Proteomics-Based Mapping of Bronchopulmonary Dysplasia-Associated Changes in Noninvasively Accessible Oral Secretions

Saima Ahmed, MS[1,2,*], Oludare A. Odumade, MD, PhD[2,3,4,*], Patrick van Zalm, MS[1,2], Benoit Fatou, PhD[1,2,4], Rachel Hansen, BS[5], Camilia R. Martin, MD, MS[6], Asimenia Angelidou, MD, PhD[2,4,5,†], and Hanno Steen, PhD[1,2,4,†]

Objective To determine if oral secretions (OS) can be used as a noninvasively collected body fluid, in lieu of tracheal aspirates (TA), to track respiratory status and predict bronchopulmonary dysplasia (BPD) development in infants born <32 weeks.

Study design This was a retrospective, single center cohort study that included data and convenience samples from week-of-life (WoL) 3 from 2 independent preterm infant cohorts. Using previously banked samples, we applied our sample-sparing, high-throughput proteomics technology to compare OS and TA proteomes in infants born <32 weeks admitted to the Neonatal Intensive Care Unit (NICU) (Cohort 1; n = 23 infants). In a separate similar cohort, we mapped the BPD-associated changes in the OS proteome (Cohort 2; n = 17 infants including 8 with BPD).

Results In samples collected during the first month of life, we identified 607 proteins unique to OS, 327 proteins unique to TA, and 687 overlapping proteins belonging to pathways involved in immune effector processes, neutrophil degranulation, leukocyte mediated immunity, and metabolic processes. Furthermore, we identified 37 OS proteins that showed significantly differential abundance between BPD cases and controls: 13 were associated with metabolic and immune dysregulation, 10 of which (eg, SERPINC1, CSTA, BPI) have been linked to BPD or other prematurity-related lung disease based on blood or TA investigations, but not OS.

Conclusions OS are a noninvasive, easily accessible alternative to TA and amenable to high-throughput proteomic analysis in preterm newborns. OS samples hold promise to yield actionable biomarkers of BPD development, particularly for prospective categorization and timely tailored treatment of at-risk infants with novel therapies. *(J Pediatr 2024;270:113774)*.

Very preterm newborns pose a unique management challenge in the neonatal intensive care unit (NICU) given the need for continuous monitoring of multiple body compartments.[1] To date, laboratory monitoring often relies on serial blood sampling, which (a) requires an invasive procedure, (b) is sometimes inaccessible, (c) increases the risk of phlebotomy-induced anemia,[2] and (d) may not provide diagnostic or prognostic information pertinent for clinical outcomes that matter most to clinicians and may affect long-term newborn health, such as bronchopulmonary dysplasia (BPD). In addition, organ- or tissue-specific human samples, such as tracheal aspirates (TA), may be inaccessible as sources of biological and/or biomedical information.

The overarching goal of this study was to demonstrate the feasibility and benefits of applying our advanced versatile sample proteomics pipeline comprising semiautomated sample processing methodology and state-of-the-art liquid chromatography/mass spectrometry (MS) to oral secretion (OS) and tracheal aspirate (TA) samples, with the latter representing the body fluid most relevant to pulmonary health and disease. In this context, we hypothesized that OS are a noninvasively

ANXA1	Annexin A1	IGKV3-20	Immunoglobulin kappa variable 3-20
ANXA8	Annexin A8		
ANXA8L1	Annexin A8-like 1	MS	Mass Spectrometry
BIDMC	Beth Israel Deaconess Medical Center	NICU	Neonatal Intensive Care Unit
		OS	Oral Secretions
BPD	Bronchopulmonary Dysplasia	PCA	Principal Component Analysis
BPI	Bactericidal permeability-increasing protein	RDS	Respiratory Distress Syndrome
		SERPINC1	Antithrombin
CSTA	Cystatin-A	TA	Tracheal Aspirates
FDR	False Discovery Rate	TKT	Transketolase
GA	Gestational Age	TPP2	Tripeptidyl-peptidase 2
		WoL	Week-of-life

From the [1]Department of Pathology, Boston Children's Hospital, Boston, MA; [2]Harvard Medical School, Boston, MA; [3]Division of Neonatology, Boston Children's Hospital and Harvard Medical School, Boston, MA; [4]Precision Vaccines Program, Department of Pediatrics, Boston Children's Hospital, Harvard Medical School, Boston, MA; [5]Department of Neonatology, Beth Israel Deaconess Medical Center, Boston, MA; and [6]Division of Neonatology, Weill Cornell Medicine, New York, NY

*Contributed equally as first authors.
†Contributed equally as last authors.

Funding Sources: The Nutrition and Infant Health Repository was supported by the Douglas K. Richardson Funding Program of the Department of Neonatology, Beth Israel Deaconess Medical Center, Boston, MA. This study was supported by an NIAID Development of Sample Sparing Assays for Monitoring Immune Responses Infrastructure and Opportunity Fund (IOF) award supplementing NIH/NIAID 5 U24 AI152179-03 (HS).

https://doi.org/10.1016/j.jpeds.2023.113774

collected body fluid alternative to TA for noninvasive respiratory monitoring in preterm infants.

Preterm infants are affected by surfactant deficiency, leading to alveolar collapse and development of neonatal respiratory distress syndrome (RDS), which precedes the development of BPD.[3] Despite decreasing rates of invasive ventilation, infants with a gestational age (GA) of <32 weeks are still at high risk for BPD, and with the transition to noninvasive ventilatory strategies, TA are less accessible. As such, easily accessible body fluid alternatives as sources of biological information are urgently needed to track lung development and better understand or predict disease risk.

In our recent pilot study, we applied proteomics to noninvasively collected urine for serial assessments of BPD risk in preterm infants <29 weeks[4] and detected proteomic biomarker candidates that were previously only detected in invasively collected body fluids, such as blood, TA, and/or bronchoalveolar lavage fluid from preterm infants.[4] Motivated by these findings, we decided to explore oral secretions (OS), a readily accessible body fluid in frail newborns, as a convenient sample in lieu of TA for monitoring respiratory health and as a potential source of proteomic biomarker discovery for BPD.

Methods

Study Sample

Newborns were recruited at Beth Israel Deaconess Medical Center (BIDMC) NICU, located in Boston, Massachusetts. BIDMC's NICU provides services to vulnerable infants as young as 22 weeks' GA and has >5000 annual births. Extensive clinical data through hospital discharge/transfer and biological samples were collected as part of a *Nutrition and Infant Health Repository* study (IRB # 2009P000014). We analyzed data and convenience samples from week-of-life (WoL) 3 from 2 independent preterm infant cohorts to compare (1) proteomic signatures between TA and OS (Cohort 1, **Table I**) and (2) the OS proteome between infants who developed BPD vs a control group who did not (Cohort 2, **Table II**). We included infants <32 weeks' GA, who are at high risk to develop BPD as defined and categorized by the 2000 *National Institute of Child Health and Human Development* definition[5] which considers the need for and type of respiratory support provided at 36 weeks' postmenstrual age.

Body Fluid Sample Collection

Samples collected from the third week of life were used for analysis. OS were collected before the infant's feeding time from the buccal mucosa with a cotton swab. Prior to processing, the samples were inspected upon collection for any visible milk and if there was evidence of milk contamination, another collection was pursued before a subsequent feed. The tip of the swab was cut off and placed in a cryovial. 0.5 ml of 10 mM Tris Solution was added and the cryovial was vortexed for 3 minutes. The tip of the swab was carefully removed from the cryovial using forceps, while sweeping the side walls to extract any extra fluid from the cotton tip, and the cryovial was stored at −80 °C until processing.

Table I. Perinatal characteristics of neonates who provided samples for retrospective cohort 1

Variable(s)	All neonates n = 23
Contributed oral secretion sample, N (%)	17 (73.9)
Timing of OS sample collection, day of life, mean (SD)	19.4 (2.9)
Contributed tracheal aspirate sample, N (%)	13 (56.5)
Timing of TA sample collection, day of life, mean (SD)	19.8 (4.5)
Neonatal variables	
Sex, female, N (%)	11 (47.8)
GA, completed weeks, mean (SD)	27.61 (2.57)
<28 wk, N (%)	14 (60.9)
28-32 wk, N (%)	9 (39.1)
Birthweight, grams, mean (SD)	998.3 (366.8)
Birthweight z-score for sex/GA	−0.14 (0.93)
VLBW (<1500g), N (%)	21 (91.3)
ELBW (<1000g), N (%)	13 (56.5)
Received surfactant, N (%)	20 (87)
Resp support at 36 wk postmenstrual age, N (%)	
None	10 (43.5)
Nasal cannula oxygen	8 (34.8)
CPAP	4 (17.4)
Mechanical ventilation	1 (4.3)
Duration of Respiratory support, days, mean (SD)	
Supplemental oxygen via low flow nasal cannula	9.9 (12)
Noninvasive positive pressure (HFNC or CPAP)	29.8 (26.2)
Invasive ventilation	23.8 (28.7)
Maternal Variables	
Age, years, mean (SD)	31.5 (6.3)
Race, N (%)	
White	11 (48)
Black	4 (17)
Latino or Hispanic	1 (4)
Asian	2 (9)
Native Hawaiian or Pacific Islander	0 (0)
Other	5 (22)
Delivery by cesarean, N (%)	17 (73.9)
Antenatal corticosteroids, N (%)	
Partial course (one of 2 doses)	7 (30.4)
Complete course (2 doses)	16 (69.6)

CPAP, continuous positive airway pressure; *ELBW*, extremely low birthweight (defined as birthweight less than 1000 g); *HFNC*, high flow nasal cannula; *VLBW*, very low birthweight (defined as birthweight less than 1500 g).

TA were collected using suction catheters connected to a mucous trap (Cardinal Health), transferred to Eppendorf tubes and centrifuged for 10 minutes, 13 500 rpm at 4 °C on the day of collection. Supernatants were transferred to cryovials and stored in −80 °C until processing.

Sample Preparation

The well-established single-pot, solid-phase-enhanced sample-preparation protocol[6] was used in conjunction with a high-throughput and robust automatic liquid handling robot instrument (Opentrons) to process all TA and OS samples. In short, using a starting volume of 100ul OS or 100ul TA we added dithiothreitol and (1:1 [w/w]) urea in 500 mM Ammonium Bicarbonate Solution pH 8.5 to a final concentration of 10 mM dithiothreitol. The resulting denatured and reduced sample was incubated for 20 minutes at 27 °C and 800 rpm in a ThermoMixer (Eppendorf). Reduced cysteine side chains were alkylated with 50 mM iodoacetamide and incubated for 20 minutes at 27 °C and 800 rpm on a ThermoMixer in the dark. All remaining consecutive incubation steps were performed with an automatic liquid

Table II. Perinatal characteristics of neonates who contributed oral secretion samples by BPD status

Characteristics	All neonates n = 17	Control n = 9	BPD n = 8	*P* value
Sex, female, N (%)	7 (41.2)	4 (44.4)	3 (37.5)	.7880
GA, completed weeks, mean (SD)	28.19 (2.54)	29.81 (2.16)	26.36 (1.47)	.0016*
<28 wk, N (%)	9 (52.9)	2 (22.2)	7 (88.5)	.0041*
28-32 wk, N (%)	8 (47.1)	7 (77.8)	1 (12.5)	
Birthweight, grams, mean (SD)	1066.7 (384.4)	1318.7 (345.4)	783.3 (167.2)	.0014*
Birthweight z-score for sex/GA	−0.23 (0.82)	−0.07 (0.72)	−0.42 (0.93)	.4051
VLBW (<1500g), N (%)	15 (88.2)	7 (77.8)	8 (100)	.1690
ELBW (<1000g, N (%)	9 (52.9)	2 (22.2)	7 (77.8)	.0041*
Weight at 36 wk postmenstrual age, grams, mean (SD)	2192.1 (261.7)	2218.9 (330.2)	2161.9 (173.0)	.6585
Apgar scores, mean (SD)				
1-min	5 (2)	7 (1)	4 (2)	.0029*
5-min	7 (1)	8 (1)	7 (1)	.0196*
Received surfactant, N (%)	14 (82)	6 (66.7)	8 (100)	.0805
Resp support at 36 weeks postmenstrual age, N (%)				
None	9 (52.9)	9 (100)	0 (0)	
Nasal cannula oxygen	5 (29.4)	0 (0)	5 (62.5)	.0008*
CPAP	2 (11.8)	0 (0)	2 (25.0)	
Mechanical ventilation	1 (5.9)	0 (0)	1 (12.5)	
Duration of Respiratory support, days, mean (SD)				
Supplemental oxygen via low flow nasal cannula	7.2 (10.7)	0 (0)	15.4 (10.9)	.0053*
Noninvasive positive pressure (HFNC or CPAP)	26.2 (27.5)	12.7 (15.7)	41.5 (30.6)	.0371*
Invasive ventilation	20.9 (32.7)	0.8 (1.3)	43.6 (36.4)	.0126*
Time to first enteral feeding, days, mean (SD)	3.6 (2.7)	2.7 (3)	4.8 (2)	.1101
Breastmilk as first feeding, N (%)	9 (52.9)	3 (33.3)	6 (75)	.0947
Neonatal morbidities, N (%)				
Intraventricular hemorrhage	6 (35.3)	2 (22.2)	4 (50)	.2658
Necrotizing enterocolitis	1 (5.9)	0 (0)	1 (12.5)	.3506
ROP	8 (47.1)	2 (22.2)	6 (75)	.0303*
Severe ROP (Required Laser treatment)	1 (5.9)	1 (11.1)	0 (0)	.5000
Patent ductus arteriosus	8 (47.1)	3 (33.3)	5 (62.5)	.2574
Late onset sepsis (sepsis after the first 72 h of life)	4 (17.6)	2 (22.2)	2 (25)	.9012
Postnatal Interventions, N (%)				
Vasopressors	1 (5.9)	0 (0)	1 (12.5)	.3506
Systemic corticosteroids	5 (29.4)	0 (0)	5 (62.5)	.0112*
Indomethacin	8 (47.1)	2 (22.2)	6 (75)	.0303*
Antibiotics (at time of sampling)	4 (23.5)	1 (11.1)	3 (37.5)	.2418
Maternal data				
Age, years, mean (SD)	31.8 (6.56)	33.8 (6.29)	29.5 (6.90)	.1939
Race, N (%)				
White	8 (47)	5 (56)	3 (38)	
Black	3 (18)	0 (0)	3 (38)	
Latino or Hispanic	1 (6)	1 (11)	0 (0)	.8565
Asian	2 (12)	2 (22)	0 (0)	
Native Hawaiian or Pacific Islander	0 (0)	0 (0)	0 (0)	
Other	3 (18)	1 (11)	2 (25)	
Delivery by cesarean, N (%)	13 (76.5)	6 (66.7)	7 (87.5)	.3339
Antenatal corticosteroids, N (%)				
Partial course (1 of 2 doses)	4 (23.5)	3 (33.3)	1 (12.5)	.3339
Complete course (2 doses)	13 (76.5)	6 (66.7)	7 (87.5)	
Maternal co-morbidities, N (%)				
Chorioamnionitis	2 (11.8)	0 (0)	2 (25)	.1705
Pregnancy-induced hypertension	1 (5.9)	1 (11.1)	0 (0)	.3466
Pre-eclampsia	4 (23.5)	2 (22.2)	2 (25)	.9012
Gestational diabetes	1 (5.9)	1 (11.1)	0 (0)	.3466

CPAP, continuous positive airway pressure; *ELBW*, extremely low birthweight (defined as birthweight less than 1000 g); *ROP*, retinopathy of prematurity; *HFNC*, high flow nasal cannula; *VLBW*, very low birthweight (defined as birthweight less than 1500 g).
*$P < .05$.

handling robot instrument (Opentrons). Samples were then incubated with 1ug of trypsin for 2 hours at 37 °C at 1000 rpm on ThermoMixer. After trypsin digestion, the samples were centrifuged 10 minutes at 3220 × g before acidification using 2% v/v formic acid.

MS and Data Analysis

All samples were processed uniformly with the approximate 0.1ug/ul protein concentration per the proteomics workflow described in previous publications[7] and as per standards to normalize in proteomics field. ∼ 200 ng of protein for each sample was injected using state-of-the-art proteomic technologies starting with liquid chromatography separation on an EVOSEP ONE LC system (EVOSEP) connected to a timsTOF Pro2 MS (Bruker, Billerica) using the high-throughput 60-samples-per-day method. All raw files were processed with MS Fragger software v3.1.1[8] using Uniport Human reference proteome

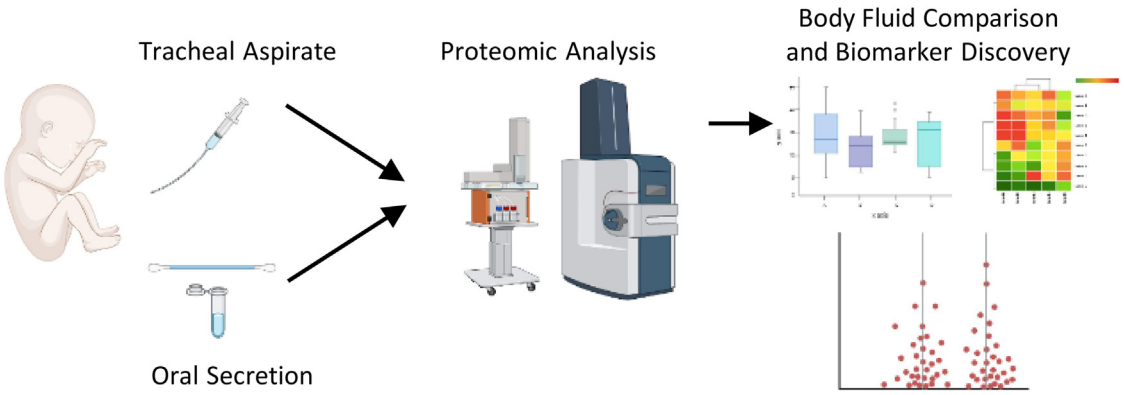

Figure 1. Schematic of proteomics platform used to analyze 100ul of noninvasively collected oral secretions and invasively collected tracheal aspirates from premature neonates at week 3 of life.

database with the following modifications: carbamidomethylated cysteine residues (fixed), oxidation of methionine (variable). All other parameters were set to default. Multiple testing correction was performed by permutation-based false discovery rate (FDR) using standard parameters in Perseus v 1.6.10.43 (FDR <0.05, s0 = 0.1) The resulting protein quantification matrix was then used for statistical analysis. For specific MS signal intensity-based normalization, we normalized our data using the median intensity value for all control samples. We analyzed the data with and without this method of normalization and observed the same robust outcome with both methods. We observed a normal distribution of all signal intensities across each sample analyzed.

Statistical Analysis

Post processing of all MS data included log-transformation of the intensity values for each protein. We then applied the parametric Student's *t*-test to identify proteins of differential abundance between BPD infants and controls. Further statistical analysis was performed using R-studio v 2022.02.0+443, Perseus v 1.6.10.43 and GraphPad Prism 9. Biological Pathway analysis was performed using string-db.org.

Results

Characteristics of the Study Samples

All OS and TA samples had previously been collected from very preterm newborns during the third week of life. The characteristics of the unpaired OS and TA samples used for the proteome comparison are provided in **Table I**. To characterize the OS proteome in BPD, we compared OS samples from 8 BPD cases and 9 controls in an independent cohort; **Table II** provides the characteristics of this second study cohort. All samples were processed and analyzed using our robust high throughput body fluid proteomic pipeline (**Figure 1**).

Comparing the Preterm Proteome between TA and OS

To understand better the similarities and differences between the OS and TA proteomes, and to highlight OS as an alternative noninvasive body fluid for monitoring the respiratory status of preterm infants, we performed a proteomic comparison of TA and OS from preterm newborns. In total, we identified 2918 and 3035 proteins in TA and OS samples, respectively. When only including proteins observed in at least half the samples, we still identified 1014 TA and 1294 OS proteins. These curated protein lists were used for all subsequent analyses. A Venn diagram shows 607 proteins unique to OS, 327 proteins unique to TA, and an overlap of 687 proteins present in both body fluids (**Figure 2**, A), highlighting a significant degree of similarity between the OS and TA proteomes.

We then performed a more detailed bioinformatic analysis of the fluid-specific and shared proteomes. Gene Ontology annotation of the biological pathways highlighted functional differences and commonalities between these 2 types of body fluids. Dominant pathways to proteins uniquely detected in OS were related to protein translation (FDR 2.08 e-37) and metabolism, such as amide biosynthesis (FDR1.03e-06) and peptide metabolism (FDR 1.32e-31). Top protein pathways associated with the proteins unique to TA were related to the immune response (FDR 3.90e-09), and exocytosis (1.37e-18) (**Figure 2**, A). Enriched biological pathways shared between OS and TA were immune effector processes (FDR 4.15e-11), neutrophil degranulation (FDR 4.47e-11), leukocyte mediated immunity (FDR 1.02 e-10), and metabolic processes (FDR 2.5 e-07).

For a more in-depth quantitative comparison, we performed a correlation analysis of the OS and TA proteomes. The bulk of the correlating OS and TA proteomes were associated with biological pathways related to immunity (FDR 6.59e-45) and viral processes (FDR 1.22e-39). Proteins showing increased abundances in OS belonged to biological pathways such as Epidermis Development (FDR 5.05e-39) and Cornification (FDR 3.10 e-46). Proteins showing higher abundance in TA belonged to biological pathways such as

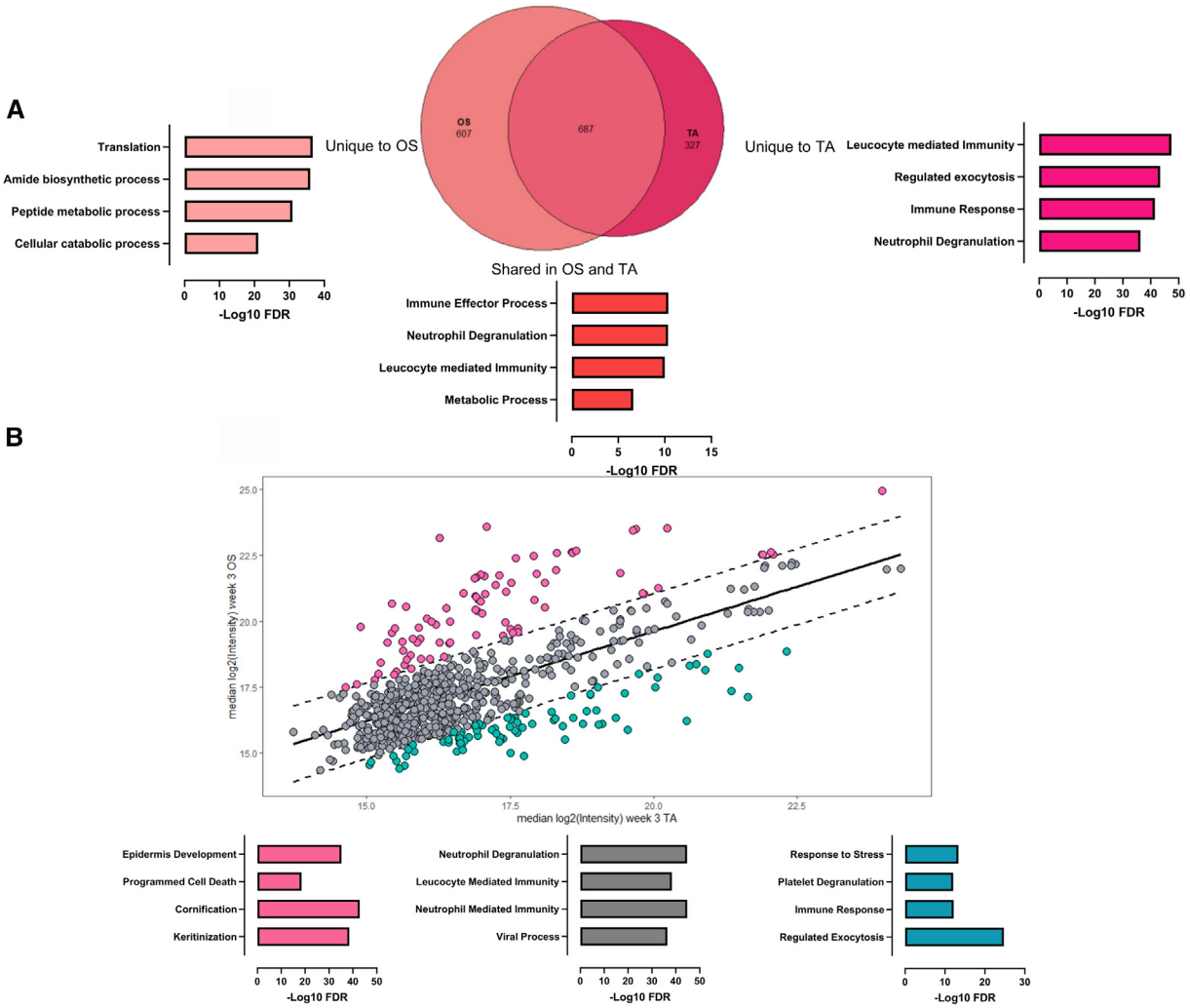

Figure 2. Proteomic comparison of noninvasively collected oral secretions (OS) and invasively collected tracheal aspirates (TA). **(A)** Venn diagram comparing the OS and TA proteomes with biological processes in relative comparisons. Some proteins may be involved in more than 1 pathway. **(B)** Correlation analysis of OS and TA using the 687 proteins shared in both body fluids and associated biological processes.

Response to Stress (FDR 1.99e-10) and Regulated Exocytosis (FDR 1.37e-18) (**Figure 2**, B).

Oral Secretion Proteomic Changes in BPD Cases and Controls

After ascertaining a significant degree of similarity between OS and TA, we continued with the second proteomic analysis: BPD-associated changes in the OS proteomes. Considering the shortcomings of clinical BPD definitions[9] and recognizing that there is a spectrum of pulmonary phenotypes in BPD, we used OS collected early in the NICU course (WoL 3) to examine molecular patterns associated with the later clinical diagnosis of BPD. This cohort is similar but distinct from the one used above for the comparison of the TA and the OS proteomes: similar, in that we used OS samples collected at WoL 3 from neonates <32 weeks' GA; distinct in that there is little overlap between the 2 cohorts (and any overlap is unintentional).

GA does influence the developmental trajectory of the newborn OS proteome. To remove GA differences as a confounder in this specific analysis, we performed a correlation analysis between protein abundance and GA and removed all 218 proteins from further analyses that showed a statistically significant GA-abundance correlation ($P < .05$). All further results described and discussed below refer to the set of proteins after removal of those proteins that show GA-dependent abundance differences.

A principal component analysis using all proteins from Cohort 2 showed a clear separation of BPD cases vs controls (**Figure 3**, A). Using a *t*-test and applying a stringent *P*-value cutoff of .01 (ie, equivalent to $-\log_{10}$ of 2.0), we identified 37 proteins that showed significant difference in abundance between BPD cases and controls (**Figure 3**, B). Hierarchical clustering of these 37 significant proteins differentiating BPD cases and controls, provides a sample-resolved view of the protein abundances revealing clear differences in

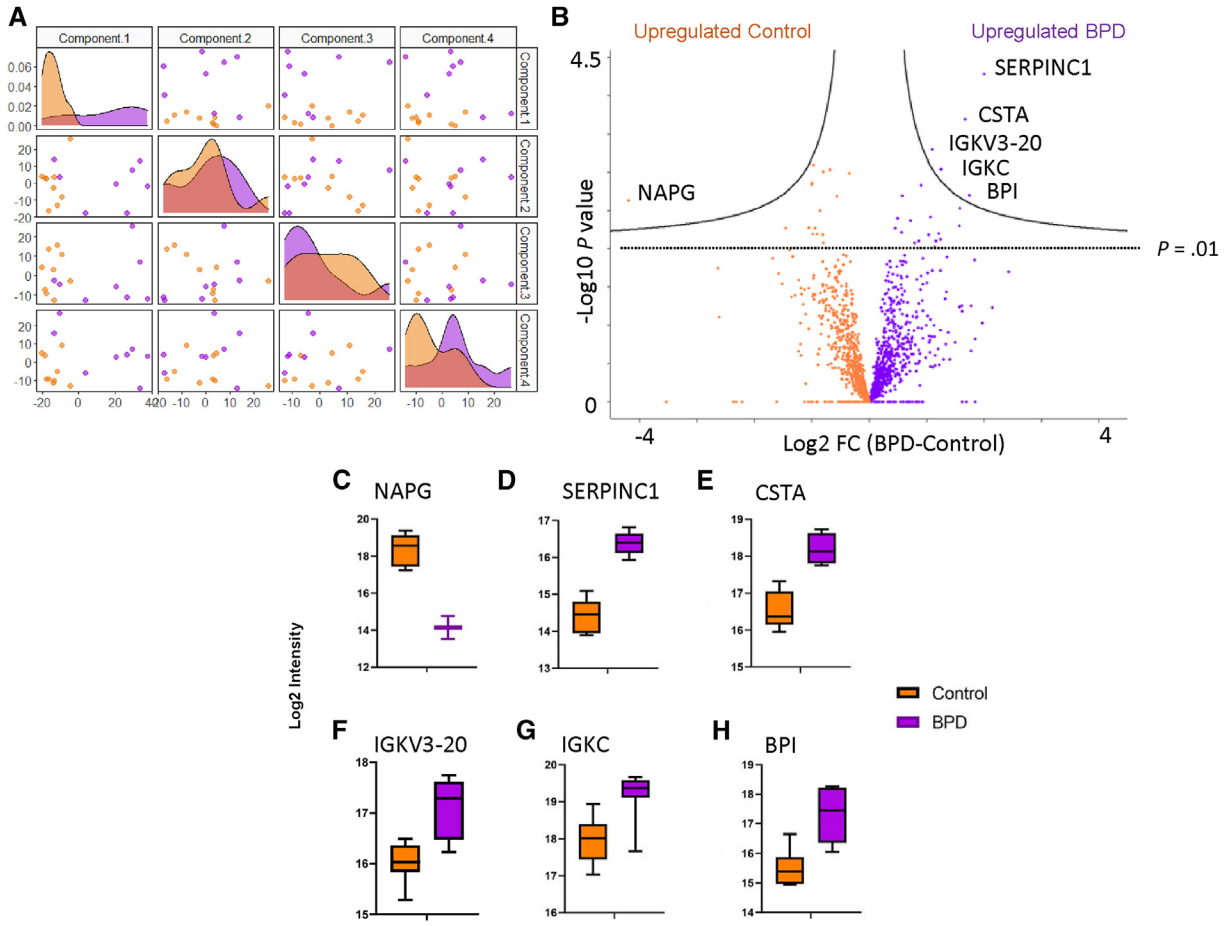

Figure 3. Proteomic changes in oral secretions of bronchopulmonary dysplasia (BPD) cases. **(A)** Principal component analysis of BPD cases vs controls for all proteins. **(B)** Volcano plot comparing BPD cases vs controls with black curve indicating statistically high significance (FDR < 0.05). **(C-H)** Boxplots of highly significant proteins (FDR < 0.05).

protein expression increased (red) or decreased (blue) in BPD (Figure 4).

Among the 37 most significant proteins, 6 proteins also passed the multiple testing correction with an FDR of <0.05 as noted beyond the black curves in the volcano plot (Figure 3, B): NAPG (*N*-ethylmaleimide-sensitive factor (NSF) attachment protein gamma; P = 2.34E-03), Antithrombin (SERPINC1; P = 5.19E-05), Cystatin-A (CSTA; P = 2.01E-04), Bactericidal permeability-increasing protein (BPI; P = 1.98E-03), Immunoglobulin kappa variable 3-20 (IGKV3-20; P = 5.02E-04), and Immunoglobulin kappa constant (P = 9.10E-04). The boxplots for these proteins are shown in Figure 3, C-H highlighting the fact that some of them show perfect separation between BPD and non-BPD cases, and that the differences can exceed 1 order of magnitude (eg, NAPG in Figure 3, C).

To determine if the 37 OS proteins, that showed differential abundances, have been previously associated with BPD, we performed a literature search using the terms "BPD," "biomarker," "premature." Interestingly, 10 of our 37 significant proteins (SERPINC1, CSTA, BPI, S100 Calcium Binding Protein A11, Annexin A1 [ANXA1], Annexin A8-like 1 [ANXA8L1], Annexin A8 [ANXA8], transketolase [TKT], tripeptidyl-peptidase 2 [TPP2], programmed cell death 6 protein) had previously been linked to BPD or other prematurity-related lung disease pathophysiology. However, these proteins had been observed in invasively collected blood or TA samples, but not in noninvasively collected OS.[4,10-15] Our identification of 6 significant metabolic proteins (6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 2, Phosphoglucomutase 2, Guanosine Monophosphate Reductase 2, Polypeptide N-Acetylgalactosaminyltransferase 7, Sorbitol Dehydrogenase, Phosphoglycerate Dehydrogenase) and 7 significant immune proteins (IGKV3-20, Immunoglobulin kappa constant, Immunoglobulin heavy constant alpha 2, Immunoglobulin heavy constant gamma 1, Immunoglobulin kappa variable 3D-20-20, Polymeric immunoglobulin receptor, Immunoglobulin heavy constant alpha 1) also support the hypotheses of metabolic[16] and immune dysregulation[17,18] in BPD.

Discussion

In this study, we compare TA and OS proteomes from very preterm neonates and highlight OS as a potential non-invasively collected body fluid alternative to TA for

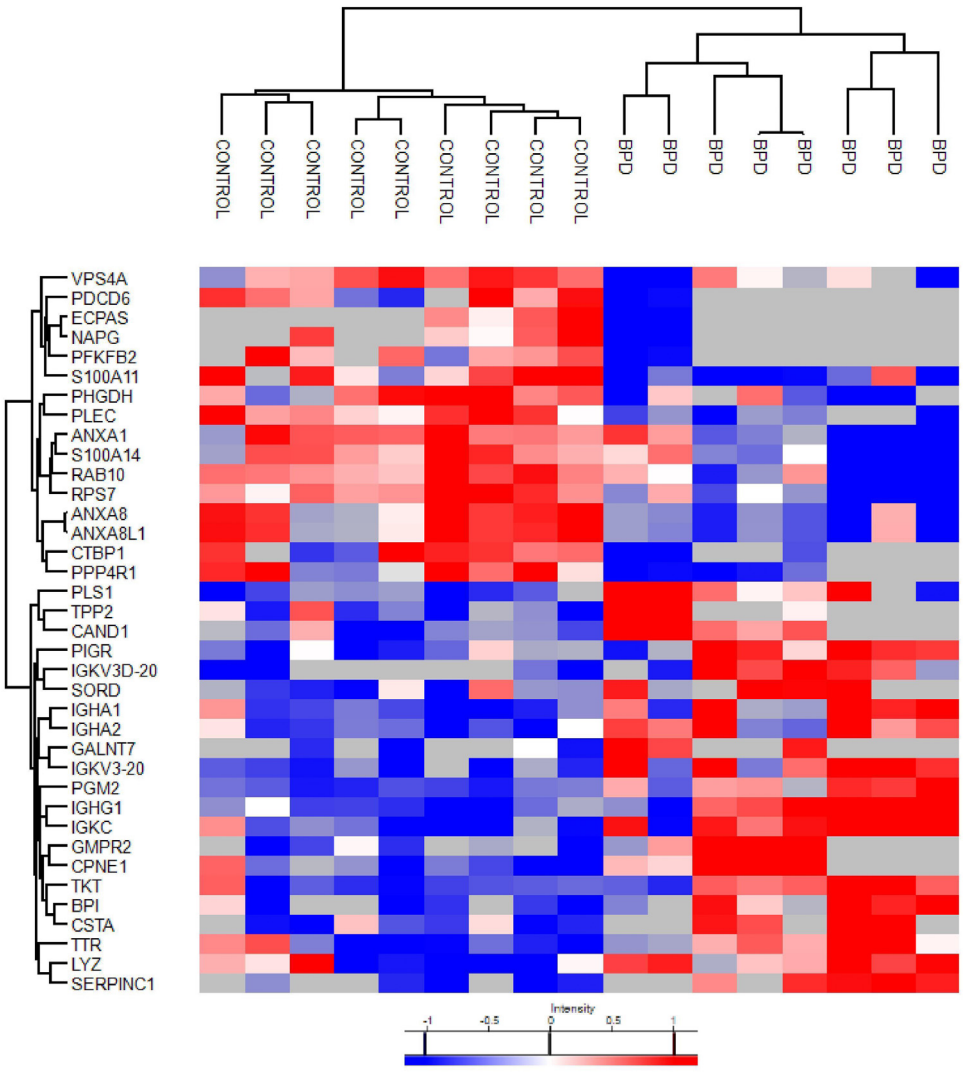

Figure 4. Hierarchical clustering of 37 significant proteins using $P = .01$ cut-off after removal of proteins correlating with gestational age to detect BPD-specific proteomic changes in OS of control vs BPD cases.

monitoring respiratory health. We show significant overlap of proteomic signatures between OS and TA and several highly correlated proteins between the 2 body fluids belonging to biological pathways related to immunity and viral processes, suggesting that the OS proteome composition, (i) is in part related to the tracheal aspirate proteome, (ii) provides a reflection of the immune state, and (iii) could be leveraged to noninvasively monitor proteomic trajectories in the lung compartment. In short, we provide strong evidence that noninvasively accessible OS can readily replace invasive samples such as TA for health/disease monitoring purposes.

Preterm infants are at high risk for developing chronic lung disease of prematurity; however, prognostic biomarkers of BPD development are lacking. Feasibility of proteomic screening for the purposes of disease-specific biomarker discovery has been shown in small preterm cohorts but has so far involved invasively collected biological fluids such as bronchoalveolar lavage fluid[19,20] or blood plasma.[21-23] Our group has previously demonstrated BPD associated proteins in noninvasively collected urine samples closely resembling the molecular changes that were described for blood, TA, and broncho-alveolar lavage from patients with BPD.[4] Motivated by this observation, we investigated OS as a noninvasively collected body fluid to detect molecular patterns associated with BPD. We identified 37 proteins that showed significant difference in abundance between BPD and controls. Among the 37 significant proteins, 10 have been described before in either BPD or other prematurity-related lung conditions but were detected in invasive body fluids like blood or TA, but never in noninvasively accessible OS.

SERPINC1, also known as antithrombin, was the most significantly differentially expressed protein in OS with a 4-fold increase in BPD cases vs controls (P value of 5.2E-05). SERPINC1concentrations are low in healthy infants compared with adults. It has been described that SERPINC1

levels correlate with GA (lower concentrations in preterm neonates); however, they reach term GA values within the first week of life.[24] Increased SERPINC1 has been previously described in lung organotypic cultures using rat fetal mixed lung cells.[25] Exposure to stretch-conditioned medium (which mimics mechanical ventilation induced stretching) led to decreased thrombin generation and decreased fibrin deposition, compared with control-conditioned medium, linking cellular stretch to increased SERPINC1 activity.[10,25] Conversely, SERPINC1 activity is decreased in the presence of neonatal RDS, characterized by alveolar collapse, and correlates with RDS severity.[26] However, a Cochrane meta-analysis of randomized control trials comparing exogenous SERPINC1 administration with placebo/no treatment in preterm infants with RDS showed no therapeutic benefit.[27] In the single included study that reported BPD rates, SERPINC1 was associated with a 25% increase in BPD incidence compared with placebo.[28] Although BPD numbers in that study were low and the criteria used to define BPD were not explicitly stated, these observations, in accordance with our findings, suggest that SERPINC1concentrations may be predictive of BPD risk.

CSTA had a ~3-fold increase in BPD OS compared with controls with a P value of $P = 2.0E-04$. Cystatin family proteins have been previously studied in BPD. For example, Cystatin M was identified as a top discriminating protein in the blood of infants who developed BPD vs those who did not.[12] In a baboon model of BPD, an imbalance between cysteine proteases and cysteine protease inhibitors cystatin B and C was noted, resulting in an overall increase in cysteine protease activity in BPD compared with control lungs.[29] Although CSTA is not specifically described in BPD literature, CSTA is a biomarker for lung cancer and is known to be upregulated in COPD,[11] ie, its increased levels are associated with pathologic lung conditions.

BPI showed a >3-fold increase in BPD OS compared with controls with a P value of 2.0E-03. BPI is an antimicrobial protein with a high affinity for the lipid A region of lipopolysaccharide, targeting Gram-negative bacteria via membrane permeabilization, opsonization, and neutralization of endotoxic activity.[30] BPI expression in cord blood supernatants of preterm infants is higher in those born in the setting of intraamniotic infection.[31] BPI is also expressed in human mucosal epithelia including the lung and intestine.[32] In TA of mechanically ventilated infants of variable GA (23-40 weeks), extracellular BPI was consistently detected and correlated with GA and TA Polymorphonuclear neutrophils, indicating endotoxin accumulation and endotoxin-directed innate immune activation.[30] However, this study focused on ontogenic differences and did not report BPD outcomes.

Additional proteins with known roles in BPD or lung pathophysiology included S100 Calcium Binding Protein A11, ANXA1, ANXA8L1, ANXA8, TKT, TPP2, and programmed cell death protein 6. The S100 family of proteins are observed in several lung diseases such as lung cancer, asthma, COPD, and cystic fibrosis.[13] Of the 3 significant annexin proteins identified in our study (ANXA1, ANXA8L1, ANXA8), ANXA1 was shown to be altered in BPD mouse models.[14,33] Due to its anti-inflammatory activity, ANXA1 is also a target for 9 Food and Drug Administration approved drugs (drugbank.com).[34] Furthermore, annexins were also identified in exosomes isolated from TA from very low birth weight neonates with BPD[15] similarly to TKT, and TPP2,[15] all of which we also identified as being upregulated in the OS of the BPD group.

Overall, our results reflect the notion that the OS proteome reflects TA proteomes confirming (1) OS can be used as a noninvasive, safe, and easily obtainable alternative to TA for respiratory health monitoring of very preterm newborns and (2) OS can be used to identify very preterm newborns at risk for BPD weeks before the clinical diagnosis is made. In addition, since the relationship between proteomics in different compartments remains poorly understood, overlaps in OS and TA proteomes in our data highlight the potential for proteomics from more than 1 source as complementary and most informative in our understanding of health and disease in vulnerable populations.

Our study had some limitations. Firstly, this was a proof-of-concept pilot involving a small number of convenience samples, collected at a later stage of life (week of life 3) with the goal of demonstrating feasibility of using OS for identifying potential markers and providing novel mechanistic insights in BPD development. While our results are congruent with findings from published literature supporting the validity of our findings, a larger, independent cohort study is still required for validation. Although we were able to show significant overlap between the OS and TA proteomes, paired analysis of OS-TA samples from the same individual was not feasible. Our study involved sample analysis at a single timepoint and follow-up studies with prospectively collected, earlier and/or longitudinal samples are needed to study OS proteome composition across time. Furthermore, larger cohorts across a wider GA range will allow for a more detailed analysis of the effects of GA on the OS proteome in the BPD context.

Following the 5-phase PRoBE design from the Early Detection Research Network, the presented data should be considered as Phase 1 and 2,[35] ie, preclinical exploratory phase upon identification of promising direction (Phase 1), and clinical assay and validation with the aim to show that the assay (here LC/MS of the OS samples) detects the established disease (here: differentiate OS from BPD patients from healthy controls) (Phase 2). Now that Phases 1 and 2 are complete, the next step will be the collection of appropriate early time point longitudinal samples so that Phase 3 can be initiated in future work.

In conclusion, we demonstrate the feasibility of using noninvasively collected OS for early detection of BPD-associated protein alterations, some of which have been previously described in BPD literature from invasively collected body fluids and highlight biologically plausible OS proteomic markers predictive of future BPD risk. In the future, we hope to expand our cohorts using noninvasively collected

body fluids not only to discover new proteins and biological pathways in BPD, but also to monitor the health status of premature neonates in the NICU to predict and select high risk cohorts for novel therapies and interventional clinical trials. ■

Statement of Ethics

The BIDMC review board approved the collection of clinical data and oral secretion and tracheal aspirate samples per our approved protocol (IRB # 2009P000014).

CRediT authorship contribution statement

SA developed the oral secretion proteomics methodology, performed the sample and data analysis, drafted the initial manuscript, and oversaw editing and finalizing the manuscript. HS conceptualized and designed the study, interpreted the proteomics findings, and revised the manuscript critically for important intellectual content. OAO and AA designed the cohort and original study, collected the clinical data, interpreted the results, and reviewed and revised the manuscript for important intellectual content. PvZ and BF statistically analyzed and interpreted the data, reviewed, and revised the manuscript for important intellectual content. CRM was the PI of the Nutrition and Infant Health Repository study and reviewed and revised the manuscript for important intellectual content. RH reviewed and revised the manuscript for important intellectual content. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Data Availability Statement

The MS proteomics data can be available upon request.

Declaration of Competing Interest

CRM receives research funding from Mead Johnson Nutrition and serves on the advisory boards of Lactalogics, Plakous Therapeutics, and Vitara Biomedical. None of these relationships have a direct conflict with this study. All other authors have no conflicts of interest to declare.

The authors would like to thank all study participants and their families, as well as the medical and nursing staff at BIDMC without whom this study would not have been possible. We thank Dr DeWayne Pursley for his support of the Nutrition and Infant Health Repository in the BIDMC NICU. We also thank the EPOCH (Exploring Pathways to Optimize Child Health) Neonatology Research Program at BIDMC, and especially Dr Angela Koo Chong, Ms Ciara Olmstead, Ms Kelly Walter and Ms Neha Talele for programmatic support.

Submitted for publication Jul 13, 2023; last revision received Sep 18, 2023; accepted Oct 9, 2023.

Reprint requests: Hanno Steen, PhD, Department of Pathology, Boston Children 's Hospital, 300 Longwood Avenue, Boston, MA 02115. E-mail: Hanno.Steen@childrens.harvard.edu

References

1. Maron JL. The neonatal salivary transcriptome. Cold Spring Harb Perspect Med 2015;6:a026369.
2. Counsilman CE, Heeger LE, Tan R, Bekker V, Zwaginga JJ, Te Pas AB, et al. Iatrogenic blood loss in extreme preterm infants due to frequent laboratory tests and procedures. J Matern Fetal Neonatal Med 2021;34:2660-5.
3. Thebaud B, Goss KN, Laughon M, Whitsett JA, Abman SH, Steinhorn RH, et al. Bronchopulmonary dysplasia. Nat Rev Dis Primers 2019;5:78.
4. Ahmed S, Odumade OA, van Zalm P, Smolen KK, Fujimura K, Muntel J, et al. Urine proteomics for noninvasive monitoring of biomarkers in bronchopulmonary dysplasia. Neonatology 2022;119:193-203.
5. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. Am J Respir Crit Care Med 2001;163:1723-9.
6. Hughes CS, Moggridge S, Muller T, Sorensen PH, Morin GB, Krijgsveld J. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nat Protoc 2019;14:68-85.
7. Lee AH, Shannon CP, Amenyogbe N, Bennike TB, Diray-Arce J, Idoko OT, et al. Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. Nat Commun 2019;10:1092.
8. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. Nat Methods 2017;14:513-20.
9. Ibrahim J, Bhandari V. The definition of bronchopulmonary dysplasia: an evolving dilemma. Pediatr Res 2018;84:586-8.
10. Menache D, Grossman BJ, Jackson CM. Antithrombin III: physiology, deficiency, and replacement therapy. Transfusion 1992;32:580-8.
11. Butler MW, Fukui T, Salit J, Shaykhiev R, Mezey JG, Hackett NR, et al. Modulation of cystatin A expression in human airway epithelium related to genotype, smoking, COPD, and lung cancer. Cancer Res 2011;71:2572-81.
12. Arjaans S, Wagner BD, Mourani PM, Mandell EW, Poindexter BB, Berger RMF, et al. Early angiogenic proteins associated with high risk for bronchopulmonary dysplasia and pulmonary hypertension in preterm infants. Am J Physiol Lung Cell Mol Physiol 2020;318:L644-54.
13. Sattar Z, Lora A, Jundi B, Railwah C, Geraghty P. The S100 protein family as Players and therapeutic targets in pulmonary diseases. Pulm Med 2021;2021:5488591.
14. Raffay TM, Locy ML, Hill CL, Jindal NS, Rogers LK, Welty SE, et al. Neonatal hyperoxic exposure persistently alters lung secretoglobins and annexin A1. Biomed Res Int 2013;2013:408485.
15. Veerappan A, Thompson M, Savage AR, Silverman ML, Chan WS, Sung B, et al. Mast cells and exosomes in hyperoxia-induced neonatal lung disease. Am J Physiol Lung Cell Mol Physiol 2016;310:L1218-32.
16. Yue L, Lu X, Dennery PA, Yao H. Metabolic dysregulation in bronchopulmonary dysplasia: implications for identification of biomarkers and therapeutic approaches. Redox Biol 2021;48:102104.
17. Bhandari A, Bhandari V. Biomarkers in bronchopulmonary dysplasia. Paediatr Respir Rev 2013;14:173-9.
18. Rivera L, Siddaiah R, Oji-Mmuo C, Silveyra GR, Silveyra P. Biomarkers for bronchopulmonary dysplasia in the preterm infant. Front Pediatr 2016;4:33.
19. Magagnotti C, Matassa PG, Bachi A, Vendettuoli V, Fermo I, Colnaghi MR, et al. Calcium signaling-related proteins are associated with bronchopulmonary dysplasia progression. J Proteomics 2013;94:401-12.
20. Vento G, Tirone C, Lulli P, Capoluongo E, Ameglio F, Lozzi S, et al. Bronchoalveolar lavage fluid peptidomics suggests a possible matrix

metalloproteinase-3 role in bronchopulmonary dysplasia. Intensive Care Med 2009;35:2115-24.
21. Forster K, Sass S, Ehrhardt H, Mous DS, Rottier RJ, Oak P, et al. Early identification of bronchopulmonary dysplasia using novel biomarkers by proteomic screening. Am J Respir Crit Care Med 2018;197:1076-80.
22. Kindt ASD, Forster KM, Cochius-den Otter SCM, Flemmer AW, Hauck SM, Flatley A, et al. Validation of disease-specific biomarkers for the early detection of bronchopulmonary dysplasia. Pediatr Res 2023;93:625-32.
23. Zasada M, Suski M, Bokiniec R, Szwarc-Duma M, Borszewska-Kornacka MK, Madej J, et al. Comparative two time-point proteome analysis of the plasma from preterm infants with and without bronchopulmonary dysplasia. Ital J Pediatr 2019;45:112.
24. Peters M, ten Cate JW, Jansen E, Breederveld C. Coagulation and fibrinolytic factors in the first week of life in healthy infants. J Pediatr 1985;106:292-5.
25. Chan AK, Baranowski B, Berry L, Liu M, Rafii B, Post M, et al. Influence of mechanical stretch on thrombin regulation by fetal mixed lung cells. Am J Respir Cell Mol Biol 1998;19:419-25.
26. Schmidt B, Vegh P, Weitz J, Johnston M, Caco C, Roberts R. Thrombin/antithrombin III complex formation in the neonatal respiratory distress syndrome. Am Rev Respir Dis 1992;145(4 Pt 1):767-70.
27. Bassler D, Millar D, Schmidt B. Antithrombin for respiratory distress syndrome in preterm infants. Cochrane Database Syst Rev 2006;2006:CD005383.
28. Fulia F, Cordaro S, Meo P, Gitto P, Gitto E, Trimarchi G, et al. Can the administration of antithrombin III decrease the risk of cerebral hemorrhage in premature infants? Biol Neonate 2003;83:1-5.
29. Altiok O, Yasumatsu R, Bingol-Karakoc G, Riese RJ, Stahlman MT, Dwyer W, et al. Imbalance between cysteine proteases and inhibitors in a baboon model of bronchopulmonary dysplasia. Am J Respir Crit Care Med 2006;173:318-26.
30. Cuenca AG, Wynn JL, Moldawer LL, Levy O. Role of innate immunity in neonatal infection. Am J Perinatol 2013;30:105-12.
31. Faust K, Gopel W, Moser K, Temole G, Bartels M, Wieg C, et al. Differential expression of antimicrobial polypeptides in cord blood samples of preterm and term infants. Acta Paediatr 2014;103:e143-7.
32. Canny G, Levy O, Furuta GT, Narravula-Alipati S, Sisson RB, Serhan CN, et al. Lipid mediator-induced expression of bactericidal/permeability-increasing protein (BPI) in human mucosal epithelia. Proc Natl Acad Sci U S A 2002;99:3902-7.
33. Dylag AM, Misra RS, Bandyopadhyay G, Poole C, Huyck HL, Jehrio MG, et al. New insights into the natural history of bronchopulmonary dysplasia from proteomics and multiplexed immunohistochemistry. Am J Physiol Lung Cell Mol Physiol 2023;325:L419-33.
34. Serres M, Viac J, Comera C, Schmitt D. Expression of annexin I in freshly isolated human epidermal cells and in cultured keratinocytes. Arch Dermatol Res 1994;286:268-72.
35. Feng Z, Pepe MS. Adding rigor to biomarker evaluations—EDRN experience. Cancer Epidemiol Biomarkers Prev 2020;29:2575-82.