Supporting Information

Spatially-targeted proteomics of the host-pathogen interface during staphylococcal abscess formation

Emma R. Guiberson[¶1,2], Andy Weiss[¶3], Daniel J. Ryan[#1,2], Andrew J. Monteith[3], Kavya Sharman[1], Danielle B. Gutierrez[1], William J. Perry[1,2], Richard M. Caprioli[1,2,4,5,6], Eric P. Skaar*[3], Jeffrey M. Spraggins*[1,2,4]

[1] Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA, 37203
[2] Department of Chemistry, Vanderbilt University, Nashville, TN, USA, 37203
[3] Department of Pathology, Microbiology, and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA, 37203
[4] Department of Biochemistry, Vanderbilt University, Nashville, TN, USA, 37203
[5] Department of Medicine, Vanderbilt University, Nashville, TN, USA, 37203
[6] Department of Pharmacology, Vanderbilt University, Nashville, TN, USA, 37203
[#] Current Address: ExxonMobil Research and Engineering Company, 1545 Route 22 East, Annandale, NJ, USA, 08801

*- Co-corresponding authors

Jeff.spraggins@vanderbilt.edu

Eric.skaar@vumc.edu

[¶] These authors contributed equally to this work.

Ethics statement

All animal experiments under protocol M1900043 were reviewed and approved by the Institutional Animal Care and Use Committee of Vanderbilt University. Procedures were performed according to the institutional policies, Animal Welfare Act, NIH guidelines, and American Veterinary Medical Association guidelines on euthanasia.

Murine model of systemic S. aureus infection

6-8 week old female C67BL/6J mice were anesthetized with tribromoethanol (Avertin) and retro-orbitally infected with ~1.5-2 x 10^7 CFUs of *S. aureus* USA300 LAC constitutively expressing sfGFP from the genome (P*sarA*-sfGFP integrated at the SaPI1 site)(1). Infections were allowed to progress for 4 or 10 days before animals were humanely euthanized and organs removed for subsequent analysis.

Chemicals

Acetonitrile, acetic acid, formic acid, trifluoracetic acid, ethanol, ammonium bicarbonate, and chloroform were purchased from Fisher Scientific (Pittsburgh, PA). Mass spectrometry sequence-grade trypsin from porcine pancreas was purchased from Sigma-Aldrich Chemical, Co. (St. Louis, MO).

Micro-Digestions

All sample preparation was completed using the method previously described by Ryan *et al.*(1). Excised murine kidneys infected with *S. aureus* USA300 LAC carrying a constitutive fluorescent reporter were fresh frozen on dry ice and stored at -80°C. Tissues were then cryosectioned at 10 µm thickness (Leica Microsystems, Buffalo Grove, IL) and thaw mounted onto custom fiducial microscope slides (Delta Technology, Loveland, CO). For all samples, an autofluorescence microscopy image (Carl Zeiss Microscopy, White Plains, NY) was acquired at 10x magnification (0.92 µm/pixel) prior to spotting using FITC and DAPI filters (FITC excitation λ, 465-495 nm; emission λ, 515-555 nm; DAPI excitation λ, 340-380 nm; emission λ, 435-485 nm). Microscopy exposure time was set to 200 ms for the DAPI filter and

100 ms for the FITC filter. High resolution H&E and autofluorescence whole slide microscopy images were aligned using the elastix library(2) following previously described methods(3) and stored in the pyramidal OME.tiff data format. Samples were washed with graded ethanol washes for 30s (70% EtOH, 100% EtOH, Carnoy's Fluid, 100% EtOH, H2O, 100% EtOH) and Carnoy's fluid for 3min (6 ethanol: 3 chloroform: 1 acetic acid) to remove salts and lipids. Samples were allowed to dry under vacuum prior to trypsin digestion. Regions targeted for digestion were annotated on microscopy images using ImageJ (U.S. National Institutes of Health, Bethesda, MD) and converted into instrument coordinates for the piezoelectric spotter using an in-house script (Figure S1). Trypsin was dissolved in ddH_2O to a final concentration of 0.048 µg/mL. A piezoelectric spotting system (sciFLEXARRAYER S3, Princeton, NJ) was used to dispense ~250 pL droplets of trypsin. Trypsin was dispensed on regions of interest (ROIs) 20 times with one drop per run in order to reduce spot size, with 4 spots per ROI. The four trypsin spots were positioned so that all tryptic peptides for an ROI could be collected using a single LESA experiment. Following trypsin deposition, samples were incubated at 37°C for three hours in 300 µL ammonium bicarbonate.

LESA

Liquid surface extraction was completed using a TriVersa NanoMate (Advion Inc., Ithaca, NY) with the LESAplusLC modification. Digested samples were scanned using a flatbed scanner and uploaded to the Advion ChipSoft software. 5 µL of extraction solvent (2:8 acetonitrile/water with 0.1% formic acid) was aspirated into the glass capillary. The capillary was then lowered to a height of approximately 0.5 mm above the sample surface and 2.5 µL of solvent was dispensed onto ROIs. Contact with the surface was maintained for 10s and 3.0 µL of solvent was re-aspirated into the capillary. The initial 5 µL volume was dispensed into a 96-well plate containing 200 µL of water/0.1% formic acid. The LESA extraction was repeated twice at the same ROI and combined. Three wash cycles of the instrument were completed between each ROI set to prevent carryover from other biological regions. Resulting extracts were then dried down under vacuum and stored at -80˚C until LC-MS/MS analysis.

LC-MS/MS

Dried peptide samples were reconstituted in 10 µL of water/0.1% formic acid prior to analysis. Tryptic peptides from tissue extracts were injected and gradient eluted on a pulled tip emitter column packed in-house with C18 material (Waters BEH C18). The column was heated to 60˚C with a flow rate of 400 nL/min using an Easy-nLC 1000 UHPLC (Thermo Scientific, San Jose, CA). Mobile phase A consisted of H_2O with 0.1% formic acid, and mobile phase B consisted of acetonitrile with 0.1% formic acid. Peptides were eluted on a linear gradient of 2-20% B for 100 minutes, followed by 20-32% B for 20 minutes, and lastly 32-95% B for one minute. Eluting peptides were analyzed using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific, San Jose, CA). MS1 scans were acquired at 120,000 resolving power at *m/z* 200 using the Orbitrap, with a mass range of *m/z* 400-1600, and an automatic gain control target of $1.0x10^6$.

Data Analysis

For peptide identification, tandem mass spectra were searched using Protalizer software (Vulcan Analytic, Birmingham, Alabama) and MaxQuant(4) against a database containing both the mouse and *S. aureus* strain USA300 LAC proteome created from the UniprotKB(5). Modifications such as glycosylation, phosphorylation, methionine, oxidation, and deamidation were included in the Protalizer database search with an FDR of 1%. For MaxQuant analysis, raw files were processed using a label-free quantification method implemented in MaxQuant version 1.6.7. Spectra were searched against mouse and *Staphylococcus aureus* (strain USA300) reference databases downloaded from UniProt KB. These were supplemented with the reversed sequences and common contaminants automatically (decoy database) and used for quality control and estimation of FDR by MaxQuant. Carbamidomethylation was set as a fixed modification and acetyl (protein N-term) and oxidation (M) were set as variable modifications. Minimal peptide length was 7 amino acids. Peptide and protein FDRs were both set to 1%. The resultant protein groups file from MaxQuant was analyzed for outliers by calculating z-scores for each sample based on number of protein

groups identified; 3 samples out of 42 were excluded as outliers. Proteins identified as “reverse”, “only identified by site”, or “potential contaminants” were also removed.

Proteins identified by Protalizer and MaxQuant were filtered based on the following criteria: 2 unique peptides contributed to the protein identification, proteins were detected in 2+ biological replicates, and proteins were detected in 2+ serial sections. Bacterial proteins also met these criteria, with the exception of those listed in Table S3 (designated as ‘lower confidence’) that only met 1 of the 2 replicate requirements criteria but did meet the 2+ unique peptide criteria. We believe these proteins did not fulfil the requirements for high confidence due to potentially relevant differences in abundance or due to heterogeneity between abscesses. Since these proteins have previously been studied in a primarily targeted manner rather than our discovery-based approach, we include these data but additional confirmation is required. The lists generated by each dataset were cross compared for all proteins, to aid in identification from the protein groups generated by MaxQuant. Tables S1 and S2 include the full list of identified proteins, and include the search algorithm(s) were used for identification

Pathway analysis was conducted using a similar workflow described previously by Gutierrez and colleagues(6). Briefly, data were uploaded into a central in-house database and organized by identifiers such as tissue location and infection time. Project data were exported into a custom data analysis and visualization tool for network construction (SIMONE) in a data driven manner. Data was entered into Reactome(7), a free online pathway mapping software, and proteins associated with metabolism were selected from the Reactome output. The genes encoding for these proteins were then input into the network construction software as root nodes. Network construction was performed as described(8) in conjunction with a custom user interface that provided visualization (through the use of Cytoscape24) of the constructed data networks.

Additional Supporting Information

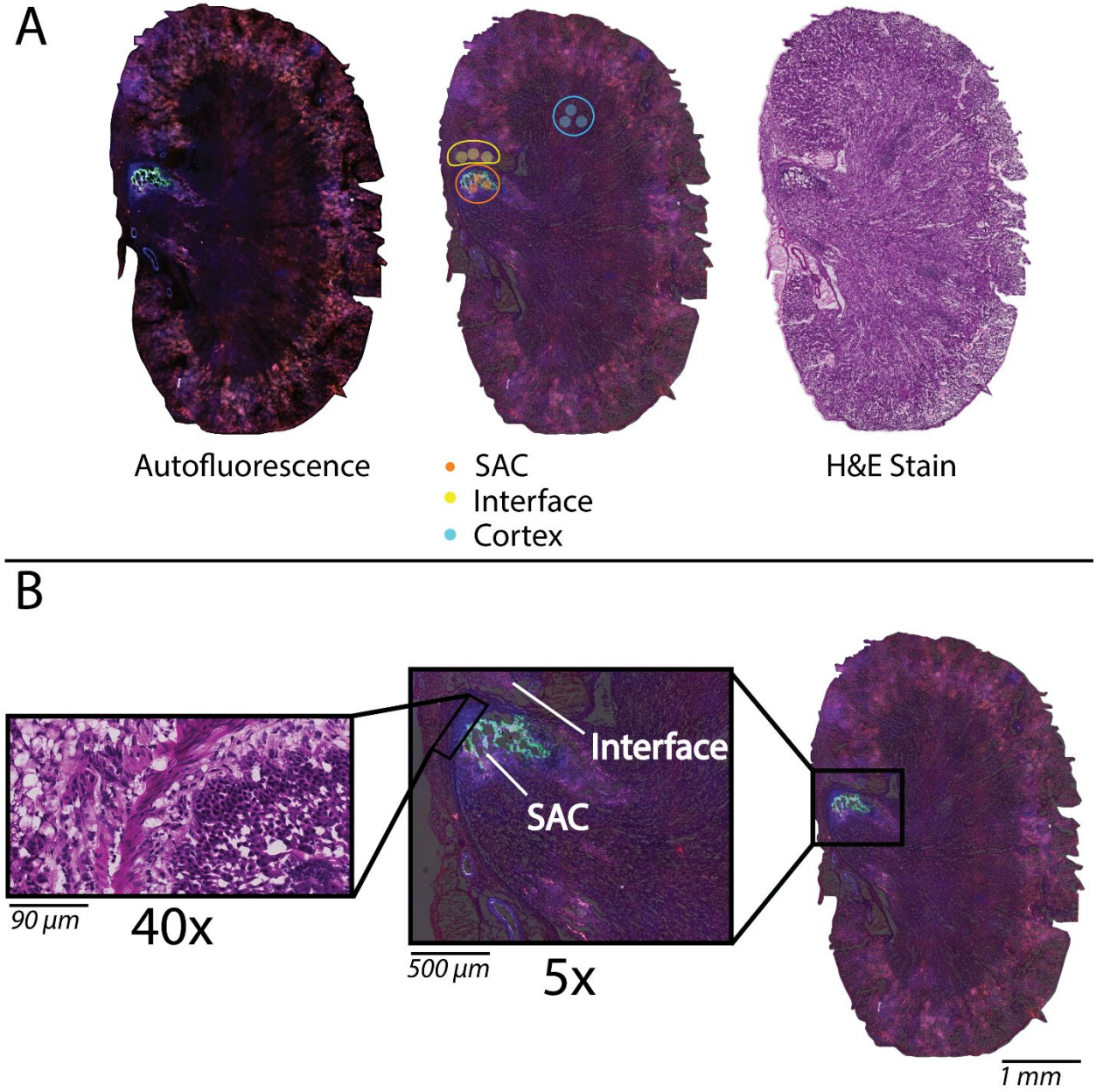

Figure S1: High resolution images of *S. aureus*-infected kidneys. a) Autofluorescence, H&E stained and overlaid image showing the locations of microLESA sampling (colored dots). Regions are shown by color-coded circles. b) 5x and 10x magnification of abscess and abscess edge regions.

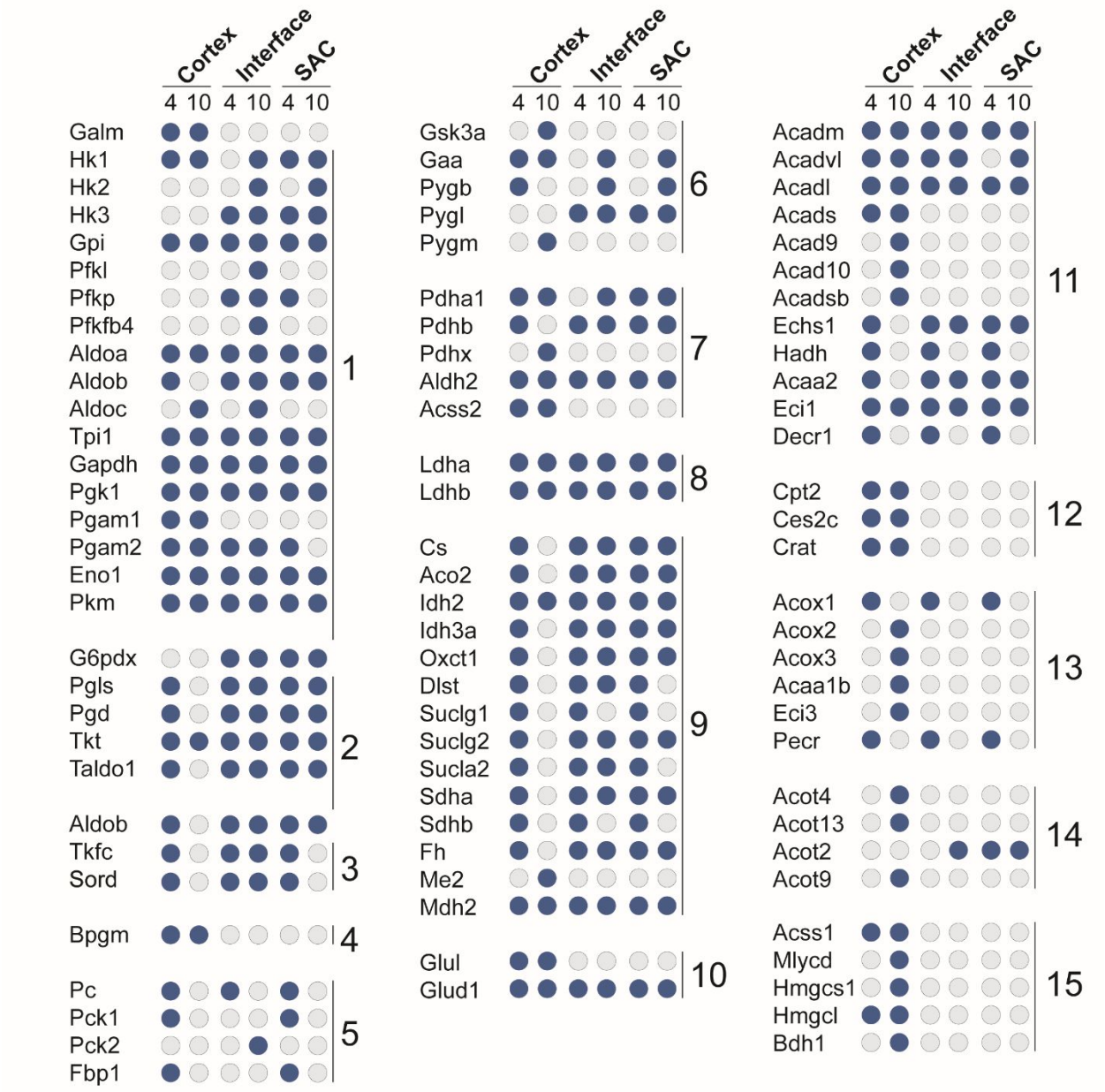

Figure S2: Spatiotemporal distribution of metabolic proteins during infection. Blue circles denote proteins that were present at the specified timepoint/region, while grey circles depict the absence of a specific protein.

A list of all host-derived proteins detected are included in Table S1, and all bacterial proteins are included in Table S2 (higher confidence identifications) and Table S3 (lower confidence identifications) which are included as separate PDF files. Descriptions of these tables are included below.

Table S1: List of host-derived proteins found by presence at time point and biological region, including platform used for identification (P = Protalizer, MQ = MaxQuant).

Table S2: List of bacterial proteins detected and their localizations based on stringent search criteria from Protalizer and MaxQuant. No bacterial proteins were detected in the cortex.

Table S3: List of bacterial proteins detected and their localizations with lower identification confidence.

References

1. Ryan DJ, Patterson NH, Putnam NE, Wilde AD, Weiss A, Perry WJ, Cassat JE, Skaar EP, Caprioli RM, Spraggins JM. MicroLESA: Integrating Autofluorescence Microscopy, In Situ Micro-Digestions, and Liquid Extraction Surface Analysis for High Spatial Resolution Targeted Proteomic Studies. Anal Chem. 2019;91(12):7578-85.
2. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging. 2010;29(1):196-205.
3. Patterson NH, Tuck M, Van de Plas R, Caprioli RM. Advanced Registration and Analysis of MALDI Imaging Mass Spectrometry Measurements through Autofluorescence Microscopy. Anal Chem. 2018;90(21):12395-403.
4. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367-72.
5. Bateman A, Martin MJ, Orchard S, Magrane M, Alpi E, Bely B, Bingley M, Britto R, Bursteinas B, Busiello G, Bye-A-Jee H, Da Silva A, De Giorgi M, Dogan T, Castro LG, Garmiri P, Georghiou G, Gonzales D, Gonzales L, Hatton-Ellis E, Ignatchenko A, Ishtiaq R, Jokinen P, Joshi V, Jyothi D, Lopez R, Luo J, Lussi Y, MacDougall A, Madeira F, Mahmoudy M, Menchi M, Nightingale A, Onwubiko J, Palka B, Pichler K, Pundir S, Qi GY, Raj S, Renaux A, Lopez MR, Saidi R, Sawford T, Shypitsyna A, Speretta E, Turner E, Tyagi N, Vasudev P, Volynkin V, Wardell T, Warner K, Watkins X, Zaru R, Zellner H, Bridge A, Xenarios I, Poux S, Redaschi N, Aimo L, Argoud-Puy G, Auchincloss A, Axelsen K, Bansal P, Baratin D, Blatter MC, Bolleman J, Boutet E, Breuza L, Casals-Casas C, de Castro E, Coudert E, Cuche B, Doche M, Dornevil D, Estreicher A, Famiglietti L, Feuermann M, Gasteiger E, Gehant S, Gerritsen V, Gos A, Gruaz N, Hinz U, Hulo C, Hyka-Nouspikel N, Jungo F, Keller G, Kerhornou A, Lara V, Lemercier P, Lieberherr D, Lombardot T, Martin X, Masson P, Morgat A, Neto TB, Paesano S, Pedruzzi I, Pilbout S, Pozzato M, Pruess M, Rivoire C, Sigrist C, Sonesson K, Stutz A, Sundaram S, Tognolli M, Verbregue L, Wu CH, Arighi CN, Arminski L, Chen CM, Chen YX, Cowart J, Garavelli JS, Huang HZ, Laiho K, McGarvey P, Natale DA, Ross K, Vinayaka CR, Wang QH,

Wang YQ, Yeh LS, Zhang J, Consortium U. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47(D1):D506-D15.
6. Gutierrez DB, Gant-Branum RL, Romer CE, Farrow MA, Allen JL, Dahal N, Nei YW, Codreanu SG, Jordan AT, Palmer LD, Sherrod SD, McLean JA, Skaar EP, Norris JL, Caprioli RM. An Integrated, High-Throughput Strategy for Multiomic Systems Level Analysis. J Proteome Res. 2018;17(10):3396-408.
7. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, Loney F, May B, Milacic M, Rothfels K, Sevilla C, Shamovsky V, Shorser S, Varusai T, Weiser J, Wu G, Stein L, Hermjakob H, D'Eustachio P. The reactome pathway knowledgebase. Nucleic Acids Res. 2020;48(D1):D498-D503.
8. James C. Pino ALRL, Leonard A. Harris, Danielle B. Gutierrez, Melissa A. Farrow, Nicole Muszynski, Tina Tsui, Jeremy L. Norris, Richard M. Caprioli, John P. Wikswo, Carlos F. Lopez. A computational framework to explore cellular response mechanisms from multi-omics datasets. 2020. doi: https://doi.org/10.1101/2020.03.02.974121.