SUPPLEMENT

Pneumonia recovery reprograms the alveolar macrophage pool

Antoine Guillon*[1,2], Emad I. Arafa*[1,3], Kimberly A. Barker[1,4], Anna C. Belkina[5,6], Ian Martin[1], Anukul T. Shenoy[1], Alicia K. Wooten[1,3], Carolina Lyon De Ana[1,4], Anqi Dai[8], Adam Labadorf[8], Jaileene Hernandez Escalante[4], Hans Dooms[3,4], Hélène Blasco[7], Katrina E. Traber[1,3], Matthew R. Jones[1,3], Lee J. Quinton[1,3,4,5], Joseph P. Mizgerd[1,3,4,9]

* These authors contributed equally.

Affiliations:

1- Pulmonary Center, Boston University School of Medicine, Boston, Massachusetts, USA
2- CHRU of Tours, service de Médecine Intensive Réanimation, Inserm u1100, University of Tours, Tours, France
3- Department of Medicine, Boston University School of Medicine
4- Department of Microbiology, Boston University School of Medicine
5- Department of Pathology and Laboratory Medicine, Boston University School of Medicine
6- Flow Cytometry Core Facility, Boston University School of Medicine
7- CHRU Tours, Medical Pharmacology Department, Inserm U1253, University of Tours, Tours, France
8- Bioinformatics Nexus, Boston University
9- Department of Biochemistry, Boston University School of Medicine

Corresponding author: Joseph P. Mizgerd, Sc.D., Boston University School of Medicine, 72 E. Concord Street, Boston, MA 02118. Email: jmizgerd@bu.edu; phone: 617-358-1156.

Methods

Mice. Female and male C57BL/6J mice were used, from Jackson Laboratories (Bar Harbor, ME). Experiments were initiated when mice were 5-7 weeks of age. Mice were maintained in specific pathogen-free conditions.

Lung infections. Bacteria were grown on blood agar plates at 37°C with 5% CO_2 and suspended in sterile saline prior to infecting mice, including *Streptococcus pneumoniae* serotype 19F (EF3030) and serotype 3 (catalog number 6303; ATCC, Manassas, VA). Mice were infected by either intranasal (i.n.) or intratracheal (i.t.) routes, after anesthesia with a ketamine (50 mg/kg) and xylazine (5 mg/kg) mixture delivered intraperitoneally (i.p.). The i.n. deliveries were performed by administering 50 µL of liquid suspension to the nares until the inoculation was inhaled completely. The i.t. deliveries were performed by placing an angiocatheter into the left bronchus through a tracheostomy, and administering 50 µL of liquid. To generate self-limiting lung infections, mice were infected with 1–3 $x10^6$ CFU of serotype 19F. To generate heterotypic immunity, mice received 2 serotype 19F infections with a 1-week interval between them, and then rested at least 4 weeks to allow elimination of infection, resolution of inflammation, and contraction of adaptive immunity (1). These *experienced* mice were compared to *naïve* mice who received sterile saline by the same route instead of serotype 19F infections. To generate severe lung infections in order to test the efficacy of heterotypic immunity, mice were infected with serotype 3 pneumococcus at doses ranging from 0.75 $x10^6$ to 15 $x10^6$ CFU, as indicated. To assess bacterial burden, mice were euthanized by isoflurane overdose, and CFU were enumerated in the lungs homogenized with a Bullet blender (Next Advance, Averill Park, NY) in water containing a protease inhibitor cocktail (Roche, Risch-Rotkreuz, Switzerland).

Bronchoalveolar lavage (BAL). The trachea was cannulated with an 18-gauge angiocatheter and the lungs were lavaged 4 times with 0.5 ml of ice-cold PBS. Cells were enumerated using LUNA-FL™ Dual Fluorescence Cell Counter (Logos Biosystems) and differentiated using cytocentrifuge preparations stained with Diff-Quick (VWR).

Alveolar macrophage depletion. For alveolar macrophage depletion experiments, mice received i.n. delivery of 100 µL of liposomes (Liposoma, Amsterdam, Netherlands) containing either clodronate (5 mg of clodronate per mL of solution) or PBS vehicle, and were studied or infected 72 hours after liposome delivery.

Lung cell suspensions. Lungs were perfused through the right ventricle with 5 mL of PBS for exsanguination. Excised lung lobes were digested in a solution containing type 2 collagenase (1 mg/ml), DNase I (150 mg/ml), and $CaCl_2$ (2.5mM) in phosphate-buffered saline while shaking at 37°C for 1 hour before being passed through a 70-mm cell strainer (Fisher, Grand Island, NY). Erythrocytes were removed using red blood cell lysis buffer (Sigma, St Louis, MO). Cells were enumerated using the LUNA-FL™ Dual Fluorescence Cell Counter (Logos Biosystems).

Flow cytometry and cell sorting. After Fc receptor blockade (FcBlock, eBioscience), cells were incubated with combinations of monoclonal antibodies for 30 minutes on ice in the dark, and then washed with PBS containing FBS and EDTA. Fluorochrome-conjugated monoclonal antibodies used include the following: CCR5/BV421 (clone C34-3448, BD Bioscience), CD4-APC (clone RM4-5, eBioscience), CD3-APC (clone 17A2, eBioscience), CD11c-PE/Cy7 (clone HL3, BD Bioscience), CD19-APC (clone 1A3, eBioscience), CD36/APC (clone HM36, Biolegend), CD45-FITC (clone 30-F11, eBioscience), CD64-PE (clone X54-5/7.1, BioLegend), CD93/BV510 (clone 493, BD Bioscience), Ly6C-eFluor 450 (clone HK1.4, eBioscience), Ly6G-APC (clone 1A8, BD Bioscience), Ly6c-APC/Cy7 (clone 1A8, BD Bioscience), MHCII-PerCP-Cy5.5 (clone M5/114.15.2, BD Bioscience), MSR-A/BUV395 (clone 268318, BD Bioscience),

and SiglecF-APC/Cy7 (clone E50-2440, BD Bioscience). Cells were stained with viability dye eFluor 506 (eBioscience) or 7-AAD Viability Staining Solution (BioLegend). Unstained, single-stained using OneComp eBeads (eBioscience), and fluorescence-minus-one (FMO) controls were used for each analysis. Data were acquired on a BD LSR II flow cytometer (BD Biosciences) using BD FACSDiva software. For cell sorting, a BD FACSARIA II SORP cell sorter was used (BD Biosciences).

Cytometry data analysis. Expert-guided/manual data analyses and data pre-processing for algorithmic analyses were performed using FlowJo 10 software (Flowjo, Inc). Equal numbers of single, live, CD45+ immune cells (3.5×10^4 per animal) were concatenated and the resulting dataset was clustered with PhenoGraph algorithm (2) and mapped into 2D opt-SNE space using Omiq.ai cloud computing platform. Embedding parameters were automatically defined by the opt-SNE algorithm (3). Statistical comparison of phenograph clusters was performed in the R environment (versions 3.3-3.6) using Wilcoxon rank-sum test with Benjamini-Hochberg correction for multiple comparisons and DepecheR package (4) for results visualization. Corrected $P < 0.05$ was considered significant for group differences. Heatmaps, phenograph cluster overlays in opt-SNE space, and manually gated population overlays in opt-SNE space were created with Omiq.ai.

Measurement of metabolites. LC/MS was used to profile and quantify the polar metabolite contents of alveolar macrophages. Because FACS has a significant impact on the cellular metabolome (5, 6), we collected alveolar macrophages for metabolite analyses by BAL without sorting. LC/MS analyses were conducted on a QExactive benchtop orbitrap mass spectrometer equipped with an Ion Max source and a HESI II probe, which was coupled to a Dionex UltiMate 3000 UPLC system (Thermo Fisher Scientific, San Jose, CA). External mass calibration was performed using the standard calibration mixture every 7 days. The metabolite extraction mix was composed of 80% methanol in water, supplemented with a mixture of 17 isotope-labeled

amino acids at 90.9 nM each, used as internal standards (Cambridge Isotope Laboratories, MSK-A2-1.2). After extraction, the protein level of each extracted sample was quantified using a BCA assay for biological normalization. The pre-determined quality control cutoffs for each metabolite were CV<0.25 for the peak area ratio across the triplicate injections, and R > 0.975 for the raw peak area across the dilution series (including the average of the triplicate injections). Samples were run in a randomized order to minimize effect of technical variation.

Metabolome analyses. Values of metabolites levels were normalized to the sum of all metabolites levels, log-transformed, and auto-scaled (mean-centered and divided by the standard deviation of each variable). Results were analyzed as paired-data for comparing left and right lungs and as unpaired data for comparing *naïve* and *experienced* mice. The univariate analysis was based on volcano plots and the non-parametric Wilcoxon test, and results were expressed as FDR-adjusted *P* values. For multivariate analysis, the classification method was based on unsupervised Principal Component Analysis (PCA) to evaluate distribution of samples. Supervised analysis based on Partial Least-Squares Discriminant Analysis (PLS-DA) was performed to highlight the most discriminant variables. Permutation tests were not performed due to the small sample size. Values of Variable Influence on Projection (VIP) represent the importance of the compound (metabolite) for the PLS-DA models. Data were analyzed using MetaboAnalyst 4.0 software (7).

Gene expression analyses. Alveolar macrophages were isolated by FACS from collagenase-digested lungs of *naïve* and *experienced* mice with 0 or 4 hours of serotype 3 infection. Four mice per group from two independent experiments were analyzed. RNA was extracted from the samples using RNeasy Micro Kit (Qiagen) and profiled using GeneChip mouse Gene 2.0 ST arrays (Affymetrix). Mouse 2.0 ST CEL files were normalized to produce gene-level expression values using the Robust Multiarray Average (RMA) (8) from the *affy* package (version 1.36.1) (9) in the Bioconductor software suite (version 2.11) (10) and an Entrez Gene-specific probeset

mapped (17.0.0) from the Molecular and Behavioral Neuroscience Institute (Brainarray) at the University of Michigan (11). Array quality was assessed using Relative Log Expression (RLE) and Normalized Unscaled Standard Error (NUSE) in the affyPLM package (version 1.34.0). Differential expression was assessed using the moderated (empirical Bayesian) t-test implemented in version 3.14.4 of the *limma* package (i.e., creating simple linear models with lmFit, followed by empirical Bayesian adjustment with eBayes). The Benjamini-Hochberg procedure was applied to generate a false discovery rate (FDR) and correct for multiple comparisons. Differences were considered significant if FDR $q < 0.05$. All microarray analyses were performed using the R environment for statistical computing (version 2.15.1). Complete datasets were deposited with NCBI GEO (Series ID GSE133975).

Gene Set Enrichment Analysis (GSEA). GSEA version 2.2.1 (12) was used to identify biological terms, pathways and processes that are coordinately up- or down-regulated within each pairwise comparison. The Entrez Gene identifiers of the human homologs of the genes interrogated by the array were ranked according to the t statistic computed for each main effect (SP19 history and SP3 infection), for each pairwise comparison, and the history:infection interaction effect. Mouse genes with multiple human homologs (or vice versa) were removed prior to ranking, so that the ranked list represents only those human genes that match exactly one mouse gene. This ranked list was then used to perform pre-ranked GSEA analyses (default parameters with random seed 1234) using the Entrez Gene versions of the Hallmark, Biocarta, KEGG, Reactome, Gene Ontology (GO), and transcription factor and microRNA motif gene sets obtained from the Molecular Signatures Database (MSigDB), version 6.0 (13). Restricting results to the Reactome pathways provided a dataset size amenable to manual inspection and arrangement; Reactome pathways reaching significance (FDR $q<0.05$) were assembled into groups defined by the authors (Table 1).

Comparisons to human datasets. Results from our studies were compared to a prior study that examined human alveolar macrophage transcriptomes with and without a 4 hour pneumococcal stimulation (14). Only expression data from healthy human subjects were included in analyses. Both mouse and human datasets were restricted to include only genes with homologs across species. Three differential expression analyses for infection *vs.* control were used to compare effects of infection in alveolar macrophages from healthy humans, experienced mice, and naïve mice. In addition to genes differentially expressed, pathways were detected using GSEA. The t statistics from the differential expression results were used to perform a GSEA employing the fgsea Bioconductor package with the MsigDB gene set database version 6.2, looking into only the c2.cp (canonical pathways), c5.all (GO gene sets) and c7.all (immunological signatures) enrichment. Degrees of similarity between the human and *naïve* or *experienced* mouse responses were calculated using the approach previously applied to compare responses of humans to those in laboratory mice or laboratory mice that had been co-housed with pet store mice (15). GSEA statistics for gene sets corresponding to the top 400 infection-induced genes in the differentially expressed gene sets from *experienced* and *naïve* mice were computed against the human differentially expressed t statistic dataset as a table and leading edge plots, and the resulting plots were used in figure panels (15). All code needed to reproduce the results comparing these data to the human data are available at https://bitbucket.org/bubfnexus/murine_alveolar_transcriptional_response.

Statistics. Statistical analyses used for each dataset including measures of central tendency and variation were specified in the Figure Legends. Unpaired data other than CFU were analyzed using parametric tests, with two groups compared using the Student's t-test and more-than-two groups compared using a one-way or two-way ANOVA followed by *post hoc* Sidak's tests for multiple comparisons. CFU data were analyzed using non-parametric tests for comparing two groups (Mann Whitney U test) or more-than-two groups varying across only one

independent variable (Kruskal-Wallis test followed by *post hoc* Dunn's tests for multiple comparisons). For CFU analyses varying across 2 independent variables, data were log-transformed prior to two-way ANOVA followed by *post hoc* Sidak's tests for multiple comparisons. Sets of paired data were analyzed using the non-parametric Sign test. Differences were considered significant if $P < 0.05$. Statistics used for bioinformatics analyses of cytometry, metabolome, or transcriptome datasets were presented individually in the description of methods for those approaches.

Study approval. Animal protocols were approved by the Boston University Institutional Animal Care and Use Committee.

References

1. Smith NM, Wasserman GA, Coleman FT, Hilliard KL, Yamamoto K, Lipsitz E, et al. Regionally compartmentalized resident memory T cells mediate naturally acquired protection against pneumococcal pneumonia. *Mucosal Immunol.* 2018;11(1):220-35.
2. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir el AD, Tadmor MD, et al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell.* 2015;162(1):184-97.
3. Belkina AC, Ciccolella CO, Anno R, Halpert R, Spidlen J, and Snyder-Cappione JE. Automated optimized parameters for t-distributed stochastic neighbor embedding improve visualization and allow analysis of large datasets. *Nature communications.* 2019;(in press).
4. Theorell A, Bryceson YT, and Theorell J. Determination of essential phenotypic elements of clusters in high-dimensional entities-DEPECHE. *PLoS One.* 2019;14(3):e0203247.
5. Binek A, Rojo D, Godzien J, Ruperez FJ, Nunez V, Jorge I, et al. Flow Cytometry Has a Significant Impact on the Cellular Metabolome. *J Proteome Res.* 2019;18(1):169-81.
6. Llufrio EM, Wang L, Naser FJ, and Patti GJ. Sorting cells alters their redox state and cellular metabolome. *Redox Biol.* 2018;16:381-7.
7. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 2018;46(W1):W486-W94.
8. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003;4(2):249-64.

9. Gautier L, Cope L, Bolstad BM, and Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics.* 2004;20(3):307-15.

10. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;5(10):R80.

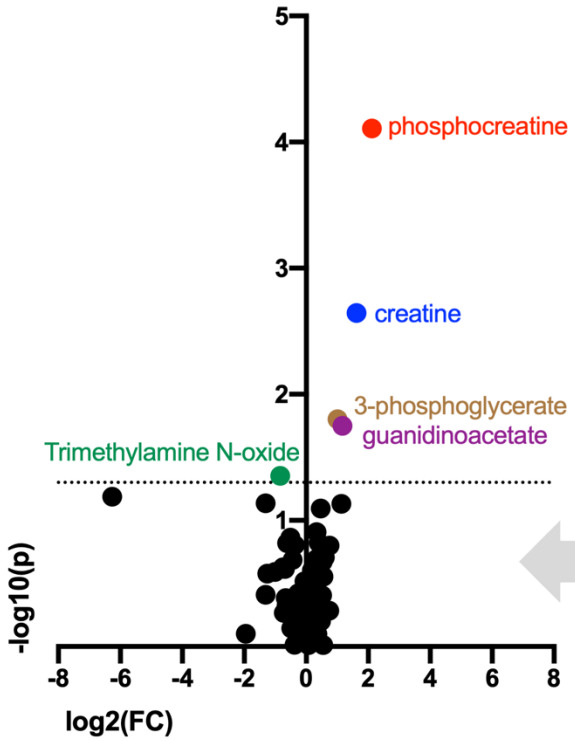
11. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* 2005;33(20):e175.

12. 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.

13. Subramanian A, Kuehn H, Gould J, Tamayo P, and Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics.* 2007;23(23):3251-3.

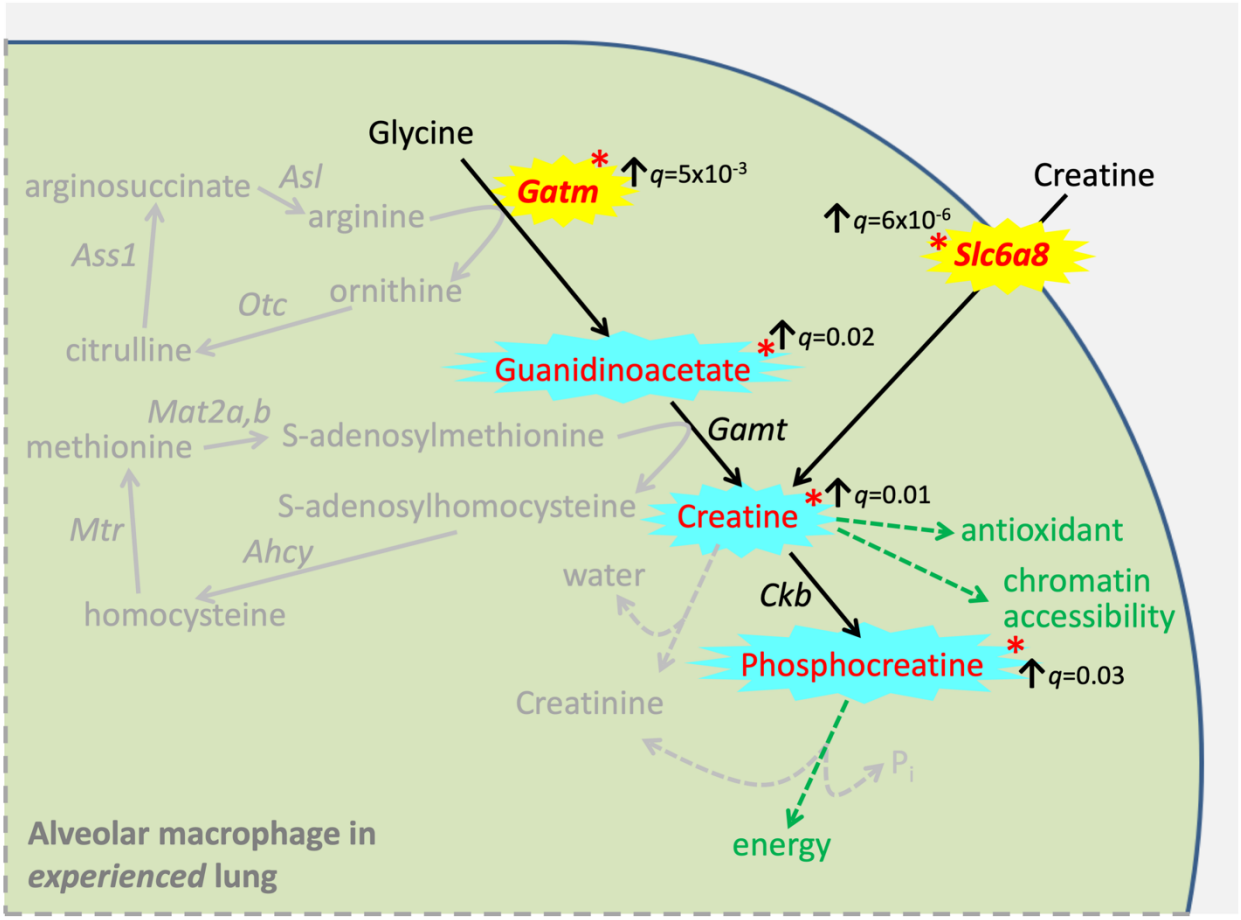
14. Bewley MA, Preston JA, Mohasin M, Marriott HM, Budd RC, Swales J, et al. Impaired Mitochondrial Microbicidal Responses in Chronic Obstructive Pulmonary Disease Macrophages. *Am J Respir Crit Care Med.* 2017.

15. Beura LK, Hamilton SE, Bi K, Schenkel JM, Odumade OA, Casey KA, et al. Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature.* 2016;532(7600):512-6.



1-methyl-adenosine	carnitine	glutamine	methionine	S-adenosyl-homocysteine
1-methyl-nicotinamide	choline	glycerol 3-phosphate	myo-Inositol	S-adenosyl-methionine
2-deoxy-cytidine	cis-urocanate	glycerophos-phocholine	NAD	serine
4-guanidino-butyric acid	citrate	glycine	NADP	taurine
acetyl-carnitine	citrulline	GMP	nicotinamide	threonine
adenosine	CMP	GSH	ornithine	trimethyl-lysine
ADMA	creatinine	GSSG	palmitate	tryptophan
ADP	cystine	histidine	phenylalanine	tyrosine
alanine	cytidine	hypoxanthine	phospho-choline	UDP-GlcNAc
AMP	dihydroxy-acetone phosphate	inosine	phospho-ethanolamine	UDP-hexose
arginine	F6P G1P	lactate	proline	UMP
ascorbic acid	g-butyro-betaine	leucine	propionyl-carnitine	urate
asparagine	glucose 6-phosphate	lysine	ribose 1-phosphate	
aspartate	glutamate	malate	ribose ribulose 5-phosphate	

Supplemental Figure S1. Metabolome analyses of alveolar macrophages after resolution of pneumococcal infections. Volcano plot of the LC/MS data of alveolar macrophages collected from whole lungs of *naïve* and *experienced* mice. The 5 metabolites that reached significance (FDR-adjusted $p<0.05$, indicated with dotted line) in comparisons between *naïve* and *experienced* mice were identified by name and color coding. The other metabolites measured were listed alongside, none of which differed by FDR-adjusted $p<0.05$ for comparisons between *naïve* and *experienced* mice.



Supplemental Figure S2. Interconnections between alveolar macrophage metabolome changes and transcriptome changes due to prior pneumonia experience. Metabolic pathways for creatine, urea cycle, and methionine cycle were depicted, including relationships to each other. Gene names for the proteins involved in these metabolic pathways were indicated in italics. The *Gatm* transcript encoding arginine:glycine amidinotransferase was significantly increased in *experienced* alveolar macrophages. The *Slc6a8* gene encoding sodium- and chloride-dependent creatine importer was significantly increased in *experienced* alveolar macrophages. None of the other genes in these pathways were significantly changed due to experience. The metabolites guanidinoacetate, creatine, and phosphocreatine were significantly increased due to experience, but no other of the depicted metabolites. Biological effects of increased creatine and phosphocreatine were depicted in green.