Lymphocyte-Specific Biomarkers Associated With Preterm Birth and Bronchopulmonary Dysplasia

Soumyaroop Bhattacharya [1], Jared A. Mereness [1], Andrea M. Baran [2], Ravi S. Misra [1], Derick R. Peterson [2], Rita M. Ryan [3,4], Anne Marie Reynolds [3], Gloria S. Pryhuber [1]* and Thomas J. Mariani [1]*

[1] Division of Neonatology, Department of Pediatrics, University of Rochester, Rochester, NY, United States, [2] Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY, United States, [3] Department of Pediatrics, University at Buffalo, Buffalo, NY, United States, [4] Department of Pediatrics, Case Western Reserve University, Cleveland, OH, United States

Many premature babies who are born with neonatal respiratory distress syndrome (RDS) go on to develop Bronchopulmonary Dysplasia (BPD) and later Post-Prematurity Respiratory Disease (PRD) at one year corrected age, characterized by persistent or recurrent lower respiratory tract symptoms frequently related to inflammation and viral infection. Transcriptomic profiles were generated from sorted peripheral blood CD8+ T cells of preterm and full-term infants enrolled with consent in the NHLBI Prematurity and Respiratory Outcomes Program (PROP) at the University of Rochester and the University at Buffalo. We identified outcome-related gene expression patterns following standard methods to identify markers for oxygen utilization and BPD as outcomes in extremely premature infants. We further identified predictor gene sets for BPD based on transcriptomic data adjusted for gestational age at birth (GAB). RNA-Seq analysis was completed for CD8+ T cells from 145 subjects. Among the subjects with highest risk for BPD (born at <29 weeks gestational age (GA); n=72), 501 genes were associated with oxygen utilization. In the same set of subjects, 571 genes were differentially expressed in subjects with a diagnosis of BPD and 105 genes were different in BPD subjects as defined by physiologic challenge. A set of 92 genes could predict BPD with a moderately high degree of accuracy. We consistently observed dysregulation of *TGFB, NRF2, HIPPO,* and *CD40*-associated pathways in BPD. Using gene expression data from both premature and full-term subjects (n=116), we identified a 28 gene set that predicted the PRD status with a moderately high level of accuracy, which also were involved in *TGFB* signaling. Transcriptomic data from sort-purified peripheral blood CD8+ T cells from 145 preterm and full-term infants identified sets of molecular markers of inflammation associated with independent development of BPD in extremely premature infants at high risk for the disease and of PRD among the preterm and full-term subjects.

Keywords: prematurity and respiratory outcomes program, bronchopulmonary dysplasia, RNA-sequencing, Pathway, Newborn Lung, Peripheral Blood Mononucleated cells, CD8+ lymphocytes

INTRODUCTION

Acute and chronic respiratory morbidities are common in extremely premature infants (1). Increased survival of very premature infants is leading to increasing numbers of children with chronic lung disease. Since the end of the last millennium, the rate of premature births <34 weeks of gestation have consistently increased in the United States, and in 2008 it was 12.3% (2). Among the extremely preterm infants, 20%–35% die before their discharge to home (1, 3). Prematurity-related deaths accounted for 35% of all infant deaths in 2010, more than any other single cause. Preterm birth cost the U.S. health care system more than $26 billion in 2005 (4). Among NICU survivors, approximately 40% develop Bronchopulmonary Dysplasia (BPD), a chronic lung disease of the newborn. BPD has both genetic and environmental risk factors. It is characterized by varying degrees of lung injury potentially due to required supplemental oxygen, exposure to inflammatory conditions *in utero*, and mechanical ventilation and is often associated with infection (5, 6). BPD results from abnormal repair and impaired lung development after acute lung injury. Airway function may even deteriorate during the first year of life in infants with BPD (7). A key component of BPD is persistent inflammation of the lung (8, 9). Infants with BPD are more likely to die than those without chronic lung disease, even if they survive the initial hospitalization. However, improved medical treatment plans have been developed that have led to a lower hospital mortality rate, however, the respiratory sequelae into childhood remain poorly defined (10). By developing a better understanding of the inflammatory process of infants with BPD, we could potentially identify biomarkers that relate to respiratory sequelae. Using such information, certain pathways could be targeted for drug development to improve the health of infants with BPD (11).

The NIH NHLBI Prematurity and Respiratory Outcomes Program (PROP) enrolled 835 extremely premature infants across the US and collected multiple biospecimens over time and extensive data including BPD and respiratory morbidity outcomes over the 1st year of life. PROP investigated the molecular mechanisms contributing to the risk of respiratory disease in premature neonates over the first year of life. A set of clinical and non-invasive respiratory assessments were performed, based on the respiratory status of the infant at the time of testing, and was used to predict the severity of respiratory outcomes in the first year of life. The primary goal of the PROP studies was to identify biomarkers (biochemical, physiological, and genetic) that are associated with, and thus potentially predictive of, respiratory morbidity in preterm infants up to 1-year corrected age. A validated, objective measure of pulmonary outcome at 1 year does not currently exist. In addition to the identification of markers of BPD, one of the primary outcomes in the PROP study was defined as presence or absence of Post-Prematurity Respiratory Disease (PRD) (12). In order to be classified as having PRD, infants were required to have a positive response in at least one of the four morbidity domains [(1) hospitalization for respiratory indication, (2) home respiratory support, (3) respiratory medication administration, and/or (4) respiratory

symptoms without a cold] during at least two separate parental interviews conducted at 3-, 6-, 9-, and 12-months corrected age (13).

High throughput sequencing for genome-wide transcriptomic analysis, by RNA-Seq or microarrays, is an unbiased approach applied to identify biomarkers that may provide predictive value. These approaches have been proven to be powerful tools capable of biomarker discovery for various disease states including BPD. We have previously presented an analysis of lung tissue gene expression in subjects with BPD (14). Application of blood-based gene expression profiling can potentially provide novel biomarkers for diagnosis and therapeutic management of BPD. Previous studies have used whole blood–derived peripheral blood mononuclear cells (PBMC) as a means of mining for novel markers for BPD (15). PBMCs are relatively easy to obtain from whole blood and can be sorted into leukocytes, including B cells, T cells, monocytes, and natural killer cells (16). The use of peripheral blood has identified changes in CD4+ T cell populations in subjects with diagnosed with BPD (17, 18). Additional studies have identified that patterns of proinflammatory cytokines in blood from subjects with BPD were associated with the phenotype of BPD (19). Our study was funded to complete transcriptomics analyses of the CD8+ T cell population.

Emerging data suggest an important role for T lymphocytes in the pathogenesis of chronic lung disease in babies born prematurely. CD8+ T cells, TNF-α, TNF receptors, and NK cells provide protection from viral infection but also contribute to the immunopathology, by contact-dependent effector functions (perforin and FasL). IFN-γ and particularly TNF-α are thought to be primary perpetrators of T-cell-mediated lung injury (20). Furthermore, one study shows that CD8+ T cells isolated from blood of infants with BPD exhibited lower levels of surface CD62L, which is consistent with an activated phenotype (21). CD8+ T cells have shown adaptive immune insufficiency in newborn mice infected with influenza A within 1 week of birth (22). RSV infected neonatal mice recruited CD8+ T cells defective in IFN-γ production in association with mild symptoms. Re-infection as adults however resulted in limited viral replication but enhanced inflammation and T cell recruitment, including Th2 cells and eosinophils (23, 24). Depletion of CD8+ T cells (but not CD4) cells during the primary neonatal infection was protective against the adult challenge. We have previously shown that CD8+ T cells appear to play a pathogenic role in subjects with BPD, and may be associated with overall risk for lung morbidity (25). In related studies, we have observed that CD8+ T cells are increased in both mouse and human lungs exposed as neonates to hyperoxia, and have a hyper-responsive, fibrotic and destructive response to subsequent viral infection (14, 26). Further, these cells have a predominant role in direct cytotoxicity in the lung, *via* interactions with epithelial cells and as regulators of macrophage responses, as well as in general human resistance to viral infection. In a separate population of premature infants, enrolled in the PROP study, phenotyped T cells at birth, at 36 weeks of adjusted gestational age, and at 12-months corrected age, were associated with a PROP-defined respiratory morbidity at 12 months (27). The goal for the current study was to demonstrate that transcriptional profiling of

CD8+ T cells, obtained from premature infants at discharge, can identify disease-related gene expression patterns informative for pathogenesis and capable of predicting risk of future respiratory distress. We hypothesized that transcriptomic analysis of sorted lymphocyte sub-populations could identify predictive markers and pathways associated with respiratory outcomes. We followed a cohort of 157 infants, ranging from 23 to 41 weeks of gestation at birth, enrolled in the Prematurity and Respiratory Outcomes Program (PROP) at the University of Rochester Medical Center and Children's Hospital of Buffalo (12).

Here, we present a novel gene expression RNA-seq data set generated from CD8+ T cells from 145 subjects with varying levels of premature birth and report the identification of disease biomarkers for BPD and PRD. In subjects who were diagnosed with BPD, we identified pathways associated with TGFB signaling (Regulation of Epithelial-Mesenchymal Transition Pathway) and T cell activation and polarization (e.g., IL-2, IL-17, IL-4, and iCOS signaling). In subjects who developed PRD at one year of life, we also identified the TGFB pathway as being important as well as the Cell2Cell pathway (which includes genes important for CD8+ T cell activation). Given that TGFB has been identified as an important factor for controlling CD8+ T cell mediated inflammation (28–32), we provide new makers in CD8+ T cells that are associated with the BPD and PRD. Such information will be of interest to researchers who are trying to develop a better understanding of factors associated with inflammatory pediatric lung disease.

METHODS

This study aimed at generating transcriptomic profiles of CD8+ T cells from newborn human blood. This study was approved by the Institutional Review Board of University of Rochester with a Memorandum of Understanding executed with the University of Buffalo IRB. Subjects were enrolled within gestational age at birth (GAB) epochs, in order to characterize the relationship among prematurity, disease risk, and gene expression. The steps involved in the process, starting from sample collection to isolation of total RNA have been described in detail in our previous publication (16). Additional steps relevant to this study are shown in **Supplemental Figure 1**.

Oxygen Exposure, BPD Diagnosis, and PRD

The traditional categorical approach of classifying BPD as absent or present is likely an oversimplification. Tooley (33) recommended that oxygen use at 28 days of age would identify preterm infants with BPD. Almost a decade later, Shennan and colleagues proposed that the best predictor of abnormal pulmonary outcomes among very low birth weight premature infants was the clinical use of oxygen at 36 weeks postmenstrual age (PMA) (34). A workshop convened by the National Institutes of Health (NIH) proposed severity-based diagnostic criteria for BPD (35) that included the use of oxygen for at least 28 days (not

necessarily consecutive) and an assessment of respiratory support at 36 weeks PMA, recognizing that some infants breathing room air at 36 weeks PMA may have residual lung disease. The majority of infants with birth weights less than 1 kg will have a diagnosis of at least mild BPD by the Consensus Conference definition based on 28 days in oxygen (35). Given clinical variations in oxygen administration, a structured trial of room air test was developed by the NICHD Neonatal Research Network (36), the frequency of BPD among the subjects was determined using two previously published definitions: the Shennan definition (34), which defines BPD as supplemental oxygen requirement at 36 weeks PMA in infants born with birth weight (BW) less than 1,500 grams, and a physiologic definition with a room-air challenge (RAC) which defines BPD as requirement of oxygen support (>21% O_2) for at least 28 days and a subsequent assessment at 36 weeks PMA or discharge, whichever comes first (13, 36).

NICU oxygen exposure was calculated, as previously reported, from the FIO2 recorded in the medical record once each noon for the first 14 days of life. FIO2 was corrected to Effective FiO2 for low nasal cannula flow using established tables (37). Oxygen utilization or $Oxygen_{AUC}$ was calculated by the formula defined in Benaron and Benitz (37) using information recorded in the daily respiratory flowsheet data (FIO2, respiratory support mode, and applied airway pressure or cannula flow) through the first 28 days of life. We chose to look at $Oxygen_{AUC}$ at 14 days of postnatal age ($Oxygen_{AUC14}$) to include the second postnatal week to capture pulmonary deterioration as presented in BPD (38).

The infants enrolled in the study were followed up periodically for up to 12 months of age, corrected by gestational age (CGA) at birth. At the 12-month CGA follow-up visit they were assessed for persistent respiratory distress based on the frequency of hospitalization due to any kind of respiratory distress. Persistent Respiratory Distress (PRD) was diagnosed if there were positive responses in at least one of the following domains: (1) hospitalization for respiratory indication, (2) respiratory support at home, (3) respiratory medication administration, and/or (4) cough or wheeze without a cold, reported on at least two caregiver post-discharge surveys completed at 3, 6, 9, and 12 months CGA, as previously reported (13). The definitions and criteria for the different diagnoses have been listed in **Supplemental Table 1**.

Sample Collection and RNA Isolation

An average of 2.5 ml of venous blood was collected into sodium heparin glass vacutainers from premature infants enrolled with consent in the Prematurity and Respiratory Outcomes Program (PROP), at the time of hospital discharge at the University of Rochester and the University at Buffalo and shipped to a central laboratory in Rochester. Freshly purified PBMCs were isolated by Ficoll gradient (Amersham Pharmacia Biotech # 17-1440-03) centrifugation, from the whole blood diluted 1:2 with 1x dPBS, and counted according to previously established protocols (39). In subjects with at least 8 million cells, 5 million cells were stained with antibodies to individual lymphocyte markers, and

sorted on a FACSAriaII sorter at the Flow Cytometry Core facility of the University of Rochester as previously reported (16). CD3+CD8+, CD3+CD4+, CD3-CD56+ (NK), or CD3-CD19+ (B) cells were collected separately. Non-marker positive and dead cells were discarded. Sorted cells were spun into pellets, which were further lysed and frozen. The steps from collection to lysis of each sample were completed within a 24-h period in order to preserve RNA quality and integrity. Frozen lysates were thawed and RNA was extracted using Agilent Absolute RNA Microprep kit (catalog #400805), with an on-column DNase digestion, as per manufacturer recommended protocol.

RNA-Seq and Data Generation

For the current study, RNA isolated from the sorted CD8+ T cells from 145 pre-term and full-term subjects, was used for transcriptomic profiling by RNA-seq. cDNA libraries were generated with 1 ng RNA, using the SMARter Ultra Low Amplification kit (Clonetech, Mountain, CA). cDNA quantity was determined with the Qubit Flourometer (Life Technologies, Grand Island, NY) and quality was assessed using the Agilent Bioanalyzer 2100 (Santa Clara, CA). Libraries were sequenced (single endreads) on the Illumina HiSeq2500 (Illumina, San Diego, CA) to generate 20 million reads/sample. Reads generated from the sequencer were aligned using the TopHat algorithm (40) and expression values were summarized using HTSeq (41). The data from this study has been provided in dbGAP. The dbGaP accession assigned to this study is phs001297.v1.p1.

Normalization and Filtering

Samples were excluded based on poor read count/mapped read numbers, or if they were extreme outliers in hierarchical clustering and Principal Components Analysis (PCA). Genes were excluded if they were not consistently identified as expressed (a count of zero in over 1/3 of subjects). Subjects with high prevalence (>75%) of zero/low reads (raw count value ≤5) were excluded. Genes with high prevalence (>75%) of low counts (normalized count value ≤3) across subjects were excluded.

The subject-specific conditional upper quartile (UQ, 75th percentile) among non-zero reads was computed. The subject-specific normalization factor was calculated by dividing the UQ for a given subject by the mean UQ across all subjects. The normalized gene values for a given subject were calculated by dividing the raw count values by the normalization factor for that subject. After adding 1 to all values to account for zeros, the normalized counts were $\log_2$ transformed.

Selection of Univariately Differentially Expressed Genes

Differences in gene expression between subject groups was assessed by SAM-Seq (42) and Likelihood Ratio Test (LRT). SAM-Seq was used to identify genes with expression patterns significantly (FDR<0.05) associated with BPD, RAC. LRT for log (normalized RNA-Seq), adjusted for GAB via logistic regression, was used to identify genes with expression patterns significantly

(FDR<0.05) associated with BPD in subjects born at less than 29 weeks of age. For quantitative analysis, the correlations between normalized counts and oxygen utilization at 14 days were estimated. Expression changes of the genes, identified as significantly different in BPD by the tests, were assessed in transcriptomic profiles of PBMCs obtained from infants with BPD, and age matched controls generated from an independent cohort (15), using the data available on Gene Expression Omnibus (GSE32472).

Prediction of Bronchopulmonary Dysplasia (BPD) *via* Screened Principal Components

The following method is our minor variant of *Screened Principal Components Analysis (sPCA)* (43), where the genes were univariately screened, and those with a nominal Wilcoxon p < 0.10 were used for further analysis. The first principal component (PC1) of the genes was derived, and genes with loadings close to 0 were removed. Genes most strongly univariately associated with BPD (with or without adjusting for gestational age) were selected using a screening threshold chosen by cross-validation (CV). The first Principal Component (PC1) of the genes passing the univariate screen was constructed and a logistic regression model was fit to predict BPD as a function of PC1 (and optionally gestational age). Receiver Operating Characteristic (ROC) curves depicting sensitivity and specificity along with associated AUC were estimated without (naïve AUC) and with an outer loop of nested CV (CV-AUC).

Prediction of Post-Prematurity Respiratory Disease (PRD) *via* Canonical Pathways

We have used canonical pathway analysis where we grouped our 13,434 genes into 1,330 biologically-relevant pathway-based gene sets of molecular signature database (mSigDB) (44), where each gene belongs to 0, 1, or more pathways. We then reduced the constituent genes in each pathway to their PC1. Genes belonging to 0 pathways were thus excluded from consideration, while CV screened logistic forward selection (or LASSO) was applied to the 1,330 pathway-based PC's. Bivariate CV was used to simultaneously select both the threshold for univariate logistic likelihood ratio test screening of pathways (with or without adjusting for gestational age) and the final number of pathways chosen by forward selection (or the LASSO penalty). The entire procedure was then nested within an outer loop of nested CV in order to estimate performance *via* the Receiver Operating Characteristic (ROC) curve and its associated Area Under the ROC Curve (CV-AUC). This method has been described in details in our previous publication (45).

Functional Classification

Genes identified as differentially expressed in individual comparisons were used for independent functional classification through canonical pathway analysis, and upstream regulators identification using Ingenuity Pathway Analysis (IPA; Qiagen, Carlsbad, CA). Pathways and upstream regulators were identified as significantly associated with the diagnoses by Fisher's exact test (p < 0.05 or -log(p-value) > 1.3) as calculated by IPA.

Quantitative Reverse Transcriptase–Polymerase Chain Reaction (qPCR)

cDNA was synthesized from 100 ng RNA using iScript cDNA synthesis kit (Biorad, HerculesCA) and quantitative reverse transcriptase–polymerase chain reaction (qPCR) was performed with a Viia7 (Applied Biosytems, Carlsbad, CA) using SYBR green chemistry as previously described (14) using noncommercial (http://pga.mgh.harvard.edu/primerbank) assays. Gene expression levels were calculated relative to GAPDH as an internal, endogenous control, according to the ddCT method.

RESULTS

Subject Demographics

Peripheral blood was collected at the time of first hospital discharge, from 145 preterm and full-term infants enrolled in the NHLBI Prematurity and Respiratory Outcomes Program (PROP) at the University of Rochester and the University at Buffalo. Of the 145 subjects, 72 were extremely premature having been born at less 29 weeks of gestation. There was insufficient evidence that the distribution of race (p=0.64) or sex (p=1) differed between BPD and non-BPD subjects, while as expected gestational age at birth was lower for those with BPD (p=0.03). Similarly, among all subjects (n=130) there was insufficient evidence of any difference in race (p=0.70) or sex (p=1) by PRD status, while as expected gestational age at birth was lower for those with PRD (p=0.01; **Table 1**). The age (in days) at the time of sample collection varied by subject, and it ranged from four days to 6 months after birth, depending on the gestational age at birth. However, age was consistent in terms of corrected gestational age, which ranged between 39 to 41 weeks. Detailed diagnostic and demographic information for the individual subjects in provided in **Supplemental Table 2**.

Transcriptomic profiles were generated from sorted and purified CD8+ T cells obtained from the blood collected at discharge. The analytical data set includes values from 13,455 genes for 130 samples, post filtering. As reported previously, the average number of sequence reads in the samples were high (9.93 ± 3.69 million sequence reads). Overall, approximately 60%

of possible genes showed detectable transcript as expected for a subset of differentiated cell type (16) as shown in **Supplemental Figure 2**.

Molecular Markers for BPD

Gene expression patterns associated with cumulative oxygen exposure (over the first 14 days of life) in subjects at greatest risk for BPD (born at GAB<29 weeks; n=72) was assessed. Rank correlation analysis identified 501 genes to be significantly associated (at pFDR<0.1) with oxygen exposure, of which 403 were upregulated in BPD, while 98 genes were downregulated in BPD. The magnitude of change, however, was not large, with only 1 gene (*GPCPD1*) induced by 2.3 fold, while all other changes were less than 2-fold, irrespective of the directionality. Twelve of these genes (*RETN, EPHX2, CD27, NOSIP, APOA1BP, TMCO6, KLHL3, B3GALNT1, SLC9A4, PRKCD, ZNF791, and B3GNT2*) were also identified as differentially expressed in BPD in an independent study studying BPD markers in PBMCs (15). The 501 genes, when further assessed for functional classification by Ingenuity Pathway Analysis (IPA), were found to be associated with 104 pathways and 300 upstream regulators (**Figure 1**). The pathways associated with oxygen utilization included TGFB signaling (epithelial-mesenchymal transition), and multiple immune signaling pathways, while tubule formation associated and immunologic molecules were present among the upstream regulators.

Data from 72 subjects born at <29 weeks CGA was used to identify gene expression associated with BPD as defined by physiologic challenge (RAC) or by Shennan criteria. Using SAM-Seq (at mFDR<0.1), 571 genes were differentially expressed in subjects receiving a diagnosis of BPD (Shennan). While all of the 571 genes were upregulated in BPD, only five genes (*GPCPD1, TMEM2, USP2, TSPYL2, and ELL2*) showed a magnitude of induction of greater than 2.0 fold. Fourteen of these genes (*RNF125, FEM1C, FAM54A, ZNF791, AXIN2, B3GNT2, ZNF565, SPON1, TIPARP, ZBTB3, FAM115C, PELO, MXD1, and PFKFB3*) were also identified as differentially expressed, and over expressed in BPD in an independent study looking at BPD markers in PBMCs (15). These 571 genes, when further assessed for functional classification by IPA, identified 113 canonical pathways and 409 upstream regulators to be associated with BPD. In addition, 105 genes were differentially expressed (SAM-

TABLE 1 | Subject demographics.

Subjects with GAB < 29 Weeks (n=72)	BPD N=34	Non BPD N=38	P value*	
GAB in Weeks, mean ± SD	26.1 ± 1.4	26.9 ± 1.5	0.03	
Male Sex, No. (%)	17 (50)	18 (47)	1.00	
White Race No. (%)	20 (59)	20 (53)	0.64	
Demographics All Subjects (n=130)	**PRD YES N=70**	**PRD NO N=46**	**PRD NA# N=14**	**P value***
GAB in Weeks, mean ± SD	28.5 ± 3.3	31.1 ± 5.2	32.2 ± 4.1	0.01
Male Sex, No. (%)	34 (48)	23 (50)	10 (71)	1.00
White Race, No. (%)	43 (61)	30 (65)	9 (64)	0.7

*Fisher's Exact or t-test P-value.
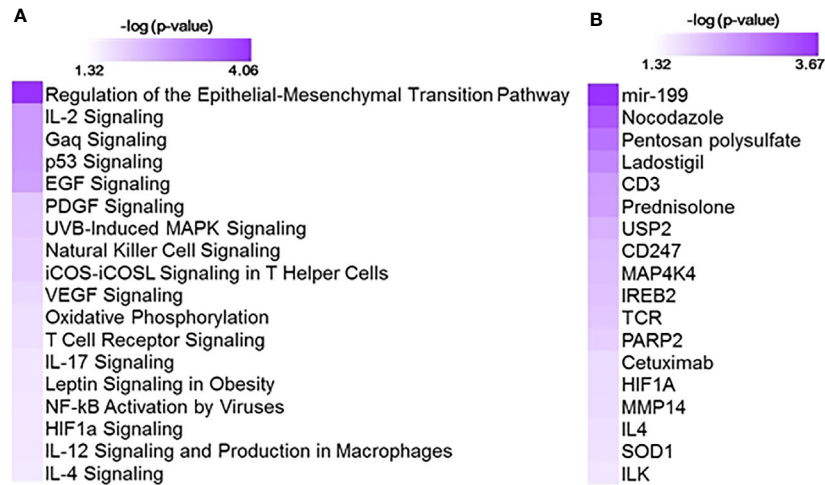GAB, gestational age at birth; PRD NA#, PRD Status Not Available.

FIGURE 1 | Functional analysis using 501 genes associated with oxygen utilization at day 14 (Oxygen$_{AUC14}$) identified with 104 pathways and 300 upstream regulators. Shown here are selected significant canonical pathways **(A)** and upstream regulators **(B)**, along with their significance level (-logP) as generated by Ingenuity Pathway Analysis (IPA).

Seq at mFDR<0.1) in subjects who failed a room air challenge. While all of the 105 genes were upregulated in BPD, only two genes (*GPCPD1 and TSPYL2*) showed a magnitude of induction of greater than 2.0 fold. These 105 genes when analyzed by IPA, identified 18 canonical pathways and 415 upstream regulators (**Figure 2**). Among the pathways associated with BPD, *PEDF, CD40, PI3K/AKT, VEGF* and *NF-kB* signaling were predicted to be activated, while p53 signaling was inhibited in BPD. Among the BPD associated upstream regulators, *CD23, CD28, PTEN,*

and *TCR*, are inhibited, while *NFκB* inhibitor, *camptothecin*, and *dexamethasone* were activated in BPD.

When adjusted for gestational age at birth, 75 genes were associated with the diagnosis of BPD (Shennan) as identified by the Likelihood Ratio Test (LRT at FDR<0.1), and were upregulated in BPD. On further analysis by IPA, the genes provided 113 canonical pathways and 409 upstream regulators (**Figure 3**), of which neuroinflammation pathway appeared to be inhibited, while upstream regulators, *CD23, CD28,* and *TCR*, are
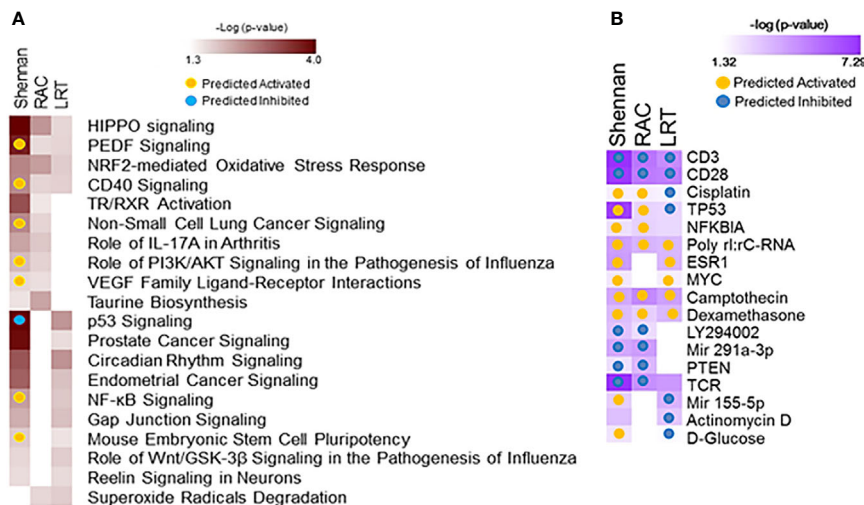

FIGURE 2 | Functional analysis of genes associated with bronchopulmonary dysplasia (BPD). Separate analyses were performed for gene sets identified from analysis of 571 genes associated with BPD (Shennan) and 101 genes associated with room-air challenge (RAC) as identified by SAM-Seq, and 92 genes identified by Screened Principal Components Analysis (sPCA) [adjusted for gestational age at birth (GAB)] to be associated with BPD (Shennan). Selected canonical pathways **(A)** and upstream regulators **(B)** identified are listed for each analysis (columns), with significance (-logP) and directionality (activated/inhibited) as generated by Ingenuity Pathway Analysis (IPA).
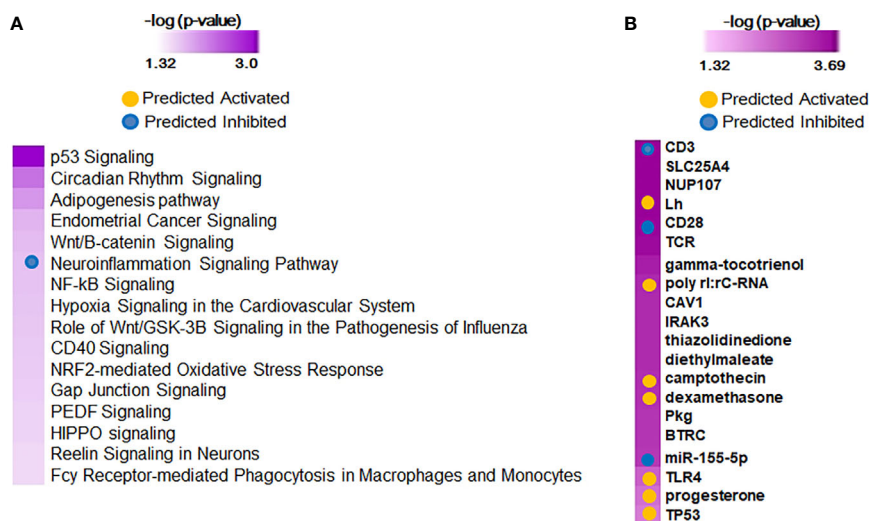
FIGURE 3 | Functional analysis of 75 predictor genes for bronchopulmonary dysplasia (BPD) defined by Likelihood Ratio Test (LRT), identified 113 canonical pathways and 409 upstream regulators. Selected significant canonical pathways **(A)** and upstream regulators **(B)** identified are shown, along with their significance level (-logP) as generated by Ingenuity Pathway Analysis (IPA).

inhibited, and *camptothecin*, and *dexamethasone* were activated in BPD. Screened Principal Components Analysis (sPCA), with screening adjusted for GAB, outperformed both screened LASSO and forward selection and identified a classifier gene set consisting of 92 genes (naïve AUC=0.86; CV-AUC=0.71) associated with BPD (Shennan), representing 21 canonical pathways and 253 upstream regulators (**Supplemental Figure 3**). All of the 92 genes were upregulated in BPD, however, none of the genes except *GPCPD1* had a magnitude of induction of greater than two-fold. These 92 genes were also inclusive of all the 75 genes identified by LRT. We subsequently assessed gene expression changes in BPD based on multiple physiologic and clinical definitions and were successful in identifying nine genes (*GPCPD1, MTSS1L, USP15, DDX24, KLF9, CLK1, ZC3H7A, ITCH,* and *PIK3R1*) that were consistently different, and upregulated in BPD, irrespective of definitions, or analytical approaches (**Supplemental Figure 4**).

Classifiers for PRD Status

Gene expression data from all subjects with PRD status (n=116), was used to identify a set of marker genes, based on our novel *canonical pathway analysis* (45), in order to classify the subjects by PRD status (PRD: n=70 and No PRD: n=46). Screened logistic forward selection outperformed both screened LASSO and sPCA, and gestational age was excluded since it did not improve performance. This process identified a set of 28 genes, derived from four canonical pathways (**Table 2**), which predicted PRD status with a moderately high degree of accuracy (naïve AUC=0.85; CV-AUC=0.70) (**Figure 4**). Interestingly, gene predictors of PRD were associated with pathways involving *TGFB* signaling, and organic ion transport. Several genes identified are also involved in T cell skewing and activation (e.g., *PRKCZ* and *FKBP1A*).

qPCR Validation

Molecular validation of the BPD-associated transcriptomic changes, was attempted on a set of eleven genes selected based on their magnitude of difference or biological relevance, by quantitative reverse transcriptase-polymerase chain reaction (qPCR) (**Table 3**). For each gene, the UQ normalized RNA-seq counts were correlated with the gene expression levels determined by qPCR of the CD8+ T cell cDNA as defined by the dCt. *GAPDH* was used as the endogenous control or housekeeping gene, whose Ct was subtracted from gene Ct to determine the dCT values for each of the genes tested. Spearman's rank correlation coefficient was estimated along with an associated p-value for each gene. We observed validation of the sequence data for nine of the 11 genes, as defined by a significant Spearman rank correlation (p<0.05) in expression between sequence-based and qPCR-based expression levels. Three genes (*KLF9, DLG5,* and *ZNF44*) differed in expression between BPD and non-BPD subjects (p<0.05), while one additional gene (*PSME4*) had borderline yet insufficient evidence of a difference (0.05<p<0.10).

DISCUSSION

Premature birth is defined as birth taking place prior to 37 weeks of GAB. Prematurity associated lung diseases have been reported to affect not only children as newborns but to also predispose to prolonged respiratory morbidity later in life. Unfortunately, not much is known about the pathophysiology of the prematurity associated lung diseases, such as BPD, and other chronic and prolonged childhood respiratory diseases. BPD is a complex disorder involving genetic–environmental interactions, with each preterm subject having a range (e.g., 1%–99%) of both hereditary

TABLE 2 | Canonical pathway based analysis selected four pathways including a total of 28 genes using cross-validated screened logistic forward selection for prediction of Post-Prematurity Respiratory Disease (PRD) status.

Source: Pathway	Pathway Coefficient (logOR)	Gene	Gene OR
Reactome: Recycling Of Bile Acids And Salts	-1.49	SLC10A1	0.61
		SLC27A5	0.66
		SLCO1A2	0.71
Reactome: Transport Of Organic Anions	1.71	SLCO1A2	0.71
		SLCO3A1	1.27
		SLCO4A1	1.58
		SLCO4C1	1.45
Reactome: TGF-Beta Receptor Signaling In Epithelial To Mesenchyme Transition	-0.49	ARHGEF18	0.93
		CGN	0.84
		F11R	0.77
		FKBP1A	0.72
		PARD3	0.58
		PARD6A	0.88
		PRKCZ	0.69
		RHOA	0.76
		RPS27A	0.85
		SMURF1	0.91
		TGFB1	1.07
		TGFBR1	0.81
		TGFBR2	1.09
		UBA52	0.87
Biocarta: Cell2Cell Pathway	-0.46	ACTN1	0.5
		CSK	0.86
		CTNNA1	0.87
		CTNNB1	1.01
		PECAM1	0.65
		PTK2	0.76
		PXN	0.76
		VCL	0.82

*Pathway logOR are log of odds ratios per standard deviation (SD) of the 1st PC of the screened genes (Wilcoxon p<0.1) within the pathway. *** Gene OR = exp (Gene Loading * log (Screened Pathway OR)/SDPathway) = (Pathway OR)(Gene Loading/(SD of Pathway)). Shown are pathway and gene names, estimated pathway coefficients (logOR), and constrained gene odds ratios (OR) factoring in the PCA loadings for each gene within a pathway.*

and environmental risks (46). We have previously published data on hereditary components to BPD risk by genetic analysis in the PROP cohort, including some of the subjects described in the current manuscript (47, 48). Transcriptomic assessment using gene-expression microarrays has previously been used to identify markers for normal lung development as well as BPD (14, 49, 50). As an alternative to lung tissues, gene expression analyses using peripheral blood have been used in lung diseases to study pathogenesis, severity, and recently as a diagnostic tool (45, 51). In this study, we have used high-throughput sequencing to explore peripheral gene-expression changes associated with prematurity and helps to add to the literature base on potential defects in the immune system of infants with BPD (11). Our analysis identified 571 genes differentially expressed in subjects with diagnosed instances of BPD when compared to extremely low birth weight (ELBW) controls born at less than 29 weeks of GAB by SAM-Seq. An independent study also that examined BPD marker genes from bulk PBMCs identified 12 genes (*RETN, EPHX2, CD27, NOSIP, APOA1BP, TMCO6, KLHL3, B3GALNT1, SLC9A4, PRKCD, ZNF791*, and *B3GNT2*) that were differentially expressed in BPD (15). Thus, our results are consistent with published literature.
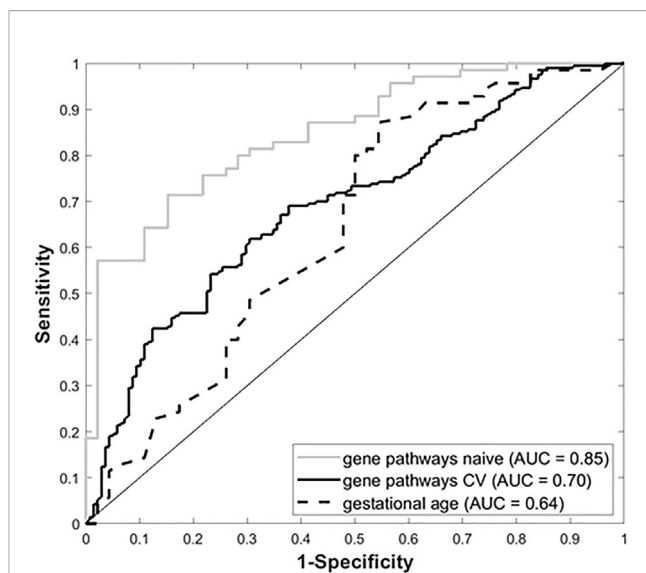

FIGURE 4 | Receiver Operating Characteristic (ROC) curves for gestational age and our four-pathway 28-gene model, with associated Area Under the ROC Curve (AUC).

TABLE 3 | Validation of bronchopulmonary dysplasia (BPD) markers.

Gene	Rank Correlation (r, GAPDH)	P-Value
USP15	0.49	<0.01
SOD2	0.35	<0.01
ITCH	0.13	0.27
AHR	0.28	0.02
*PSME4***	0.34	<0.1
STAT1	0.60	<0.01
GFI1	0.18	0.14
*KLF9**	0.29	0.02
*DLG5**	0.27	0.02
*ZNF44**	0.25	0.04
NEAT1	0.42	<0.01

*KLF9, DLG5, and ZNF44 differed in expression between BPD and non-BPD subjects (p<0.05),
**PSME4 had borderline yet insufficient evidence of a difference between BPD and non-BPD (0.05<p<0.10).

Our gene list is further restricted to 92, when we adjust for gestational age at birth (sPCA) as shown in **Supplemental Figure 4**. In addition, we have identified markers, pathways and upstream regulators putatively associated with cumulative oxygen utilization. The pathway most closely associated with oxygen utilization involves several *TGFB* signaling genes (epithelial-mesenchymal transition pathway), which is important for helping to diminish CD8+ T cell activation (28–32). Additional pathways involved with T cell activation (e.g., T cell receptor signaling) and differentiation (e.g., IL-12 signaling). Thus, we have identified pathways that could be useful for identifying new therapeutic targets to treat the postnatal inflammation of preterm infants and to improve the health of children with BPD (52).

Among the differentially expressed genes associated with BPD, the *PFKFB3* gene was not only consistently identified as

differentially expressed in BPD subjects irrespective of the approach used, but was also identified as a BPD marker in an independent transcriptomic analysis of PBMCs derived from BPD subjects (15). Through murine studies *PFKFB3* has been identified as potential therapeutic target for the treatment of Pulmonary Hypertension (PH) (53). Increased expression *PFKFB3* in BPD is consistent with PH associated with BPD which is characterized by abnormal vascular remodeling, and vascular growth arrest, which are well documented pathophysiology associated with BPD (54). *In vitro* studies have reported that *PFKFB3-kinase* activity attenuates the activation of T cells, and demonstrated the effectiveness of *PFKFB3* antagonists, even in small amounts, as T cell immunosuppressive agents (55). One of the differentially expressed genes, *KLF9*, has been previously identified as differentially expressed in T cells from patients with autoimmune rheumatoid arthritis (56), while another, *DLG5*, is involved in the *HIPPO* pathway and modulated *TGFB* signaling (57–59).

Pathway analyses indicate dysregulation of *NRF2, HIPPO* and *CD40* pathways to be consistently associated with BPD. Another marker, *PSME4*, has been associated with tuberculosis through *in vitro* and *in vivo* cultures, and is also related to *CD4, IFNβ1*, and *TGFBI* pathways (60). Interestingly, *NRF2* has been linked to the generation of reactive oxygen species that contributes to inflammation in a variety of diseases (61). The *HIPPO* has been shown to play a role in T cell receptor signaling and in Th17 differentiation (62, 63). Systemic administration of agonistic *CD40* antibody has been shown to increase CD8+ T cell responses in the lungs of non-human primates (64). While *IL2* signaling is expected, *Gnaq* is known to play a role in survival of immune cells (B and T-cells) (65, 66). In addition, *iCOS-iCOSL* are known to be involved in T-cell skewing, and reduced SOD expression is related to impaired CD8+ T cell responses in tumor infiltrating lymphocytes (67). Interestingly, regulators associated by anti-survival, dexamethasone and camptothecin, appear to be activated in BPD, while T cell co-receptors, *CD3* and *CD28*, appear to be inhibited in BPD. One study shows that T cells isolated from patients with chronic viral infection rely on topoisomerase activity to maintain DNA stability and inhibit apoptosis (68). It is well established that chronically activated T cells become hyporesponsive to T cell receptor mediated stimulation (69). These may indicate potential arrest in lung development as a consequence of BPD and indicate induced immune and stress response as a result of therapeutic responses to BPD, due to either oxidative stress, or surfactant treatment (70). Thus, we have identified significant differences in the expression of several genes and pathways in BPD subjects that are related to T cell signaling and effector function.

In addition to studying CD8+ T cell gene expression in relationship to BPD, we also explored multiple approaches to identify a set of genes whose expression may be useful for classification of markers associated with Post-Prematurity Respiratory Disease (PRD) respiratory morbidity. This included approaches to leverage biological priors as a means of identifying the most robust predictors since (1) expression

changes at the individual gene level alone may not be sufficient to identify biologically meaningful data and (2) there exists substantial statistical advantages to dimension reduction strategies in the analysis of genome-wide data. We used a curated list of genes (71) to partition our transcriptomic data into 1,330 biologically-relevant gene sets. Using a novel data reduction approach, we identified a 28 gene set classifier that groups subjects according to their PRD status with a moderately high degree of accuracy. Similar to BPD, the *TGFB* pathway was also found to be associated with PRD status. Several genes identified control T cell function. For instance, *SMURF1* accumulation in cells infected with RNA viruses leads to the downregulation of Type I Interferons (72). *RhoA* is a GTPase plays an essential role in the migration and activation of T cells (73). *PRKCZ* is a protein kinase C family member that is highly expressed in Th2 CD4+ T cells (74). *PARD3* and *PARD6A* have been shown regulate the *RhoA* signaling pathway (75). *PARD3* also modulates *HIPPO* signaling pathways (75). *FKBP1A* is a signaling molecule in the *mTOR* pathway that regulates memory CD8+ T cell formation (76). Expression of *F11Rm* which encodes for the protein junctional adhesion molecule A, plays a role in T cell adhesion and migration that is upregulated in T cells of lupus patients (77). Their association with persistent respiratory disease indicates propensity to future immunological complications, which may result in chronic lung disorders. Even with the limited set of predictive markers, it does appear that gene expression changes, in peripheral blood at the time of initial discharge after birth, are indicative of future respiratory diseases later in childhood.

One of the limitations of this study is the use of CD8+ T-cells in identifying biomarkers of a lung disease. While peripheral markers have been widely used in identifying expression based markers, it needs to be acknowledged that these hematopoietic cells are from different cell lineage from the pulmonary system, which is the primary organ system affected in BPD and subsequently in PRD. However, even with only one cell type of a different lineage, we have been able to identify previously known, as well as novel molecular markers, and pathways, associated with pulmonary disease due to premature birth. In a longitudinal study involving a subset of subjects from the current cohort, we were able to identify differences in T cell development, post birth, which were able to predict respiratory outcome at 1 year of age (25). Another limitation of the study is the usage of multiple definitions of BPD in order to assess transcriptomic changes caused by it. Despite the various definitions and analytical approaches used, we were able to identify a set of nine marker genes that were observed to be increased in BPD, irrespective of the diagnostic definition and approach used. By using approaches similar to those presented here, researchers will be able to develop better correlations between clinical courses of preterm infants with alterations in the immune system. Previously published data indicate that the patterns of cytokines in the blood of BPD patients relates to the subtype of disease (19). Through the current study we have established proof-of-principle that gene expression provides value for predicting respiratory morbidity following pre-term birth.

In conclusion, we have successfully generated genome-wide transcriptomic data from sort-purified peripheral CD8+ T cells obtained from early pre-term, late preterm, and term infants. We have identified molecular markers, pathways and upstream regulators putatively associated with cumulative oxygen utilization, BPD diagnosis, and PRD prediction. Further studies are needed to determine if the findings are unique to the circulating T cells sampled in this study, or reflective of similar effects in other cells including lung parenchymal cells.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in dbGap (https://www.ncbi.nlm.nih.gov/gap/) with accession number phs001297.v1.p1.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the (Institutional Review Boards of the University of Rochester (IRB# 00037933) and the Children's Hospital of Buffalo (IRB# 612707). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

SB, JM, AB, and DP analyzed the data. SB, RM, and DP wrote the paper. RR, AR, GP, and TM conceived and designed the study. RM, TM, and GP oversaw the project. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

PROP Investigators

Vanderbilt University:

Judy Aschner, MD, Amy B Beller BSN, Candice Fike, MD, Scott Guthrie, MD, Tina Hartert, MD, Nathalie Maitre, MD, Paul Moore, MD, Mark O' Hunt, Theresa J. Rogers, RN, Odessa L. Settles, RN, MSN, CM, Steven Steele, RN, Marshall Summar, MD, Sharon Wadley, BSN, RN, CLS

University of Rochester-University at Buffalo:

Investigators:

Carl D'Angio, MD, Vasanth Kumar, MD, Tom Mariani, PhD, Gloria Pryhuber, MD, Clement Ren, MD, Anne Marie Reynolds, MD, MPH, Rita M. Ryan, MD*, Kristin Scheible, MD, Timothy Stevens, MD, MPH

Technical Staff:

Heidie Huyck, BS, Valerie Lunger, MS

Study Staff:

Shannon Castiglione, RN, Aimee Horan, LPN, Deanna Maffet, RN, Jane O'Donnell, PNP, Michael Sacilowski, MAT, Tanya Scalise, RN, BSN, Elizabeth Werner, MPH, Jason Zayac, BS Respiratory Therapists and Nurses:

Kim Bordeaux, RRT, Pam Brown, RRT, Julia Epping, AAS, RT, Lisa Flattery-Walsh, RRT, Donna Germuga, RRT, CPFT, Nancy Jenks, RN, Mary Platt, RN, Eileen Popplewell, RRT, Sandra Prentice, CRT

*Present address, Case Western Reserve University, Cleveland, OH

Duke University:

Kim Ciccio, RN, C. Michael Cotten, M.D., Kim Fisher, Ph.D., Jack Sharp, M.D., Judith A. Voynow, M.D.*

*Present address, Virginia Commonwealth University

Indiana University:

Charles Clem, RRT, Stephanie Davis, M.D., Susan Gunn, NNP, CCRC, Lauren Jewett, RN, CCRC Brenda Poindexter, M.D., M.S.#

#Present address, University of Cincinnati

Steering Committee Chair:

Lynn M. Taussig, MD, University of Denver

NHLBI Program Officer:

Carol J. Blaisdell, MD

University of Pennsylvania Data Coordinating Center:

Scarlett Bellamy, ScD, Maria Blanco, BS, Denise Cifelli, MS, Sara DeMauro, MD,

Jonas Ellenberg, PhD, Rui Feng, PhD, Melissa Fernando, MPH, Howard Panitch, MD,

Barbara Schmidt, MD, MSc Pamela Shaw, PhD, Ann Tierney, BA, MS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2020.563473/full#supplementary-material

REFERENCES

1. Ehrenkranz RA, Walsh MC, Vohr BR, Jobe AH, Wright LL, Fanaroff AA, et al. Validation of the National Institutes of Health consensus definition of bronchopulmonary dysplasia. *Pediatrics* (2005) 116(6):1353–60. doi: 10.1542/peds.2005-0249

2. Mathews TJ, Minino AM, Osterman MJ, Strobino DM, Guyer B. Annual summary of vital statistics: 2008. *Pediatrics* (2011) 127(1):146–57. doi: 10.1542/peds.2010-3175

3. Schmidt B, Asztalos EV, Roberts RS, Robertson CM, Sauve RS, Whitfield MF, et al. Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months: results from the trial of indomethacin prophylaxis in preterms. *JAMA J Am Med Assoc* (2003) 289(9):1124–9. doi: 10.1001/jama.289.9.1124

4. (CDC) C.f.D.C. Preterm Birth. 2014 12/23/2014 02/04/2015].

5. Villamor-Martinez E, Alvarez-Fuente M, Ghazi AMT, Degraeuwe P, Zimmermann LJI, Kramer BW, et al. Association of Chorioamnionitis With Bronchopulmonary Dysplasia Among Preterm Infants: A Systematic Review, Meta-analysis, and Metaregression. *JAMA Netw Open* (2019) 2(11):e1914611. doi: 10.1001/jamanetworkopen.2019.14611

6. Wang SH, Tsao PN. Phenotypes of Bronchopulmonary Dysplasia. *Int J Mol Sci* (2020) 21(17):1–21. doi: 10.3390/ijms21176112

7. Jacob SV, Coates AL, Lands LC, MacNeish CF, Riley SP, Hornby L, et al. Long-term pulmonary sequelae of severe bronchopulmonary dysplasia. *J Pediatr* (1998) 133(2):193–200. doi: 10.1016/S0022-3476(98)70220-3

8. Collins JJP, Tibboel D, de Kleer IM, Reiss IKM, Rottier RJ. The Future of Bronchopulmonary Dysplasia: Emerging Pathophysiological Concepts and Potential New Avenues of Treatment. *Front Med (Lausanne)* (2017) 4:61. doi: 10.3389/fmed.2017.00061

9. Balany J, Bhandari V. Understanding the Impact of Infection, Inflammation, and Their Persistence in the Pathogenesis of Bronchopulmonary Dysplasia. *Front Med (Lausanne)* (2015) 2:90. doi: 10.3389/fmed.2015.00090

10. Islam JY, Keller RL, Aschner JL, Hartert TV, Moore PE. Understanding the Short- and Long-Term Respiratory Outcomes of Prematurity and Bronchopulmonary Dysplasia. *Am J Respir Crit Care Med* (2015) 192(2):134–56. doi: 10.1164/rccm.201412-2142PP

11. Kolls JK. Commentary: Understanding the Impact of Infection, Inflammation and Their Persistence in the Pathogenesis of Bronchopulmonary Dysplasia. *Front Med (Lausanne)* (2017) 4:24. doi: 10.3389/fmed.2017.00024

12. Pryhuber GS, Maitre NL, Ballard RA, Cifelli D, Davis SD, Ellenberg JH, et al. Prematurity and respiratory outcomes program (PROP): study protocol of a prospective multicenter study of respiratory outcomes of preterm infants in the United States. *BMC Pediatr* (2015) 15:37. doi: 10.1186/s12887-015-0346-3

13. Dylag AM, Kopin HG, O'Reilly MA, Wang H, Davis SD, Ren CL, et al. Early Neonatal Oxygen Exposure Predicts Pulmonary Morbidity and Functional Deficits at 1 Year. *J Pediatr* (2020) 223:20–8.e2. doi: 10.1016/j.jpeds.2020.04.042

14. Bhattacharya S, Go D, Krenitsky DL, Huyck HL, Solleti SK, Lunger VA, et al. Genome-wide transcriptional profiling reveals connective tissue mast cell accumulation in bronchopulmonary dysplasia. *Am J Respir Crit Care Med* (2012) 186(4):349–58. doi: 10.1164/rccm.201203-0406OC

15. Pietrzyk JJ, Kwinta P, Wollen EJ, Bik-Multanowski M, Madetko-Talowska A, Gunther CC, et al. Gene expression profiling in preterm infants: new aspects of bronchopulmonary dysplasia development. *PloS One* (2013) 8(10):e78585. doi: 10.1371/journal.pone.0078585

16. Misra RS, Bhattacharya S, Huyck HL, Wang JC, Slaunwhite CG, Slaunwhite SL, et al. Flow-based sorting of neonatal lymphocyte populations for transcriptomics analysis. *J Immunol Methods* (2016) 437:13–20. doi: 10.1016/j.jim.2016.07.001

17. Wright CJ, Kirpalani H. Targeting inflammation to prevent bronchopulmonary dysplasia: can new insights be translated into therapies? *Pediatrics* (2011) 128(1):111–26. doi: 10.1542/peds.2010-3875

18. Pagel J, Twisselmann N, Rausch TK, Waschina S, Hartz A, Steinbeis M, et al. Increased Regulatory T Cells Precede the Development of Bronchopulmonary Dysplasia in Preterm Infants. *Front Immunol* (2020) 11:565257. doi: 10.3389/fimmu.2020.565257

19. D'Angio CT, Ambalavanan N, Carlo WA, McDonald SA, Skogstrand K, Hougaard DM, et al. Blood Cytokine Profiles Associated with Distinct Patterns of Bronchopulmonary Dysplasia among Extremely Low Birth

Weight Infants. *J Pediatr* (2016) 174:45–51.e5. doi: 10.1016/j.jpeds.2016.03.058

20. Bruder D, Srikiatkhachorn A, Enelow RI. Cellular immunity and lung injury in respiratory virus infection. *Viral Immunol* (2006) 19(2):147–55. doi: 10.1089/vim.2006.19.147

21. Turunen R, Vaarala O, Nupponen I, Kajantie E, Siitonen S, Lano A, et al. Activation of T cells in preterm infants with respiratory distress syndrome. *Neonatology* (2009) 96(4):248–58. doi: 10.1159/000220764

22. You D, Ripple M, Balakrishna S, Troxclair D, Sandquist D, Ding L, et al. Inchoate CD8+ T cell responses in neonatal mice permit influenza-induced persistent pulmonary dysfunction. *J Immunol* (2008) 181(5):3486–94. doi: 10.4049/jimmunol.181.5.3486

23. Tregoning JS, Yamaguchi Y, Harker J, Wang B, Openshaw PJ. The role of T cells in the enhancement of respiratory syncytial virus infection severity during adult reinfection of neonatally sensitized mice. *J Virol* (2008) 82(8):4115–24. doi: 10.1128/JVI.02313-07

24. Tasker L, Lindsay RW, Clarke BT, Cochrane DW, Hou S. Infection of mice with respiratory syncytial virus during neonatal life primes for enhanced antibody and T cell responses on secondary challenge. *Clin Exp Immunol* (2008) 153(2):277–88. doi: 10.1111/j.1365-2249.2008.03591.x

25. Scheible KM, Emo J, Laniewski N, Baran AM, Peterson DR, Holden-Wiltse J, et al. T cell developmental arrest in former premature infants increases risk of respiratory morbidity later in infancy. *JCI Insight* (2018) 3(4). doi: 10.1172/jci.insight.96724

26. Bhattacharya S, Zhou Z, Yee M, Chu CY, Lopez AM, Lunger VA, et al. The genome-wide transcriptional response to neonatal hyperoxia identifies Ahr as a key regulator. *Am J Physiol Lung Cell Mol Physiol* (2014) 307(7):L516–23. doi: 10.1152/ajplung.00200.2014

27. Keller RL, Feng R, DeMauro SB, Ferkol T, Hardie W, Rogers EE, et al. Bronchopulmonary Dysplasia and Perinatal Characteristics Predict 1-Year Respiratory Outcomes in Newborns Born at Extremely Low Gestational Age: A Prospective Cohort Study. *J Pediatr* (2017) 187:89–97 e3. doi: 10.1016/j.jpeds.2017.04.026

28. Mishra S, Liao W, Liu Y, Yang M, Ma C, Wu H, et al. TGF-beta and Eomes control the homeostasis of CD8+ regulatory T cells. *J Exp Med* (2021) 218(1):1–15. doi: 10.1084/jem.20200030

29. Zhuang H, Zhang C, Hou B. GTF2IRD1 overexpression promotes tumor progression and correlates with less CD8+ T cells infiltration in pancreatic cancer. *Biosci Rep* (2020) 40(9):1–12. doi: 10.1042/BSR20202150

30. Wang Y, Gao Z, Du X, Chen S, Zhang W, Wang J, et al. Co-inhibition of the TGF-beta pathway and the PD-L1 checkpoint by pH-responsive clustered nanoparticles for pancreatic cancer microenvironment regulation and anti-tumor immunotherapy. *Biomater Sci* (2020) 8(18):5121–32. doi: 10.1039/D0BM00916D

31. de Streel G, Bertrand C, Chalon N, Lienart S, Bricard O, Lecomte S, et al. Selective inhibition of TGF-beta1 produced by GARP-expressing Tregs overcomes resistance to PD-1/PD-L1 blockade in cancer. *Nat Commun* (2020) 11(1):4545. doi: 10.1038/s41467-020-17811-3

32. Bromley SK, Akbaba H, Mani V, Mora-Buch R, Chasse AY, Sama A, et al. CD49a Regulates Cutaneous Resident Memory CD8(+) T Cell Persistence and Response. *Cell Rep* (2020) 32(9):108085. doi: 10.1016/j.celrep.2020.108085

33. Tooley WH. Epidemiology of bronchopulmonary dysplasia. *J Pediatr* (1979) 95(5 Pt 2):851–8. doi: 10.1016/S0022-3476(79)80451-5

34. Shennan AT, Dunn MS, Ohlsson A, Lennox K, Hoskins EM. Abnormal pulmonary outcomes in premature infants: prediction from oxygen requirement in the neonatal period. *Pediatrics* (1988) 82(4):527–32.

35. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* (2001) 163(7):1723–9. doi: 10.1164/ajrccm.163.7.2011060

36. Walsh MC, Yao Q, Gettner P, Hale E, Collins M, Hensman A, et al. Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics* (2004) 114(5):1305–11. doi: 10.1542/peds.2004-0204

37. Benaron DA, Benitz WE. Maximizing the stability of oxygen delivered via nasal cannula. *Arch Pediatr Adolesc Med* (1994) 148(3):294–300. doi: 10.1001/archpedi.1994.02170030064015

38. Wai KC, Kohn MA, Ballard RA, Truog WE, Black DM, Asselin JM, et al. Early cumulative supplemental oxygen predicts bronchopulmonary dysplasia in high risk extremely low gestational age newborns. *J Pediatr* (2016) 177:97–102.e2. doi: 10.1016/j.jpeds.2016.06.079

39. Scheible K, Secor-Socha S, Wightman T, Wang H, Mariani TJ, Topham DJ, et al. Stability of T cell phenotype and functional assays following heparinized umbilical cord blood collection. *Cytometry A* (2012) 81p(11):937–49. doi: 10.1002/cyto.a.22203

40. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* (2013) 14(4):R36. doi: 10.1186/gb-2013-14-4-r36

41. Anders S, Pyl PT, Huber W. HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* (2015) 31(2):166–9. doi: 10.1093/bioinformatics/btu638

42. Li J, Tibshirani R. Finding consistent patterns: a nonparametric approach for identifying differential expression in RNA-Seq data. *Stat Methods Med Res* (2013) 22(5):519–36. doi: 10.1177/0962280211428386

43. Bair E, Hastie T, Paul D, Tibshirani R. Prediction by supervised principal components. *J Am Stat Assoc* (2006) 101(473):119–37. doi: 10.1198/016214505000000628

44. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* (2011) 27(12):1739–40. doi: 10.1093/bioinformatics/btr260

45. Bhattacharya S, Rosenberg AF, Peterson DR, Grzesik K, Baran AM, Ashton JM, et al. Transcriptomic Biomarkers to Discriminate Bacterial from Nonbacterial Infection in Adults Hospitalized with Respiratory Illness. *Sci Rep* (2017) 7(1):6548. doi: 10.1038/s41598-017-06738-3

46. Bhandari V, Lodha A. Is bronchopulmonary dysplasia decided before birth? *Pediatr Res* (2020) 87(5):809–10. doi: 10.1038/s41390-020-0819-4

47. Ambalavanan N, Cotten CM, Page GP, Carlo WA, Murray JC, Bhattacharya S, et al. Integrated genomic analyses in bronchopulmonary dysplasia. *J Pediatr* (2015) 166(3):531–7 e13. doi: 10.1016/j.jpeds.2014.09.052

48. Hamvas A, Feng R, Bi Y, Wang F, Bhattacharya S, Mereness J, et al. Exome sequencing identifies gene variants and networks associated with extreme respiratory outcomes following preterm birth. *BMC Genet* (2018) 19(1):94. doi: 10.1186/s12863-018-0679-7

49. Kho AT, Bhattacharya S, Mecham BH, Hong J, Kohane IS, Mariani TJ, et al. Expression profiles of the mouse lung identify a molecular signature of time-to-birth. *Am J Respir Cell Mol Biol* (2009) 40(1):47–57. doi: 10.1165/rcmb.2008-0048OC

50. Kho AT, Bhattacharya S, Tantisira KG, Carey VJ, Gaedigk R, Leeder JS, et al. Transcriptomic analysis of human lung development. *Am J Respir Crit Care Med* (2010) 181(1):54–63. doi: 10.1164/rccm.200907-1063OC

51. Bhattacharya S, Tyagi S, Srisuma S, Demeo DL, Shapiro SD, Bueno R, et al. Peripheral blood gene expression profiles in COPD subjects. *J Clin Bioinform* (2011) 1(1):12. doi: 10.1186/2043-9113-1-12

52. Davidson LM, Berkelhamer SK. Bronchopulmonary Dysplasia: Chronic Lung Disease of Infancy and Long-Term Pulmonary Outcomes. *J Clin Med* (2017) 6(1):1–20. doi: 10.3390/jcm6010004

53. Cao Y, Zhang X, Wang L, Yang Q, Ma Q, Xu J, et al. PFKFB3-mediated endothelial glycolysis promotes pulmonary hypertension. *Proc Natl Acad Sci U S A* (2019) 116(27):13394–403. doi: 10.1073/pnas.1821401116

54. Hansmann G, Sallmon H, Roehr CC, Kourembanas S, Austin ED, Koestenberger M, et al. Pulmonary hypertension in bronchopulmonary dysplasia. *Pediatr Res* (2020). doi: 10.1038/s41390-020-0993-4

55. Telang S, Clem BF, Klarer AC, Clem AL, Trent JO, Bucala R, et al. Small molecule inhibition of 6-phosphofructo-2-kinase suppresses t cell activation. *J Trans Med* (2012) 10(1):95. doi: 10.1186/1479-5876-10-95

56. Ye H, Zhang J, Wang J, Gao Y, Du Y, Li C, et al. CD4 T-cell transcriptome analysis reveals aberrant regulation of STAT3 and Wnt signaling pathways in rheumatoid arthritis: evidence from a case-control study. *Arthritis Res Ther* (2015) 17:76. doi: 10.1186/s13075-015-0590-9

57. Kwan J, Sczaniecka A, Heidary Arash E, Nguyen L, Chen CC, Ratkovic S, et al. DLG5 connects cell polarity and Hippo signaling protein networks by linking PAR-1 with MST1/2. *Genes Dev* (2016) 30(24):2696–709. doi: 10.1101/gad.284539.116

58. Venugopal P, Veyssiere H, Couderc JL, Richard G, Vachias C, Mirouse V, et al. Multiple functions of the scaffold protein Discs large 5 in the control of growth, cell polarity and cell adhesion in Drosophila melanogaster. *BMC Dev Biol* (2020) 20(1):10. doi: 10.1186/s12861-020-00218-0

59. Sezaki T, Tomiyama L, Kimura Y, Ueda K, Kioka N, et al. Dlg5 interacts with the TGF-beta receptor and promotes its degradation. *FEBS Lett* (2013) 587(11):1624–9. doi: 10.1016/j.febslet.2013.04.015

60. Kaewseekhao B, Naranbhai V, Roytrakul S, Namwat W, Paemanee A, Lulitanond V, et al. Comparative Proteomics of Activated THP-1 Cells Infected with Mycobacterium tuberculosis Identifies Putative Clearance Biomarkers for Tuberculosis Treatment. *PloS One* (2015) 10(7):e0134168. doi: 10.1371/journal.pone.0134168

61. Mohan S, Gupta D. Crosstalk of toll-like receptors signaling and Nrf2 pathway for regulation of inflammation. *BioMed Pharmacother* (2018) 108:1866–78. doi: 10.1016/j.biopha.2018.10.019

62. Stampouloglou E, Cheng N, Federico A, Slaby E, Monti S, Szeto GL, et al. Yap suppresses T-cell function and infiltration in the tumor microenvironment. *PloS Biol* (2020) 18(1):e3000591. doi: 10.1371/journal.pbio.3000591

63. Zhou J, Zhang N, Zhang W, Lu C, Xu F. The YAP/HIF-1α/miR-182/EGR2 Axis is Implicated in Asthma Severity by Control of Th17 Cell Differentiation. *Res Square* (2020) 1–28. doi: 10.21203/rs.3.rs-61653/v1

64. Thompson EA, Liang F, Lindgren G, Sandgren KJ, Quinn KM, Darrah PA, et al. Human Anti-CD40 Antibody and Poly IC:LC Adjuvant Combination Induces Potent T Cell Responses in the Lung of Nonhuman Primates. *J Immunol* (2015) 195(3):1015–24. doi: 10.4049/jimmunol.1500078

65. Misra RS, Shi G, Moreno-Garcia ME, Thankappan A, Tighe M, Mousseau B, et al. G alpha q-containing G proteins regulate B cell selection and survival and are required to prevent B cell-dependent autoimmunity. *J Exp Med* (2010) 207(8):1775–89. doi: 10.1084/jem.20092735

66. Sun Y, Wang Y, Chen S, Fan G, Zhang J, Dai F, et al. Expression of Galphaq Is Decreased in Lymphocytes from Primary Sjogren's Syndrome Patients and Related to Increased IL-17A Expression. *J Immunol Res* (2018) 2018:8212641. doi: 10.1155/2018/8212641

67. Siska PJ, Beckermann KE, Mason FM, Andrejeva G, Greenplate AR, Sendor AB, et al. Mitochondrial dysregulation and glycolytic insufficiency functionally impair CD8 T cells infiltrating human renal cell carcinoma. *JCI Insight* (2017) 2(12):1–14. doi: 10.1172/jci.insight.93411

68. Ji Y, Dang X, Nguyen LNT, Nguyen LN, Zhao J, Cao D, et al. Topological DNA damage, telomere attrition and T cell senescence during chronic viral infections. *Immun Ageing* (2019) 16:12. doi: 10.1186/s12979-019-0153-z

69. Saeidi A, Zandi K, Cheok YY, Saeidi H, Wong WF, Lee CYQ, et al. T-Cell Exhaustion in Chronic Infections: Reversing the State of Exhaustion and Reinvigorating Optimal Protective Immune Responses. *Front Immunol* (2018) 9:2569. doi: 10.3389/fimmu.2018.02569

70. Wang J, Dong W. Oxidative stress and bronchopulmonary dysplasia. *Gene* (2018) 678:177–83. doi: 10.1016/j.gene.2018.08.031

71. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U.S.A.* (2005) 102(43):15545–50. doi: 10.1073/pnas.0506580102

72. Zhang L, Liu J, Qian L, Feng Q, Wang X, Yuan Y, et al. Induction of OTUD1 by RNA viruses potently inhibits innate immune responses by promoting degradation of the MAVS/TRAF3/TRAF6 signalosome. *PloS Pathog* (2018) 14 (5):e1007067. doi: 10.1371/journal.ppat.1007067

73. Bros M, Haas K, Moll L, Grabbe S. RhoA as a Key Regulator of Innate and Adaptive Immunity. *Cells* (2019) 8(7):1–30. doi: 10.3390/cells8070733

74. Wang Z, Yoo YJ, De La Torre R, Topham C, Hanifin J, Simpson E, et al. Inverse correlation of TRIM32 and PKCzeta in Th2 biased inflammation. *J Invest Dermatol* (2020). doi: 10.1016/j.jid.2020.09.021

75. Lv XB, Liu CY, Wang Z, Sun YP, Xiong Y, Lei QY, et al. PARD3 induces TAZ activation and cell growth by promoting LATS1 and PP1 interaction. *EMBO Rep* (2015) 16(8):975–85. doi: 10.15252/embr.201439951

76. Araki K, Turner AP, Shaffer VO, Gangappa S, Keller SA, Bachmann MF, et al. mTOR regulates memory CD8 T-cell differentiation. *Nature* (2009) 460 (7251):108–12. doi: 10.1038/nature08155

77. Tsou PS, Coit P, Kilian NC, Sawalha AH. EZH2 Modulates the DNA Methylome and Controls T Cell Adhesion Through Junctional Adhesion Molecule A in Lupus Patients. *Arthritis Rheumatol* (2018) 70(1):98–108. doi: 10.1002/art.40338