Discovery of novel proteomic biomarkers for the early diagnosis of Bronchopulmonary Dysplasia in preterm infants

Fernando Garrido Muñoz[1], Alejandro Fernández Vega[1], Rebeca Gregorio-Hernández[2], Victoria Aldecoa-Bilbao[3], Marta Padín-Fontán[4], Alberto Trujillo-Fagundo[5], Paula Alonso-Quintela[6], Mónica de las Heras-Martín[7], Almudena Alonso-Ojembarrena[1,8]

[1] Perinatal Brain Damage Group. Biomedical Research and Innovation Institute of Cadiz (INiBICA). Puerta del Mar University Hospital, Cadiz, Spain [2] Neonatal Intensive Care Unit, Gregorio Marañón University Hospital, Madrid, Spain [3] Neonatology Department, Hospital Clínic Barcelona. BCNatal - Barcelona Center for Maternal-Fetal and Neonatal Medicine. Barcelona, Spain [4] Neonatal Intensive Care Unit, Alvaro Cunqueiro Hospital, Vigo, Spain.Sanitary investigation institute Galicia Sur (IIS Galicia Sur) [5] Neonatal and Pediatric Intensive Care Unit, Doctor Josep Trueta University Hospital, Girona, Spain [6] Neonatal Intensive Care Unit, Complejo Asistencial Universitario de León, León, Spain. Biomedicine Institute of León (IBIOMED). University of León, León, Spain [7] Neonatal Intensive Care Unit, Basurto University Hospital, Bilbao, Spain
[8] Neonatology Unit. Puerta del Mar University Hospital, Cadiz, Spain

INTRODUCTION

Bronchopulmonary Dysplasia (BPD) is a pulmonary human disease patent in preterm neonates born before 30 weeks of gestation. It severely compromises patients' respiratory airways development and integrity, in some cases causing death. Early intervention of BPD is crucial as it allows its prediction as well as the initiation of therapies when they are most effective. Samples collected from nasopharyngeal aspirates (NPA) – a novel non-invasive sampling method – were analyzed to search for possible BPD biomarkers.

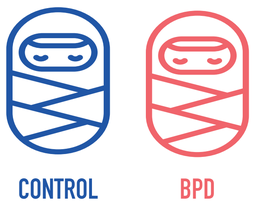
OBJECTIVES

In this preliminary study, we wanted to prove that proteomics applied to NPA could help for BPD early detection and prevention. To do so we:

- **Assess NPA quality** as an useful sample collection method for BPD
- **Search for possible BPD biomarkers** using a novel non-invasive collection method
- **Use of high-throughput mass spectrometry-based proteomics** applied to NPA
- **Bioinformatic approach** for protein differential expression, functional enrichment, exploratory and predictive analyses

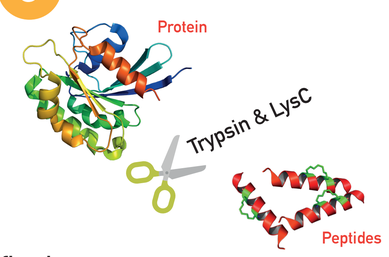
METHODOLOGY

1 Patients



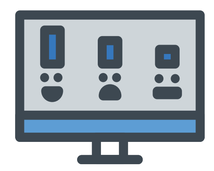
2 Nasopharyngeal aspirate samples

3 Protein extraction



4 Nano HPLC-MS/MS

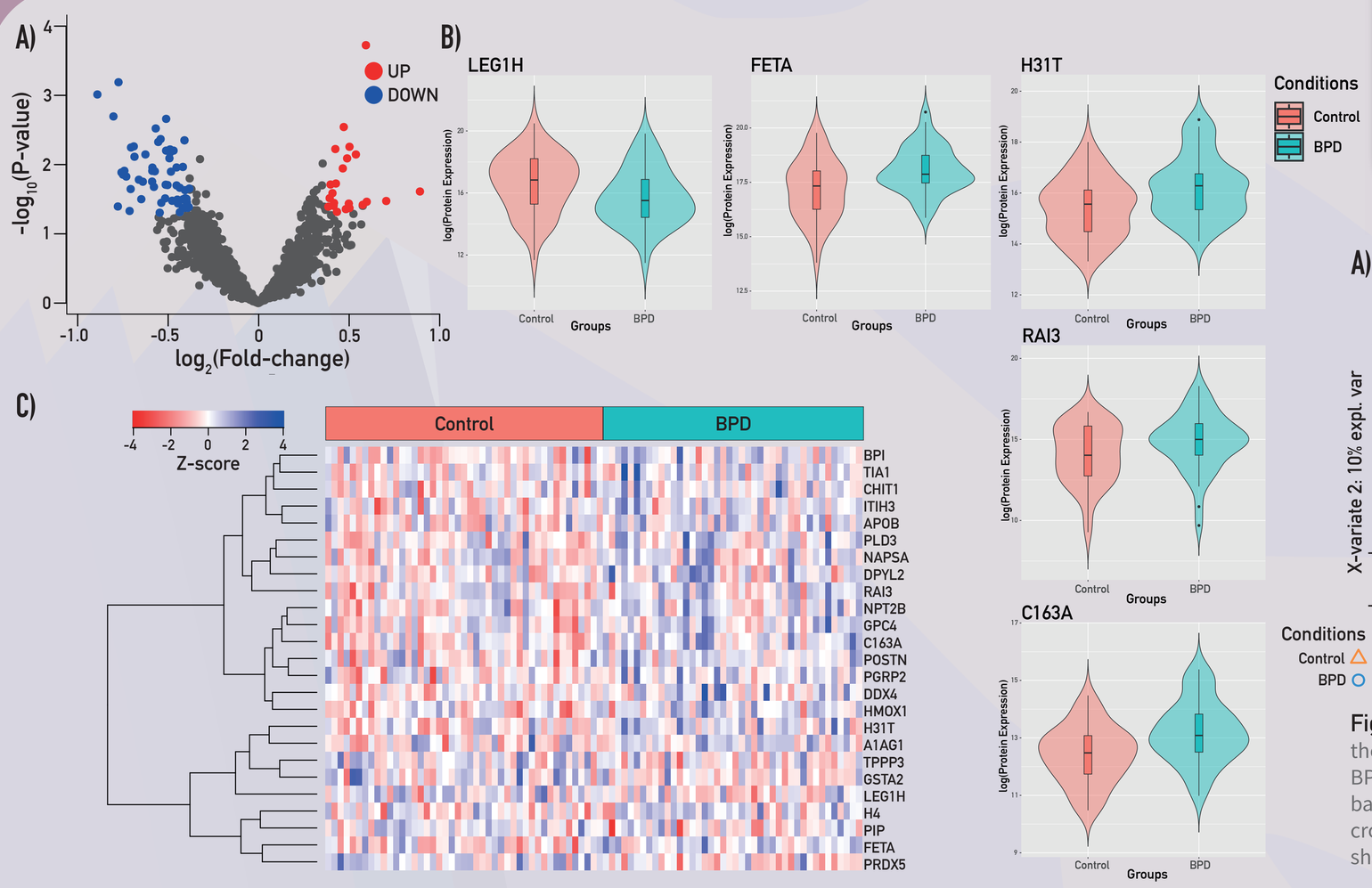
5 Bioinformatics: protein quantification and analysis



R Studio · mixOmics · STRING

DATA. A total of 87 samples was analyzed in this study; samples were collected from preterm neonates born before 30 weeks of gestation at a week post-delivery. It consisted of a cohort of 45 healthy patients and 42 suffering from BPD. Samples were gathered from several spanish hospitals which participate in the multiparametric BPD national proyect PARADYS (clinical trials ID: NCT04785859). Protein extracts were obtained and analyzed with nano-liquid chromatography coupled to mass spectrometry (nLS-MS) working in diaPASEF acquisition. **The software RStudio was selected for bioinformatic analyses. The packages limma, mixOmics and clusterProfiler were mainly used. The online software tool STRING was applied for protein network visualization.**

RESULTS



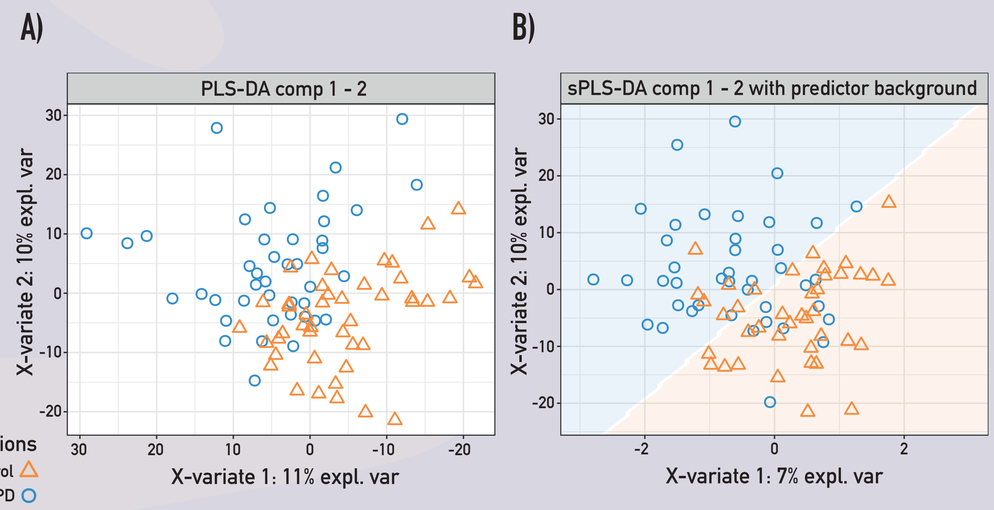
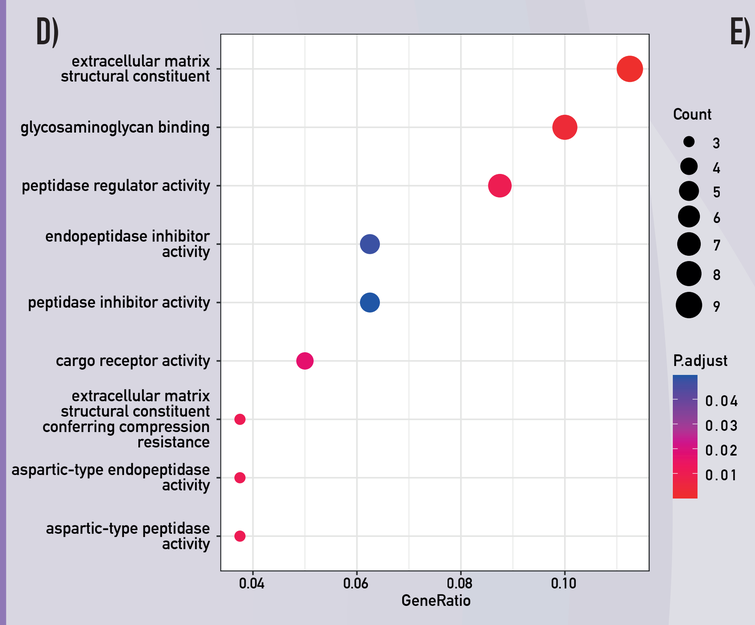


Figure 2. Exploratory and predictive analyses. The RStudio package mixOmics was used for these purposes. A) Partial Least Square discriminant analysis (PLS-DA) between control and BPD groups; B) sparse PLS-DA (sPLS-DA) model for sample discrimination with a predictive background; model was assesed with the "Mfold" algorith, given folds of length 5; internal cross-validation was performed with a number of iterations of 100. A total of 2 components are shown, which explain 17% variability (total of 8 predictive variables).



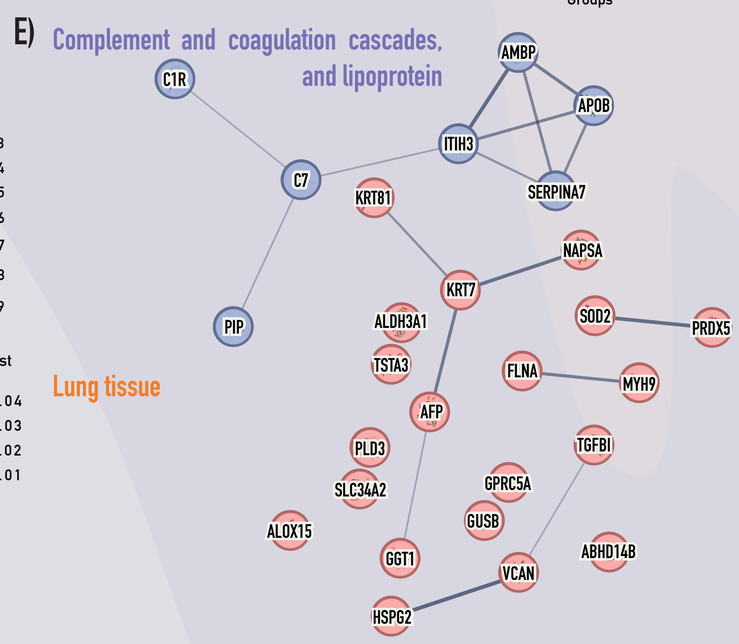


Figure 1. Differential expression and functional enrichment analyses carried out in this study. A) Volcano plot of differentially expressed proteins (DEP); cutoff |fold change| > 1.3, P-value < 0.05); 23 upregulated and 59 downregulated proteins were found (control vs. BPD). B) Violin plots of the top 5 DEP found in this study. C) Heatmap of the top 25 DEP; clustering was performed with them, but samples were grouped given their condition. D) Dotplot of the functional enrichment analysis showing the top 10 functional hits found for all of our DEP from this study. E) STRING networks (string-db.org) highlighting two of the most related functions of our DEP to BPD.

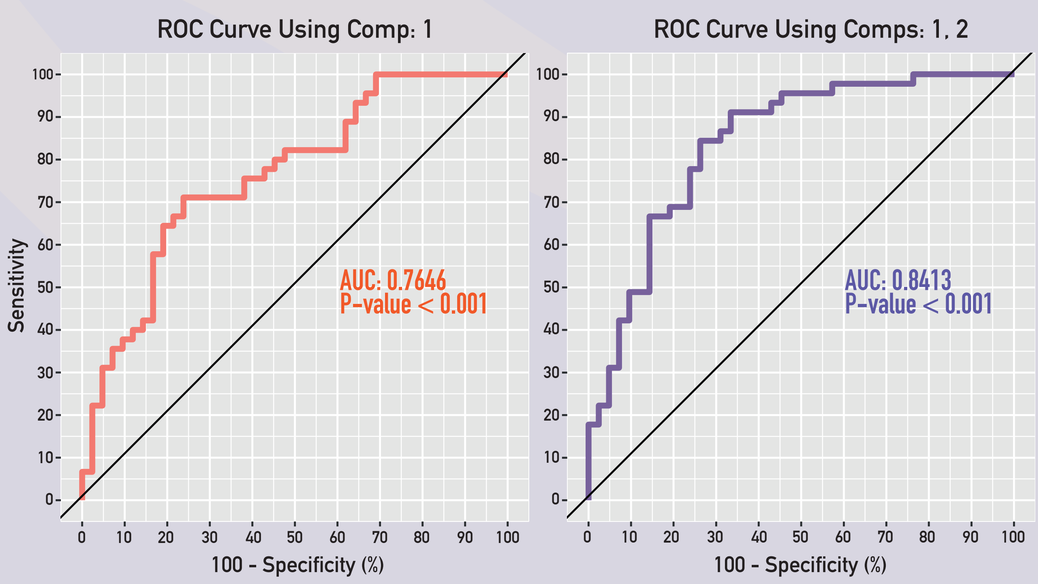


Figure 3. Receiver operating characteristic curves (ROC). ROC curves were built with the previous sPLS-DA predictive model for sample discrimination given 1 and 1, 2 components; a set of 50 randomized samples from this study was used for training, while the rest were for model performance testing. Area Under the Curve (AUC) values are repressented for each of them, as well as their approximated P-values.

CONCLUSION

These preliminary results show how high throughput proteomics has showed the potential of NPA as a non-invasive method capable for prediction and early detection of BPD disease. This represents a step forward in BPD research, which may improve its early detection, prevention and treatment. We spect to analyze a larger sample cohort in the near future, which will again help us to discover BPD at a much deeper level.

PARADYS PROTEOMICS
MultiPARAmetric Bronchopulmonary DYSplasia Prediction

INiBICA Investigación Innovación Biomédica · Hospital Universitario Puerta del Mar · Gobierno de España · Ministerio de Economía y Competitividad · Instituto de Salud Carlos III · Proyect Code: PI21/01172

JABI Jornadas Andaluzas de Bioinformática Sevilla, 3-5 de julio 2023