathway in modulating the immune interactions between mother and fetus. While many studies have focused on the fetus as the "risk case" in studies of PTB, in this study we used the mother as the risk case and a family-based approach using DNA from the mother and her parents to improve the chance for identifying risk alleles acting through the mother.

[1]Department of Pediatrics, University of Iowa, Iowa City, Iowa; [2]Department of Genetics, Centro de Educación Médica e Investigaciones Clínicas, Buenos Aires, Argentina; [3]Department of Neonatology, Instituto de Maternidad y Ginecología Nuestra Señora de las Mercedes, San Miguel de Tucumán, Argentina; [4]Department of Epidemiology, University of Southern Denmark, Odense, Denmark; [5]Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, Iowa. Correspondence: Dinushan C. Kaluarachchi (dinushan-kaluarachchi@uiowa.com)

Received 20 May 2015; accepted 4 November 2015; advance online publication 10 February 2016. doi:10.1038/pr.2016.7

## RESULTS

We analyzed the genotype data of 728 maternal triads from United States, Argentina, and Denmark. Maternal triad refers to mother and maternal grandparents of a preterm infant. A description of the study population is provided in **Table 1**. Data on ethnicity were only available for US triads (Caucasian 97%, African American 2%, and other 1%). A total of 24 SNPs in four candidate genes (*NR5A2*, *FSHR*, *FOXP3*, and *SERPINH1)* were studied (**Table 2**).

All SNPs tested had a minor allele frequency greater than 20%. Genetic predisposition for PTB was associated ($P < 0.05$) in multiple SNPs in our study as shown in **Table 3**. Analysis of all maternal triads found a significant association of PTB with rs2737667 in *NR5A2* ($P = 0.02$). Analysis of the US triads found a significant association of PTB with rs2737667 ($P = 0.01$) and rs603647 in *SERPINH1* ($P = 0.02$). There were no significant associations in the maternal triads from Argentina or Denmark (see **Supplementary Table S1** online).

When stratifying by gestational age, three SNPs in *NR5A2* (rs12131233, $P = 0.007$; rs2737667, $P = 0.04$; and rs2816949, $P = 0.02$) and rs6165 in *FSHR* ($P = 0.01$) had $P$ values <0.05 in the <32-wk gestational age group. There were no significant findings in the 32–36-wk group (see **Supplementary Table S2** online).

When cases of PPROM were excluded rs2737667 in *NR5A2* showed the strongest association with PTB ($P = 0.0002$) for US and Argentina maternal triads. This association was the only one to remain significant after correction for multiple testing (Bonferroni-corrected $P$ value <0.0006). In the PPROM group, rs6165 in *FSHR* had a $P$ value of 0.02 (**Supplementary Table S3** online). Results are summarized in **Table 3**.

## DISCUSSION

PTB is a common complex trait. The etiology of PTB depends upon the interplay between genetics and environmental factors such as nutrition, infection, stress, trauma, and drug use. In this study, we investigated the association of PTB and four candidate genes with different biological mechanisms. To the best of our knowledge, this is the first candidate gene study investigating the role of *NR5A2* gene in the etiology of PTB.

A SNP in *NR5A2*, rs2737667 showed an association with PTB in all maternal triads, <32-wk gestation age group and in the spontaneous preterm with intact membranes group. Of these, association with spontaneous preterm with intact membranes reached formal level of significance ($P = 0.0002$). Two other SNPs in the same gene (rs12131233 and rs2816949) were associated with PTB in <32-wk group but did not reach formal level of significance.

The *NR5A2* gene encodes an orphan nuclear receptor named LRH1. Orphan nuclear receptors regulate transcription independent of known ligands. LRH1 is expressed in high levels in the ovary with highest expression in granulosa cells and corpus lutea. It is also expressed in endometrium, liver, intestine, and pancreas. LRH1 plays a vital role in bile acid synthesis, cholesterol metabolism, and steroidogenesis (22). It also has a role in the regulation of progesterone synthesis (22). LRH1 mediates progesterone production in the ovary and its actions in the uterus. Lack of LRH1 activity deregulates the expression of genes involved in progesterone synthesis and metabolism. In one study, mice lacking NR5A2 in granulosa cells had extremely low progesterone levels (23). LRH1 in the uterus has shown to be essential for implantation, decidualization, and placentation in mouse models. Zhang *et al.* (18) concluded that LRH 1 is necessary for maintenance of the corpus luteum, for promotion of decidualization, and for formation of the placenta, therefore playing multiple roles in establishing and sustaining pregnancy. Progesterone is critical for maintaining pregnancy and decline of progesterone action is implicated in the onset of parturition. Prenatal administration of progesterone has significantly decreased the PTB in women considered to be at increased risk (24,25). We can speculate that women with polymorphisms in NR5A2 gene and resulting functional variations of LRH1 are at risk of PTB due to low progesterone levels and lack of progesterone-dependent actions on the uterus.

SNP rs2737667 is an intronic variant in NR5A2 gene located 15 bp from a noncoding regulatory element. There are no

**Table 1.** Demographic characteristics of study population

| | SNP genotyping | | |
|---|---|---|---|
| | US | Argentina | Denmark |
| Number of maternal triads | 176 | 372 | 180 |
| Number of pedigrees | 168 | 371 | 123 |
| Number of individuals | 512 | 1,114 | 423 |
| Gestational age (wk)[a] | 31.8 (±3.8) | 32.5 (±3.1) | 34 (±2) |
| Births <32 wk (%) | 45 | 35 | 9 |
| Infant birth weight (g)[a] | 1957 (±778) (Unknown 5) | 1762 (±569) (Unknown 2) | 2434 (±651) |
| Infant gender (%male) | 46 | 53 | 51 |
| PPROM (%) | 39 | 51 | NA |
| Maternal age at delivery (y)[a] | 30.1 (±4.8) (Unknown 16) | 23.6 (±5.6) (Unknown 28) | 27.5 (±3.9) |

For mothers who had more than one preterm delivery, data on first preterm delivery was used.
[a]Data are presented as mean ± SD.

**Table 2.** List of SNPs

| Gene | SNP | Chromosome | Position | Allele | MAF |
|---|---|---|---|---|---|
| *NR5A2* | rs2821330 | 1 | 200082464 | A/G | 0.416 |
| *NR5A2* | rs12133107 | 1 | 200089274 | G/A | 0.25 |
| *NR5A2* | rs2737670 | 1 | 200071286 | G/A | 0.482 |
| *NR5A2* | rs2363573 | 1 | 200108427 | T/C | 0.293 |
| *NR5A2* | rs12131233 | 1 | 200135175 | A/T | 0.264 |
| *NR5A2* | rs6658424 | 1 | 200049302 | T/A | 0.319 |
| *NR5A2* | rs2690036 | 1 | 200141801 | G/A | 0.314 |
| *NR5A2* | rs2246209 | 1 | 200145533 | A/G | 0.357 |
| *NR5A2* | rs2246923 | 1 | 200139350 | C/T | 0.254 |
| *NR5A2* | rs2821367 | 1 | 200016146 | T/C | 0.32 |
| *NR5A2* | rs2737667 | 1 | 200062858 | T/G | 0.25 |
| *NR5A2* | rs10919806 | 1 | 200037727 | C/T | 0.192 |
| *NR5A2* | rs3790844 | 1 | 200007432 | T/C | 0.212 |
| *NR5A2* | rs2821312 | 1 | 200050204 | G/A | 0.467 |
| *NR5A2* | rs2816949 | 1 | 199997778 | A/G | 0.228 |
| *FSHR* | rs11686474 | 2 | 49287983 | T/C | 0.442 |
| *FSHR* | rs11680730 | 2 | 49288060 | G/T | 0.432 |
| *FSHR* | rs12473870 | 2 | 49292341 | G/A | 0.442 |
| *FSHR* | rs12473815 | 2 | 49292362 | C/T | 0.438 |
| *FSHR* | rs6165 | 2 | 49191041 | A/G | 0.403 |
| *SERPINH1* | rs667531 | 11 | 75272716 | G/C | 0.155 |
| *SERPINH1* | rs681390 | 11 | 75274630 | T/C | 0.296 |
| *SERPINH1* | rs603647 | 11 | 75276721 | T/C | 0.46 |
| *FOXP3* | rs2280883 | X | 49109128 | T/C | 0.412 |

SNPs, single nucleotide polymorphisms.

**Table 3.** Significant SNPs

| Gene | SNP | *P* value |
|---|---|---|
| All triads | | |
| *NR5A2* | rs2737667 | 0.02 |
| US triads | | |
| *NR5A2* | rs2737667 | 0.01 |
| *SERPINH1* | rs603647 | 0.02 |
| Gestation age <32 wk | | |
| *NR5A2* | rs12131233 | 0.007 |
| *NR5A2* | rs2737667 | 0.04 |
| *NR5A2* | rs2816949 | 0.02 |
| *FSHR* | rs6165 | 0.01 |
| Spontaneous preterm birth with intact membranes | | |
| *NR5A2* | rs2737667 | 0.0002* |
| Spontaneous preterm birth with PPROM | | |
| *FSHR* | rs6165 | 0.02 |

PPROM, preterm premature rupture of membranes; SNPs, single nucleotide polymorphisms.
*Remained significant after correction for multiple testing (Bonferroni-corrected *P* value < 0.0006)

reports of clinical significance for this variant in the current literature. We can hypothesize that this intronic variant may affect the function of noncoding regulatory element and affect expression of LRH1. Changes in LRH1 expression may result in variations of cholesterol metabolism, steroidogenesis, and progesterone synthesis contributing to PTB.

One SNP in *FSHR* gene (rs6165) showed a *P* value <0.05 in the gestation age <32-wk group and PPROM group, but did not reach formal levels of significance once corrected for multiple comparisons. Notably our results did not replicate the previously described associations of *FSHR* polymorphisms and PTB (19,20). In addition, SNPs in *SERPINH1* did not show an association with PPROM cases as shown previously (21). Previously described associations for both of these genes were in Finnish and African-American cohorts. It is possible that the causative alleles have different frequencies in different populations, which explains why we could not replicate the previous result in our study cohort. Genetic polymorphisms vary greatly between different populations, so as the etiologies of PTB. This explains why the same associations are not found across all populations.

Our study has several limitations. We did not study PTB in relation to different ethnicities as the data on ethnicity were only available for US triads. We had to exclude Denmark maternal triad from PPROM analysis due to the unavailability of PPROM status.

Notable strengths of the study include large sample size across three different regions of the world. We studied the true spontaneous PTB excluding indicated PTB and multiple gestations, leaving a clean cohort of PTBs. We further divided the cohort into spontaneous PTB with intact membranes and PPROM in our sub group analysis, as the mechanisms of PTB can be different in these two groups.

In conclusion, this study suggests a potential association between intronic SNPs in the *NR5A2* gene and PTB. *NR5A2* gene encodes for the LRH1 protein, which plays a critical role in regulation of cholesterol metabolism and steroidogenesis. It also has a role in the regulation of progesterone synthesis. These findings suggest that *NR5A2* may be important in the pathophysiology of PTB and exploring noncoding regulators of *NR5A2* is warranted.

## METHODS

### Study Population

The study population consisted of 728 maternal triads (662 pedigrees and 2,049 individuals) from United States (four sites: University of Iowa Hospitals and Clinics in Iowa City, IA; Magee-Women's Hospital in Pittsburgh, PA; University of Rochester Medical Center in Rochester, NY; and Wake Forest University in Wake Forest, NC), Argentina (two sites: Instituto de Maternidad y Ginecología Nuestra Señora de las Mercedes in Tucumán and Hospital Provincial de Rosario in Rosario), and Denmark (Island of Funen and the Danish National Birth Cohort) (**Table 1**). Inclusion criteria were spontaneous PTB before 37 wk of gestation. Gestational age was based on best obstetrical estimate (last menstrual period or ultrasound examination). Multiple gestation and major fetal chromosomal or structural anomalies were excluded.

All individuals provided signed informed consent for study enrollment in accordance with the protocols approved by research ethics committees in the US (the University of Iowa Institutional Review Board (IRB), University of Pittsburgh IRB, University of Rochester Research

Subjects Review Board, Wake Forest University Health Sciences IRB), Argentina (the Research Ethics Committee of Centro de Educación Médica e Investigaciones Clínicas), and Denmark (the Scientific-Ethical Committee of the Southern Danish Region and the Biomedical Research Ethics Committee of the Capital City Region of Denmark). DNA samples were obtained from the DNA repository at the University of Iowa. Patients were recruited between 1999 and 2013. Neonates and families in this DNA repository were recruited as part of an initiative to create a research bio bank of biologic material for subsequent use in investigations into the genetic contributions of PTB and neonatal diseases. DNA was originally extracted from venous blood, buccal swabs or saliva from the mothers and maternal grandparents. Demographic information was collected through medical chart abstraction.

### SNP Genotyping

A total of 24 SNPs in four candidate genes (*NR5A2, FSHR, FOXP3, SERPINH1*) were studied (**Table 2**). SNPs for *FSHR* and *SERPINH1* were chosen based on previously reported associations with PTB. SNPs for *FOXP3* and *NR5A2* were selected using HapMap database to get 80% coverage of the gene with a minor allele frequency ≥20% for each SNP. SNPs were genotyped using Taqman probes (Applied Biosystems, Foster City, CA) and the Dynamic Array Integrated Fluidic Circuits (Fluidigm, San Francisco, CA). All SNP genotyping assays were available and ordered using the Assay-on-Demand service from Applied Biosystems. These genotyping assays included primers to amplify the region containing the SNP of interest and two TaqMan Minor Groove Binder probes that are specific to the polymorphic variant alleles at the site labeled with different fluorescent reporter dyes, FAM, and VIC. All reactions were performed using standard conditions supplied by Fluidigm. Following thermocycling, fluorescence levels of the FAM and VIC dyes were measured using the EP1 Reader and genotypes were scored using the Fluidigm Genotyping Analysis software.

Genotyping was completed on 728 maternal triads from Denmark (180 triads), Argentina (372 triads), and the United States (176 triads). Genotyping efficiency was >96% for all markers. Greater than 98% of samples tested provided usable genotyping information. Genotypes were entered into a laboratory database (Progeny, South Bend, IN) to generate datasets for analysis.

### Statistical Analysis

Genotyping data were analyzed by transmission disequilibrium test using the Family Based Association Test software (FBAT, Cambridge, MA) (26,27) to identify nonrandom allele transmission from parents to offspring and *P* values were obtained from the maternal effect analysis. We were able to test for maternally mediated genetic effects using the maternal triad as the analysis unit. We analyzed the data for all maternal triads together as well as for each of the populations individually (US, Argentina, and Denmark). We then performed subgroup analysis for gestational age categories (<32 and 32–36 wk) and PPROM status (preterm labor with PPROM and without PPROM). The earliest gestational age was used to determine the gestation age for any mothers who experienced multiple PTB s. Mothers who experienced multiple PTBs and had discrepant PPROM status were excluded from analysis when conducting PPROM analyses. PPROM analysis was only completed for the US and Argentina maternal triads as PPROM data were not available for Denmark triads. To correct for the multiple tests performed in this study, a conservative Bonferroni correction would place significance at $P < 0.0006$ using a standard $\alpha$ of 0.05. However, given the exploratory nature of this initial study, less stringent values are also of interest.

### SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/pr

### ACKNOWLEDGMENTS

We express our sincere thanks to the families that participated in this study. We also thank all the lab members in the Murray, Dagle, and Ryckman laboratories at University of Iowa for their support. This work was supported by March of Dime Foundation (White Plains, NY) #21-FY13-19.

### STATEMENT OF FINANCIAL SUPPORT

The authors have no financial relationships relevant to this article to disclose.

Disclosure: The authors have no conflicts of interest to disclose.

### REFERENCES

1. Beck S, Wojdyla D, Say L, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. Bull World Health Organ 2010;88:31–8.
2. Hamilton BE, Martin JA, Osterman MJ, Curtin SC. Births: final data for 2013. Natl Vital Stat Rep. 2015;64:1–65.
3. Lawn JE, Cousens S, Zupan J; Lancet Neonatal Survival Steering Team. 4 million neonatal deaths: when? where? why? Lancet 2005;365:891–900.
4. Serenius F, Källén K, Blennow M, et al.; EXPRESS Group. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. JAMA 2013;309:1810–20.
5. Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. N Engl J Med 2008;359:262–73.
6. Plunkett J, Muglia LJ. Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. Ann Med 2008;40:167–95.
7. Muglia LJ, Katz M. The enigma of spontaneous preterm birth. N Engl J Med 2010;362:529–35.
8. Adams MM, Elam-Evans LD, Wilson HG, Gilbertz DA. Rates of and factors associated with recurrence of preterm delivery. JAMA 2000;283:1591–6.
9. Winkvist A, Mogren I, Högberg U. Familial patterns in birth characteristics: impact on individual and population risks. Int J Epidemiol 1998;27:248–54.
10. Porter TF, Fraser AM, Hunter CY, Ward RH, Varner MW. The risk of preterm birth across generations. Obstet Gynecol 1997;90:63–7.
11. Kistka ZA, Palomar L, Lee KA, et al. Racial disparity in the frequency of recurrence of preterm birth. Am J Obstet Gynecol 2007;196:131.e1–6.
12. Romero R, Velez Edwards DR, Kusanovic JP, et al. Identification of fetal and maternal single nucleotide polymorphisms in candidate genes that predispose to spontaneous preterm labor with intact membranes. Am J Obstet Gynecol 2010;202:431.e1–34.
13. Myking S, Myhre R, Gjessing HK, et al. Candidate gene analysis of spontaneous preterm delivery: new insights from re-analysis of a case-control study using case-parent triads and control-mother dyads. BMC Med Genet 2011;12:174.
14. Haataja R, Karjalainen MK, Luukkonen A, et al. Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. PLoS Genet 2011;7:e1001293.
15. Plunkett J, Doniger S, Orabona G, et al. An evolutionary genomic approach to identify genes involved in human birth timing. PLoS Genet 2011;7:e1001365.
16. Kim J, Stirling KJ, Cooper ME, et al. Sequence variants in oxytocin pathway genes and preterm birth: a candidate gene association study. BMC Med Genet 2013;14:77.
17. Bream EN, Leppellere CR, Cooper ME, et al. Candidate gene linkage approach to identify DNA variants that predispose to preterm birth. Pediatr Res 2013;73:135–41.
18. Zhang C, Large MJ, Duggavathi R, et al. Liver receptor homolog-1 is essential for pregnancy. Nat Med 2013;19:1061–6.
19. Plunkett J, Doniger S, Orabona G, et al. An evolutionary genomic approach to identify genes involved in human birth timing. PLoS Genet 2011;7:e1001365.
20. Chun S, Plunkett J, Teramo K, Muglia LJ, Fay JC. Fine-mapping an association of FSHR with preterm birth in a Finnish population. PLoS One 2013;8:e78032.
21. Wang H, Parry S, Macones G, et al. A functional SNP in the promoter of the SERPINH1 gene increases risk of preterm premature rupture of membranes in African Americans. Proc Natl Acad Sci USA 2006;103:13463–7.
22. Lazarus KA, Wijayakumara D, Chand AL, Simpson ER, Clyne CD. Therapeutic potential of liver receptor homolog-1 modulators. J Steroid Biochem Mol Biol 2012;130:138–46.
23. Bertolin K, Gossen J, Schoonjans K, Murphy BD. The orphan nuclear receptor Nr5a2 is essential for luteinization in the female mouse ovary. Endocrinology 2014;155:1931–43.

24. Romero R, Nicolaides K, Conde-Agudelo A, et al. Vaginal progesterone in women with an asymptomatic sonographic short cervix in the midtrimester decreases preterm delivery and neonatal morbidity: a systematic review and metaanalysis of individual patient data. Am J Obstet Gynecol 2012;206:124.e1–19.
25. Dodd JM, Jones L, Flenady V, Cincotta R, Crowther CA. Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth. Cochrane Database Syst Rev 2013;7:CD004947.
26. Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype–phenotype associations. Eur J Hum Genet 2001;9:301–6.
27. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol 2000;19:Suppl 1:S36–42.