DR. JEROME F STRAUSS III (Orcid ID : 0000-0001-6199-0480)

Article type : Original Article

Corresponding author mail id: jerome.strauss@vcuhealth.org

MUTATIONS IN FETAL GENES INVOLVED IN INNATE IMMUNITY AND HOST DEFENSE AGAINST MICROBES INCREASE RISK OF PRETERM PREMATURE RUPTURE OF MEMBRANES (PPROM)

Bhavi P. Modi[1], Maria E. Teves[2], Laurel N. Pearson[5], Hardik I. Parikh[3], Hannah Haymond-Thornburg[2], John L. Tucker[2], Piya Chaemsaithong[6], Nardhy Gomez-Lopez[6-9], Timothy P. York[1,2], Roberto Romero[6-9] and Jerome F. Strauss, III[1,2]*

[1]Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA

[2]Department of Obstetrics and Gynecology, Virginia Commonwealth University, Richmond, VA, USA

[3]Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA, USA

[4]Center for Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA, USA

[5]Department of Anthropology, Pennsylvania State University, University Park, PA, USA

[6]Perinatology Research Branch, *Eunice Kennedy Shriver* National Institute for Child Health and Human Development, NIH, Bethesda, MD and Detroit, MI, USA

[7]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, MI, USA

[8]Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI, US

[9]Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI, USA.

Running head: Fetal innate immunity genes and PPROM

Communicating author contact information:
Jerome F. Strauss, III, M.D., Ph.D.
11-029 Sanger Hall
1101 East Marshall Street
Richmond, VA 23298
Phone (804) 828-5598
FAX: (804) 828-5076

ABSTRACT

Background: Twin studies have revealed a significant contribution of the fetal genome to risk of preterm birth. Preterm premature rupture of membranes (PPROM) is the leading identifiable cause of preterm delivery. Infection and inflammation of the fetal membranes is commonly found associated with PPROM.

Methods: We carried out whole exome sequencing (WES) of genomic DNA from neonates born of African-American mothers whose pregnancies were complicated by PPROM (76) or were normal term pregnancies (N=43) to identify mutations in 35 candidate genes involved in innate immunity and host defenses against microbes. Targeted genotyping of mutations in the candidates discovered by WES was conducted on an additional 188 PPROM cases and 175 controls.

Results: We identified rare heterozygous nonsense and frameshift mutations in several of the candidate genes, including *CARD6, CARD8, DEFB1, FUT2, MBL2, NLP10, NLRP12,* and *NOD2*. We discovered that some mutations (*CARD6, DEFB1, FUT2, MBL2, NLRP10, NOD2*) were present only in PPROM cases.

Conclusions: We conclude that rare damaging mutations in innate immunity and host defense genes, the majority being heterozygous, are more frequent in neonates born of pregnancies complicated by PPROM. These findings suggest that the risk of preterm birth in African-Americans may be conferred by mutations in multiple genes encoding proteins involved in

dampening the innate immune response or protecting the host against microbial infection and microbial products.

Key Words: Preterm birth, preterm premature rupture of membranes, innate immunity, anti-microbial peptides, inflammasome, mannose-binding lectin protein, fucosyltransferase, defensins, chorioamnionitis

INTRODUCTION

Preterm birth, especially among African-Americans, has challenged the U.S. health care system for decades (Behrman & Butler, 2007; Kempe et al., 1992; Aveyard et al., 2002; Ahern et al., 2003; Shen et al., 2003). The disparities in prematurity among U.S. populations is thought to be the result of multiple biological and environmental factors (Meis et al., 2000; Anum et al., 2009b; Moutquin, 2003). Preterm premature rupture of membranes (PPROM) is the leading identifiable cause of preterm birth, and more common among African-Americans. Our research has been focused on understanding the pathophysiology of PPROM, and the factors that contribute to population-specific risk (Parry and Strauss, 1998; Strauss, 2013).

The notion that heritable factors play an important role in preterm birth is supported by studies based on twins (Boyd et al., 2009; Svensson et al., 2009; York et al., 2009; York et al. 2010, 2013, 2014, 2015). These studies demonstrated that both the fetal and maternal genomes contribute to the timing of parturition. In addition, there is increasing evidence that gene-environment interactions amplify the effect of specific alleles (Anum et al., 2009b; Wang et al., 2002; Macones et al., 2004). However, the search for maternal and fetal genes linked to preterm birth has yet to produce robust and reproducible candidates. Although association studies have found significant relationships for some candidate genes, the primary reports and available meta-analyses indicate that these associations are weak or population specific (e.g., Genc et al., 2002; Fujimoto et al., 2002; Ferrand et al., 2002b; Lorenz et al., 2002; Moore et al., 2004; Roberts et al., 1999; Romero et al., 2010; Simhan et al., 2003; Witkin et al., 2003; Wang et al., 2004; 2006; 2008; see Sheikh et al., 2016 for a recent review). Moreover, attempts to identify loci contributing to prematurity through genome-wide association studies (GWAS) have not delivered strong candidates (Parets et al., 2015), prompting investigators to pursue alternative approaches to identify genes contributing to preterm birth (Brubaker et al., 2016; Bacelis et al., 2016). Recently, we took a different approach based on the hypothesis that rare mutations or damaging variants in multiple genes (which might escape detection by GWAS or standard association studies, especially with small sample sizes) make significant contributions

to PPROM (Modi et al., 2017). The approach was based on mutation/damaging variant detection using whole exome sequencing (WES), which we applied in this study to explore fetal gene mutations in the innate immune system and PPROM.

Innate immunity encompasses recognition systems that detect molecules derived from bacteria and viruses (Pathogen-Associated Molecular Patterns (PAMPs)) and endogenous alarmins (Damaged-Associated Molecular Patterns (DAMPs)). Pattern recognition receptors (PRRs) responsible for the initiation of innate immune response induced by PAMPs and DAMPs include NOD-like receptor family pyrin domain containing proteins and toll-like receptors (TLR). The response triggered by the PRRs includes activation of transcription of genes that encode cytokines and factors that resolve infection/inflammation (Brubaker et al., 2015). Enhanced production of pro-inflammatory cytokines has been postulated to play a central role in preterm birth and PPROM (Parry & Strauss, 1998; Murtha & Menon, 2015;Gomez-Lopez et al., 2017). The pro-inflammatory cytokines induce expression of matrix metalloproteinases which degrade fetal membrane extracellular matrix leading to rupture (Parry and Strauss, 1998; Strauss, 2013).

The innate immune system is modulated by a number of molecules that dampen/inhibit the inflammatory response triggered by “activating” toll-like receptors and inflammasomes. Bacterial lipids and proteins derived from Gram negative and Gram positive bacteria (PAMPs) reaching the fetal membranes are potent activators of the innate immune response leading to inflammation. Numerous animal studies have shown that Gram negative bacterial lipopolysaccharide (LPS) precipitates preterm birth, and that the fetal membranes possess molecules that recognize bacterial products and trigger an inflammatory response, usually involving the activation of the transcription factor, NFkB (Courtois, 2005). Endogenous enzymes (e.g., acyloxyacyl hydrolase, alkaline phosphatase) protect the host from the potent actions of LPS by altering LPS structure.

A number of endogenous proteins with anti-microbial activity like lactoferrin, mannose-binding lectin 2, and fucosyltransferase 2 help protect exposed surfaces including mucosa, and the fetal membranes. The *FUT2 (OMIM: (* **+**182100) and *MBL2 (OMIM:* ***** 154545) genes are both expressed in the fetal membranes. The defensin family of genes expressed maternally and by the fetus probably combat bacteria ascending from the vagina, but possibly from other sources. Several defensins are known to be produced by fetal membranes including (Avila, 2016).

We analyzed WES data from neonatal DNA from 76 PPROM cases and 43 term controls born of African-American mothers to identify damaging mutations in innate immunity genes and discovered that there was an overrepresentation of these damaging alleles in PPROM cases.

MATERIALS AND METHODS

Study Population: WES was performed on 76 PPROM cases and 43 healthy term control neonatal DNA samples all obtained in Richmond, Virginia. Additional genotyping of select variants was performed on an independent cohort of 188 case and 175 control fetal/neonatal DNA samples collected in Richmond, Virginia and Detroit, Michigan. DNA was isolated from cord blood or umbilical cords. Subjects were self-reported African-American women and their neonates receiving obstetrical care at MCV Hospitals, Richmond, VA (all samples in the initial WES) and Hutzel Hospital in Detroit, MI. The study was approved by the Institutional Review Boards of MCV Hospitals, Richmond, VA (IRB Number: HM15009); Wayne State University (IRB Numbers: 103897MP2F (5R), 082403MP2F (5R), 110605MP4F, 103108MP2F, 052308MP2F) as well as NICHD (National Institute of Child Health and Human Development) (IRB Numbers: 0H97-CH-N065, OH98-CH-N001, OH97-CH-N067, OH99-CH-N056, OH09-CH-N014). Subjects from Hutzel Hospital, Detroit, MI were enrolled under both Wayne State University as well as NICHD protocols and thus respective IRB numbers for both institutes are provided. Written informed consent was obtained from mothers before sample collection. Demographic and clinical data were obtained from surveys and medical records. Control DNA samples (n = 43 + 175) were obtained from neonates of singleton pregnancies delivered at term (> 37 weeks of gestation) of mothers with no prior history of PPROM or preterm labor. Cases of PPROM (n = 76 + 188) were defined as neonates from pregnancies complicated by spontaneous rupture of membranes prior to 37 weeks of gestation. The diagnosis of membrane rupture was based on pooling of amniotic fluid in the vagina, amniotic fluid fern ing patterns and a positive nitrazine test. Women with multiple gestations, fetal anomalies, trauma, connective tissue diseases and medical complications of pregnancy requiring induction of labor were excluded. A DNA biobank at Virginia Commonwealth University and Hutzel Hospital of PPROM cases and term controls collected using the same criteria as those used for for the WES cohort was employed for subsequent genotyping of selected mutations identified by WES (Modi et al., 2017).

Ancestry Estimates: Genetic ancestry was estimated to investigate population structure in the cases and control cohorts. Genetic ancestry estimates were generated in a two-way model of admixture, European and West African, for the neonates of each self-reported African-American study subject using 102 ancestry informative markers (AIMs) single nucleotide polymorphisms with large allele frequency differences between ancestral populations. (Modi et al., 2017). The mean allele frequency difference between ancestral populations for the AIMs panel was delta (δ)= 0.733. The AIMs panel was derived from the overlap of the WES and the Illumina African American Admixture Mapping Panel (Illumina, San Diego, CA) and genotyped using a custom

iPLEX assay (Agena Biosciences, San Diego, CA) for study subjects who were not part of the WES discovery set (Modi et al., 2017). Prior allele frequencies derived from the HapMap West Africans (YRI, Yoruba in Ibadan, Nigeria) and Europeans (CEU, CEPH Utah residents with ancestry from northern and western Europe) were used to estimate individual genetic ancestry estimates following a maximum-likelihood approach.

Whole Exome Sequencing Analysis: Whole exome capture and sequencing was performed at BGI (BGI, Cambridge, MA) using the SureSelect Target Enrichment System Capture Process followed by high-throughput sequencing on an Illimina HiSeq2000 platform with 50-100X coverage. The bioinformatics analysis for variant discovery and annotation was performed as described earlier (Modi et al., 2017). In brief, sequences were mapped to the human reference genome (build hg19) using BWA, followed by marking PCR duplicates using Picard tools and base quality recalibration using GATK (Modi et al., 2017) GATK-HaplotypeCaller was used to identify variants in individual sample, followed by joint genotyping of all samples in the cohort for population-level analysis. The raw SNPs and INDELs were filtered for high quality and annotated for their functional effects using SnpEff tool and known variant databases like dbSNP, ClinVar and the 1000 Genomes Project. Damaging missense variants were selected on the basis of most deleterious predictions in both Polyphen2 (HumDiv - probably damaging) as well as SIFT (damaging) platforms. PCR and Sanger sequencing was used to validate mutations detected by WES (Supplemental Table 1) or mutations were confirmed by custom genotyping.

Custom Genotyping: The variants identified and selected for further analysis from Whole Exome Sequencing were validated and additional samples (an independent cohort of additional 188 cases and 175 controls) were genotyped for the selected variants. Genotyping was performed on the Agena (previously Sequenom) MassArray iPLEX platform following manufacturer's instructions at the University of Minnesota Genomics Center (Modi et al., 2017). The primer sets used for iPlex genotyping are presented in Supplemental Table 2.

Statistical analysis: Mean levels of demographic variables were tested using a 2-tailed Student's t-test. Count data (for gravidity and parity) was square root transformed before performing tests. P-values < 0.05 were considered statistically significant. The paired Wilcoxon rank-sum test were used to assess significant differences in minor allele frequencies.

RESULTS

WES was performed on 76 PPROM and 43 healthy term control neonatal DNA samples. The demographic characteristics of the WES study population is presented in Table 1. The characteristics of the follow-up cohort have been previously reported (Modi et al., 2017). With

152 chromosomes, the probability of detecting a variant with an allele frequency of 0.005 was 78%.

The WES PPROM cases and term controls had similar West African and European ancestry based on genotyping of 102 ancestry informative markers (Means ± S.D.; West African ancestry: PPROM cases: 0.695 ± 0.073 (mean ± S.D.); Term controls 0.698 ± 0.087 ($p>0.10$)).

A total of 35 candidate genes were selected for investigation of nonsense mutations and insertions/deletions causing damaging frameshift mutations (Table 2) based on their involvement in the innate immune response and host defense against microbes. Mutations identified through WES were validated by direct sequence analysis or specific genotyping assays. The mutations were evaluated in an independent cohort of an additional 188 PPROM cases and 175 controls.

Mutations in genes negatively regulating innate immunity: We detected mutations in the *CARD6, CARD8, NLRP10, NLRP12, NOD2,* and *TLR10* genes (Table 3). Several of these were only found in PPROM cases (*CARD6, NLRP10,* and *NOD2*) in both WES and the follow-up genotyping cohorts. The SNP for the *CARD6* nonsense mutation has two alternative alleles C or G. We confirmed by DNA sequence analysis that the PPROM case had the G allele creating the stop codon TAG, which truncates the 1037 amino acid protein at position 560, which retains the caspase activation and recruitment (CARD) domain, but deletes the IMPDH (inosine 5'-monophosphate dehydrogenase/GMP reductase) domain and C-terminal proline-rich domain. This nonsense mutation was detected in 2 PPROM cases (combined WES and follow-up genotyping) and none of the combined term pregnancy controls. The one heterozygous *NLRP10* nonsense mutation detected only in a PPROM case truncates the 655 amino acid protein at position 103. The *NOD2* frameshift mutation truncates the C-terminal 33 amino acids from the 1040 amino acid protein, disrupting a leucine-rich repeat. Mutations in *CARD8, NLRP12* and *TLR10* were found in both PPROM cases and controls.

Mutations in LPS detoxifying enzymes: A nonsense mutation was found in *AOAH*, which encodes an enzyme that catalyzes the hydrolysis of acyloxylacyl-linked fatty acyl chains from LPS. The nonsense mutation disrupts the 688 amino acid protein at position 556, retaining the lipase consensus sequence. This mutation was found in both PPROM cases and term controls.

Mutations in anti-microbial protein genes: A heterozygous nonsense mutation was found in *DEFB1*, which encodes beta-defensin 1, an anti-microbial factor that is produced by amnion epithelial cells. The rs5743490 SNP reference allele is C with two reported alternatives: T, which

results in a synonymous codon change that is functionally not significant, and A which creates a stop codon (TGA). We sequence verified that the allele in our PPROM cases was an A. This stop codon truncates the mature beta defensin 1 peptide sequence after 4 amino acids, so no active peptide made (Porto et al., 2016). Additionally, the translated truncated N-terminal peptide could serve as a dominant negative, competing for the intact signal peptide or processing protease of intact beta-defensin 1 peptide encoded by the other *DEFB1* allele. The heterozygous *DEFB1* mutation was found in 6 PPROM cases (WES and follow-up genotyping combined) and no term controls.

A heterozygous nonsense mutation in *MBL2* was identified which deletes the 38 terminal amino acids in the C-type lectin carbohydrate recognition domain of the 248 amino acid protein. The reference allele of this SNP is a G, with alternate alleles of C, producing a benign missense variant or a T, creates a TAG stop codon. We confirmed by DNA sequence analysis that the minor allele in our PPROM cases was a T. This nonsense mutation was detected in 6 of the total PPROM cases and none of the total term controls. Using RT-PCR, we demonstrated that the *MBL2* gene is expressed in fetal membranes (Supplemental data Figure 1).

Three mutations were discovered in the *FUT2* gene, which encodes a fucosyltransferase involved in protecting epithelium from bacterial infection. One of the nonsense mutations (rs143482452) was found in one PPROM case (combined WES and follow-up genotyping cohort) only and not in the combined term controls. Another one (rs601338) has a relatively high minor allele frequency and was detected in PPROM cases and term controls. The *FUT2* gene is expressed in amnion epithelial cells, and mutations that disrupt the protein cause the "non-secretor" phenotype, which is associated with absent ABH blood groups (Goto et al., 2016).

All of the mutations described above were heterozygous, except for *FUT2* rs601338. In the case of this common mutation, there were 16 homozygous PPROM cases (21%) out of the 76 cases, and 4 homozygous controls (9.3%) out of the 43 term pregnancies. Among this cohort, 7 subjects had di-genic mutations, two with *TLR10* rs62617795 mutation and the *CARD8* mutation; 2 with *AOAH* mutations, one with a *TLR10* rs62617795 mutation, and one with the *CARD8* mutation; and 3 with the *FUT2* rs601338 mutation in combination with either the CARD6 mutation, *MBL2* mutation, and *NLRP12* nonsense mutation.

We found no nonsense or damaging frameshift mutations in *ALPP, BPI, CAMP, DEFA1, DEFB4A, DEFB103A, IL10, IL10RA, IL10RB, LBP, LTF, LYZ, NLRP3, SOCS1, SOCS2, SOCS3, SOCS4, SOCS5, SOCS6, NFKBIA, NFKBIB, NFKBID, NFKBIE, NFKBIZ* and *NOD1* Therefore, these genes did not undergo further interrogation.

Of the 14 mutations identified through WES, 10 had minor allele frequencies in the combined WES and follow-up genotyping cohort that were nominally greater in PPROM cases than term controls. The allele frequency of two mutant alleles were similar in cases and controls, and two mutations were more frequent in controls than PPROM cases. A paired Wilcoxon rank sum test estimated that across loci, variants were overrepresented at PPROM case loci were compared to term controls (Empirical P-value from 10K permutations = 0.0416).

In addition to nonsense and damaging frameshift mutations, a number of rare predicted damaging or known pathogenic missense mutations (e.g., *NOD2* rs34936594) were identified through WES in the candidate genes (Supplemental Table 3). The allele frequencies of these missense mutations were higher in the 76 PPROM cases than the 43 term controls. The association of these predicted rare missense variants with PPROM needs to be replicated with a larger sample size.

DISCUSSION

Our working hypothesis of whether neonatal genes that negatively regulate innate immunity or help the host combat microbes and their noxious products would be more likely to harbor rare, damaging mutations in PPROM cases was supported by our findings. Interestingly, there were a number of important negative regulators of innate immunity and the host defense system that were not mutated (e.g., *IL10, IL10RA, IL10RB NLRP3, SOCS1, SOCS2, SOCS3, SOCS4, SOCS5, SOCS6, NFKBIA, NFKBIB, NFKBID, NFKBIE, NFKBIZ,* and *NOD1*). Of course, the limited WES sample size may have precluded the detection of very rare alleles in these genes.

Inflammasomes and toll-like receptors are critical to host defense mechanisms during the physiological and pathological inflammatory processes in the chorioamniotic membranes that accompany labor. Thus, it is not unexpected that mutations in genes that negatively regulate the inflammasome as well as the toll-like receptors were detected in PPROM cases (Gotsch et al., 2008; Eisenbarth et al., 2012; Oosting et al., 2014).

Mutations in genes encoding host defense mechanisms against microbes had been anticipated based on studies documenting differential expression of the proteins in fetal membranes associated with labor with ruptured and non-ruptured membranes (Erez et al., 2009) Notable in this regard are the rare heterozygous damaging mutations in *DEFB1*, *FUT2* and *MBL2* that were found only in PPROM cases. Variation in these genes have been previously associated with increased risk of infection and in some cases preterm birth (Annells et al., 2005; Gibson et al., 2011; Jaffe et al., 2013).

The discovery of a rare nonsense mutation in the *DEFB1* gene is of interest in that variation in this gene (rs1047031, a SNP in the 3'-UTR) has been associated with chronic and aggressive periodontitis, a condition associated with preterm birth (Schaefer et al., 2010). However, the functional significance of the rs1047031 minor allele has not been established.

Polymorphisms in the *MBL2* gene are more frequent in African-Americans and multiple studies have suggested an association between *MBL2* genetic variants that result in dimished MBL2 protein levels and preterm birth, and conditions commonly found in preterm pregnancies including chorioamnionitis (Annells et al., 2004; Annells et al., 2005; Capece et al., 2014; Gibson et al., 2011; Jaffe et al., 2013; Nedovic et al., 2014). Our discovery of a nonsense mutations that significantly truncates the MBL2 protein is thus consistent with the notion that loss of this anti-microbial protein increases risk of prematurity.

Given the distribution of allele frequencies of *FUT2* mutations we identified, we speculate that the "non-secretor" type is not a strong risk factor for PPROM since the more common mutation was found at allele frequencies that were similar in PPROM cases and controls. It is possible, however, that if both mother and fetus harbor mutations in *FUT2* that there could be an increased risk of PPROM, a possibility that we did not explore.

It is noteworthy that genes associated with inflammatory bowel disease also appear to have an association with PPROM, including *CARD* and *NLRP* genes, *NOD2* and *BRIC2* (Hugot et al., 2001; Andreoletti et al., 2017). Although not included in the 35 candidate genes, a novel heterozygous nonsense mutation in *BIRC2* (NC_000011.10: g.102248476T>G), creating a stop codon at position 539 in this 618 amino acid protein, which deletes the C-terminal zinc finger domain), a gene that negatively regulates the NOD1/NOD2 signaling pathway, and has been recently found to be associated with pediatric inflammatory bowel disease, was discovered in the WES of one PPROM case and no term controls (Andreoletti et al., 2017). A heterozygous damaging frameshift mutation (rs779381525, NC_000010.10 g.49440248_49440249insA) was detected in *FRMPD2*, another gene associated with the NOD2 pathway, in one WES PPROM case.

Chorioamnionitis is often found in PPROM fetal membrane specimens, and the pathways that lead to an accentuated bowel inflammation in Crohn's disease and ulcerative colitis may also contribute to the severity of chorioamnionitis and therefore risk of PPROM. Preterm birth is associated with maternal inflammatory bowel disease but there are no reports that we are aware of that link inflammatory disease in offspring to increased risk of preterm birth and PPROM (Broms et al., 2016; Caruso et al., 2014; Getahun et al., 2014; Palomba et al., 2014; Shand et al., 2016).

We previously examined the association between 2936insC (rs2066847) in the *CARD15/NOD2* gene and PPROM in African-Americans and reported that this frameshift mutation was only found in term controls (Ferrand et al., 2002). This study used genotyping by restriction length polymorphism (RFLP) with digestion with *Nla* IV which cuts the sequence: GGNNCC. We re-evaluated the putative mutations in the control samples previously analyzed using DNA sequencing and discovered that none of them harbored the frameshift mutation, indicating that the RFLP genotyping was flawed. The genotyping methods employed in the present study can distinguish these frameshift mutations, and therefore provides evidence that 2936insC is a risk allele for PPROM.

The mutations that we identified could be spontaneous, or inherited from the father or mother (Li et al., 2017). We speculate that maternal inheritance may be most likely in the setting of PPROM, since an enhanced maternal reproductive tract inflammatory response to bacteria or viruses, or deficiency in endogenous anti-microbial defenses would presumably act in synergy with similar defects in the fetus when both mother and fetus are heterozygous for damaging mutations (Plunkett et al., 2009).

In conclusion, our WES studies, supplemented with additional target genotyping, revealed a number of rare damaging mutations, the majority being heterozygous, that are more frequent in neonates born of pregnancies complicated by PPROM. These findings suggest that the increased risk of preterm birth in African-Americans may be conferred by mutations and damaging missense variants in genes encoding proteins involved in dampening the innate immune response and protecting the host against microbial infection.

ACKNOWLEDGEMENTS

We thank Ms. Sonya Washington for enrolling subjects for the study at MCV Hospitals, Richmond VA. This research was funded by National Institutes of Health Grants R01 HD073555 and P60 MD002256. This research was also supported, in part, by the Perinatology Research Branch, Division of Intramural Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services (NICHD/ NIH); and, in part, with Federal funds from NICHD, NIH under Contract No. HSN275201300006C.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

Ahern J, Pickett KE, Selvin S, Abrams B (2003) Preterm birth among African-American and white women: a multilevel analysis of socioeconomic characteristics and cigarette smoking. J Epidemiol Community Health
57:606-11.

Andreoletti G, Shakhnoich V, Christenson K et al. (2017) Exome anlysis of rare and common variants within the NOD signaling pathway. Sci Reports 7:46454 doi: 10.1038/srep46454.

Annells MF, Hart PH, Mullighan CG et al. (2004) Interleukins-1,-4,-6,-10, tumor necrosis factor, transforming growth factor-β, FAS, and mannose-binding protein C gene polymorphisms in Australian women: Risk of preterm birth. Am J Obstet Gynecol 191, 2056-67.

Annells MF, Hart PH, Mullighan CG, Heatley SL, Robinson JS, McDonald HM (2005). Polymorphisms in immunoregulatory genes and the risk of histologic chorioamnionitis in Caucasoid women: a case control study. BMC Pregnancy Childbirth. 5(1):4.

Anum EA, Hill LD, Pandya A, Strauss JF 3rd. (2009a) Connective tissue and related disorders and preterm birth: clues to genes contributing to prematurity. Placenta. 30(3):207-15.

Anum EA, Springel EH, Shriver MD, Strauss JF 3rd. (2009b) Genetic contributions to disparities in preterm birth. Pediatr Res. 65(1):1-9.

Aveyard P, Cheng KK, Manaseki S, Gardosi J (2002) The risk of preterm delivery in women from different ethnic groups. BJOG 109:894-9.

Avila EE (2016) Functions of antimicrobial peptides in vertebrates. Curr Protein Pept Sci Aug 13, in press.

Bacelis J, Juodakis J, Sengpiel V, Zhang G, Myhre R, Muglia LJ, Nilsson S, Jacobsson B. Literature-Informed Analysis of a Genome-Wide Association Study of Gestational Age in Norwegian Women and Children Suggests Involvement of Inflammatory Pathways. PLoS One. 2016 Aug 4;11(8):e0160335. doi: 10.1371/journal.pone.0160335. Erratum in: PLoS One. 2016 Oct 19;11(10):e0165328.

Behrman RE, Bulter AS (2007) Preterm Birth: Causes, consequences and prevention. National Academy Press, Washington, D.C.

Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. (2009) Maternal contributions to preterm delivery.Am J Epidemiol. 170(11):1358-64.

Bröms G, Granath F, Stephansson O, Kieler H. (2016) Preterm birth in women with inflammatory bowel disease - the association with disease activity and drug treatment. Scand J Gastroenterol. 51(12):1462-1469.

Brubaker SW, Bonham KS, Zanoni I, Kagan JC. (2015) Innate immune pattern recognition: a cell biological perspective. Annu Rev Immunol. 33:257-90.

Brubaker D, Liu Y, Wang J, Tan H, Zhang G, Jacobsson B, Muglia L, Mesiano S, Chance MR. (2016) Finding lost genes in GWAS via integrative-omics analysis reveals novel sub-networks associated with preterm birth. Hum Mol Genet. 25(23):5254-5264.

Capece A, Vasieva O, Meher S, Alfirevic Z, Alfirevic A. (2014) Pathway analysis of genetic factors associated with spontaneous preterm birth and prelabor preterm rupture of membranes. PLoS One Sep 29; 9(9):e108578.doi: 10.1371/journal.pone.01085778.

Caruso R, Warner N, Inohara N, Núñez G. (2014) NOD1 and NOD2: signaling, host defense, and inflammatory disease. Immunity. 41(6):898-908.

Collins-Schramm HE, Chima B, Operario DJ et al. (2003) Markers informative for ancestry demonstrate consistent megabase-length linkage disequilibrium in the African American population. Hum Genet 113:211-219.

Courtois G. (2005) The NF-kappaB signaling pathway in human genetic diseases. Cell Mol Life Sci. 62(15):1682-91.

Eisenbarth SC, Williams A, Colegio OR, Meng H, Strowig T, Rongvaux A, Henao-Mejia J, Thaiss CA, Joly S, Gonzalez DG, Xu L, Zenewicz LA, Haberman AM, Elinav E, Kleinstein SH, Sutterwala FS, Flavell RA. (2012) NLRP10 is a NOD-like receptor essential to initiate adaptive immunity by dendritic cells. Nature. 25;484(7395):510-3. Erratum in: Nature. 2016 Feb 25;530(7591):504.

Erez O, Romero R, Tarca AL, Chaiworapongsa T, Kim YM, Than NG, Vaisbuch E, Draghici S, Tromp G. (2009) Differential expression pattern of genes encoding for anti-microbial peptides in The fetal membranes of patients with spontaneous preterm labor and intact membranes and those with preterm prelabor rupture of the membranes. J Matern Fetal Neonatal Med. 2009 22(12):1103-15.

Ferrand PE, Fujimoto T, Chennathukuzhi V, Parry S, et al. (2002a) The CARD15 2936insC mutation and TLR4 896 A>G polymorphism in African Americans and risk of preterm premature rupture of membranes (PPROM). Mol Human Reprod 8, 1031-34.

Ferrand PE, Parry S, Sammel et al. (2002b) A polymorphism in the matrix metalloproteinase-9 promoter is associated with increased risk of preterm premature rupture of membranes in African Americans. Mol Hum Reprod 8:494-501.

Fujimoto T, Parry S, Urbanek M et al. (2002) A single nucleotide polymorphism in the matrix metalloproteinase-1 (MMP-1) promoter influences amnion cell MMP-1 expression and risk for preterm premature rupture of the fetal membranes. J Biol Chem 277:6296-6302.

Genc MR, Gerber S, Nesin M, Witkin SS (2002) Polymorphism in the interleukin-1 gene complex and spontaneous preterm delivery. Am J Obstet Gynecol 187:157-63.

Getahun D, Fassett MJ, Longstreth GF, Koebnick C, Langer-Gould AM, Strickland D, Jacobsen SJ.(2014) Association between maternal inflammatory bowel disease and adverse perinatal outcomes. J Perinatol. 34(6):435-40.

Gibson CS, Maclennan AH, Haan EA, Priest K, Dekker GA. (2011) Fetal MBL2 haplotypes combined with viral exposure are associated with adverse pregnancy outcomes. J Maternal Fetal Med 24:6, 847-854.

Gomez-Lopez N, Romero R, Xu Y, Garcia-Flores V, Leng Y, Panaitescu B, Miller D, Abrahams VM, Hassan SS. (2017) Inflammasome assembly in the chorioamniotic membranes during spontaneous labor at term. Am J Reprod Immunol. May;77(5). doi: 10.1111/aji.12648. Epub 2017 Feb 24.

Gomez-Lopez N, Romero R, Xu Y, Plazyo O, Unkel R, Leng Y, Than NG, Chaiworapongsa T, Panaitescu B, Dong Z, Tarca AL, Abrahams VM, Yeo L, Hassan SS. (2017) A Role for the Inflammasome in Spontaneous Preterm Labor With Acute Histologic Chorioamnionitis. Reprod Sci. Jan 1:1933719116687656.

Gotsch F, Romero R, Chaiworapongsa T, Erez O, Vaisbuch E, Espinoza J, Kusanovic JP, Mittal P, Mazaki-Tovi S, Kim CJ, Kim JS, Edwin S, Nhan-Chang CL, Hamill N, Friel L, Than NG, Mazor M, Yoon BH, Hassan SS. (2008) Evidence of the involvement of caspase-1 under physiologic and pathologic cellular stress during human pregnancy: a link between the inflammasome and parturition. J Matern Fetal Neonatal Med. 21(9):605-16.

Goto Y, Uematsu S, Kiyono H. (2016) Epithelial glycosylation in gut homeostasis and inflammation. Nat Immunol. 17(11):1244-1251. doi: 10.1038/ni.3587.

Hugot J-P, Chamaillard M, Zouall H et al. (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Chron's disease. Nature 411, 599-603.

Jaffe S, Norman N, Jayaram A, et al. (2013) Unique variation in genetic selection among Black North American women and its potential influence on pregnancy outcomes. Medical Hypotheses 81:919-922.

Kempe A, Wise PH, Barkan SE, Sappenfield WM, Sachs B, Gortmaker SL, Sobol AM, First LR, Pursley D, Rinehart H, et al. (1992) Clinical determinants of the racial disparity in very low birth weight. N Engl J Med. 1;327(14):969-73.

Li J, Oehlert J, Snyder M, Stevenson DK, Shaw GM (2017). Fetal de novo mutations and preterm birth. PLoS Genet. 13(4):e1006689. doi: 10.1371/journal.pgen.1006689.

Liu S, Wen SW, Demissie K, Marcoux S, Kramer MS (2001) Maternal asthma and pregnancy outcomes: a retrospective cohort study. Am J Obstet Gynecol 184:90-6.

Lorenz E, Hallman M, Marttila R, et al. (2002) Association between the Asp299Gly polymorphisms in the Toll- like receptor 4 and premature births in the Finnish population. Pediatr Res 52:373-6.

Macones GA, Parry S, Elkousy M et al. (2004) A polymorphism in the promoter region of TNF and bacterial vaginosis: Preliminary evidence of gene-environment interaction in the etiology of spontaneous preterm birth. Am J Obstet Gynecol 190:1504-8.

Meis PJ, Goldenberg RL, Mercer BM et al. (2000) Preterm prediction study: is socioeconomic status a risk factor for bacterial vaginosis in Black or in White women? Am J Perinatol 17:41-5.

Modi BP, Teves ME, Pearson LN, Parikh HI, Chaemsaithong P, Sheth NU, York TP, Romero R, Strauss JF 3rd. (2017b) Rare mutations and potentially damaging missense variants in genes encoding fibrillar collagens and proteins involved in their production are candidates for risk for preterm premature rupture of membranes. PLoS One. Mar 27;12(3):e0174356. doi: 10.1371/journal.pone.0174356. eCollection 2017 Mar 27.

Moore S, Ide M, Randhawa M et al. (2004) An investigation into the association among preterm birth, cytokine gene polymorphisms and periodontal disease. BJOG 111:125-32.

Moutquin JM (2003) Socio-economic and psychosocial factors in the management and prevention of preterm labour. BJOG 110 Suppl 20:56-60.

Murtha AP, Menon R. (2015) Regulation of fetal membrane inflammation: a critical step in reducing adverse pregnancy outcome. Am J Obstet Gynecol Oct; 213(4): 447-448.

Nedovic B, Posteraro B, Leoncini E, Ruggeri A, Amore R, Sanguinetti M, Ricciardi W, Boccia S. (2014) Mannose-binding lectin codon 54 gene polymorphism and vulvovaginal candidiasis: a systematic review and meta-analysis. Biomed Res Int. 2014:738298.

Oosting M, Cheng SC, Bolscher JM, Vestering-Stenger R, Plantinga TS, Verschueren IC, Arts P, Garritsen A, van Eenennaam H, Sturm P, Kullberg BJ, Hoischen A, Adema GJ, van der Meer JW, Netea MG, Joosten LA. (2014) Human TLR10 is an anti-inflammatory pattern-recognition receptor. Proc Natl Acad Sci U S A. 21;111(42):E4478-84.

Palomba S, Sereni G, Falbo A, et al. (2014) Inflammatory bowel diseases and human reproduction: A comprehensive evidence-based review. World J Gastroenterol 20, 7123-7136.

Parets SE, Knight AK, Smith AK. (2015) Insights into genetic susceptibility in the etiology of spontaneous preterm birth.Appl Clin Genet. 2015 Dec 14;8:283-90. doi: 10.2147/TACG.S58612.

Parry S, Strauss III JF (1998) Premature rupture of the fetal membranes. NEJM 338:663-670.

Plunkett J, Feitosa MF, Trusgnich M, Wangler MF, Palomar L, Kistka ZA, DeFranco EA, Shen TT, Stormo AE, Puttonen H, Hallman M, Haataja R, Luukkonen A, Fellman V, Peltonen L, Palotie A, Daw EW, An P, Teramo K, Borecki I, Muglia LJ. (2009) Mother's genome or maternally-inherited genes acting in the fetus influence gestational age in familial preterm birth. Hum Hered. 68(3):209-19.

Porto WF, Nolasco DO, Pires AS, Pereira R, Franco OL, Alencar SA (2016) Biopolymers (Peptide Science) 106: 633-644

Roberts AK, Monzon-Bordonaba F, Van Deerlin PG et al. (1999) Association of polymorphism within the promoter of the tumor necrosis factor alpha gene with increased risk of preterm premature rupture of the fetal membranes. Am J Obstet Gynecol 180:1297-302.

Romero R, Friel LA, Velez Edwards DR, Kusanovic JP, Hassan SS, Mazaki-Tovi S, Vaisbuch E, Kim CJ, Erez O, Chaiworapongsa T, Pearce BD, Bartlett J, Salisbury BA, Anant MK, Vovis GF, Lee MS, Gomez R, Behnke E, Oyarzun E, Tromp G, Williams SM, Menon R. (2010) A genetic association study of maternal and fetal candidate genes that predispose to preterm prelabor rupture of membranes (PROM). Am J Obstet Gynecol. 203(4):361.e1-361.e30.

Romero R, Xu Y, Plazyo O, Chaemsaithong P, Chaiworapongsa T, Unkel R, Than NG, Chiang PJ, Dong Z, Xu Z, Tarca AL, Abrahams VM, Hassan SS, Yeo L, Gomez-Lopez N. (2016) A Role for the Inflammasome in Spontaneous Labor at Term. Am J Reprod Immunol. Mar 8. doi: 10.1111/aji.12440. [Epub ahead of print]

Schaefer AS, G M Richter GM, Nothnagel M, et al. (2010) A 3′ UTR transition within *DEFB1* is associated with chronic and aggressive periodontitis Genes and Immunity 11**,** 45–54.

Shand AW, Chen JS, Selby W, Solomon M, Roberts CL. (2016) Inflammatory bowel disease in pregnancy: a population-based study of prevalence and pregnancy outcomes. BJOG. 123(11):1862-70.

Sheikh IA, Ahmad E, Jamal MS, Rehan M, Assidi M, Tayubi IA, AlBasri SF, Bajouh OS, Turki RF, Abuzenadah AM, Damanhouri GA, Beg MA, Al-Qahtani M. Spontaneous preterm birth and single nucleotide gene polymorphisms: a recent update. BMC Genomics. 2016 Oct 17;17(Suppl 9):759.

Shen TT, DeFranco EA, Stamilio DM, Chang JJ, Muglia LJ. (2008) A population-based study of race-specific risk for preterm premature rupture of membranes. Am J Obstet Gynecol. 199(4):373.e1-7.

Simhan HN, Krohn MA, Roberts JM, et al. (2003) Interleukin-6 promoter -174 polymorphism and spontaneous preterm birth. Am J Obstet Gynecol. 189:915-8.

Strauss JF 3rd. (2013) Extracellular matrix dynamics and fetal membrane rupture. Reprod Sci. 20(2):140-53.

Svensson AC, Sandin S, Cnattingius S, Reilly M, Pawitan Y, Hultman CM, Lichtenstein P. (2009) Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families. Am J Epidemiol. 170(11):1365-72.

Wang H, Ogawa M, Wood JR, Bartolomei MS, Sammel MD, Kusanovic JP, Walsh SW, Romero R, Strauss JF 3rd. (2008) Genetic and epigenetic mechanisms combine to control MMP1 expression and its association with preterm premature rupture of membranes. Hum Mol Genet. 2008 Apr 15;17(8):1087-96.

Wang H, Parry S, Macones G, Sammel MD, Ferrand PE, Kuivaniemi H, Tromp G, Halder I, Shriver MD, Romero R, Strauss JF 3rd. (2004) Functionally significant SNP MMP8 promoter

haplotypes and preterm premature rupture of membranes (PPROM). Hum Mol Genet. 13(21):2659-69.

Wang H, Parry S, Macones G, Sammel MD, Kuivaniemi H, Tromp G, Argyropoulos G, Halder I, Shriver MD, Romero R, Strauss JF 3rd. (2006) 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 U S A. 103(36):13463-7.

Wang X, Zuckerman B, Pearson C et al. (2002) Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. JAMA 287:195-202.

Witkin SS, Vardhana S, Yih M, et al. (2003) Polymorphism in intron 2 of the fetal interleukin-1 receptor antagonist genotype influences midtrimester amniotic fluid concentrations of interleukin-1beta and interleukin-1 receptor antagonist and pregnancy outcome. Am J Obstet Gynecol. 189:1413-7.

York TP, Eaves LJ, Lichtenstein P, Neale MC, Svensson A, Latendresse S, Långström N, Strauss JF 3rd. Fetal and Maternal Genes' Influence on Gestational Age in a Quantitative Genetic Analysis of 244,000 Swedish Births. Am J Epidemiol. 2013 Apr 7;

York TP, Eaves LJ, Neale MC, Strauss JF 3rd. The contribution of genetic and environmental factors to the duration of pregnancy. Am J Obstet Gynecol. 2014 May;210(5):398-405.

York TP, Strauss JF 3rd, Eaves LJ. A narrow heritability evaluation of gestational age at birth. Hum Genet. 2015 Jul;134(7):809-11.

York TP, Strauss JF 3rd, Neale MC, Eaves LJ. Racial differences in genetic and environmental risk to preterm birth. PLoS One. 2010 Aug 25;5(8):e12391.

York TP, Strauss JF 3rd, Neale MC, Eaves LJ. (2009) Estimating fetal and maternal genetic contributions to premature birth from multiparous pregnancy histories of twins using MCMC and maximum-likelihood approaches. Twin Res Hum Genet. 12(4):333-42.

609 **Supporting Material**

610

611 Supplemental Table 1. Primers used for mutation verification by DNA sequence analysis

612 Supplemental Table 2. iPLEX Genotyping Design

613

614 Supplemental Table 3. Predicted Damaging SNPs in Innate Immunity Genes

615 Supplemental Fig 1. *MBL2* mRNA expression in fetal membrane samples from normal term
616 pregnancy.

617

Table 1: Study subject characteristics for WES

Characteristic	Cases Mean (SD)	Controls Mean (SD)	p-value
Maternal Age (years)	27.18 (5.33)	26.02 (5.32)	0.256
Gestational Age at Delivery (weeks)	30.05 (4.17)	38.93 (1.16)	<0.001
Neonatal Weight (kgs)	1.69 (1.59)	3.14 (0.46)	<0.001
Gravidity	3.53 (2.04)	3.25 (2.57)	0.555
Parity	1.47 (1.57)	1.35 (1.41)	0.657

PPROM cases, N=76; Term controls, N=43

Table 2. Candidate genes selected for analysis

Category	Gene IDs and (OMIM number)

	CARD6 (* 609986), *CARD8* (* 609051)
	IL10 (* 124092), *IL10RA* (* 146933)
	IL10RB (* 123889), *NFKBIA* (* 164008)
	NFKBIB (* 604495), *NFKBID,*
	NFKBIE (* 604548), *NFKBIZ* (* 608004)
Innate immune response modulators	*NLRP3* (* 606416), *NLRP10* (* 609662)
	NLRP12 (* 609648), *NOD1*(* 605980)
	NOD2 (* 605956), *TLR10* (* 606270)
	SOCS1(* 603597), *SOCS2* (* 605117)
	SOCS3 (* 604176), *SOCS4* (* 616337)
	SOCS5 (* 607094), *SOCS6* (* 605118)
LPS detoxification	*ALPP* (* 171800), *AOAH* (* 102593)

Table 3. Damaging mutations identified in genes involved in modulation of the innate immune response in PPROM cases.

Gene ID	Chromosome	Position	SNP ID	Ref-Allele	Alternate-Allele	Effect	Minor Allele	AA Position (residue change)	Minor Allele Frequency Cases/Controls	Sequence Variant
AOAH	7p14.2	36514524	rs145455591	C	T	Nonsense	T	556	0.036/0.026	NC_000007.14:g.36514524C>T
CARD6	5p13.1	40853011	rs150487186	T	G	Nonsense	G	560	0.004/0.000	NC_000005.10:g.40853011T>G
CARD8	19q13.33	48231760	rs140826611	-	AA	Frameshift	AA	148	0.016/0.006	NC_000019.10:g.48231760_48231761insAA
DEFB1	8p23.1	6870777	rs5743490	C	A	Nonsense	A	37	0.011/0.000	NC_000008.11:g.6870777C>A
FUT2	19q13.3	48703417	rs601338	G	A	Nonsense	A	154	0.374/0.376	NC_000019.10:g.48703417G>A
FUT2	19q13.3	48703041	rs143482452	C	T	Nonsense	T	29	0.002/0.000	NC_000019.10:g.48703041C>T
FUT2	19q13.3	48703767	rs1799761	C	-	Frameshift	C	271	0.007/0.012	NC_000019.10:g.48703767delC
MBL2	10q21.1	52768256	rs74754826	G	T	Nonsense	T	210	0.011/0.000	NC_000010.11:g.52768256G>T
NLRP10	11p15.4	7961305	rs765522475	C	T	Nonsense	T	103	0.002/0.000	NC_000011.10:g.7961305C>T
NLRP12	19q13.42	53795911	rs35064500	C	T	Nonsense	T	1017	0.021/0.007	NC_000019.10:g.53795911C>T
NLRP12	19q13.42	53795917	rs776426826	AG	-	Frameshift	-	1015	0.002/0.003	NC_000019.10:g.53795917_53795918delAG
NOD2	16q12.1	50729867	rs2066847	-	C	Frameshift	C	1007	0.004/0.000	NC_000016.10:g.50729867_50729868insC
TLR10	4p14	38774483	rs62617795	C	T	Nonsense	T	370	0.020/0.016	NC_000004.12:g.38774483C>T
TLR10	4p14	38775590	rs140873456	A	G	Start loss	G	1	0.003/0.003	NC_000004.12:g.38775590A>G

Mutations identified through WES (76 PPROM, 43 term controls) were validated by direct sequence analysis or genotyping using TaqMan reagents. The mutations were evaluated in an independent cohort of an additional 188 PPROM cases and 175 controls. Genotyping was performed on the Agena MassArray iPLEX platform. All allele frequencies were based on called genotypes excluding missing samples or those samples without a genotype call. MAF=minor allele frequency.