SUPPLEMENTAL MATERIAL

Proteomics identifies a convergent innate response to infective endocarditis and extensive proteolysis in vegetation components

Daniel R. Martin, Ph.D, James C. Witten, MD, Carmela Tan, MD, E. Rene Rodriguez, MD, Eugene H. Blackstone, MD, Gosta B. Pettersson, MD Ph.D, Deborah E. Seifert, Belinda Willard, Ph.D, Suneel Apte, M.B.B.S. D.Phil.

Contents:

Detailed Methods:

Vegtation collection

Visible blood clots on vegetations were removed with forceps and residual blood was minimized by extensive rinsing in phosphate buffered saline (PBS, 2.7 mM KCl, 1.5 mM KH_2PO_4, 136.9 mM NaCl, 8.9 mM $Na_2HPO_4•7H_2O$). Vegetations to be used for proteomics were snap-frozen in liquid nitrogen, stored at -80 °C and analyzed between one week and six-months postoperatively. For large vegetations, a portion was used for proteomics and another was embedded in paraffin for histology and immunohistochemistry as described in the Online Supplement. Where smaller vegetations were used in entirety for proteomics, separate vegetations from the same patient were used for histology unless unavailable (SA3).

Histopathology and Immunofluorescence

6 μm-thick vegetation sections were stained with hematoxylin & eosin (H&E), Movat pentachrome, phosphotungstic acid-haematoxylin (PTAH), and the Gram stain. Slides were viewed and scanned using a Leica SCN400 automated scanner. Histopathology features were reviewed by cardiac pathology specialists (authors E.R.R. and C.T.). Serial sections of each vegetation sample were selected for indirect immunofluorescence using antibodies against CD66b, a membrane marker for neutrophils (AB_2728422, Biolegend) and citrullinated histone H3 (ab5103, Abcam). Sections were heated in citrate buffer, pH 6.0 for antigen retrieval and allowed to cool to room temperature, treated with chondroitinase ABC (0.05 U/μl, diluted 1:200 in PBST (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na2HPO4, 10 mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20, pH 7.2) with 5% normal goat serum), blocked for 1 hour at room temperature, and stained with the primary antibody followed by Alexa Fluor secondary goat anti-mouse or goat anti-rabbit antibodies (Invitrogen) diluted 1:200 in PBST and incubated on the tissue sections for 1 hour at room temperature in the dark. Slides were washed thrice for 5 minutes using PBST and once for 5 minutes using PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na2HPO4, 10 mM Tris-HCl, 150 mM NaCl). Coverslips were mounted using Prolong-Gold with the DNA label 4′, 6-diamidino-2-phenylindole (DAPI,

Thermo Fisher Scientific) and imaged on an Olympus BX51 microscope with a Leica DFC7000 T camera using Leica Application Suite software. ImageJ (NIH) or Adobe Photoshop software was used for post-processing of images, specifically, white/black balancing and adjustment of brightness and contrast as needed for clarity, with the adjustments being applied equally to the entire image for all images.

Protein extraction for mass spectrometry

100-150 mg of each vegetation was homogenized in 1ml of T-Per (ThermoFisher Scientific) extraction buffer supplemented with protease inhibitor (cOmplete tablets, Roche) using a T10 ultra Turrax (IKA) homogenizer, boiled at 95° C for 5 minutes and cooled to room temperature before adding 1 µl of benzonase (EMD Millipore Corp.). The homogenates were tip probe-sonicated at 20% amperage with 3 sec on/off intervals a total of 6 times using a Q-500 sonicator (Qsonica), centrifuged at 15,000 g for 10 minutes, and the supernatant was retained for analysis (T-Per extract). The pellet was washed twice with cold PBS and resuspended in a chaotropic buffer (5 M GuHCL, 1% CHAPS, 25 mM $NaC_2H_3O_2$, 50 mM aminocaproic acid, 5 mM Na_2EDTA) supplemented with protease inhibitor (cOmplete tablets, Roche) at 4° C with rotation for 72 h, centrifuged at 17,000 g for 10 minutes and the supernatant was retained for analysis (matrix extract). Each extract was analyzed separately on the mass spectrometry instrument and the data from both was combined for the database search.

TAILS workflow

The extracted proteins were reduced with 5 mM dithiothreitol (DTT) for 30 minutes at 37° C followed by alkylation with 20 mM iodoacetamide in the dark for 20 minutes. The reaction was quenched by adding DTT to a final concentration of 15 mM. Proteins were chloroform-methanol precipitated, the pellet was dried and resolubilized in 50 µl of 100 mM NaOH followed by 50 µl of 1 M HEPES, pH 7.5 and 100 µl of 6 M GuHCL, vortexed and kept at 4° C until no precipitate remained. Extracted protein was quantified using a 660 nm Bradford assay (ThermoFisher Scientific) and standardized to 275 µg of protein per sample. Extracts from the same vegetation were combined when insufficient protein was available from each extract (NSA1, 2, 3, and 4 TAILS preparations). Otherwise, extractions were

processed separately and combined for the database search (TAILS SA1, 2, 3, both NSA5, and all shotgun preperations). Proteins were labeled overnight with 40 mM light formaldehyde, which binds specifically to free N- termini and lysine side-chains and (α and ε amines respectively) in the presence of 20 mM sodium cyanoborohydride. They were treated with an additional fresh 20 mM formaldehyde and 10 mM sodium cyanoborohydride for two hours the next morning and the reaction was quenched with 100 mM Tris for 1 h. The samples then underwent buffer exchange on a 3 kDa molecular weight cut-off (MWCO) column (EMD Millipore corp.) into 50 mM ammonium bicarbonate, followed by digestion with mass spectrometry grade trypsin gold (Promega) at a 1:50 trypsin: protein ratio overnight at 37° C. 30 µg of this digest was reserved for shotgun analysis and the remaining protein was mixed with hyperbranched polyglycerol-aldehydes (HPG-ALD, Flintbox, https://www.flintbox.com/public/project/1948/) at a 5:1 polymer: protein ratio. HPG-ALD binds and retains unblocked (i.e. trypsin-generated) amino acid termini, thus allowing enrichment of naturally and experimentally blocked (e.g., by reductive dimethylation) N-termini.[1] The HPG-ALD-peptide mixture was filtered through a 10 kDa MWCO filter (EMD Millipore corp.) to obtain peptides with blocked N-termini (TAILS peptides) in the flow-through. The TAILS and shotgun peptides were desalted on a C18 Sep-Pak (Waters) column and eluted in 60:40 ACN: 1% TFA. Samples were vacuum centrifuged until dry and resuspended in 1% acetic acid for mass spectrometry.

Peptide fractionation

Peptide fractionation was performed on the two vegetations with sufficiently available material, SA2 and NSA5T. Protein extraction was performed as described above and 100 µg of protein from each extraction was denatured, labeled, digested, and desalted as described above. Peptides were lyophilized following desalting and resuspended in 100 µl of mobile phase A, 10 mM ammonium formate pH 10. High pH reversed phase fractionation was undertaken using mobile phase A and mobile phase B (90% ACN: 10% 10 mM ammonium formate pH 10) on a Waters XBridge BEH C18 column, 130 Å, 3.5 µm, and 2.1 mm X 150 mm. The mobile phase B was increased from 2% to 90% over 27 minutes with a 5-minute hold at 90% B and then equilibrated back to 2% between samples. 1 mL fractions were collected

using a 250 µl/min flow rate and consolidated into 8 bins. Maximum phase separation between bins was ensured by combining every 8th fraction (i.e., fraction 1 combined with 9, 17, 26 and fraction 2 combined with 10, 18, 27 etc). Fractions were vacuum-centrifuged until dry and resuspended in 1% acetic acid for mass spectrometry.

LC-MS/MS

Peptides were analyzed on a ThermoFisher Scientific Fusion Lumos tribrid mass spectrometer system interfaced with a Thermo Ultimate 3000 nano-UHPLC. The HPLC column was a Dionex 15 cm x 75 µm id Acclaim Pepmap C18, 2 µm, 100 Å reversed phase capillary chromatography column. 5 µL volumes of the trypsin-digested extract were injected, peptides were eluted from the column by an acetonitrile/ 0.1% formic acid gradient at a flow rate of 0.3 µL/min and introduced in-line into the mass spectrometer over a 120 minute gradient. The nanospray ion source was operated at 1.9 kV. The digest was analyzed using a data-dependent method with 35% collision-induced dissociation fragmentation of the most abundant peptides every 3 seconds and an isolation window of 0.7 m/z for ion-trap MS/MS. Scans were conducted at a maximum resolution of 120,000 for full MS. Dynamic exclusion was enabled with a repeat count of 1 and ions within 10 ppm of the fragmented mass were excluded for 60 seconds.

Proteomics data analysis

Peptides were identified using a precursor mass tolerance of 10 ppm, and fragment mass tolerance of 0.6 Da. The only static modification was carbamidomethyl (C), whereas dynamic modifications included the light (28.03 Da) dimethyl formaldehyde (N-terminal, K), oxidation (M), deamidation (N, R (citrullination)), acetylation (N-terminal), and Gln to pyro-Glu N-terminal cyclization. Peptides were validated using a false discovery rate (FDR) of 1% against a decoy database. Chromatographic retention time alignment was used across like (e.g., all TAILS) samples for accurate label-free quantitation comparison and increased peptide identifications. Only high confidence proteins (containing peptides at a 99% confidence level or higher) were recorded from each sample for data

analysis. Shotgun data required a minimum of two high-confidence peptides for protein identification and TAILS required a single peptide. When a peptide corresponding to multiple proteins was identified, all protein accession numbers were reported. Peptides corresponding to multiple accession numbers were assigned to one master protein accession and termed a protein group (based on the number of unique peptides in each accession and overall sequence length of the protein, e.g., ACTG2, ACTB, ACTA1, and ACTA2 have peptides with the same sequence, but only the protein which has a unique peptide(s) found is considered a master protein, and all others are combined into the same protein group).

The human database included all reviewed proteins with experimental evidence. Common contaminant proteins (albumin, serotransferrin, keratin, etc.) were filtered out based on the Global Proteome Machine common Repository of Adventitious Proteins (cRAP) database (version 2012.01.01). Pathogen databases included all predicted proteins and were searched on a human database background to compensate for their limited number of proteins. The protein and peptide data from each extraction were pooled to obtain the proteome of each sample. Data was analyzed using Microsoft Excel to determine which proteins were common between individual samples and between the staphylococcal and non-staphylococcal cohorts. Proteins were filtered by their presence in either all vegetations from the staphylococcal group, or in at least 3 of the 4 *Streptococcus* strains and in both the *Cutibacterium acnes* and *Candida parapsilosis* infected vegetations in the non-staphylococcal group. Protein accession numbers were submitted to the PANTHER software (pantherdb.org) for pathway analysis and GO annotations. Proteins were sorted based on annotations and cross-referenced with previously reported proteomes (for blood, platelets and neutrophils). N-terminal peptides were identified by the presence of an N-terminal dimethyl label or other N-terminal blocking entity. These were analyzed with CLIPPER V 1.0 to assign their position with respect to the complete protein sequence.

References

1. Kockmann T, Carte N, Melkko S and Keller UAD. Identification of Protease Substrates in Complex Proteomes by iTRAQ-TAILS on a Thermo Q Exactive Instrument. *Neuromethods*. 2015;114:187-207.

Supplemental Table 1. Histopathological observations of vegetation tissue

Patient ID	Vegetation Location	H&E	Movat	PTAH	Gram	GMS	Blood culture	Valve culture	Valve sequencing	Final microbial diagnosis
SA1	Base of mitral valve	Fibrinous vegetation with acute inflammation	Fibrinous vegetation material without organization	N/A	-	-	*S. aureus*	*S. aureus*	*S. aureus*	*S. aureus*
SA2	Tricuspid valve	Fibrinous thrombotic material with acute inflammation.	Large segments of fibrinous vegetation. Small fragment of valve composed of fibrous and elastic tissue. Granulation tissue and fibrinous vegetation present.	Fibrin present in segments of vegetations and on edge of valvular tissue.	+	-	*S. aureus*	*S. aureus*	*S. aureus*	*S. aureus*
SA3	Right atrial fistula	N/A	N/A	N/A	N/A	N/A	No growth	No growth	*S. aureus*	*S. aureus*
NSA1	Native aortic valve	Acute inflammatory exudate with calcification, regions of organizing vegetations.	Fibrous tissue with acute inflammation and necrosis. Granulation tissue was present in proximity to fibrous tissue.	Large regions of vegetation composed of fibrin, interspersed with segments of fibrinoid exudate devoid of fibrin.	-	+	No growth	No growth	*S. dysgalactiae*	*S. dysgalactiae*
NSA2	Native aortic valve	Valve tissue without inflammatory infiltrates	Semilunar valve without inflammatory infiltrates or granulation tissue.	N/A	-	-	No growth	No growth	*S. bovis*	*S. bovis*

NSA3	Prosthetic aortic valve	Bioprosthetic valve with large formation	Bovine pericardial bioprosthetic valve.	N/A	-	+	*C. parapsilosis*	*C. parapsilosis*	*C. parapsilosis*	*C. parapsilosis*
NSA4	Prosthetic aortic valve	Bland fibrinous exudate with large bacterial colonies and no acute inflammatory cells.	Fibrinous exudate with large bacterial colonies. No connective/valvular tissue present.	Layers of fibrin and interspersed segments of fibrinoid exudate devoid of fibrin.	+	+	No growth	*C. acnes*	*C. acnes*	*C. acnes*
NSA5	**1.** Native aortic valve **2.** Native tricuspid valve	**1. Aortic valve:** Fibrinous and organizing vegetation, with sparse acute inflammation. **2. Tricuspid valve:** Large fibrinous vegetation with acute inflammation, micro abscesses, and necrosis.	**1. Aortic valve**: Trilaminar aortic valve tissue distorted by large region of calcification. **2. Tricuspid Valve:** No evidence of valvular tissue. Large segment of fibrinous exudate with trapped erythrocytes and acute inflammation.	**1. Aortic Valve:** Vegetation composed primarily of fibrin. Thin coating of fibrin thrombus on valve surface. **2. Tricuspid Valve:** Vegetation composed of large segments of fibrin, with trapped erythrocytes, platelets, and acute inflammation. Interspersed regions of fibrinoid exudate devoid of fibrin.	-	+	*S. parasanguinis*	No growth	*S. parasanguinis*	*S. parasanguinis*

N/A, indicates tissue not available for staining; GMS, Grocott methanamine silver; SA, staphylococcal; NSA, non-staphylococcal; PTAH, phosphotungstic acid haematoxylin;

Supplemental Table 2. Summary of non-fractionated mass spectrometry data

Sample (n)	Type	1% FDR proteins	Protein groups	PSMs	Peptide groups	Proteins in all samples	Labeling efficiency	Matrix/secreted in all samples	Pull down retention	Unique internal peptides	Unique internal peptides in 3+	Proteins with unique internal peptides
SA (3)	Shotgun	2256 (1729±24)	1073	33410 (4248±463)	5999	605	77.3%	79	N/A	N/A	N/A	N/A
NSA (5)	Shotgun	2269 (828±60.7)	1093	40707 (3392±898)	6946	216	79%	34	N/A	N/A	N/A	N/A
SA (5)	TAILS	1807 (861±155)	1128	10050 (1675±310)	3308	348	N/A	45	56%	2099	503	193
NSA (5)	TAILS	1534 (563±236)	957	13402 (1675±603)	3694	76	N/A	10	70%	2850	617 (127 all)	374

N/A indicates data that was not measurable or relevant for that particular sample; PSM, peptide spectral matches; SA, staphylococcal; NSA, non-staphylococcal; "3+", 3 or more samples from that group, "all", every sample in that group (Aorta and tricuspid valve vegetation data from NSA5 were combined).

Supplemental Table 3. Proteins shared between each vegetation and a reference neutrophil proteome (PXD010701)

Vegetation	Proteins	Number of proteins shared with neutrophils	Vegetation proteins shared with neutrophils	Neutrophil proteins shared with vegetations
SA1	782	545	70%	13%
SA2	851	611	72%	15%
SA3	894	628	70%	15%
NSA1	669	464	69%	11%
NSA2	625	439	70%	11%
NSA3	656	455	69%	11%
NSA4	335	189	56%[A]	4.56%[A]
NSA5A	615	426	69%	10%
NSA5T	725	544	75%	13%

[A]Values fall below two standard deviations of the mean and are outliers determined using a two-sided Grubbs' test for a single outlier with *P* values of <.01 for 56% and <.05 for 4.56%.

SA, staphylococcal; NSA, non-staphylococcal

Supplemental Table 4. Ragging in fibrin and fibronectin TAILS peptides

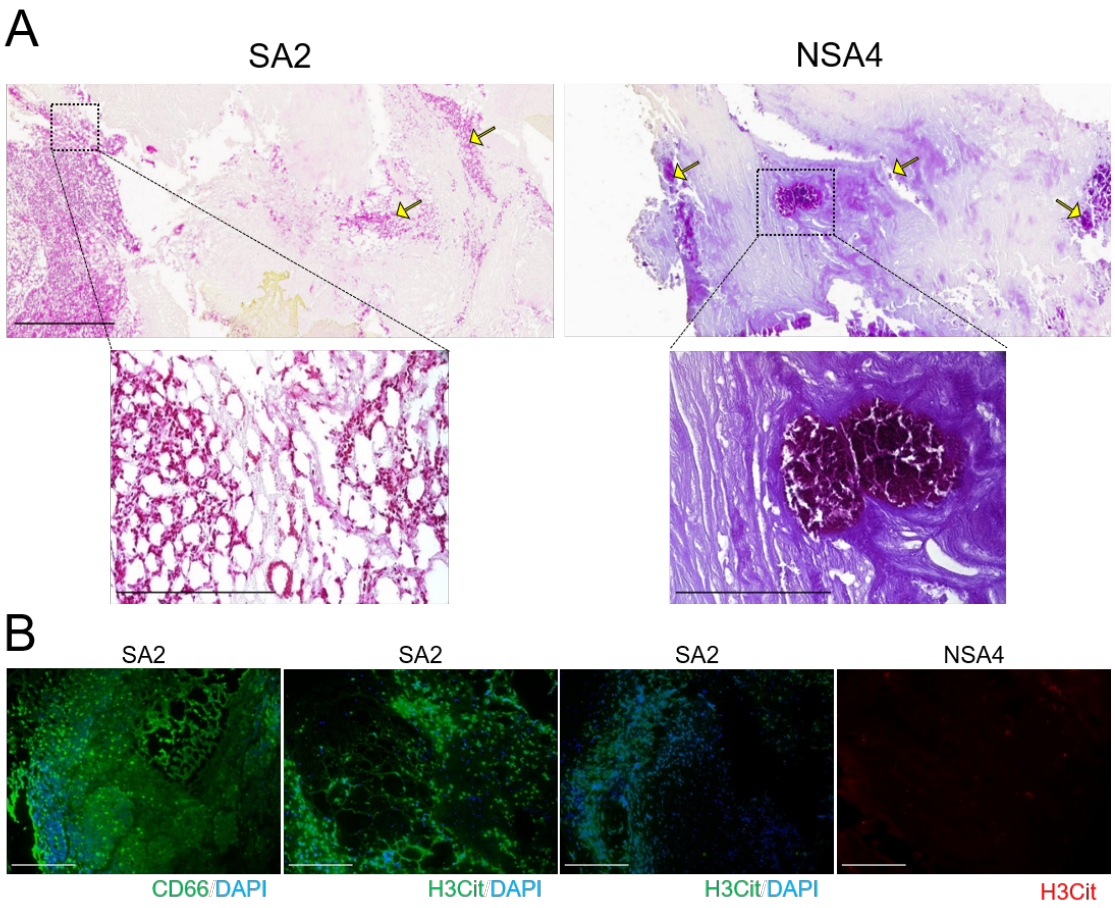
Protein	Position	Sequence
Fibrinogen α	E268	EITRGGSTSYGTGSETESPR
		TRGGSTSYGTGSETESPR
		RGGSTSYGTGSETESPR
		GGSTSYGTGSETESPR
		GGSTSYGTGSETESPRN
		GSTSYGTGSETESPR
		STSYGTGSETESPR
		TSYGTGSETESPR
		SYGTGSETESPR
		YGTGSETESPR
		GTGSETESPR
Fibrinogen β	T396	TMTIHNGMFFSTYDR
		MTIHNGMFFSTYDR
		TIHNGMFFSTYDR
		HNGMFFSTYDR
		GMFFSTYDR
Fibrinogen γ	A286	ADYAMFKVGPEADKYR
		DYAMFKVGPEADKYR
		FKVGPEADKYR
		KVGPEADKYR
		VGPEADKYR
Fibronectin	Q1354	QKTGLDSPTGIDFSDITANSFTVHWIAPR
		QKTGLDSPTGIDFSDITANSFTVH
		QKTGLDSPTGIDFSDITANSF
		QKTGLDSPTGIDFSDITAN
	T1836	TKTETITGFQVDAVPANGQTPIQR
		TKTETITGFQVDAVPANGQTPI
		TGFQVDAVPANGQTPIQR
		QVDAVPANGQTPIQR
		DAVPANGQTPIQR
		AVPANGQTPIQR

Supplemental Table 5. Human proteases identified in vegetations

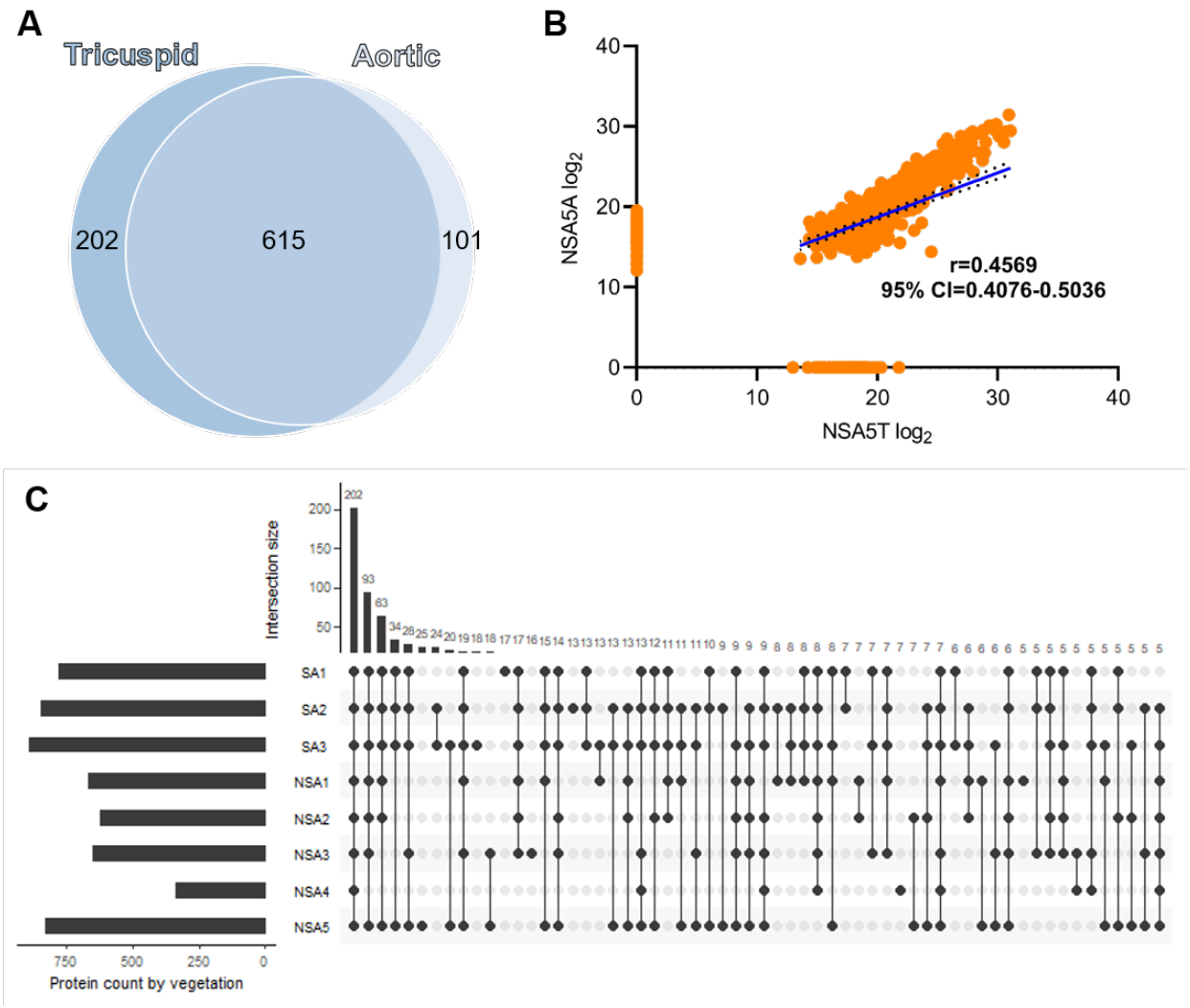
Accession number	Protein name	Gene	Protease type	Endo/Exo	Average abundance	NSA	SA
P00747	Plasminogen	*PLG*	Serine protease	Endo	4.96E+07	6	3
P00734	Prothrombin	*F2*	Serine protease	Endo	3.67E+07	6	3
P08246	Neutrophil elastase	*ELANE*	Serine protease	Endo	2.74E+07	6	3
P07858	Cathepsin B	*CTSB*	Cysteine protease	Endo/exo	2.58E+07	3	3
P24158	Myeloblastin	*PRTN3*	Serine protease	Endo	2.48E+07	5	3
P02790	Hemopexin	*HPX*	Metalloprotease	Uk	1.84E+07	6	3
P14780	Matrix metalloproteinase-9	*MMP9*	Metalloprotease	Exo	1.37E+07	5	3
P20160	Azurocidin	*AZU1*	Serine protease	Endo	1.25E+07	6	3
P15144	Aminopeptidase N	*ANPEP*	Metalloprotease	Exo	7.41E+06	5	3
Q9UBR2	Cathepsin Z	*CTSZ*	Cysteine protease	Exo	6.36E+06	5	3
Q96IY4	Carboxypeptidase B2	*CPB2*	Metalloprotease	Endo	4.08E+06	5	3
P08519	Apolipoprotein(a)	*LPA*	Serine protease	Endo	2.96E+06	4	3
P22894	Neutrophil collagenase	*MMP8*	Metalloprotease	Endo	2.65E+06	4	3
P39900	Macrophage metalloelastase	*MMP12*	Metalloprotease	Endo	2.48E+06	2	0
P67812	Signal peptidase complex catalytic subunit SEC11A	*SEC11A*	Serine protease	Endo	2.40E+06	3	3
P00736	Complement C1r subcomponent	*C1R*	Serine protease	Endo	2.11E+06	4	1
O14773	Tripeptidyl-peptidase 1	*TPP1*	Serine protease	Endo/exo	2.02E+06	5	3
P09871	Complement C1s subcomponent	*C1S*	Serine protease	Endo	1.98E+06	6	2
P45974	Ubiquitin carboxyl-terminal hydrolase 5	*USP5*	Cysteine protease	Exo	1.66E+06	3	3
P54578	Ubiquitin carboxyl-terminal hydrolase 14	*USP14*	Cysteine protease	Endo	1.52E+06	4	1
P07384	Calpain-1 catalytic subunit	*CAPN1*	Cysteine protease	Endo	1.49E+06	2	3
P53634	Dipeptidyl peptidase 1	*CTSC*	Cysteine protease	Endo/exo	1.12E+06	1	2
P55786	Puromycin-sensitive aminopeptidase	*NPEPPS*	Metalloprotease	Endo	6.12E+05	3	3
P00748	Coagulation factor XII	*F12*	Serine protease	Endo	5.87E+05	2	1
Q9H3G5	Probable serine carboxypeptidase CPVL	*CPVL*	Serine protease	Exo	5.16E+05	3	0
P03952	Plasma kallikrein	*KLKB1*	Serine protease	Endo	5.01E+05	2	3
Q9NZ08	Endoplasmic reticulum aminopeptidase 1	*ERAP1*	Metalloprotease	Exo	4.00E+05	4	3
Q5TDH0	Protein DDI1 homolog 2	*DDI2*	Aspartic protease	Endo	3.80E+05	5	2
Q9HB40	Retinoid-inducible serine carboxypeptidase	*SCPEP1*	Serine protease	Exo	3.66E+05	2	0
P29466	Caspase-1	*CASP1*	Cysteine protease	Endo	2.61E+05	1	0
P28838	Cytosol aminopeptidase	*LAP3*	Metalloprotease	Exo	2.39E+05	2	1
P08253	72 kda type IV collagenase	*MMP2*	Metalloprotease	Endo	2.08E+05	2	3
Q92743	Serine protease HTRA1	*HTRA1*	Serine protease	Endo	1.65E+05	1	2
	Proteases found only after vegetation fractionation						
Q16740	ATP-dependent Clp protease proteolytic subunit, mitochondrial	*CLPP*	Serine protease	Endo	1.25E+06	0	1
P29122	Proprotein convertase subtilisin/kexin type 6	*PCSK6*	Serine protease	Endo	2.43E+06	0	1

P07477	Trypsin-1	*PRSS1*	Serine protease	Endo	6.16E+06	0	1
P35030	Trypsin-3	*PRSS3*	Serine protease	Endo	1.20E+07	1	0
Q9Y4W6	AFG3-like protein 2	*AFG3L2*	Metalloprotease	Endo	2.12E+06	1	1
P00746	Complement factor D	*CFD*	Serine protease	Endo	1.65E+06	1	1
O00187	Mannan-binding lectin serine protease 2	*MASP2*	Serine protease	Endo	8.14E+06	1	1
P49768	Presenilin-1	*PSEN1*	Aspartic protease	Endo	8.68E+06	1	1
P00742	Coagulation factor X	*F10*	Serine protease	Endo	3.59E+07	1	1
P08473	Neprilysin	*MME*	Metalloprotease	Endo	1.02E+07	1	1
P04070	Vitamin K-dependent protein C	*PROC*	Serine protease	Endo	6.75E+06	1	1
P42574	Caspase-3	*CASP3*	Aspartic protease	Endo	5.92E+06	1	1
P00740	Coagulation factor IX	*F9*	Serine protease	Endo	2.10E+07	1	1
Q6UWY2	Serine protease 57	*PRSS57*	Serine protease	Endo	1.23E+07	1	1
P48740-2	Mannan-binding lectin serine protease 1	*MASP1*	Serine protease	Endo	1.03E+07	1	1

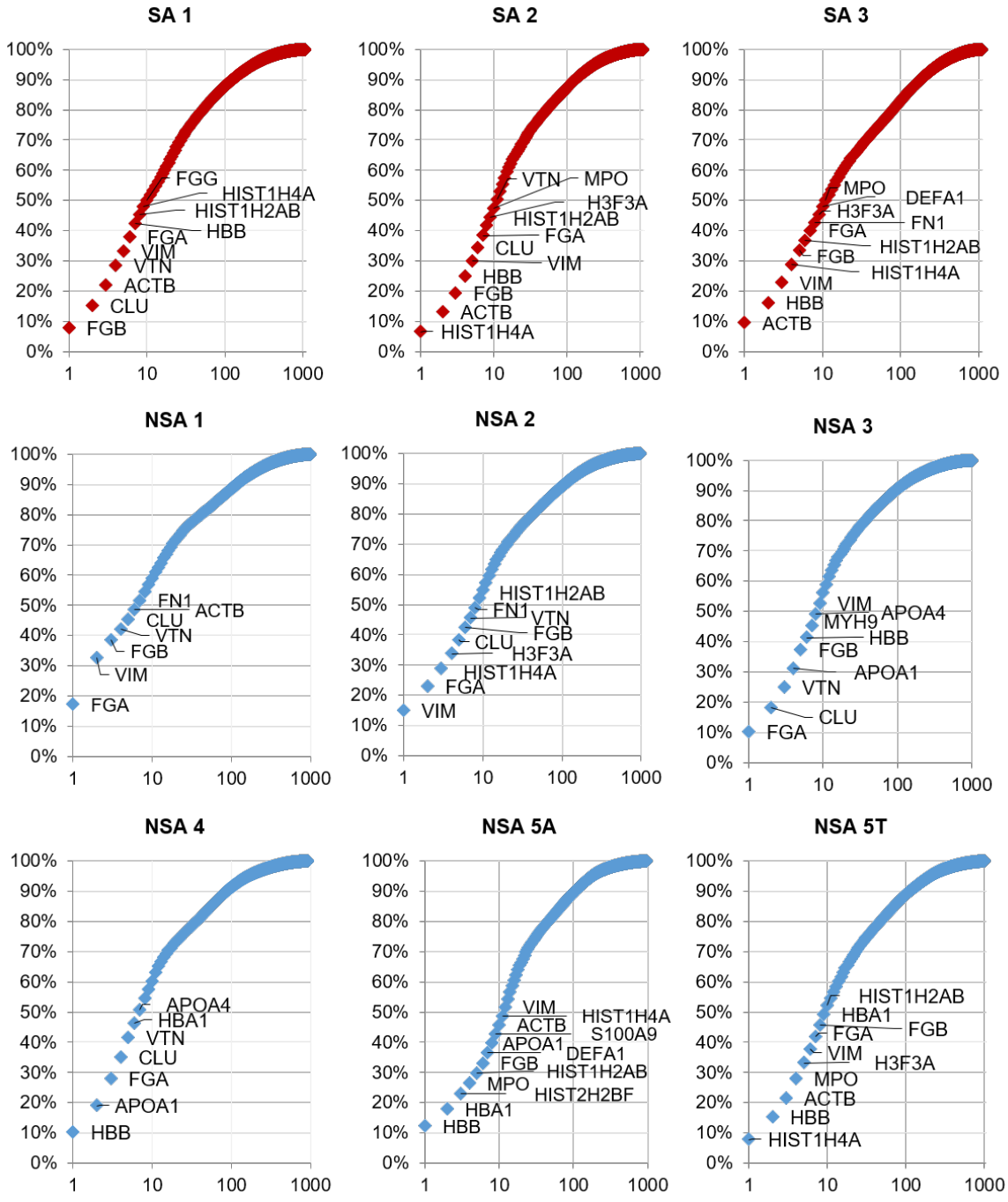
Proteases are listed in order of their average abundance in all the shotgun samples. Uk (unknown) indicates that the specific protease activity of this protein has not been determined, Endo, endoprotease, Exo, exoprotease.



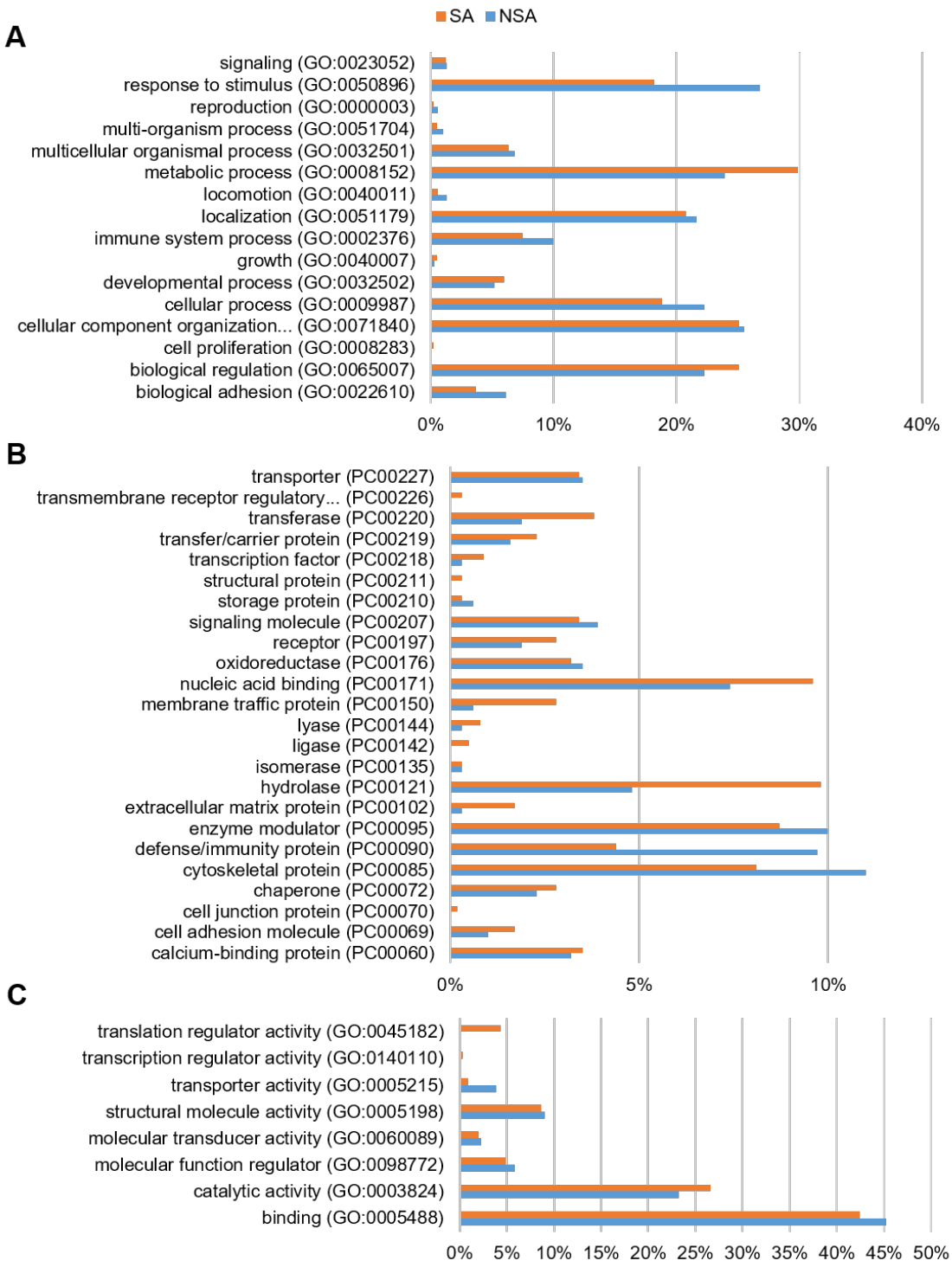
Supplemental Figure 1. Histology and immunofluorescence reveal key features of vegetations. A. Positive Gram stain (yellow arrows) showing the presence of the pathogen (SA2- *S. aureus*, and NSA4- *C. acnes*). **B**. Neutrophil infiltration and NEtosis in vegetation SA2 are shown by immunostaining for CD66 and histone H3 citrullination, respectively (green signal). Note their absence in the acellular NSA4 vegetation. Nuclei were stained using DAPI. Scale bars in **A** are 500 µM (50 µM for zoomed images), and 200 µM in **B**.



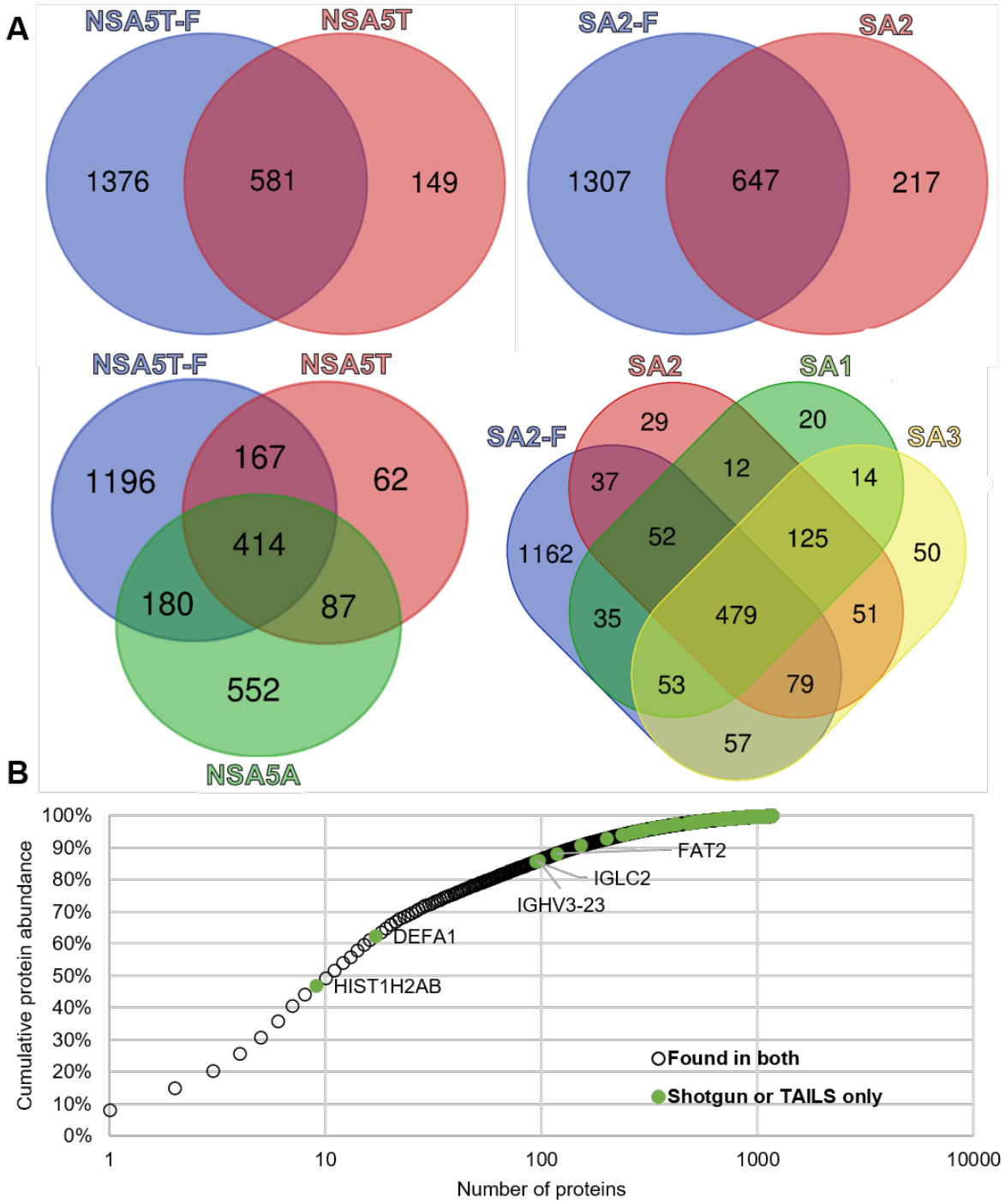
Supplemental Figure 2. Comparison of vegetations from the same patient and from all patients. A. The proteomes of a tricuspid and aortic valve vegetation from the same patient (NSA5) showed a 67% overlap in identified proteins. **B**. A Pearson correlation plot shows a moderate positive correlation (significantly non-zero P< .0001) between the log2 transformed protein intensities in the aortic valve (NSA5A) and tricuspid valve (NSA5T) vegetations. **C**. UpSet plot aggregating the overlap of proteins found in the SA and NSA vegetations. The top 60 overlapping groups of 255 possible overlaps are shown. The total number of proteins in each vegetation is identified in the horizontal bar chart next to each vegetation name (numbered by SA (staphylococcal) and NSA (non-staphylococcal)designations). Each vertical bar shows the number of shared proteins unique to the grouped vegetation(s) indicated by solid black circles.



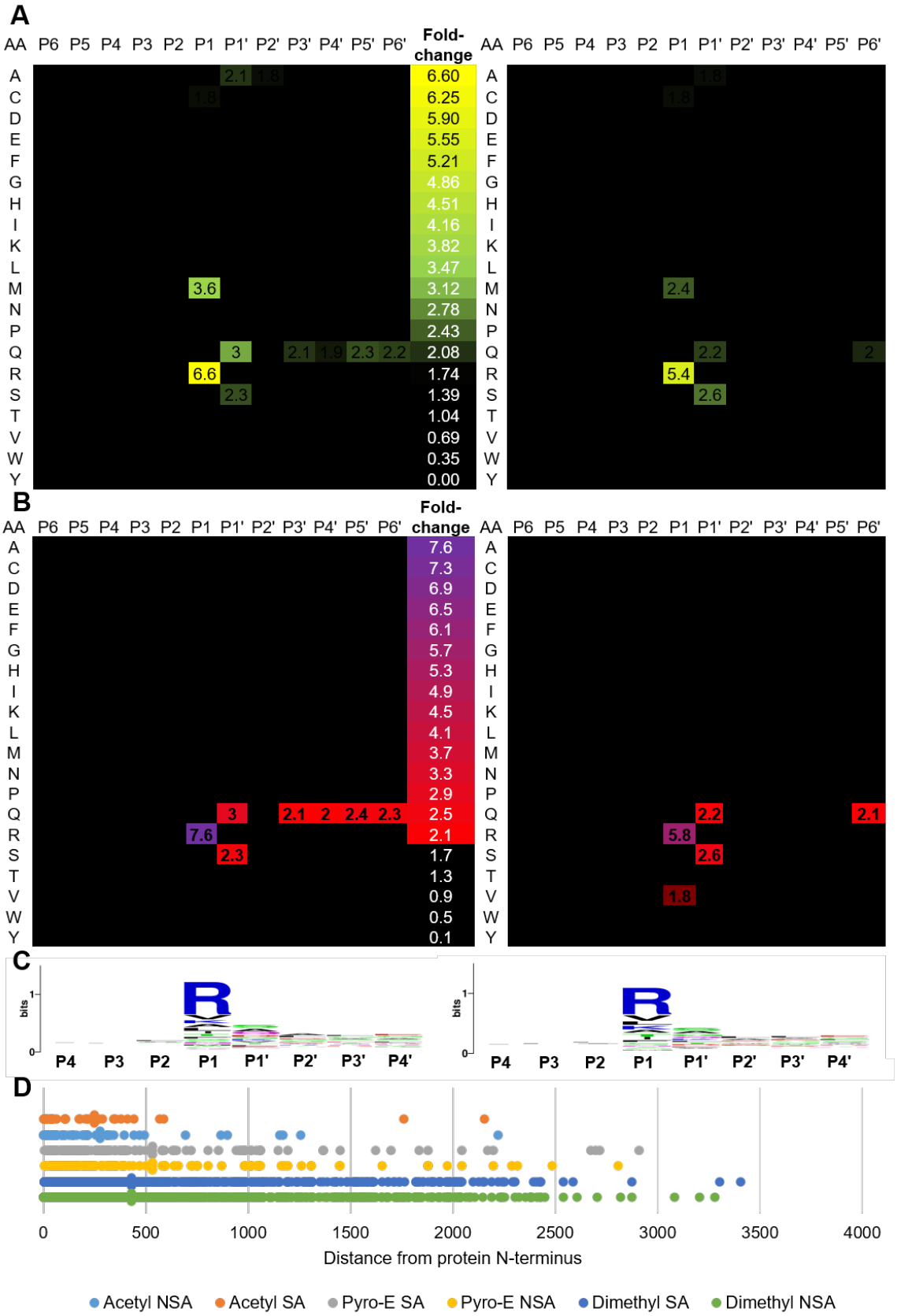
Supplemental Figure 3. Normalized protein ion abundance plot for each vegetation. Proteins were ranked in order of normalized abundance (x-axis) and graphed based on their additive percent abundance contribution (y-axis) to the overall protein abundance recorded in that vegetation. Proteins that fell within the top 50th percentile of protein abundance in the respective vegetation are labeled.



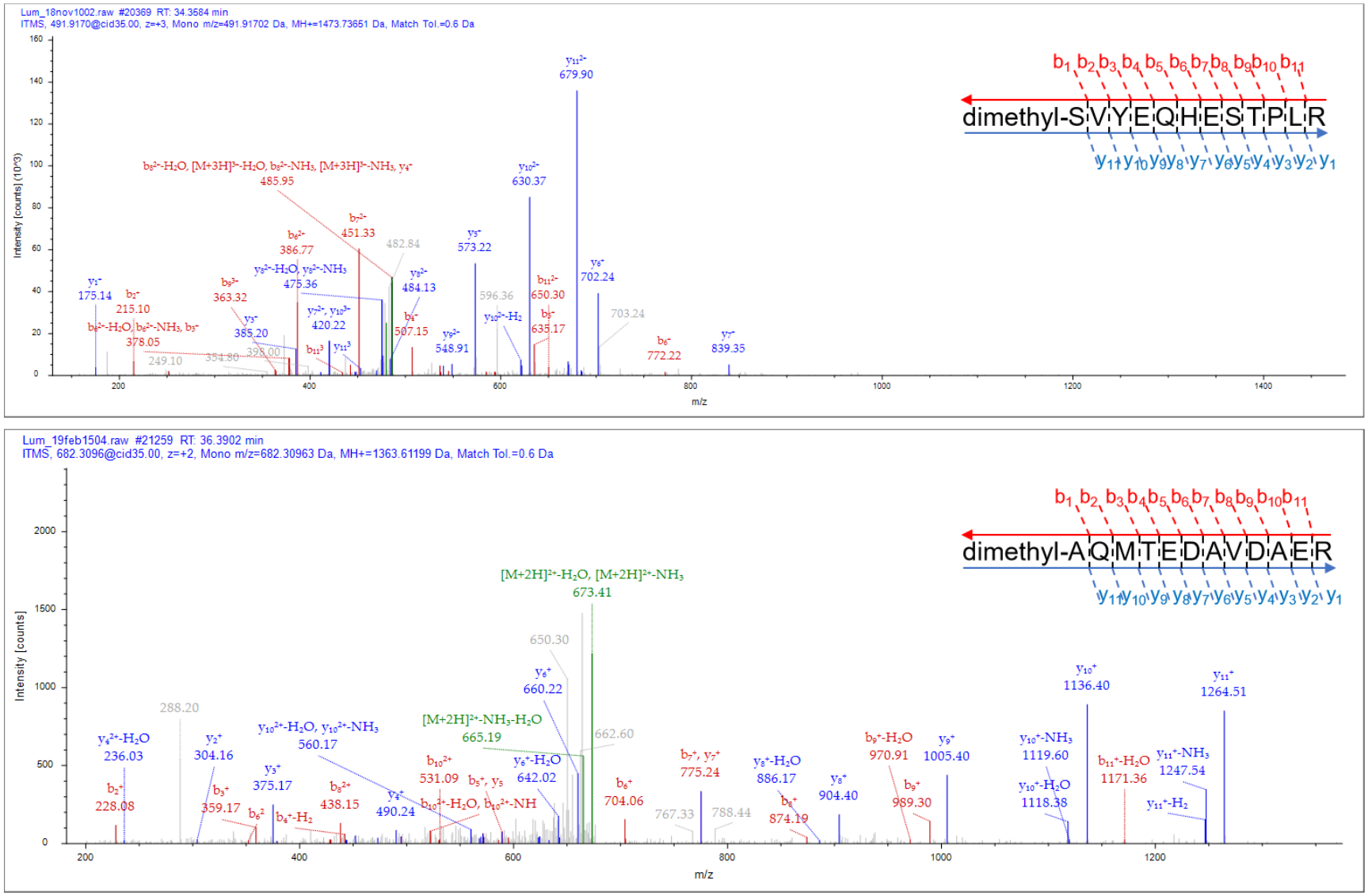
Supplemental Figure 4. PantherDB pathway analysis shows no significant differences between staphylococcal (SA) and non-staphylococcal (NSA) proteomes. Pathway analysis of the proteins identified in staphylococcal and non-staphylococcal vegetations revealed only minor differences in (**A**) biological processes, (**B**) protein classes, and(**C**) molecular functions of the identified proteins.



Supplemental Figure 5. Vegetation proteome coverage was enhanced by fractionation. A. Venn diagrams of proteins identified following fractionation (NSA5T-F or SA2-F) and those originally found without fractionation. **B**. Normalized protein ion abundance plot of the proteins originally identified without fractionation (all points) highlighting the proteins that were exclusively identified in the non-fractionated analysis (green). Proteins within the 90th percentile were labeled with their gene symbols.



Supplemental Figure 6. Similar protease signatures are observed in staphylococcal (SA) and non-staphylococcal (NSA) vegetations. A. CLIP-PICS analysis of amino acid frequency around the scissile bond of all identified N-termini shows a preference for Met and Arg at the P1 position in both groups (left SA, right NSA) indicating an enrichment of natural N-termini and trypsin-like cleavages, respectively. **B.** CLIP-PICS diagram of internal peptides shows similar preferences in SA (left) and NSA (right) groups. **C.** WebLogo plots of residues around the scissile bonds identified in SA (left) and NSA (right) internal peptides also indicate the same cleavage site preference in each group. **D.** Localization of cleavage sites relative to the original protein N-terminus does not show a significant positional difference between the SA and NSA groups. Peptides are categorized by N-terminal modification and the average position is denoted by the vertical line.



Supplemental Figure 7. Annotated MS/MS spectra of biomarker candidate peptides. MS/MS spectra show identification of the fibronectin peptide (Top, dimethyl-SVYEQHESTPLR) and the complement C3 peptide (bottom, dimethyl-AQMTEDAVDAER) using a high mass accuracy Orbitrap Fusion Lumos tribrid instrument and CID fragmentation.