Supplementary Information for:

Breathprints of Model Murine Bacterial Lung Infections are Linked with Immune Response

Authors: Heather D. Bean, Jaime Jiménez-Díaz, Jiangjiang Zhu, and Jane E. Hill

MATERIAL AND METHODS

Portions of the methods described below were previously detailed in [1, 2].

Sample size estimation

The estimated sample size for each treatment group (PAL, SAL, or PBS exposure) at each time point (6, 12, 24, 48, 72, or 120 h) to achieve 80% power to discriminate the groups by SESI-MS breathprinting was calculated assuming a type I error rate of 0.05, and included a post-hoc analysis of the inter- and intra-class variation we previously measured for the SESI-MS breathprints of 125 mice (C57BL/6J) exposed to seven different bacterial species and to PBS controls [1, 2, 4]. A total of 87 mice were used in this study – 5 mice per time point for SAL and PAL and 4 mice per time point for PBS, with the exception of 24 h PAL and PBS, which included 6 mice each. No method of randomization was used to assign mice to treatment groups, and all treated mice were included in the analysis. The breath, BAL, and lung homogenates for all three treatment groups within a given time point were collected at the same time, with the exception of the 24 h PBS controls, which were collected over a span of several days in order to measure inter-day variability of the breathprinting method. Based on Spearman's rank correlations between the 24 h PBS breathprints, and on previously published data from our group [1, 4], we found no significant inter-day variation within a treatment group and time point.

Airway exposure, mice ventilation, and breath collection

For airway exposure, mice were briefly anesthetized (isoflurane by inhalation) and exposed to the *P. aeruginosa* or *S. aureus* bacterial lysates by oropharyngeal aspiration of 40 μL PBS containing 5 μg of protein. For breath collection, the mice were anesthetized with pentobarbital 24 h after infection and their tracheas were cannulated. The mice were placed on the ventilators (flexivent, SCIREQ, Montreal, QC, Canada) at 180 breaths/min with a positive end-expiratory pressure (PEEP) of 3 cm H_2O for 1 h. While on the ventilators the mice were paralyzed with intraperitoneal pancuronium bromide (0.5 mg/kg) and their heart rates were monitored by electrocardiography to ensure proper anesthesia. Breath was collected in Tedlar bags (SKC, Eighty Four, PA) as it exited the ventilator.

Bronchoalveolar lavage fluid (BALF) collection and analysis

After breath collection, 1 mL of cold PBS with 5% fetal bovine serum (FBS) was instilled into the lungs and the bronchoalveolar lavage fluid (BALF) was collected through the cannula installed previously for ventilation. BALF cells were pelleted, decanted, and immediately resuspended in PBS + 5% FBS. Total cells were counted using an ADVIA cell counter (Bayer, Terrytown, NY). BALF cells were then fixed onto glass slides (2×10^4 cells/slide) and stained with Hema-3 (Biochemical Sciences, Swedesboro, NJ). The leukocytes (or white blood cells, WBCs) were counted (300/slide) and categorized as macrophages, eosinophils, polymorphonuclear neutrophils (PMNs), or lymphocytes based on characteristic morphology and staining. *In vivo* lung tissue damage was determined by measuring lactate dehydrogenase (LDH) activity in BALF using the CytoTox 96 Non-Radioactive Cytotoxicity Assay (Promega, Madison, WI, USA).

Cytokine panel

Cytokine concentrations in the BAL fluid were analyzed using customized Milliplex assays (Millipore, Billerica, MA). The cytokine panel consisted of granulocyte colony stimulating factor (G-CSF), interferon-γ (IFN- γ), interleukin 1β (IL-1β), interleukin 6 (IL-6), keratinocyte chemoattractant protein (KC), macrophage chemotactic protein 1 (MCP-1), macrophage inflammatory proteins 1β and 2 (MIP-1β, MIP-2), RANTES, and tumour necrosis factor α (TNFα).

Secondary electrospray ionization-mass spectrometry (SESI-MS) breath analysis

Breath volatiles analyses were performed within one hour of collection using SESI-MS. The instrumental setup has been previously described [7-9]. Briefly, we replaced the original ionization source of an API 3000 mass spectrometer (Sciex, Concord, ON, Canada) with a stainless steel SESI-MS reaction chamber equipped with an electrospray capillary and a gas transfer line through which the breath volatiles are introduced into the reaction chamber (for a detailed schematic of the SESI-MS system, please see reference [8]. Gas flow of 5 L/min was driven by a mechanical pump that was connected to the sampling gas outlet of the SESI-MS reaction chamber. Breath samples were introduced into the reaction chamber for 30 s at a flow rate of 3 L/min, supplemented with 2 L/min CO_2 (99.99 %) at ambient temperature. Formic acid (0.1 % (v/v)) was used as the electrospray solution, delivered at a flow rate of 5 nL/s through a non-conductive silica capillary (40 μm ID) with a sharpened needle tip. The operation voltage was ~ 3.5 kV. Spectra were collected within 30 s as an accumulation of 10 scans in positive-ion mode. The system was flushed with CO_2 between samples until the spectrum returned to background levels.

Spearman's rank correlation analyses

Spearman's rank analyses were used to identify correlations between breathprint peaks and cytokine concentrations, total WBCs, PMNs, and/or LDH activity at 6, 12, 24, and 48 h after exposure to PAL, SAL, or PBS. Correlations were calculated on a time point-by-time point basis by aggregating and ranking every sample replicate (PAL, SAL, and PBS) based on each spectral and immunological measurement. Pair-wise correlations (ρ) between each breathprint peak and each immunological marker were calculated, resulting in four values per pair (one value for each time point). The correlations reported

in table 1 correspond to the average Spearman's rank coefficients of all four time points when two or more of the time points were considered significant (i.e., average $\rho > 0.7$ or < -0.7, and $p < 0.05$).

Partial least squares-discriminant analyses (PLS-DA)

PLS-DA is a supervised classification method that generates a linear model to describe predicted variables (test groups) in terms of observable variables (SESI-MS breathprint peaks or cytokines). The SESI-MS peaks with a mass-to-charge ratio (*m/z*) 20 – 200 Da and with signal-to-noise ratios (S/N) greater than two were used as observable variables for PLS-DA in this study. For the cytokines PLS-DA, the concentrations of all cytokines were used as observable variables. We took all biological replicates into consideration for PLS-DA, using leave-one-out cross validation to optimize the results, and calculated the prediction sum of squares (PRESS) residuals to evaluate the model [10]. The first two PLS factors, which explain the largest percentage of variation, are plotted.

RESULTS

Table S1 (in separate Excel file): Mann-Whitney U-tests comparing SESI-MS breathprints at 6 - 120 h after murine airway exposures to *P. aeruginosa* lysate (PAL), *S. aureus* lysate (SAL), live *P. aeruginosa* (PA) or *S. aureus* (SA) pathogens, or phosphate buffered saline (PBS) control (see Materials & Methods for calculation details). To facilitate comparisons to the other statistical methods employed for data analysis, the PLS-DA VIP scores (Figs. S1-S3) and immune correlations (Table 1) are summarized in the four left-most columns.

Table S2 (in separate Excel file): Average cytokine concentrations (pg/mL) in bronchoalveolar lavage fluid at six time points 6 – 120 h after murine airway exposure to *P. aeruginosa* cell lysate, *S. aureus* cell lysate, or phosphate buffered saline (PBS) control.

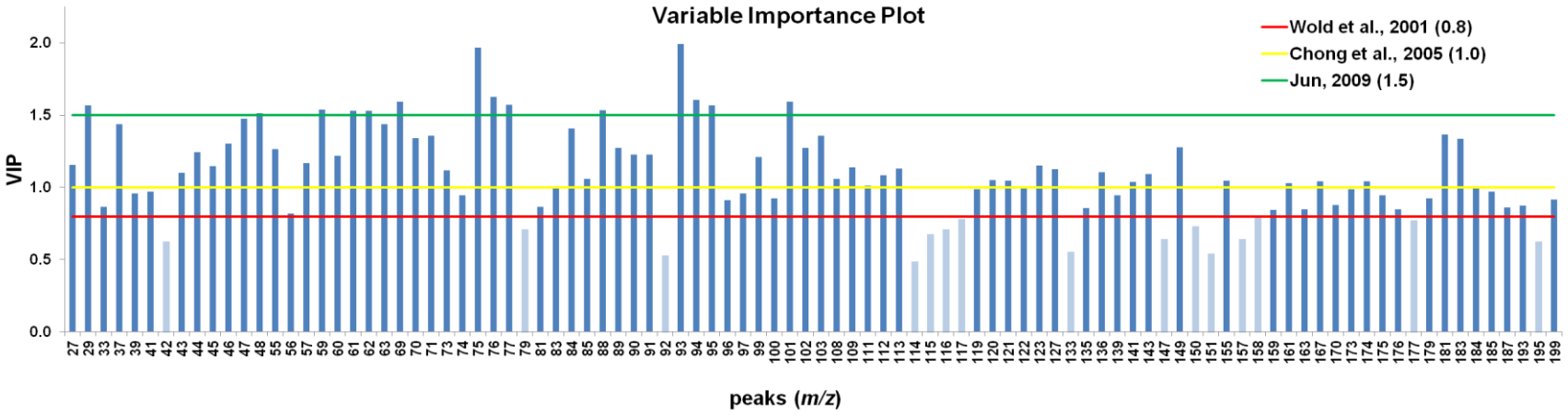

Figure S1: Variable importance plot (VIP) showing each individual peak's contribution for projecting the data to PLS scores (Fig. 3a). SESI-MS breathprint peaks from *P. aeruginosa* infections, *P. aeruginosa* cell lysate exposure, and PBS controls are included (6 – 120 h). The VIP summarizes the contribution that a variable makes to the model. Three commonly-accepted minima for significant contributions to the model are indicated: Red for VIP > 0.8,[3] yellow for VIP > 1.0,[5] and green for VIP > 1.5.[6]

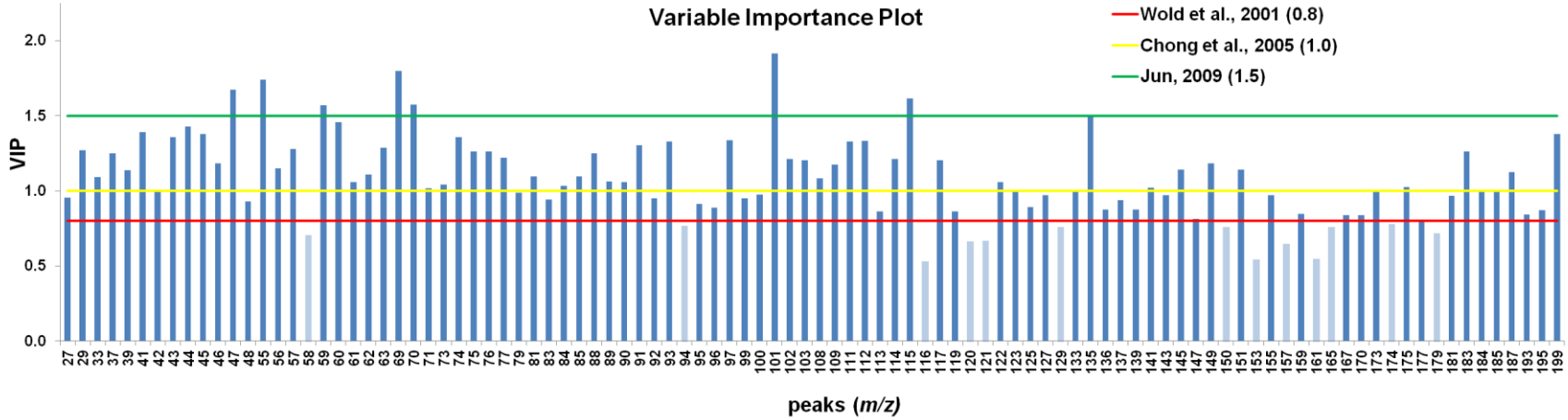

Figure S2: Variable importance plot (VIP) showing each individual peak's contribution for projecting the data to PLS scores (Fig. 3b). SESI-MS breathprint peaks from *S. aureus* infections, *S. aureus* cell lysate exposure, and PBS controls are included (6 – 120 h). The VIP summarizes the contribution that a variable makes to the model. Three commonly-accepted minima for significant contributions to the model are indicated: Red for VIP > 0.8,[3] yellow for VIP > 1.0,[5] and green for VIP > 1.5.[6]

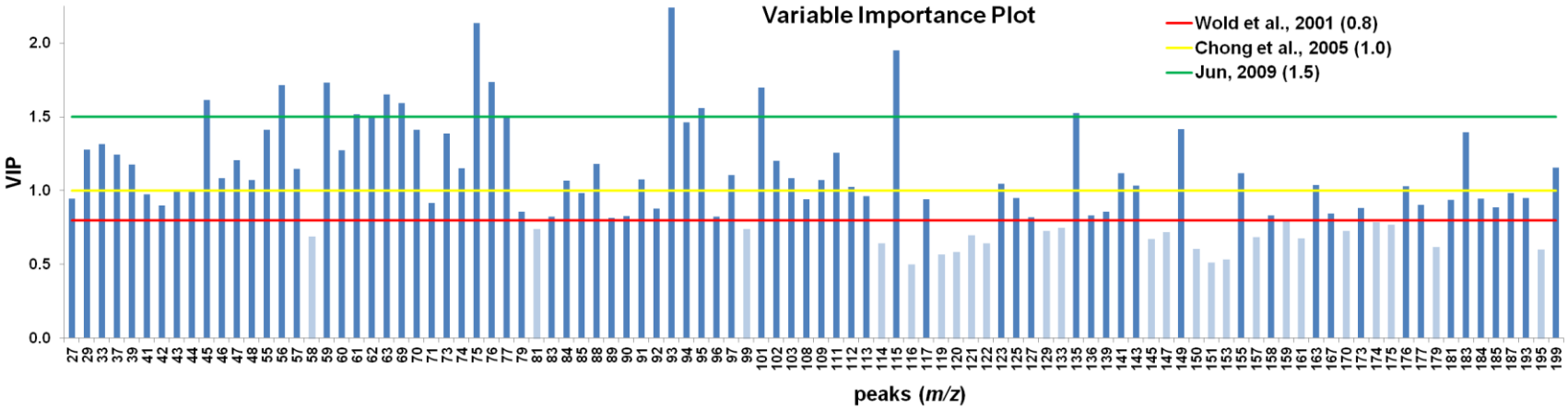

Fig.ure S3: Variable importance plot (VIP) showing each individual peak's contribution for projecting the data to PLS scores (Fig. 4). SESI-MS breathprint peaks from *P. aeruginosa* cell lysate and *S. aureus* cell lysate and PBS controls are included (6 – 120 h). The VIP summarizes the contribution that a variable makes to the model. Three commonly-accepted minima for significant contributions to the model are indicated: Red for VIP > 0.8,[3] yellow for VIP > 1.0,[5] and green for VIP > 1.5.[6]

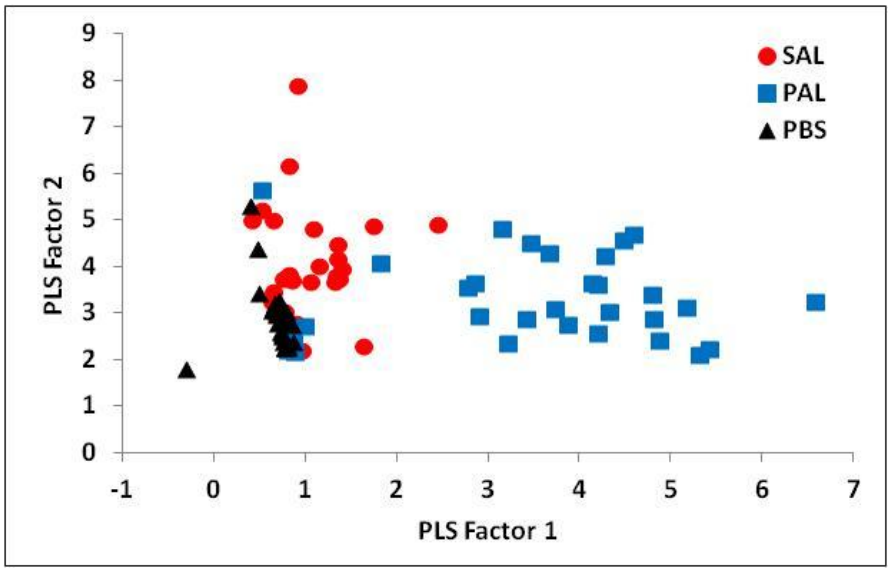

Figure S4: Canonical plots for partial least squares-discriminant analyses (PLS-DA) for the separation of mice treated with *P. aeruginosa* lysate (PAL), *S. aureus* lysate (SAL), or the phosphate buffered saline (PBS) control using cytokine profiles in bronchoalveolar lavage fluid (BALF). All replicates for the six time points (i.e., 6, 12, 24, 48, 72 and 120 h) for each group were included, resulting in 87 biological replicates in the analysis. The first two PLS factors explain the largest percentage of the variation (55.57 %), which did not provide a meaningful separation of all three groups.

Table S3: Summary of the correct classification rate for the PLS-DA validation tests using cytokine concentrations up to 120 h after *P. aeruginosa* lysate (PAL), *S. aureus* lysate (SAL), or phosphate buffered saline (PBS) exposure.

		Actual Treatment		
		PBS	PAL	SAL
Predicted Treatment	Round 1 *			
	PBS	92%	16%	37%
	PAL	0	71%	0
	SAL	8%	13%	63%
	Round 2 **			
	PBS	0	74%	40%
	PAL	85%	16%	0
	SAL	15%	10%	60%

*Round 1: PLS-DA using 90% training data and 10% testing data
**Round 2: PLS-DA using 70% training data and 30% testing data

References

1. Zhu J, Bean HD, Jiménez-Díaz J, Hill JE. Secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting of multiple bacterial lung pathogens, a mouse model study. *J Appl Physiol* 2013; 114: 1544-1549.
2. Zhu J, Bean HD, Wargo MJ, Leclair LW, Hill JE. Detecting bacterial lung infections: *In vivo* evaluation of *in vitro* volatile fingerprints. *J Breath Res* 2013; 7: 016003.
3. Wold S, Sjöström M, Eriksson L. PLS-regression: A basic tool of chemometrics. *Chemometrics Intellig Lab Syst* 2001; 58: 109-130.
4. Zhu J, Jiménez-Díaz J, Bean HD, Daphtary NA, Aliyeva MI, Lundblad LKA, Hill JE. Robust detection of *P. aeruginosa* and *S. aureus* acute lung infections by secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting: From initial infection to clearance. *J Breath Res* 2013; 7: 037106.
5. Chong I-G, Jun C-H. Performance of some variable selection methods when multicollinearity is present. *Chemometrics Intellig Lab Syst* 2005; 78: 103-112.
6. Jun C-H, Sang-Ho L, Hae-Sang P, Jeong-Hwa L. Use of partial least squares regression for variable selection and quality prediction. In: International Conference on Computers & Industrial Engineering; 2009. p. 1302-1307.
7. Martinez-Lozano P, de la Mora JF. On-line detection of human skin vapors. *J Am Soc Mass Spectrom* 2009; 20: 1060-1063.
8. Bean HD, Zhu J, Hill JE. Characterizing bacterial volatiles using secondary electrospray ionization mass spectrometry (SESI-MS). *J Vis Exp* 2011; 52: http://www.jove.com/details.php?id=2664.
9. Zhu J, Bean HD, Kuo Y-M, Hill JE. Fast detection of volatile organic compounds from bacterial cultures by SESI-MS. *J Clin Microbiol* 2010; 48: 4426-4431.
10. Tarpey T. A note on the prediction sum of squares statistic for restricted least squares. *Amer Statist* 2000; 54: 116-118.