HHS Public Access

Author manuscript

Arch Dis Child Fetal Neonatal Ed. Author manuscript; available in PMC 2017 September 01.

Published in final edited form as:

Arch Dis Child Fetal Neonatal Ed. 2017 September ; 102(5): F439–F445. doi:10.1136/archdischild-2016-311545.

Genome Wide Association Study of Sepsis in Extremely Premature Infants

Lakshmi Srinivasan, MBBS, MTR[1], **Grier Page, PhD**[2], **Haresh Kirpalani, BM, MSc**[1], **Jeffrey C. Murray, MD**[3], **Abhik Das, PhD**[4], **Rosemary D. Higgins, MD**[5], **Waldemar A. Carlo, MD**[6], **Edward F. Bell, MD**[3], **Ronald N. Goldberg, MD**[7], **Kurt Schibler, MD**[8], **Beena G. Sood, MD, MS**[9], **David K. Stevenson, MD**[10], **Barbara J. Stoll, MD**[11], **Krisa P. Van Meurs, MD**[10], **Karen J. Johnson, RN**[3], **Joshua Levy, MS**[2], **Scott A. McDonald, BS**[2], **Kristin M. Zaterka-Baxter, RN, BSN**[2], **Kathleen A. Kennedy, MD, MPH**[12], **Pablo J. Sánchez, MD**[13], **Shahnaz Duara, MD**[14], **Michele C. Walsh, MD, MS**[15], **Seetha Shankaran, MD**[16], **James L. Wynn, MD**[17], and **C. Michael Cotten, MD, MHS**[7] **for the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network**

[1]Department of Pediatrics, The Children's Hospital of Philadelphia and The University of Pennsylvania, Philadelphia, PA

[2]Social, Statistical and Environmental Sciences Unit, RTI International, Research Triangle Park, NC

[3]University of Iowa, Department of Pediatrics, Iowa City, IA

[4]Social, Statistical and Environmental Sciences Unit, RTI International, Rockville, MD

[5]*Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD

[6]Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL

[7]Department of Pediatrics, Duke University, Durham, NC

[8]Department of Pediatrics, University of Cincinnati, Cincinnati, OH

[9]Department of Pediatrics, Wayne State University, Detroit, MI

[10]Department of Pediatrics, Division of Neonatal and Developmental Medicine, Stanford University School of Medicine and Lucile Packard Children's Hospital, Palo Alto, CA

[11]Emory University School of Medicine, Department of Pediatrics, Children's Healthcare of Atlanta, Atlanta, GA

[12]Department of Pediatrics, University of Texas Medical School at Houston, Houston, TX

[13]Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX

Contact information for the corresponding author (and reprint request author): Lakshmi Srinivasan, MBBS MTR, Clinical Associate in Pediatrics, The Children's Hospital of Philadelphia, Room 2NW17, 34th and Civic Center Boulevard, Philadelphia, PA 19104, srinivasanl@email.chop.edu, Phone: 215-590-1653, Fax: 215-590-3051.

CONFLICTS OF INTEREST STATEMENT:

The authors do not have any conflicts of interest to report.

[14]University of Miami Miller School of Medicine, Miami, FL

[15]Department of Pediatrics, Rainbow Babies & Children's Hospital, Case Western Reserve University, Cleveland, OH

[16]Department of Pediatrics, Wayne State University, Detroit, MI

[17]Department of Pediatrics, University of Florida, Gainesville, FL

Abstract

Objective—To identify genetic variants associated with sepsis (early and late-onset) using a genome wide association (GWA) analysis in a cohort of extremely premature infants.

Study Design—Previously generated GWA data from the Neonatal Research Network's anonymized genomic database biorepository of extremely premature infants were used for this study. Sepsis was defined as culture-positive early-onset or late-onset sepsis or culture-proven meningitis. Genomic and whole genome amplified DNA was genotyped for 1.2 million single nucleotide polymorphisms (SNPs); 91% of SNPs were successfully genotyped. We imputed 7.2 million additional SNPs. P values and false discovery rates were calculated from multivariate logistic regression analysis adjusting for gender, gestational age and ancestry. Target statistical value was $p<10^{-5}$. Secondary analyses assessed associations of SNPs with pathogen type. Pathway analyses were also run on primary and secondary end points.

Results—Data from 757 extremely premature infants were included: 351 infants with sepsis and 406 infants without sepsis. No SNPs reached genome-wide significance levels (5×10^{-8}); two SNPs in proximity to FOXC2 and FOXL1 genes achieved target levels of significance. In secondary analyses, SNPs for ELMO1, IRAK2 (Gram positive sepsis), RALA, IMMP2L (Gram negative sepsis) and PIEZO2 (fungal sepsis) met target significance levels. Pathways associated with sepsis and Gram negative sepsis included gap junctions, fibroblast growth factor receptors, regulators of cell division and Interleukin-1 associated receptor kinase 2 (p values<0.001 and FDR<20%).

Conclusions—No SNPs met genome-wide significance in this cohort of ELBW infants; however, areas of potential association and pathways meriting further study were identified.

List of keywords

infection; ELBW; extreme prematurity; genetics

BACKGROUND

Sepsis is a potentially life threatening illness, affecting 10–20% of neonates worldwide, with a higher incidence in preterm and very low birth weight infants.[12] Up to one-third of extremely premature infants develop sepsis, which carries a high mortality and morbidity.[1–3]

Susceptibility to neonatal sepsis is mediated by complex interactions between environmental and maternal factors and the neonatal immune response, which may be modified by differences in genetic composition or function.[3–5] Twin studies and ethnic variations suggest that host genetic factors may contribute to susceptibility to sepsis.[4–8] Genetic studies of

sepsis in newborns have previously used candidate gene approaches, with varied findings. [9–14]

While genome-wide association studies (GWAS) have been performed in adult sepsis, no GWAS has reported genetic variation in risk of sepsis in extremely preterm infants.[15–17] An agnostic GWAS approach rather than a study on known variant candidate genes has potential to identify gene variants across many genes (1 million or more variants), including disease associations with previously unknown genes. In addition, investigators do not possess sufficient knowledge of the function of all genes to predict candidate genes well. The high rate of sepsis in extremely low birth weight (ELBW) neonates establishes them as a unique population for unbiased study of genetic associations with sepsis.[4]

METHODS

Study population and data source

We performed a secondary analysis utilizing an anonymized DNA biorepository from a prior study of cytokines and neurodevelopmental outcomes in ELBW infants conducted by the Eunice Kennedy Shriver NICHD Neonatal Research Network (NRN).[18,19] Enrolled infants were <1000 grams birth weight and <72 hours old at enrollment; infants with major congenital anomalies were excluded. The NRN's anonymized DNA database and biorepository was previously approved by all site institutional review boards (IRBs) for candidate gene analyses and genome wide analyses of diseases associated with extreme prematurity, including sepsis.[20] We received a waiver of IRB authorization for use of this anonymized data set. In this multi-racial and multi-ethnic study, blood spot samples were collected on more than 800 ELBW infants and analyzed using the Illumina Omni-1 Quad array.[20]

Primary Outcome

Our primary outcome was the identification of single nucleotide polymorphisms (SNPs) significantly associated with the occurrence of all sepsis in extremely preterm infants.

Secondary Outcomes

Given that EOS and LOS have distinct etiologies and likely differences in mechanisms underpinning their pathogenesis, we planned *a priori* to identify SNPs significantly associated with late-onset sepsis (LOS) alone.[3,21] Due to the relatively low incidence of early-onset sepsis (EOS), it was not feasible to examine SNPs associated with EOS alone. We also identified SNPs associated with Gram positive, Gram negative, and fungal pathogens. While infants who experienced multiple episodes of sepsis with different bacterial pathogens were included only once in the primary analysis, they were included in each pathogen category that they fit in secondary analyses. Finally, we also reviewed SNPs tested in previous candidate gene association studies in neonatal sepsis.

Study definitions

Cases consisted of infants with one or more episodes of culture-proven sepsis (EOS or LOS) or meningitis, treated with antibiotics for ≥ 5 days, or, who died before treatment was

completed.[12] Controls were infants not classified as cases. Infants with positive blood or CSF cultures were defined as having culture proven sepsis. EOS comprised infants with culture proven sepsis in the first 72 hours of life, while LOS comprised cases beyond that timeframe. All subjects included in the study were followed through the duration of their hospital course for occurrence of sepsis.

Sample source and storage

Our study utilized DNA data obtained from samples originally collected as part of a prospective multicenter cohort study of cytokines in premature infants (the Cytokine Study, performed in 17 centers of the National Institute of Child Health and Human Development Neonatal Research Network).[18] In the parent study, whole blood spots were obtained on filter paper within 4 hours after birth and at serial timepoints (days 3, 7, 14, 21). These filter paper blood spots were frozen as soon as possible. For our study and other secondary genomic analyses performed utilizing this dataset, DNA was extracted from filter paper blood spots for the earliest age available for each subject, amplified and analyzed using the Illumina Omni-1 Quad array, as described below. Additional details regarding sample methods are available in prior studies utilizing this sample set.[1820]

Genomic analysis

Genomic and whole genome amplified DNA were genotyped on the Human-OMNI1-Quad_v1-0_B BeadChip platform (Illumina, San Diego, California). Initially, 1.2 million SNPs were tested, of which over 900,000 (91%) were successfully genotyped. We employed standard quality control thresholds, and included samples and SNPs with > 97% call rate, while excluding minor allele frequencies < 1%. Additionally, 7.2 million SNPs were imputed against the HapMap 3 cosmopolitan reference panel with Impute2.[20] All SNP-gene assignments were made based on the UCSC gene tracks within the GRCh37/hg19 genome build.[2223] Genotypes, and a subset of subject information inclusive of key characteristic and outcome phenotype data have been stored, with no identifying links, in the National Human Genome Research Institute Database of Genotypes and Phenotypes (dbGaP Study Accession: phs000353.v1.p1).[1920].

Data Analysis

Following standard quality checks, single and multiple logistic regression models were developed for the primary and secondary analyses. Covariates included in the multiple regression model included gestational age, small for gestational age (SGA) status, ancestry, surgery, C-section, and the first four eigenvalues for ancestry (described below). In previous analyses of this dataset, these covariates were determined to be the most important predictors of sepsis.[8] To minimize false discoveries, we established a conservative significance level of 10^{-5} as the ‘target’ for reporting, while a p value of 5×10^{-8} is regarded as a “genome-wide” association. The latter value is widely accepted to be the threshold in genome wide association studies that would adequately adjust for the multiple comparisons involved, and is estimated to correspond to a p value of 0.05 in a classical epidemiological study testing a single hypothesis.[22–25]

We used this two-tiered approach in order to screen for genes of interest, which may be missed if only the more stringent threshold were applied. When SNPs met threshold significance, we did not limit reporting based on FDR thresholds, in order to better identify any genes of putative biologic importance that should be examined closely in future validation studies. [25] P values were calculated per SNP for a one degree of freedom additive model. For the secondary analysis of previously identified candidate SNPs, we reported p values <0.05 as trends of interest.

Correction for population stratification

Spurious associations may be obtained, or true associations overlooked, if allele frequencies differ between subpopulations and are represented unequally between cases and controls due to systematic ancestry differences. [26–28] The first four eigenvalues derived from linkage disequilibrium pruned for genotyped SNPs were included as covariables in primary and secondary analyses, to adjust for individual ancestry (eigenvalues generated using GWAS-Tools). [20]

Exploratory pathway analysis

We assigned genes to biological pathways using the Reactome database (www.reactome.org), using the adaptive rank truncated product (ARTP) method (pathway tools package ARTP).[29–31] SNPs were notated to genes based on their location within 50 kb of the gene models defined in Refseq, and p-values assigned based on the results of the GWA analysis. A summary of the association between each gene and sepsis was generated, based on the multiple-testing adjusted P-value associated with the most significant SNP within that gene. The ARTP method was used to combine gene-level P-values into a test statistic for the pathway association with sepsis.[31] Given the exploratory nature of this analysis, we designated pathways with p values <0.005 as the cut-off for reporting, and reported false discovery rates (FDR) of less than 20% to facilitate assessment of the strength of the evidence.[25]

RESULTS

Data from 751 ELBW infants passed quality control checks and were included in the analysis. This cohort included 345 infants with sepsis (Supplementary Table 1). Most infants with sepsis had experienced LOS (320 infants, 393 episodes). Ninety-nine infants experienced more than one episode of sepsis. Gram positive pathogens were present in more than three-fourths of the LOS events (252 episodes), while Gram negative and fungal sepsis occurred with lower frequency (72 and 69 episodes respectively, Supplementary Table 1). Infants could be positive for more than one organism per sepsis event. Of the infants with sepsis, 42 were diagnosed with culture proven meningitis. Eighty-one infants in the overall cohort developed necrotizing enterocolitis (NEC), including 40 cases of surgical NEC. NEC events occurred more frequently in infants with sepsis compared to uninfected infants (14.1% versus 8.7%, p = 0.02, Chi square test); however the dataset lacked information on timing of NEC in relation to sepsis event.

Primary Outcome

In the primary analysis of overall sepsis, no SNPs reached a priori genome-wide significance levels (5×10^{-8}); however, several achieved p values of 10^{-5} (Table 1). While these SNPs were generally noted to be in intergenic regions, several were located on chromosome 16, in proximity to genes for FOXC2 (forkhead box protein C2) and FOXL1 (forkhead box protein L1) (Supplementary Figure 1). The analysis limited to LOS demonstrated similar results.

Secondary Outcomes

Examining Gram positive LOS only, SNPs that achieved target level of associations ($<10^{-5}$) included introns for ELMO1 (engulfment and cell motility 1) on chromosome 7, and introns for VIL1 (villin 1, chromosome 2) (Table 2, Supplementary Figures 2a–b). For Gram negative LOS, SNPs in introns for RALA (Ras-like protein A) and IMMP2L (inner mitochondrial membrane peptidase-2 like), located on chromosome 7, and several SNPs in the intergenic area of chromosome 2 achieved the highest levels of association (Table 3). In the analysis of fungal LOS, over-represented SNPs included introns for PIEZO2 (piezo-type mechanosensitive ion channel component 2, chromosome 18), and several SNPs in the intergenic area of chromosome 10 (Table 4). In the analysis of previously tested candidate gene SNPs, we identified 62 SNPs from previous gene association studies of neonatal sepsis. Several SNPs demonstrated p values <0.05, but none reached target or genome wide levels of significance (Supplementary Table 2).

Findings from pathway analyses

We tested 1308 Reactome pathways. For Gram negative sepsis, 30 pathways had a p value <0.005 and FDR<20% (Supplementary Table 3). Pathways identified involved gap junction genes, fibroblast growth factor receptor and insulin receptor signaling. No pathways for Gram positive sepsis or fungal sepsis met FDR thresholds for reporting.

DISCUSSION

We report the first GWA study of sepsis in ELBW infants. In this unique cohort, we identified several SNPs of potential association, although none met genome-wide significance. The results contain potential clues to newborn genetic susceptibility to sepsis, with several SNPs identified related to innate immunity. Over-represented SNPs were found in FOXC2, FOXL1 (all sepsis), ELMO1, VIL1 (Gram positive sepsis), RALA, IMMP2L (Gram negative sepsis) and PIEZO2 (fungal sepsis).

Sepsis is a major co-morbidity in the ELBW population, with up to 36% of infants experiencing at least one episode during their hospital stay.[1,3,23,3] Immature innate and adaptive immune defenses, and environmental factors, contribute to a complex multifactorial etiology.[18] Recent evidence also suggests genetic susceptibility to infection in preterm neonates. [6–8]

Epidemiologic and twin studies indicate a heritable component in certain infections and their complications in adult populations.[3,4,35] In neonates, a twin study suggested that genetic risk factors were responsible for 49% of variability in sepsis risk;[7] however this was not

confirmed in a subsequent study of multiple births.[8] Several candidate gene studies examined the association of neonatal sepsis with biologically plausible SNPs in cytokines, cytokine receptors, pathogen recognition receptors, cluster differentiation molecules, endothelial and hemostasis markers.[10–1336–38] These studies demonstrated varying results, described in a recent systematic review.[14]

We performed this study to explore genetic associations in neonatal sepsis with an unbiased genome wide approach. The NRN's anonymized genomic database provides a sizable cohort of ELBW infants with genotype and phenotype data. In the analysis of all sepsis, we found over-represented SNPs in proximity to genes encoding forkhead box protein family transcription factors FOXC2 and FOXL1. These genes are involved in intestinal mucosal structure and function, including development of Peyer's patches; alterations could plausibly contribute to risk of LOS via microbial translocation, and NEC.[39–42]

There are likely differences in diagnosis, and in the interaction between the environment and genome in EOS as compared to LOS.[21] To assess LOS as a distinct disease entity, we performed a secondary analysis to identify SNPs associated with LOS alone. As LOS comprised the majority of sepsis cases, the findings of this analysis were very similar to those previously mentioned in all sepsis. We also examined SNPs associated with known specific pathogen categories in LOS. In the analysis of Gram positive LOS, over-represented SNPs were found in engulfment and cell motility 1 (ELMO1) (chromosome 7), which promotes phagocytosis and cell migration (Table 2).[43–45]

Our exploratory analysis of SNPs that have been previously tested in candidate gene association studies yielded trends of interest in gene variants located in TNF, IL-6, IL-10RA, TLR-2 and TLR-4 (Supplementary Table 2). While none reached target levels of significance in our study, these trends warrant further exploration in replication cohorts.

Traditional GWAS and candidate gene analyses examine significance at the level of individual SNPs and genes. This approach risks missing modest but more meaningful changes in multiple genes involved in a biological pathway.[46] In contrast, a pathway analysis approach uncovers enrichment of 'significant' SNPs in groups of related genes.[2930] The strongest pathway associations were noted in Gram negative sepsis, where pathways involving regulation of gap junctions and connexin function had low FDR estimates (Supplementary Table 3). Connexin proteins provide the framework for gap junction hemichannels, which play an important role in migration and intracellular signaling in immune cells.[47] Downregulation of gap junction channels results from inflammation and infection.[48–51] Several pathways regulating fibroblast growth factor receptors were also significantly enriched (Supplementary Table 3), and are implicated in attachment and replication of certain pathogens.[52–55] Pathways related to gap junction channels and fibroblast growth factor receptors may play important roles in mediating respiratory and gut mucosal integrity.[505155] On analyzing expression quantitative trait loci (eQTL), ELMO1 levels were noted to be regulated mostly in the gastrointestinal tract, in addition to vasculature. This finding preliminarily suggests a role for alteration in gastrointestinal ELMO1 expression in sepsis. Lastly, IRAK-2, a known component of the classical pro-

inflammatory signaling pathway activated by infection, was also identified as enriched in patients that developed Gram positive sepsis.[57]

Our moderately powered study did not identify SNPs that achieved significance at the genome-wide level, indicating that genetic susceptibility to sepsis may be mediated by more than just one or a few SNPs. We did find several over-expressed SNPs at our targeted 'discovery' association levels (10^{-6} to 10^{-7}) in genes related to cell migration, phagocytosis and cytokine function. We also identified pathways related to innate immune receptors, maintenance of mucosal homeostasis and barrier defense mechanisms, and cellular signaling, that were enriched in patients who developed sepsis. Adaptive immunity is poorly developed in extremely premature infants; the immune response to infectious insults is therefore reliant on innate immune pathways.[58] We speculate that the identified variants, and pathways that they are involved in, may regulate key aspects of the premature neonate's innate immune response, leading to alterations in susceptibility to sepsis. Alteration of gut and respiratory epithelial integrity mediated by these pathways may be an important mechanism of increased sepsis and NEC risk.

The relatively modest size of our cohort imposes limitations on the statistical power of this GWAS. While adult GWA studies have been larger, this is the largest ELBW neonatal population with complete genotyping and sepsis-specific phenotype data available. Our study provides novel insights into potential heritable factors for sepsis predilection in this extremely vulnerable population. The smaller sample sizes of subgroups limited conclusions that could be drawn from subgroup analyses, and also restricted our ability to perform analyses of specific phenotypic subgroups such as meningitis without bacteremia. We limited our analysis to a stringent definition of sepsis (culture proven infection that was treated with a prolonged course of antibiotics). Other limitations include the multi-ethnic population (fairly representative of the US preterm birth population). To address this, we used eigenvalues to represent ancestry in the analysis, which allowed more complete utilization of the dataset; but precluded any inferences related to specific racial or ethnic subgroups. This dataset contained 70 identical twins, with separate clinical data on each; SNPs were run on the first sample in an identical twin set (with blinding to clinical phenotype), and SNP results were ascribed to both twins. As we used anonymized clinical data collected in a general database, we possessed limited information on the temporality of comorbidities in relation to the sepsis event. As a result, NEC and other co-morbidities could not be reliably accounted for as covariates in the analyses.

We provide data on the role of genetic variants in predisposition to sepsis in extremely premature neonates. While no genes achieved genome-wide significance, several SNPs in genes and pathways relevant to innate immunity were identified with our screening approach using 'target' levels of significance. Additional pathway analyses and gene expression profiling studies could help better understand interactions and downstream effects of these genetic variations in neonatal sepsis. It is essential to ascertain whether these SNPs and pathways related to them, continue to demonstrate an association with sepsis in additional cohorts of ELBW infants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (General Clinical Research Center grants M01 RR30, M01 RR32, M01 RR39, M01 RR70, M01 RR80, M01 RR633, M01 RR750, M01 RR997, M01 RR6022, M01 RR7122, M01 RR8084, M01 RR16587, UL1 RR24979) and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (grants U01 HD36790, U10 HD21364, U10 HD21373, U10 HD21385, U10 HD21397, U10 HD21415, U10 HD27851, U10 HD27853, U10 HD27856, U10 HD27871, U10 HD27880, U10 HD27881, U10 HD27904, U10 HD34216, U10 HD40461, U10 HD40492, U10 HD40498, U10 HD40689, U10 HD53109, U10 HD68244) provided grant support for the Neonatal Research Network's Genomics Study. In addition, Dr. Murray received assistance for the GENEVA study from the National Human Genome Research Institute (U01 HG4423). The funding agencies provided overall oversight for study conduct, but all data analyses and interpretation were independent of the funding agencies. Participating NRN sites collected data and transmitted it to RTI International, the data coordinating center (DCC) for the network, which stored, managed and analyzed the data for this study. On behalf of the NRN, Drs. Abhik Das (DCC Principal Investigator) and Grier Page (DCC Statistician) had full access to all of the data in the study, and with the NRN Center Principal Investigators, take responsibility for the integrity of the data and accuracy of the data analysis.

We are indebted to our medical and nursing colleagues and the infants and their parents who agreed to take part in this study. The following investigators, in addition to those listed as authors, participated in this study:

NRN Steering Committee Chair: Alan H. Jobe, MD PhD, University of Cincinnati.

Case Western Reserve University, Rainbow Babies & Children's Hospital (U10 HD21364, M01 RR80) – Avroy A. Fanaroff, MD; Nancy S. Newman, RN; Bonnie S. Siner, RN.

Cincinnati Children's Hospital Medical Center, University Hospital, and Good Samaritan Hospital (U10 HD27853, M01 RR8084) – Edward F. Donovan, MD; Vivek Narendran, MD MRCP; Barbara Alexander, RN; Cathy Grisby, BSN CCRC; Jody Hessling, RN; Marcia Worley Mersmann, RN CCRC; Holly L. Mincey, RN BSN.

Duke University School of Medicine University Hospital, Alamance Regional Medical Center, and Durham Regional Hospital (M01 RR30, U10 HD40492) – Kathy J. Auten, MSHS.

Emory University, Children's Healthcare of Atlanta, Grady Memorial Hospital, and Emory Crawford Long Hospital (U10 HD27851, M01 RR39) – Ellen C. Hale, RN BS CCRC.

Eunice Kennedy Shriver National Institute of Child Health and Human Development – Linda L. Wright, MD; Sumner J. Yaffe, MD; Elizabeth M. McClure, Med; Stephanie Wilson Archer, MA.

RTI International (U10 HD36790) – W. Kenneth Poole, PhD (deceased); Betty K. Hastings; Jeanette O'Donnell Auman, BS.

Stanford University, Lucile Packard Children's Hospital (U10 HD27880, M01 RR70) – M. Bethany Ball, BS CCRC.

University of Alabama at Birmingham Health System and Children's Hospital of Alabama (U10 HD34216, M01 RR32) – Namasivayam Ambalavanan, MD; Monica V. Collins, RN BSN MaEd; Shirley S. Cosby, RN BSN.

University of California – San Diego Medical Center and Sharp Mary Birch Hospital for Women (U10 HD40461) – Neil N. Finer, MD; Maynard R. Rasmussen, MD; David Kaegi, MD; Kathy Arnell, RNC; Clarence Demetrio, RN; Wade Rich, BSHS RRT.

University of Iowa (U10 HD53109) – Edward F. Bell, MD.

University of Miami Holtz Children's Hospital (U10 HD21397, M01 RR16587) – Charles R. Bauer, MD; Ruth Everett-Thomas, RN MSN.

University of Pennsylvania (U10 HD68244) – Barbara Schmidt, MD MSc.

University of Tennessee (U10 HD21415) – Sheldon B. Korones, MD (deceased); Henrietta S. Bada, MD; Tina Hudson, RN BSN.

University of Texas Southwestern Medical Center at Dallas, Parkland Health & Hospital System, and Children's Medical Center Dallas (U10 HD40689, M01 RR633) – Abbot R. Laptook, MD; Walid A. Salhab, MD; Susie Madison, RN; Nancy A. Miller, RN; Gaynelle Hensley, RN; Alicia Guzman.

University of Texas Health Science Center at Houston Medical School, Children's Memorial Hermann Hospital, and Lyndon B. Johnson General Hospital (U10 HD21373) – Jon E. Tyson, MD MPH; Kathleen A. Kennedy, MD MPH; Esther G. Akpa, RN BSN; Patty A. Cluff, RN; Claudia I. Franco, RNC MSN; Anna E. Lis, RN BSN; Georgia E. McDavid, RN; Patti Pierce Tate, RCP.

Wayne State University, Hutzel Women's Hospital, and Children's Hospital of Michigan (U10 HD21385) – G. Ganesh Konduri, MD; Rebecca Bara, RN BSN; Geraldine Muran, RN BSN.

SOURCE OF FUNDING:

This work was supported by the National Institutes of Health (General Clinical Research Center grants M01 RR30, M01 RR32, M01 RR39, M01 RR70, M01 RR80, M01 RR633, M01 RR750, M01 RR997, M01 RR6022, M01 RR7122, M01 RR8084, M01 RR16587, UL1 RR24979) and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (grants U01 HD36790, U10 HD21364, U10 HD21373, U10 HD21385, U10 HD21397, U10 HD21415, U10 HD27851, U10 HD27853, U10 HD27856, U10 HD27871, U10 HD27880, U10 HD27881, U10 HD27904, U10 HD34216, U10 HD40461, U10 HD40492, U10 HD40498, U10 HD40689, U10 HD53109, U10 HD68244) provided grant support for the Neonatal Research Network's Genomics Study. In addition, Dr. Murray received assistance for the GENEVA study from the National Human Genome Research Institute (U01 HG4423). The funding agencies provided overall oversight for study conduct, but all data analyses and interpretation were independent of the funding agencies. Participating NRN sites collected data and transmitted it to RTI International, the data coordinating center (DCC) for the network, which stored, managed and analyzed the data for this study. On behalf of the NRN, Drs. Abhik Das (DCC Principal Investigator) and Grier Page (DCC Statistician) had full access to all of the data in the study, and with the NRN Center Principal Investigators, take responsibility for the integrity of the data and accuracy of the data analysis.

References

1. Stoll BJ, Hansen NI, Bell EF, et al. Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. Pediatrics. 2010; 126(3):443–56. [PubMed: 20732945]
2. Stoll BJ, Hansen NI, Adams-Chapman I, et al. Neurodevelopmental and growth impairment among extremely low-birth-weight infants with neonatal infection. JAMA. 2004; 292(19):2357–65. [PubMed: 15547163]
3. Wynn JL, Cvijanovich NZ, Allen GL, et al. The influence of developmental age on the early transcriptomic response of children with septic shock. Mol Med. 2011; 17(11–12):1146–56. [PubMed: 21738952]
4. Hartel C, Schultz C, Herting E, et al. Genetic association studies in VLBW infants exemplifying susceptibility to sepsis--recent findings and implications for future research. Acta Paediatr. 2007; 96(2):158–65. [PubMed: 17429897]
5. Hill AV. The immunogenetics of human infectious diseases. Annu Rev Immunol. 1998; 16:593–617. [PubMed: 9597143]
6. Bhandari V, Bizzarro MJ, Shetty A, et al. Familial and genetic susceptibility to major neonatal morbidities in preterm twins. Pediatrics. 2006; 117(6):1901–6. [PubMed: 16740829]
7. Bizzarro MJ, Jiang Y, Hussain N, et al. The impact of environmental and genetic factors on neonatal late-onset sepsis. The Journal of pediatrics. 2011; 158(2):234–8. e1. [PubMed: 20850766]
8. Boghossian NS, Page GP, Bell EF, et al. Late-onset sepsis in very low birth weight infants from singleton and multiple-gestation births. The Journal of pediatrics. 2013; 162(6):1120–4. 24 e1. [PubMed: 23324523]
9. Chauhan M, McGuire W. Interleukin-6 (−174C) polymorphism and the risk of sepsis in very low birth weight infants: meta-analysis. Archives of disease in childhood Fetal and neonatal edition. 2008; 93(6):F427–9. [PubMed: 18375611]
10. Arcaroli J, Fessler MB, Abraham E. Genetic polymorphisms and sepsis. Shock. 2005; 24(4):300–12. [PubMed: 16205313]
11. Abu-Maziad A, Schaa K, Bell EF, et al. Role of polymorphic variants as genetic modulators of infection in neonatal sepsis. Pediatric research. 2010; 68(4):323–9. [PubMed: 20463618]
12. Sutherland AM, Walley KR, Russell JA. Polymorphisms in CD14, mannose-binding lectin, and Toll-like receptor-2 are associated with increased prevalence of infection in critically ill adults. Crit Care Med. 2005; 33(3):638–44. [PubMed: 15753758]

13. Ahrens P, Kattner E, Kohler B, et al. Mutations of genes involved in the innate immune system as predictors of sepsis in very low birth weight infants. Pediatric research. 2004; 55(4):652–6. [PubMed: 14739370]
14. Systematic review and meta-analysis of gene association studies in neonatal sepsis. EPAS. 2013:1540.643.
15. Yu W, Gwinn M, Clyne M, et al. A navigator for human genome epidemiology. Nat Genet. 2008; 40(2):124–5. [PubMed: 18227866]
16. Man M, Close SL, Shaw AD, et al. Beyond single-marker analyses: mining whole genome scans for insights into treatment responses in severe sepsis. Pharmacogenomics J. 2013; 13(3):218–26. [PubMed: 22310353]
17. Rautanen A, Mills TC, Gordon AC, et al. Genome-wide association study of survival from sepsis due to pneumonia: an observational cohort study. Lancet Respir Med. 2015; 3(1):53–60. [PubMed: 25533491]
18. Ambalavanan N, Carlo WA, D'Angio CT, et al. Cytokines associated with bronchopulmonary dysplasia or death in extremely low birth weight infants. Pediatrics. 2009; 123(4):1132–41. [PubMed: 19336372]
19. Cotten, CM., Page, G., Carlo, W., Higgins, R., Murray, JC. PAS. Boston, MA: 2012. Genome-Wide Association Study of Morbidities of Extreme Prematurity.
20. Ambalavanan N, Cotten CM, Page GP, et al. Integrated genomic analyses in bronchopulmonary dysplasia. The Journal of pediatrics. 2015; 166(3):531–7. e13. [PubMed: 25449221]
21. Wynn JL, Hansen NI, Das A, et al. Early sepsis does not increase the risk of late sepsis in very low birth weight neonates. The Journal of pediatrics. 2013; 162(5):942–8. e1–3. [PubMed: 23295144]
22. Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. Nature reviews Genetics. 2009; 10(5):318–29.
23. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science. 2008; 322(5903): 881–8. [PubMed: 18988837]
24. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2016
25. Panagiotou OA, Ioannidis JP. Genome-Wide Significance P. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. Int J Epidemiol. 2012; 41(1):273–86. [PubMed: 22253303]
26. Reiner AP, Ziv E, Lind DL, et al. Population structure, admixture, and aging-related phenotypes in African American adults: the Cardiovascular Health Study. American journal of human genetics. 2005; 76(3):463–77. [PubMed: 15660291]
27. Hoggart CJ, Parra EJ, Shriver MD, et al. Control of confounding of genetic associations in stratified populations. American journal of human genetics. 2003; 72(6):1492–504. [PubMed: 12817591]
28. Barnholtz-Sloan JS, Chakraborty R, Sellers TA, et al. Examining population stratification via individual ancestry estimates versus self-reported race. Cancer Epidemiol Biomarkers Prev. 2005; 14(6):1545–51. [PubMed: 15941970]
29. Croft D, Mundo AF, Haw R, et al. The Reactome pathway knowledgebase. Nucleic Acids Res. 2014; 42(Database issue):D472–7. [PubMed: 24243840]
30. Milacic M, Haw R, Rothfels K, et al. Annotating cancer variants and anti-cancer therapeutics in reactome. Cancers (Basel). 2012; 4(4):1180–211. [PubMed: 24213504]
31. Yu K, Li Q, Bergen AW, et al. Pathway analysis by adaptive combination of P-values. Genet Epidemiol. 2009; 33(8):700–9. [PubMed: 19333968]
32. Stoll BJ, Hansen N. Infections in VLBW infants: studies from the NICHD Neonatal Research Network. Seminars in perinatology. 2003; 27(4):293–301. [PubMed: 14510320]
33. Stoll BJ, Hansen N, Fanaroff AA, et al. Late-onset sepsis in very low birth weight neonates: the experience of the NICHD Neonatal Research Network. Pediatrics. 2002; 110(2 Pt 1):285–91. [PubMed: 12165580]
34. Cooke GS, Hill AV. Genetics of susceptibility to human infectious disease. Nature reviews Genetics. 2001; 2(12):967–77.

35. Sorensen TI, Nielsen GG, Andersen PK, et al. Genetic and environmental influences on premature death in adult adoptees. The New England journal of medicine. 1988; 318(12):727–32. [PubMed: 3347221]
36. Gopel W, Hartel C, Ahrens P, et al. Interleukin-6-174-genotype, sepsis and cerebral injury in very low birth weight infants. Genes Immun. 2006; 7(1):65–8. [PubMed: 16208404]
37. Hartel C, Hemmelmann C, Faust K, et al. Tumor necrosis factor-alpha promoter -308 G/A polymorphism and susceptibility to sepsis in very-low-birth-weight infants. Crit Care Med. 2011; 39(5):1190–5. [PubMed: 21317641]
38. Baier RJ, Loggins J, Yanamandra K. IL-10, IL-6 and CD14 polymorphisms and sepsis outcome in ventilated very low birth weight infants. BMC medicine. 2006; 4:10. [PubMed: 16611358]
39. Fukuda K, Yoshida H, Sato T, et al. Mesenchymal expression of Foxl1, a winged helix transcriptional factor, regulates generation and maintenance of gut-associated lymphoid organs. Developmental biology. 2003; 255(2):278–89. [PubMed: 12648490]
40. Katz JP, Perreault N, Goldstein BG, et al. Foxl1 null mice have abnormal intestinal epithelia, postnatal growth retardation, and defective intestinal glucose uptake. American journal of physiology Gastrointestinal and liver physiology. 2004; 287(4):G856–64. [PubMed: 15155178]
41. Takano-Maruyama M, Hase K, Fukamachi H, et al. Foxl1-deficient mice exhibit aberrant epithelial cell positioning resulting from dysregulated EphB/EphrinB expression in the small intestine. American journal of physiology Gastrointestinal and liver physiology. 2006; 291(1):G163–70. [PubMed: 16469829]
42. Madison BB, McKenna LB, Dolson D, et al. FoxF1 and FoxL1 link hedgehog signaling and the control of epithelial proliferation in the developing stomach and intestine. J Biol Chem. 2009; 284(9):5936–44. [PubMed: 19049965]
43. Caron E, Hall A. Identification of two distinct mechanisms of phagocytosis controlled by different Rho GTPases. Science. 1998; 282(5394):1717–21. [PubMed: 9831565]
44. Park D, Tosello-Trampont AC, Elliott MR, et al. BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. Nature. 2007; 450(7168):430–4. [PubMed: 17960134]
45. Das S, Owen KA, Ly KT, et al. Brain angiogenesis inhibitor 1 (BAI1) is a pattern recognition receptor that mediates macrophage binding and engulfment of Gram-negative bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108(5): 2136–41. [PubMed: 21245295]
46. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(43):15545–50. [PubMed: 16199517]
47. Eugenin EA. Role of connexin/pannexin containing channels in infectious diseases. FEBS Lett. 2014; 588(8):1389–95. [PubMed: 24486013]
48. Esen N, Shuffield D, Syed MM, et al. Modulation of connexin expression and gap junction communication in astrocytes by the gram-positive bacterium S. aureus. Glia. 2007; 55(1):104–17. [PubMed: 17029244]
49. Karpuk N, Burkovetskaya M, Fritz T, et al. Neuroinflammation leads to region-dependent alterations in astrocyte gap junction communication and hemichannel activity. J Neurosci. 2011; 31(2):414–25. [PubMed: 21228152]
50. Ey B, Eyking A, Gerken G, et al. TLR2 mediates gap junctional intercellular communication through connexin-43 in intestinal epithelial barrier injury. J Biol Chem. 2009; 284(33):22332–43. [PubMed: 19528242]
51. Bou Saab J, Losa D, Chanson M, et al. Connexins in respiratory and gastrointestinal mucosal immunity. FEBS Lett. 2014; 588(8):1288–96. [PubMed: 24631537]
52. Blackburn SD, Steadman RA, Johnson FB. Attachment of adeno-associated virus type 3H to fibroblast growth factor receptor 1. Arch Virol. 2006; 151(3):617–23. [PubMed: 16195782]
53. Kim JH, Jiang S, Elwell CA, et al. Chlamydia trachomatis co-opts the FGF2 signaling pathway to enhance infection. PLoS Pathog. 2011; 7(10):e1002285. [PubMed: 21998584]

54. Liu X, Lai C, Wang K, et al. A Functional Role of Fibroblast Growth Factor Receptor 1 (FGFR1) in the Suppression of Influenza A Virus Replication. PLoS One. 2015; 10(4):e0124651. [PubMed: 25909503]

55. Dignass AU, Tsunekawa S, Podolsky DK. Fibroblast growth factors modulate intestinal epithelial cell growth and migration. Gastroenterology. 1994; 106(5):1254–62. [PubMed: 7513666]

56. Fehr AR, Yu D. Control the host cell cycle: viral regulation of the anaphase-promoting complex. J Virol. 2013; 87(16):8818–25. [PubMed: 23760246]

57. Liew FY, Xu D, Brint EK, et al. Negative regulation of toll-like receptor-mediated immune responses. Nat Rev Immunol. 2005; 5(6):446–58. [PubMed: 15928677]

58. Levy O, Wynn JL. A prime time for trained immunity: innate immune memory in newborns and infants. Neonatology. 2014; 105(2):136–41. [PubMed: 24356292]

What is known about this topic

- Sepsis is a major co-morbidity in extremely low birth weight infants.
- Preterm infants may have genetic susceptibility to infection.

What this study adds

- In a a genome wide association study in extremely low birth weight infants with sepsis, we identified several over-represented single nucleotide polymorphisms (SNPs) related to innate immunity, although none met genome-wide significance.
- Alteration of gut and respiratory epithelial integrity mediated by these SNPs and related pathways may be an important mechanism of increased sepsis risk.

Table 1

SNPs with highest association with all sepsis *

Chr	*SNP*	*Location***	*P value*	*Gene*
16	rs13380717	86904135	1.08E-07	Intergenic
16	rs9933616	86903194	3.64E-07	Intergenic
16	rs59876150	86918734	6.80E-07	Intergenic
16	rs34528289	86907297	6.95E-07	Intergenic
4	rs2412930	59588228	1.16E-06	Intergenic
4	rs6837629	59583956	1.49E-06	Intergenic
6	rs9456883	1.64E+08	1.63E-06	Intergenic
2	rs41461846	2.19E+08	1.93E-06	Intergenic
2	chr2:219344165	2.19E+08	1.99E-06	Intergenic
2	rs6717433	2.19E+08	2.04E-06	Intergenic

* Top ten SNPs for each analysis listed; covariates in analysis: gestational age, SGA, eigenvalues for ancestry, surgery, and Cesarean section; infants who died before 72 hours excluded from controls

** Location in Base pairs from p telomere

Table 2

SNPs with highest association with Gram positive LOS *

Chr	*SNP*	*Location*	*P value*	*Gene*	*Gene function*
7	rs6462728	37071352	8.42E-07	Intron ELMO1	Encodes cell motility and engulfment protein. Interacts with cytokinesis proteins to promote phagocytosis, cell migration.[4344] Animal studies of Gram negative sepsis: bacterial internalization, phagocytosis mediated via ELMO1/Dock180 pathway, which triggers activation of Rac1, causing membrane invagination and pseudopod formation.
7	rs13227735	37073437	8.83E-07	Intron ELMO1	See above
16	rs13380717	86904135	9.37E-07	Intergenic	
3	rs3844280	10212258	1.24E-06	Intron IRAK2	Encodes Interleukin-1 receptor associated-kinase 2. Upon stimulation, this kinase associates with IL-1 receptor and participates in IL-1 mediated up-regulation of NF-KB
1	rs3100127	2.02E+08	1.30E-06	Intergenic	
13	rs11840143	93284661	1.76E-06	Intron GPC5	Encodes glypican 5, a cell surface proteoglycan
2	rs41461846	2.19E+08	1.83E-06	Intron VIL1	Encodes villin 1, a calcium regulated actin binding protein
2	chr2:219344165	2.19E+08	1.89E-06	Intron VIL1	See above
2	rs6717433	2.19E+08	1.93E-06	Intron VIL1	See above
1	rs2361422	2.02E+08	1.97E-06	5′ PTPRVP	

*Top ten SNPs for each analysis listed; covariates in analysis: gestational age, SGA, eigenvalues for ancestry, surgery and Cesarean section; infants who died before 72 hours excluded from controls and LOS

**Location in base pairs from p telomere

Table 3

SNPs with highest association with Gram negative LOS *

Chr	*SNP*	*Location* **	*P value*	*Gene*	*Gene function*
2	chr2:146369734	1.46E+08	1.13E-07	Intergenic	
8	rs513793	20774064	5.82E-07	Intergenic	
2	chr2:146358394	1.46E+08	8.84E-07	Intergenic	
2	chr2:146356963	1.46E+08	8.85E-07	Intergenic	
7	rs2237499	39740846	1.06E-06	Intron RALA	Encodes protein belonging to Ras family of GTP binding proteins, implicated in filopodia formation for cell migration
2	rs58162439	1.46E+08	1.55E-06	Intergenic	
16	rs59876150	86918734	1.99E-06	Intergenic	
7	rs4730486	1.11E+08	3.10E-06	Intron IMMP2L	Encodes protein important for function of mitochondrial inner membrane peptidase (IMP) complex
7	rs6950974	37075828	3.20E-06	Intron RALA	
7	rs7811308	1.11E+08	3.92E-06	Intron IMMP2L	

*Top ten SNPs for each analysis listed; covariates in analysis: gestational age, SGA, eigenvalues for ancestry, surgery and Cesarean section; infants who died before 72 hours excluded from controls and LOS

**Location in base pairs from p telomere

Arch Dis Child Fetal Neonatal Ed. Author manuscript; available in PMC 2017 September 01.

Table 4

SNPs with highest association with fungal LOS*

Chr	*SNP*	*Location***	*P value*	*Gene*	*Gene function*
18	rs645505	10682720	6.35E-07	Intron PIEZO2	Encodes protein which functions as part of mechanically-activated cation channels, connecting mechanical forces to biological signals
10	rs11597285	8589813	7.36E-07	Intergenic	
10	rs2031558	8589408	7.49E-07	Intergenic	
10	rs11596835	8589969	7.53E-07	Intergenic	
10	chr10:113662222	1.14E+08	1.28E-06	Intergenic	
10	rs11593542	8600339	1.37E-06	Intergenic	
10	rs1325884	8599820	1.43E-06	Intergenic	
12	rs16913666	18134955	1.73E-06	Intergenic	
4	rs17599816	47164556	1.85E-06	Intron GABRB1	Encodes gamma-aminobutyric acid (GABA) A receptor, beta 1
3	rs12490944	50059651	1.90E-06	Intron RBM6	Encodes RNA binding motif protein 6 (previously implicated in various cancers)

*Top ten SNPs for each analysis listed; covariates in analysis: gestational age, SGA, eigenvalues for ancestry, surgery and Cesarean section; infants who died before 72 hours excluded from controls and LOS

**Location in base pairs from p telomere

Arch Dis Child Fetal Neonatal Ed. Author manuscript; available in PMC 2017 September 01.