**Mass Spectrometry-based proteomic analysis of lung tissue from ventilated preterm lambs**

***Supplementary Material to accompany ‘First description of ventilation strategy-specific phenotypes involved in initiating preterm lung injury’***

Prue Pereira-Fantini\*1,2, Kristin Ferguson3, Karen McCall1, Regina Oakley1, Elizabeth Perkins1, Sean Byars4, Nicholas Williamson5, Shuai Nie5 and David G Tingay1,2,3

1Neonatal Research Group, Murdoch Children’s Research Institute, Parkville, Victoria, Australia

2Department of Paediatrics, University of Melbourne, Parkville, Victoria, Australia

3Neonatal Intensive Care Unit, The Royal Children’s Hospital, Parkville, Victoria, Australia

4Howard Florey Institute, Parkville, Victoria, Australia.

5Melbourne Mass Spectrometry and Proteomics Facility (MMSPF), The Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria, Australia

\*Corresponding author

**Abstract**

The associated publication reports proteomic analysis of preterm lung tissue, where we identified protein subtypes associated with different aeration strategies applied at birth in the preterm lung with acute RDS. This protocol describes the sample preparation and mass spectrometry (MS)-based analysis of lung samples using TMT labelling, followed by LC-MS data-dependent acquisition. Lung samples were from two distinct regions; the gravity non-dependent (n=49) and dependent lung regions (n=51). The reported protocol achieved analytical depth of 3754 proteins which was reduced to an analysis data set of 2373 proteins following removal of proteins which were missing from all samples in a 10-plex set.

**Keywords:** *preterm, lung, mass spectrometry, proteomics*

**Procedure**

*Sample acquisition and selection:* Immediately after animals were killed, the lungs were removed *en bloc*. Lung tissue samples from the gravity-non dependent and dependent zones of the right upper lobe were snap frozen and stored at -80°C until analysis. Tissue samples were obtained as part of a larger prior study(*1*), therefore additional data quality checks were performed prior to protein extraction and tissue samples which were not accompanied by a complete functional dataset were excluded.

*Protein extraction and digestion:* Protein extraction was performed on the same day using the same batch of reagents. 100mg of lung tissue was homogenized on ice in 1 ml homogenization buffer (RIPA buffer containing protease inhibitor Merck, Kenilworth, NJ). Homogenized samples were spun at 13,000 rpm for 20 minutes at 4°C and the supernatant transferred to a fresh tube and overnight acetone precipitation. The pellet was washed once with ice-cold acetone and reconstituted in 8M urea in 50mM triethylammonium bicarbonate (TEAB, pH 8.0) at 37°C with vortexing to aid in protein solubilization. Protein concentration was determined using the Pierce bicinchoninic acid (BCA) protein assay with a BSA standard (Thermo Fisher Scientific, Rockford IL). The following steps were carried out with 100 µg protein of each sample. Disulphide bonds were reduced with 10 mM tris(2-carboxyethyl) phosphine (TCEP) for 45 minutes at 37°C. Cysteines were alkylated with 55 mM iodoacetamide (IAA) for 45 minutes at 37°C. Samples were diluted with 5- mM TEAB to a final urea concentration of 1M. Protein was digested with Trypsin (Cat # V5071, Promega, Madison, WI) at a 1:50 w/w enzyme/protein ratio overnight 37 °C. Digestion was terminated by the addition of trifluoroacetic acid to 1.0% and peptides were solid phase extracted using Waters Oasis HLB cartridges (Cat # 186000383, Waters, Taunton, MA) according to manufacturer’s instructions, dried down, and stored at -80°C.

*TMT-labelling:* Peptides were reconstituted in 100 mM TEAB and their concentration was determined by the BCA assay described above. Samples with an insufficient peptide quantity (< 50 ug?) were excluded from TMT labelling. A pooled sample for normalization between runs was prepared by combining 1 µg of peptide from each individual sample. TMT labelling was carried out on 10 µg of protein digest from each individual sample and the pooled sample. Samples were randomly distributed between thirteen 10-plex label sets, along with 1 pooled sample per set. TMT 10-plex labelling reagents (0.8 mg) were each dissolved in 41 µl anhydrous acetonitrile (ACN). Each sample was combined with 4 µl of its respective 10-plex TMT reagent and incubated at 400 rpm, 25°C for 1 hour. 1 µl of 5% hydroxylamine was added, and the sample incubated for a further 15 minutes, 25°C and 400 rpm to quench the reaction. Nine samples and an aliquot of pooled sample for each run were then combined in 1:1:1:1:1:1:1:1:1:1 ratio, the mixture was then dried down, dissolved in 2% ACN/0.05% TFA, and 1 µg of peptide was loaded as described below.

*LC-MS/MS Analysis:* Samples were analysed by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). The nano-LC system, Ultimate 3000 RSLC (Thermo Fisher Scientific, San Jose, CA) was equipped with an Acclaim Pepmap nano-trap column (C18, 100 Å, 75 μm× 2 cm, Thermo Fisher Scientific, San Jose, CA) and an Acclaim Pepmap RSLC analytical column (C18, 100 Å, 75 μm × 50 cm, Thermo Fisher Scientific, San Jose, CA) maintained at a temperature of 50 ℃. Typically for each LC-MS/MS experiment, 1 μg of the peptide mixture was loaded onto the trap column at an isocratic flow of 5 μL/min of 3% CH3CN containing 0.05% trifluoroacetic acid for 6 min before the enrichment column was switched in-line with the analytical column. The eluents used for the LC were water with 0.1% v/v formic acid and 5% v/v dimethyl sulfoxide (DMSO) for solvent A and acetonitrile with 0.1% v/v formic acid and 5% DMSO for solvent B. The gradient used at 300 nL/min was from 3% B to 23% B for 134 min, 23% B to 40% B in 20 min, 40% B to 80% B in 10 min and maintained at 80% B for the final 5 min before dropping to 3% B in 1 min and equilibration for 9 min at 3% B prior to the next analysis. The MS experiments were performed using a nano electrospray ionization source at positive mode and Eclipse Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The spray voltages, capillary temperature and S-lens RF level were set to 1.9 kV, 275°C and 30%. The mass spectrometry data was acquired with a 3 s cycle time for one full scan MS spectra and as many data dependent higher-energy collisional dissociation (HCD)-MS/MS spectra as possible. Full scan MS spectra had a m/z of 375-1500, a resolution of 120,000 at m/z 200, an auto gain control (AGC) target value of 4e5, a maximum ion trapping time of 50 milliseconds. The data dependent HCD-MS/MS of precursor ions (charge states from 2 to 7) was performed using an m/z isolation window of 0.7, a first mass at m/z of 110, a AGC target value of 1.25e5, a normalized collision energy (NCE) of 38%, a resolution of 30,000 at m/z 200 and a maximum ion trapping time of 54 milliseconds. All mass spectrometry data were acquired using Orbitrap mass analyser. Dynamic exclusion was used for 30 s.

*TMT Data Analysis:* TMT-labelled data was analysed by MaxQuant (1.6.17.0). The precursor MS and MS/MS tolerance were set to 20 ppm. The peptides were searched against Ovis aries (Sheep) database from Uniprot with 23111 protein entries. Oxidation of methionine and Protein N-terminal acetylation were set as variable modifications. 10-plex TMT tags on the peptide N terminus and lysine residues and the carbamidomethylation of cysteine were set as fixed modifications. Trypsin was set as enzyme allowing up to two missed cleavage. False discovery rate (FDR) at peptide and protein levels were all at 1%. Quantification was performed based on intensity of the reporter ion, with a mass tolerance of 0.003 Da and correction for isotopic impurities.

Search results and TMT reporter ion intensities were exported as text files. Within the data set unreliable protein groups (contaminants, only identified by site, reversed) and proteins which were missing from all samples in a 10-plex set were removed. Two normalization procedures were then employed as per the method of Plubell *et al* (*2*). The first was applied within each 11-plex experiment. The grand total reporter ion intensity for each channel was multiplied by global scaling factors to adjust its total intensity to the average total intensity across the 10 channels. This corrects for small sample loading and labelling reaction efficiency differences. Secondly, common, pooled internal standards were used to normalize reporter ion intensities of proteins between different TMT experiments. This allowed preservation of individual intensity-scale measurements and avoided calculation of relative intensity measures such as ratios or percentages within each TMT experiment. To accomplish this, scaling factors which were calculated from the internal standards in each run and used to adjust the summed reporter ion intensities for each protein in the remaining eight experimental samples in each TMT experiment.

**References**

1. D. G. Tingay *et al.*, Gradual Aeration at Birth Is More Lung Protective Than a Sustained Inflation in Preterm Lambs. *Am J Respir Crit Care Med* **200**, 608-616 (2019).

2. D. L. Plubell *et al.*, Extended Multiplexing of Tandem Mass Tags (TMT) Labeling Reveals Age and High Fat Diet Specific Proteome Changes in Mouse Epididymal Adipose Tissue. *Mol Cell Proteomics* **16**, 873-890 (2017).