Supplementary Materials

Materials and Methods

Bacterial culture and purification of LtxA

Aa strains HK1651 (ATCC 700685), SUNYab 75 (ATCC 43717), and CU1000 (Daniel H. Fine; Rutgers School of Dental Medicine, Newark, NJ) were grown aerobically in 2× YT medium at 37°C in CO_2-enriched (5%) air. *S intermedius* PK2821 and *S oralis* 34, both from the NIDCR culture collection, were grown in brain-heart infusion broth at 37°C, 5% (v/v) CO_2. *A defectiva* (ATCC 49176) was grown at 37°C, 5% (v/v) CO_2 in brain-heart infusion broth supplemented with 10% (v/v) horse serum and 10 µg/mL filter-sterilized pyridoxal hydrochloride.

The following organisms were grown in static 10-mL tubes at 37°C in $N_2/H_2/CO_2$ (90%/5%/5%). Filter-sterilized supplements were added to the media after autoclaving, and tubes were placed in anaerobic conditions overnight before inoculation. *P gingivalis* W83 was grown in Todd Hewitt broth supplemented with 5 µg/mL hemin and 1 µg/mL vitamin K. *P intermedia* (ATCC 15032) and *P micra* (ATCC 33270) were grown in Schaedler broth supplemented with hemin and vitamin K. *T forsythia* (ATCC 43037) was grown in Tryptic Soy broth supplemented with 10 µg/mL N-acetylmuramic acid. *T denticola* (ATCC 35405) was grown in GM-1 medium. *F nucleatum* (ATCC 10953) was grown in brain-heart infusion broth supplemented with 0.25% ammonium glutamate.

LtxA was purified as previously described (*53*). Briefly, *Aa* strain HK921 (serotype b) was grown in liquid culture, and LtxA was purified from the supernatant using an Ächta 900 Purifier system and gel permeation chromatography. LPS was removed using Polymyxin B activated

agarose. Lytic activity was determined using K562 cells with recombinant expression of β2 integrin in the form of $\alpha_L\beta2$, $\alpha_M\beta2$, or $\alpha_X\beta2$. Studies were done in triplicate with LtxA at 37°C for 1 h, at which point cell viability (median, SD) was determined by the CellTiter-GloTM assay. Under these conditions, viability was decreased by 50% using LtxA at 0.3 μg/mL.

In vitro stimulation of human leukocytes

Peripheral blood neutrophils were isolated from healthy donor whole blood as previously described (*2*). Briefly, neutrophils were recovered by Ficoll gradient centrifugation. Erythrocytes were removed by hypotonic lysis, and cells were resuspended in HBSS 1x, HEPES 10 mM, $CaCl_2$ 1.5 mM (HBSS-Ca) at a final density of $10x10^6$ cells/mL. Peripheral blood mononuclear cells (PBMCs) were resuspended in HBSS-Ca. To obtain macrophages, PBMCs were fractionated by plastic adherence. Non-adherent lymphocytes were removed by washing plates, and monocytes differentiated for 14 days in RPMI-1640 medium. Leukocytes were incubated with 0-1000 ng/mL purified LtxA at 37°C and lysed in SDS sample buffer. In some experiments, neutrophils were incubated in parallel with PMA (Cayman), LPS from *P gingivalis* (InvivoGen), or purified LtxA at a final concentration of 500 nM, 1000 ng/mL, and 300 ng/mL, respectively. After incubation for 3 h at 37°C, neutrophils were lysed in SDS sample buffer. For inhibitor treatment, neutrophils were preincubated with 10 μM DPI (Cayman), 200 μM Cl-amidine (Millipore), or 1.5 mM EDTA for 30 min before stimulation. For analysis of extracellular material released from lytic (LtxA-treated) and NETotic (PMA or LPS-treated) cells by SDS-PAGE, stimulated neutrophils were incubated with fresh buffer containing 10 U/mL micrococcal nuclease (Thermo Scientific) for 10 min at 37°C. Supernatants were collected in 1.5 mM EDTA,

centrifuged at 20,000 x *g* for 15 min to remove cell debris, and proteins were precipitated in 20% trichloroacetic acid (TCA). Precipitates were washed in acetone and resuspended in SDS sample buffer.

Incubation of human neutrophils with bacterial cells

Bacterial cells were grown in solution, pelleted by centrifugation, and washed in HBSS. Bacteria were resuspended in HBSS-Ca and adjusted by optical density at 600 nm wavelength (OD_{600}). PMNs ($10x10^6$ cells/mL) in HBSS-Ca were co-incubated with increasing amounts of bacteria at dilutions of 1-10 µL of bacterial suspension/ 200 µL HBSS-Ca. The final optical densities at a dilution of 3 µL/ 200 µL HBSS-Ca were OD_{600}=0.350 for all bacterial species tested. Bacterial numbers were estimated using the OD_{600} as previously described (*59*) and expressed as bacterial cells per neutrophil. After incubation at 37°C, samples were immediately lysed in SDS sample buffer. Similarly, bacterial cells alone were incubated in parallel and lysed in sample buffer. For inhibition experiments, neutrophils were incubated with *Aa* HK1651 cells in the presence or absence of a blocking antibody directed against the C-terminal domain of LtxA (DAKO) or with antibody alone (*53*).

SDS-PAGE and immunoblotting

GCF was collected using absorbent filter paper strips, and proteins were recovered by washing strips in SDS sample buffer containing β-mercaptoethanol. Leukocyte lysates and GCF samples were analyzed by SDS-PAGE. Immunoblotting was performed with antibodies to histone H3 (Abcam), citrullinated histone H3 (Abcam), myeloperoxidase (MPO) (R&D),

proteinase 3 (PR3) (Santa Cruz), or peptidylarginine deiminase type 4 (PAD4) (Sigma). Total protein citrullination was detected by anti-modified citrulline (AMC) immunoblotting as previously described (*60*). SuperSignal West Pico Chemiluminescent substrate (Thermo Fisher Scientific) was used for chemiluminescence detection. To ensure linear detection of total protein citrullination, the chemiluminescence signal was quantified electronically using a chemiluminescent imager (ProteinSimple).

Immunofluorescence staining

Neutrophils ($5x10^5$ cells/50 µL serum-free HBSS, HEPES 10 mM, $CaCl_2$ 1.5 mM) were plated onto standard microscope slides coated with poly-D-lysine (Sigma), and incubated for 30 min at 37°C to allow for cell attachment. Neutrophils were then incubated in the absence or presence of purified LtxA at 0.3 µg/mL for 1-3 h at 37°C and fixed in 4% paraformaldehyde (PFA). Fixed cells were carefully washed in PBS and permeabilized in ice-cold acetone. Before antibody probing, cells were blocked in 5% normal goat serum (GS) and incubated with rabbit anti-histone H3 (citrulline R2 + R8 + R17) (Abcam) or with human serum in PBS 5% GS for 2 h at RT. Secondary antibodies against human (conjugated with Alexa Fluor 488) and rabbit IgG (conjugated with Alexa Fluor 495) (Life Technologies) were used to visualize antibody binding. Cells were mounted with ProLong Gold (Molecular Probes) plus DAPI. To visualize extracellular DNA, PMNs were fixed in 4% PFA without permeabilization and stained with SYTOX Green (Life Technologies).

Mass spectrometric analysis

LtxA-treated human primary neutrophils were analyzed by AMC immunoblotting to confirm hypercitrullination. LtxA-treated neutrophils and untreated controls incubated in parallel were lysed in 1% SDS, 2 mM EDTA, and 1 mM PMSF, sonicated on ice for 20 cycles of 5 seconds, and centrifuged for 20 minutes at 20,000 x *g*. Sample supernatants were recovered, and proteins were precipitated in TCA and analyzed by liquid chromatography–mass spectrometry (LC-MS/MS) (Mass Spectrometry and Proteomics Facility, Johns Hopkins University School of Medicine) to identify deiminated residues (Proteome Discoverer). GCF protein reduced in SDS sample buffer was partially resolved by 1D SDS-PAGE, fixed in methanol and acetic acid, and subjected to in-gel digestion for LC-MS/MS characterization. GCF samples from individuals with periodontitis (n=4) and healthy controls without periodontitis (n=4) were processed in parallel and subjected to MS independently. Tryptic peptides were basic reverse phase (bRP) fractionated and analyzed by nanoLCMSMS-FTFT using reverse-phase chromatography on a nanoAcquity (Waters) interfaced with QExactive Plus (Thermo). Precursor isolation and higher-energy collisional dissociation (HCD) fragmentation were performed at 70000 and 35000 resolutions, respectively. Spectra were searched against a custom bacteria and human database, using Mascot (v2.5.1) through Proteome Discoverer (v1.4) applying MS2Processor. Mass spectra were analyzed by Scaffold PTM to confirm citrullination sites, and localization probabilities were calculated using the Ascore algorithm as previously described (*49*). Localization probabilities ≥98% were considered significant. The data were used to define the proteome and citrullinome contained in each sample. The information was then

aggregated to present the general spectra of citrullinated proteins and citrullinated autoantigens detected in GCF.

Hierarchical clustering of citrullinated proteins detected by MS in GCF samples from individuals with periodontitis and oral health as compared to RA SF (published datasets) (*2, 8-10*) was performed in Cluster 3.0 and visualized using Java TreeView. Values were normalized by protein and subjected to unsupervised hierarchical clustering (centroid linkage) to define patterns of protein modification (citrullinated vs. native expression) in individual compartments. Protein expression, modification status, and autoantigen classification were considered. GCF samples were also searched for peptides unique to periodontitis-associated bacterial species and commensal controls to define the microbial composition of the periodontal pocket in patients with periodontitis and controls without periodontitis.

Anti-LtxA, anti-*Aa* serotype b, and anti-*P gingivalis* antibody ELISA

RA patients and controls (healthy individuals without clinical evidence of periodontal disease) were assayed for IgG serum antibodies to purified LtxA. Polystyrene plates were coated with 100 ng/well purified LtxA in PBS, pH 7.4, or PBS alone. Plates were blocked in PBS 0.1% Tween, 3% non-fat dry milk (PBS–TM). Sera were diluted at 1:1000 in PBS–TM 1% and assayed in duplicate. HRP-conjugated anti-human IgG was used as a secondary antibody (diluted at 1:20,000 in PBS–TM 1%). Antibody concentrations were calculated in reference to wells coated with a serial dilution of purified human IgG (Athens Research) and a serial dilution of high titer patient serum that served as standards. *Aa* strain HK1651 (serotype b) was propagated until smooth-colony, non-adherent cultures were obtained. Of note, the ability of the fixed-cell assay

to discriminate between periodontitis and periodontally healthy controls was lost when adherent cultures (rough phenotype) were used. *P gingivalis* W83 was grown as outlined above. Bacterial cells were pelleted by centrifugation, washed in PBS 1x, pH 7.4, and adjusted to OD_{600}=0.350. Bacteria were fixed on ice in 4% PFA and diluted at 1:100 in PBS 1x, pH 7.4 to coat ELISA plates.

DNA isolation and PCR for *LTXA* gene detection

Microbial DNA was isolated from subgingival plaque samples collected from patients with periodontitis and healthy controls (no periodontal disease) using the MasterPure DNA Extraction Kit (Epicentre Biotechnologies) as previously described (*61*). Polymerase chain reaction was conducted to amplify the *LTXA* gene using 0.4 µM of primer (forward ltx3: 5'-GCCGACACCAAAGACAAAGTCT-3' and reverse ltx4 5'-GCCCATAACCAAGCCACATAC-3') as previously described (*62*). DNA isolated from *Aa* HK1651 and *Aa* SUNY ab75 cultures was used as a reference for JP2 and non-JP2 clones, respectively.

Cloning, expression, and immunoprecipitation (IP) of radiolabeled full-length and truncated LtxA.

LtxA-encoding DNA was amplified from *Aa* SUNY ab75 and cloned into pcDNA3.1 (+). Truncated forms of LtxA were generated by PCR and cloned into pcDNA3.1 (+). ^{35}S-methionine-labeled proteins were generated by in vitro coupled transcription/translation (IVTT) using the TnT Quick system (Promega). ^{35}S-methionine-labeled LtxA products were diluted in IP buffer (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40 Alternative) and incubated with

patient or control sera. Immune complexes were purified using Protein A agarose beads (Pierce). Samples were washed in IP buffer, and immunoprecipitates visualized by SDS-PAGE/ sodium salicylate fluorography.

Supplementary Figures

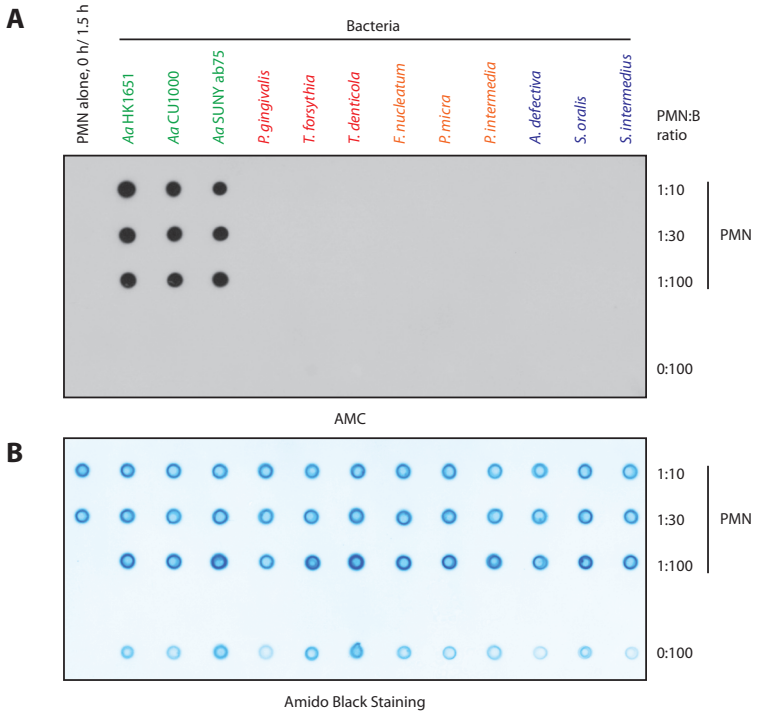

Fig. S1. Protein citrullination is only detected in neutrophils co-incubated with *Aa* strains. **(A)** Dot blot analysis of neutrophils incubated alone (PMN alone, 0 h/ 1.5 h), of neutrophils incubated with increasing amounts of bacteria (10-100 bacterial cells per neutrophil), and of bacteria incubated alone to detect protein citrullination (AMC). Citrullination was only observed in *Aa*-exposed neutrophils (PMN), but not in neutrophils incubated with other bacterial species or in bacteria alone. (**B**) Amido black staining is shown to visualize protein loading. The experiments were performed on two separate occasions, with similar results.

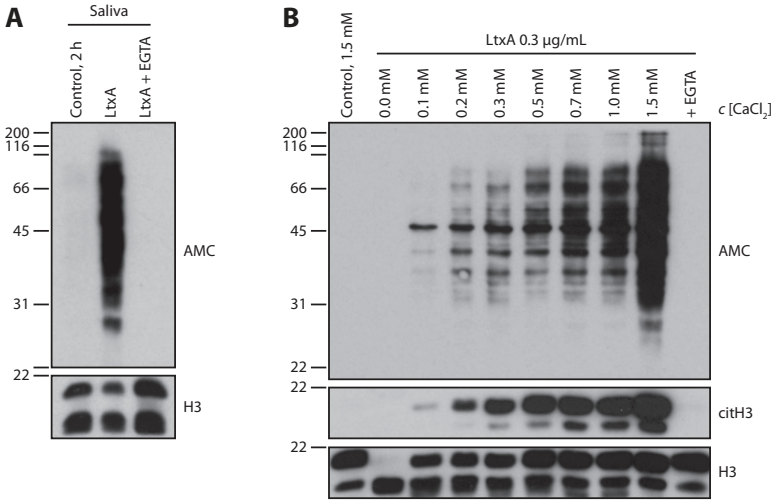

Fig. S2. LtxA induces neutrophil hypercitrullination in human saliva and across a range of calcium concentrations. (A) Neutrophils were incubated in human saliva (cleared by centrifugation) in the presence or absence of 0.3 µg/mL purified LtxA ± 5 mM EGTA. Total citrullinated protein (AMC) and histone H3 (H3) (loading control) were detected by immunoblotting. **(B)** Human neutrophils were incubated with purified LtxA at increasing concentrations of calcium (0.0-1.5 mM) or 5 mM EGTA (last lane). As control, neutrophils were incubated with 1.5 mM calcium in the absence of LtxA. Total citrullinated protein (AMC), citrullinated H3 (citH3), and histone H3 (H3) (loading control) were detected by immunoblotting. The experiments were performed on three (**A**) and two (**B**) separate occasions, with similar results.

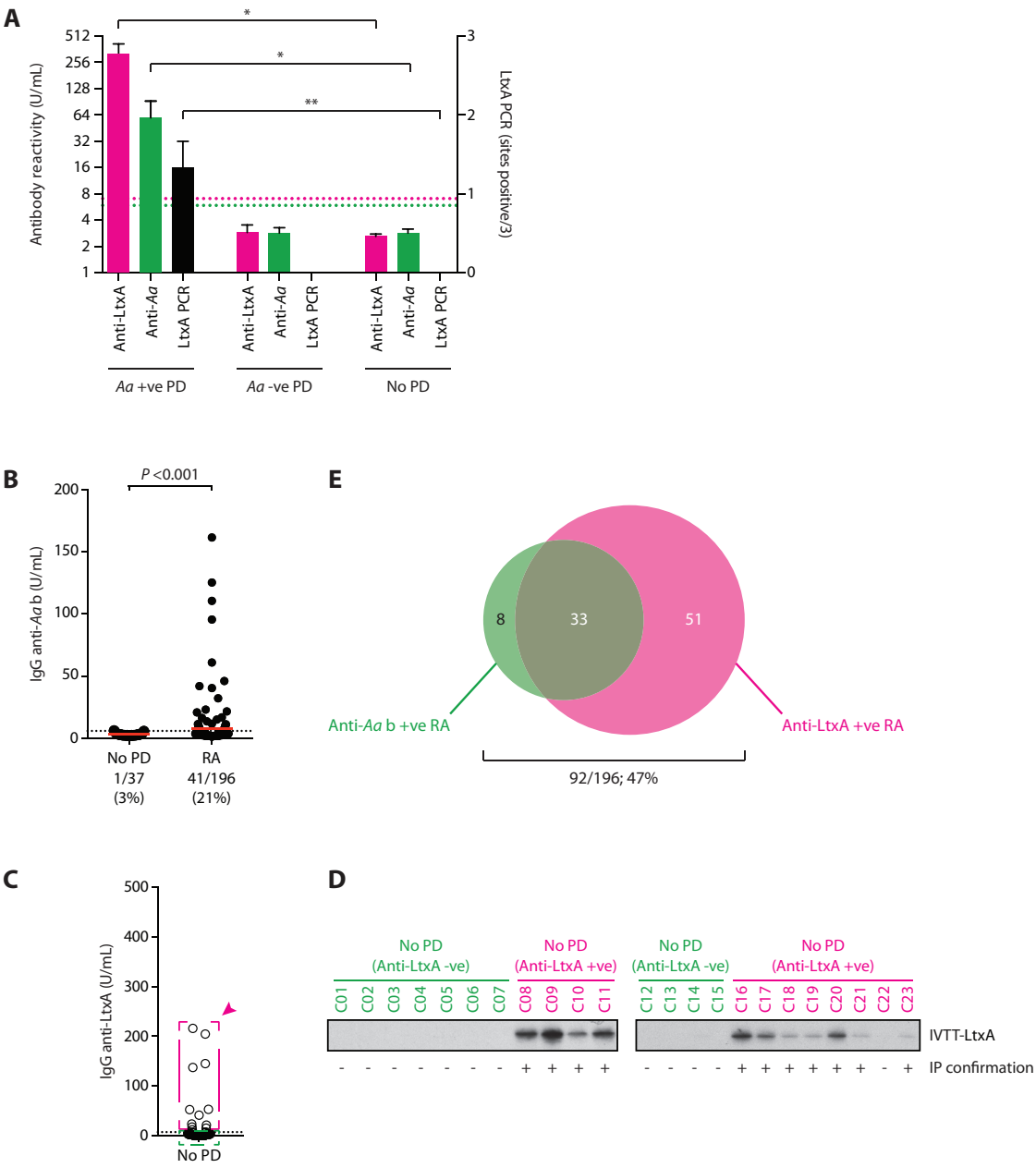

Fig. S3. Antibodies against LtxA and *Aa* serotype b identify patients with active periodontitis associated with leukotoxic strains of *Aa*. (A) Serum antibodies to LtxA and *Aa* (serotype b) were assessed by ELISA, and subgingival plaque *Aa* DNA was amplified by PCR in 6 patients with periodontitis (PD) and 8 controls without periodontitis (No PD). Antibodies to *Aa* and LtxA were only positive in a subgroup of periodontitis patients who were also positive for *Aa* DNA in subgingival plaque (*Aa*+ve PD group, $n = 4$). Mean values ±SEM are shown (Kruskal–Wallis test; *p=0.02, **p=0.006). The pink and green dotted lines mark the cut-offs for positivity as determined by ELISA for antibodies against LtxA and *Aa*, respectively. **(B)** Antibodies against

PFA-fixed Aa serotype b in patients with RA (n = 196) and healthy controls without periodontitis (No PD; n = 37) as detected by ELISA. Red line indicates mean IgG concentration. The dotted line marks the 95th percentile of controls (cut-off for positivity) (Mann-Whitney test). **(C)** Reactivity of No PD control sera against purified LtxA by ELISA. Twelve No PD controls (pink box) demonstrated antibody reactivity. These sera were further interrogated in **D** by IVTT-immunoprecipitation (IP) to confirm antibody binding. **(D)** Reactivity against LtxA in No PD control sera positive for anti-LtxA antibodies by ELISA (n = 12) was determined by IP using radiolabelled IVTT-LtxA. Anti-LtxA antibody negative individuals without periodontitis (No PD; n = 11) were used as negative controls. 11/12 of No PD controls reactive by ELISA were also positive for anti-LtxA antibodies by IP, confirming that these healthy individuals were indeed exposed to leukotoxic strains of *Aa*. One sample with low-titer anti-LtxA antibodies by ELISA (15.1 U/mL) was negative by immunoprecipitation. **(E)** Venn diagram of anti-*Aa* serotype b and anti-LtxA antibody positive RA as determined by ELISA.

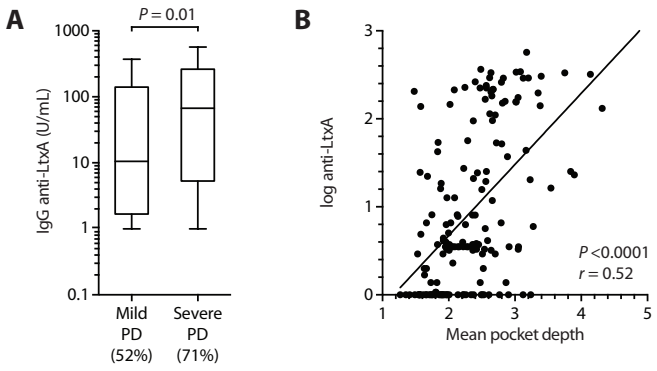

Fig. S4. Anti-LtxA antibody concentrations correlate with periodontitis severity. (A) Serum antibodies to LtxA in patients with mild versus severe periodontitis (PD). Box and whiskers plot (box indicates the median, 25th and 75th percentiles; whiskers mark the minimum to maximum values); *P* value calculated using Wilcoxon Rank-Sum test. **(B)** Scatter plot of log anti-LtxA antibodies versus mean pocket depth in patients with periodontitis ($\beta = 1.12$; $P < 0.0001$).

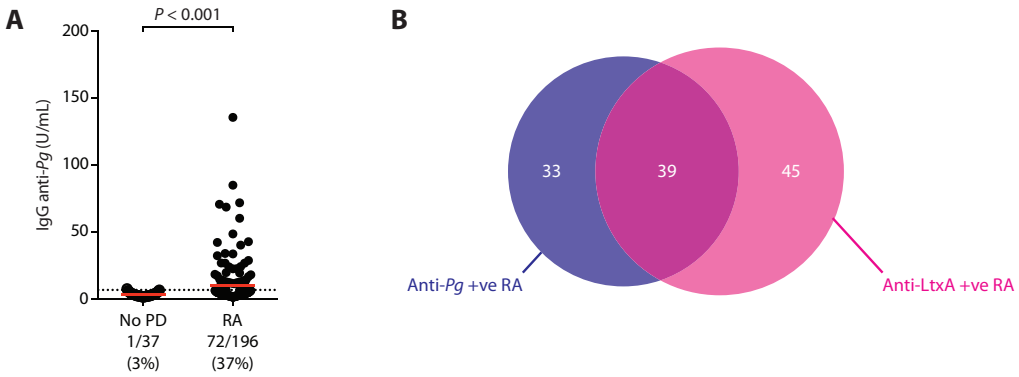

Fig. S5. Antibodies to *P. ginvivalis* (*Pg*) are enriched in a subset of RA that largely differs from patients exposed to LtxA. **(A)** Antibodies against PFA-fixed *Pg* in patients with RA (n = 196) and healthy controls without periodontitis (No PD; n = 37) as detected by ELISA. Red line indicates mean IgG concentration. The dotted line marks the 95th percentile of controls (cut-off for positivity) (Mann-Whitney test). **(B)** Venn diagram of anti-LtxA and anti-*Pg* antibody positive RA as determined by ELISA. Anti-*Pg* and anti-LtxA antibodies overlapped in 39 patients with RA (39/196; 20%).

Table S1. Specific bacterial peptides identified in GCF.

S1A - *Aggregatibacter actinomycetemcomitans*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
Branched-chain amino acid transport system II carrier protein	WP_005589245.1.	LLQVTIPALLLIYPVAIMLVVLQMVRTK	99.50%	96.10%	1
Cell division protein FtsX	WP_025298256.1.	KFATLLTVLVIAVSFTIPTVSYLLR	99.40%	97.20%	1
Iron ABC transporter permease	WP_005578275.1.	ASSFQVLRFILLPLLKLALLSALVTGFVR	99.60%	98.10%	1
Iron-dicitrate transporter subunit FecD	WP_053330367.1.	LVGGRHRTLLPAALLIGALLLQISDILAR	99.60%	98.00%	2
Membrane protein	WP_005565029.1.	VSPVFAGGVEYAILPELALR	99.70%	99.80%	6
Membrane protein	WP_005557524.1.	QFITELQEK	99.40%	97.40%	1
Molybdate ABC transporter permease [hypothetical protein]	WP_005578990.1.	KLLIAMLLPPFNVLILWLLSLFFSLIR	99.70%	98.60%	1
Na(+)-translocating NADH-quinone reductase subunit E [hypothetical protein]	WP_005542966.1.	MKLFLLTLGIFLLIILAMALGYIVKR	99.20%	97.20%	1
Protease modulator HflC	WP_005576047.1.	MRKLLLPVILVIVAIVYSSIVVVTEGTR	99.60%	98.10%	1
Sel1 domain protein repeat-containing protein [hypothetical protein]	WP_014702465.1.	MKKCLKILLVLVLVAIIGLVWLHR	99.30%	97.70%	1
Tyrosine-specific transport protein 1	[ACX82440.1]	VLGGNLTLAVALVLGVIITSHPLRHPR	99.50%	97.70%	1
UDP-phosphate N-acetylglucosaminyl 1-phosphate transferase	WP_005538677.1.	LLTLIVTFWGAFLTLLVMRPVAIR	99.30%	97.70%	2
XylR family transcriptional regulator	WP_005594847.1.	LLHRLLTGQKVATTTPLLIPPVKVIER	99.50%	97.60%	1

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1B - *Porphyromonas gingivalis*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
Calcineurin-like phosphoesterase	WP_005873839.1.	RGLMISLLLLVALGIWIAKRYR	98.70%	100.00%	2
Hydrolase Nlp/P60	WP_053444281.1.	TVKRAVRIALLTLIGILFSSPSLVR	99.70%	98.60%	1
Hypothetical protein	WP_012457215.1.	KQVIVILLLVSLAYLGSVQVARFLYGR	99.00%	96.50%	1
Hypothetical protein HMPREF1555_01833	[ERJ64501.1]	MQLIKVFLVQLLLLPIFFYKR	99.30%	97.70%	1
Hypothetical protein HMPREF1989_00976	[ERJ86808.1]	RGKVSLENLLPR	99.40%	97.60%	1
Hypothetical protein HMPREF1990_02184	[ERJ85621.1]	TGSKIIIILFILIRNILLLYCLSVR	99.70%	98.60%	1
Lactate permease	WP_054191174.1.	LTAKKAADTSVALEK	98.80%	95.40%	2
Sugar isomerase	WP_005875317.1.	EYKNVLIVKLLLAIKIIAQIIALR	99.70%	98.60%	1
TonB-dependent receptor	WP_021677449.1.	MKRLISRTLIWLLLILCIPISLAAQR	99.50%	97.60%	1
TonB-dependent receptor	WP_058018781.1.	RKRILQLFLTTLLLALGSSLAIAQTVVTGK	99.10%	95.30%	1
Transporter, small conductance mechanosensitive ion channel MscS family	[EOA11822.1]	VVIALVAFYLGKLLLNALVKWLDRIMVR	99.30%	97.40%	1
UDP-N-acetylglucosamine 1-carboxyvinyltransferase	WP_061156788.1.	GSVLLVGPLISR	99.70%	98.40%	2

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1C - *Tannerella forsythia*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
AraC family transcriptional regulator	WP_060828652.1; WP_046825421.1; WP_014225969.1.	TEKRVILIAGIALVAVLLLVMAVIIYMGR	99.00%	99.80%	2
Bacterial group 2 Ig-like protein [hypothetical protein]	WP_060831565.1; WP_060828264.1; WP_014225980.1.	MKRLILLTLLTLIGWIGATNTLR	98.90%	96.10%	1
Collagen-binding protein	WP_052449125.1; WP_014225582.1.	VYWQHSPYNTPTVR	99.70%	98.90%	2
DNA starvation/stationary phase protection protein	WP_014225577.1; WP_046826198.1.	IAMLGGHPENR	99.70%	98.60%	2
DNA-directed RNA polymerase subunit alpha	WP_014224923.1.	IEGVDHEFATIPGVMEDVTNIILNLK	99.70%	98.80%	2
Ferritin	WP_060828637.1; WP_014225642.1.	IVTGLINDLYTLALQESDYATQSMLK	99.70%	98.80%	2
Flavodoxin	WP_014225386.1.	IAEQMNIASEDVYEVAK	99.70%	99.00%	4
		TAIIYGSTTGTTENIAGR	99.70%	100.00%	11
Forsythia cell detaching factor protein [hypothetical protein]	WP_060830860.1; WP_046825757.1; WP_060827711.1; WP_014224776.1.	LTSSVKPIVDTQK	99.70%	98.40%	1
Fructose-bisphosphate aldolase	WP_041590974.1.	LVPPPLAFDILEAVEK	99.70%	100.00%	23
Glycosyl hydrolase	WP_046824980.1; WP_014225135.1.	TQPSYISAPEYQILDNANHPDAK	99.70%	99.40%	5
Hypothetical protein	WP_014225175.1.	ITEQDQQLHR	99.70%	98.90%	2

Hypothetical protein	WP_014225890.1.	APVMGSIYTDVTSGTAVTSNSLASK	99.70%	98.80%	2
Hypothetical protein	WP_014226123.1.	IIYMDSNNSTK	99.70%	98.40%	1
Hypothetical protein	WP_041590592.1.	EMANLTPEQSQK	99.70%	100.00%	6
		VAAGDAEATQK	99.70%	100.00%	11
		VQDLAADLQK	99.50%	100.00%	3
Hypothetical protein	WP_041591050.1.	AFADAITDLLDGDLLVEKF	99.70%	98.80%	2
Hypothetical protein	WP_046825066.1; WP_060827931.1; WP_014225330.1.	LSPYIGGDIGLAYQTQSSSSKPK	99.70%	98.50%	1
Hypothetical protein	WP_046826079.1; WP_014224394.1.	IPEAANYHYIR	99.70%	98.40%	1
Hypothetical protein	WP_060830884.1.	VNKPTNAGVLNIK	99.70%	100.00%	7
		HADNIYLYQDANHPHPFASAFGGLGEFAHTGITK	99.70%	100.00%	7
Hypothetical protein	WP_060831299.1.	MTHEVANVDMAEMRR	99.50%	99.70%	4
Hypothetical protein	WP_060831356.1; WP_014225605.1.	LNNFLTLTDEQK	98.30%	99.50%	2
Hypothetical protein	WP_060831494.1; WP_046824471.1.	SLIILRNIDSEGIILTLVKKNR	99.00%	96.70%	1
Hypothetical protein	WP_060831874.1.	ASTDPSGIDAEASLKK	99.30%	96.30%	1
		VPAGVYEASATDTR	99.50%	97.40%	1
Hypothetical protein TF3313_1200/ hypothetical protein BFO_1336	[BAR48735.1; AEW20591.1]	VNNAFGAGESSYITPGNHAEQK	99.70%	100.00%	7
Hypothetical protein TF3313_2048/ hypothetical protein TF3313_2046	[BAR49513.1; BAR49511.1]	TYKILLILSLILGVMQLVQAQHILK	99.30%	97.00%	1
Integration host factor subunit beta	WP_041590922.1.	NAVLASVESFMSVVK	99.70%	100.00%	6

Iron ABC transporter	WP_046472598.1; WP_037983921.1.	QKLLLLVPGSKPRELIVPREPLK	98.00%	98.40%	2
Lipid A biosynthesis (KDO)2-(lauroyl)-lipid IVA acyltransferase	WP_060827663.1; WP_014224679.1.	ITFSVVNVR	99.60%	97.90%	1
Malate dehydrogenase	WP_014223968.1	GNAEVAEQLGK	99.60%	100.00%	4
Membrane protein	WP_041590550.1.	EFEVLEPMDYDLLNDLNSQINALR	99.70%	100.00%	16
		FIPFVGLGYALRPGYSDK	99.70%	100.00%	27
		IAPSGAISIGK	99.70%	100.00%	23
		IDKHQEVSIFNTAEYAK	99.70%	100.00%	20
		KEFEVLEPMDYDLLNDLNSQINALR	99.50%	100.00%	10
		VVIMNTDDK	99.60%	100.00%	20
		WYNPYMAFR	99.70%	100.00%	13
MFS transporter	WP_046824671.1; WP_060827113.1; WP_014223608.1.	DFLALAADVIK	99.70%	100.00%	14
		EYEDIAER	98.80%	96.00%	1
		FYQEALKR	99.00%	100.00%	5
		GSASVDEHSTFGQGGGNVTK	99.70%	100.00%	9
Molecular chaperone GroEL	WP_046825748.1; WP_060827723.1; WP_014224796.1.	DLLPILEQVVQSGR	99.70%	100.00%	33
		ISANGDEGIGK	99.70%	100.00%	16
		VTVNKDNTTIVK	99.70%	100.00%	9
		VVESIASQSEAVGTNMDR	99.70%	100.00%	16
Peptidase M16	WP_014225081.1.	MIVPNIPVEAINQIMQQIITDK	99.70%	100.00%	9
Peptidase M16 inactive domain protein [hypothetical protein]	WP_046824956.1; WP_060831131.1; WP_060827826.1.	GSITNETTDAER	99.70%	100.00%	8

Peptidylprolyl isomerase	WP_041591316.1.	VITEGTGAKPTAEDQVK	99.70%	99.00%	5
Phosphoglycerate kinase	WP_014223696.1.	GLADDASEQEK	99.70%	99.00%	2
Phosphopyruvate hydratase	WP_046825706.1; WP_014225384.1.	GLSTAVGDEGGFAPALNGTEDALDSIIQAIK	99.70%	100.00%	8
Putative internalin-J	WP_014223509.1.	VIDVRQNLK	99.70%	97.80%	1
Putative lipoprotein [hypothetical protein]	WP_014225175.1.	AQSTTELNEMLSTLNDIQTDIQAIR	99.70%	100.00%	6
		DGDGNLVLNILDPQVFWSLSK	99.70%	100.00%	6
		IQELDELVTSLNENVENLSVTTAAQSQK	99.40%	100.00%	3
Putative lipoprotein [hypothetical protein]	WP_046825761.1; WP_014224769.1.	IAPGFFQISDFLGGYYAQR	99.70%	98.90%	2
RNA polymerase subunit sigma-54	WP_014225268.1.	VNDTFEEAVDDAVVALEK	99.70%	99.00%	2
S26 family signal peptidase	WP_060827209.1.	SQLCKIVNIALNVFNGISIIIIGILLLR	98.00%	99.70%	2
Single-stranded DNA-binding protein	WP_060831578.1; WP_014226022.1.	LANGTEIPER	99.70%	98.60%	2
Sodium bile acid symporter family protein	WP_060831660.1.	KVFPLLILPLFVALAVRKWTPR	99.70%	98.60%	1
Surface layer protein A [hypothetical protein]	WP_041590756.1; WP_046824918.1; WP_060827793.1.	DLPLVGGNLGIWK	99.70%	100.00%	112
		GAYTDWIAEAETWNVK	99.70%	100.00%	54
		IEYVQIYNDYELFK	99.70%	100.00%	75
		LDENEVANMTPSYQWYVEQVNEGSSISK	99.70%	100.00%	85
		LVVAHKDEFDVNADFYKK	99.30%	100.00%	21
		MQDLSAGER	99.60%	100.00%	36
		TTVEGDLYTIK	99.60%	100.00%	30
		VPGFAFTNK	99.70%	100.00%	99
		VSAGQVNPFEQPLMAEDVTGTTGDK	99.70%	100.00%	158
		VSAGQVNPFEQPLMAEDVTGTTGDKQYLNLK	99.70%	100.00%	93

		VTLQTATGPVDK	99.70%	100.00%	178
		YLVVPMYAGNFTPQWLR	99.70%	100.00%	131
		YVTALAGNTEENYVNDLGYR	99.70%	100.00%	178
Surface layer protein A [hypothetical protein]	WP_041590757.1; WP_046824918.1; WP_060827793.1.	EADGTETEYVNYK	99.70%	100.00%	36
		EADGTETEYVNYKK	99.70%	100.00%	86
		EYLYEDAHSIYSYDIK	99.70%	100.00%	30
		FTEVGGGSNKDFMIESETTR	99.70%	100.00%	129
		FTEVGGGSNKDFMIESETTRR	99.70%	100.00%	43
		FTLVNARPR	99.70%	100.00%	73
		GAYTDWIAEAETWNVK	99.70%	100.00%	25
		GTGHIKPQYLLAVDPK	99.70%	100.00%	142
		HKDEVFSFR	99.70%	100.00%	17
		KEADGTETEYVNYK	99.70%	100.00%	49
		KEADGTETEYVNYKK	99.70%	100.00%	49
		KGTDDYIAPLER	99.70%	100.00%	36
		KKEADGTETEYVNYKK	99.30%	100.00%	50
		LAGHPDGTIVDKDYIWDAK	99.70%	100.00%	178
		MIAPMNGGWVK	99.00%	100.00%	50
		NEYYLVLDDNGR	99.70%	100.00%	73
		NHSSVNFLGVQNEDDAIANK	99.70%	100.00%	36
		NHSSVNFLGVQNEDDAIANKDKDQVFK	99.70%	100.00%	68
		NHSSVNFLGVQNEDDAIANKDKDQVFKR	99.70%	100.00%	30
		NNSFGVVHASVDDLNNYVR	99.70%	100.00%	142
		RNEYYLVLDDNGR	99.70%	100.00%	49
		YGIGWTSNSPLK	99.70%	100.00%	142

		YGYTPEVVANAR	99.70%	100.00%	93
Surface layer protein B [hypothetical protein]	WP_041590758.1.	EHLGAQLTSLIMGNLK	99.70%	100.00%	210
		EHLGAQLTSLIMGNLKK	99.70%	100.00%	55
Surface layer protein B [hypothetical protein]	WP_060827794.1.	AASGNIIPSSTGAAK	99.70%	100.00%	17
		ADAGTPNENDYLWGTAWNR	99.70%	100.00%	108
		AGAESPSVWK	99.70%	100.00%	136
		AGEFNPGVFVAAVEEQAPMK	99.70%	100.00%	235
Surface layer protein B [hypothetical protein]	WP_060827794.1; WP_046824919.1; WP_041590758.1.	AWDVAPISQVGTAGNPWFSVDVR	99.70%	100.00%	217
		EAMPFALDKK	99.70%	100.00%	140
		EGGDFLIESETQDR	99.70%	100.00%	42
		FNQNATFPSDGSPMIGQFVWR	99.70%	100.00%	161
		FNTELEGDKPGDTPETVK	99.70%	100.00%	235
		FSPNGTPNPFTDQK	99.70%	100.00%	178
		IGAVTPGAGAPIDYSYYHR	99.70%	100.00%	193
		INGMLLFTLK	99.70%	100.00%	217
		KAGEFNPGVFVAAVEEQAPMK	99.20%	100.00%	42
		KIALNNNFHK	99.70%	100.00%	80
		KIGAVTPGAGAPIDYSYYHR	99.70%	100.00%	80
		KLDAASDSSPAAPK	99.70%	100.00%	122
		KNYLGLVSINSNPNVK	99.70%	100.00%	80
		LDAASDSSPAAPK	99.70%	100.00%	98
		LVFTDAIHANDALYILGHADLK	99.70%	100.00%	160
		LYVSLQNLTVDGKPGTR	99.70%	100.00%	137
		NAEASPINIINR	99.70%	100.00%	235
		NQVADLVQLK	99.70%	100.00%	235

		NVHPEELPSFQWIVEK	99.70%	100.00%	235
		NYLFENTNALK	99.60%	100.00%	108
		NYLFENTNALKDQQAYYR	99.70%	100.00%	137
		NYLGLVSINSNPNVK	99.70%	100.00%	137
		QIVNSEVIDFEK	99.70%	100.00%	84
		RFNTELEGDKPGDTPETVK	99.70%	100.00%	99
		SDVLQAIADAEIFNVEK	99.70%	100.00%	235
		SFETVLGTQASQEAR	99.70%	100.00%	235
		SSSNPNAQTPGSISAWEFSETYK	99.70%	100.00%	235
		SWYTFTAYPTPK	99.70%	100.00%	137
		VKPYDNAKPTLK	99.70%	100.00%	132
		VSIDPDLYLIR	99.70%	100.00%	235
		YNEVIGTPLFFLR	99.70%	100.00%	217
		YNQNGLSHPALQVPVDER	99.70%	99.00%	8
		YYQGQNPTFDFTNK	99.70%	100.00%	187
SusC/RagA family TonB-linked outer membrane protein	WP_041590933.1; WP_046826132.1.	HWQKPGDNAEYAR	99.70%	100.00%	4
		TQEAVAQPDMK	98.20%	100.00%	2
		LNVLLGHSFQR	99.70%	100.00%	8
		NANMQNHQISVTSGGETTK	99.70%	100.00%	4
SusC/RagA family TonB-linked outer membrane protein	WP_046825525.1; WP_060830558.1; WP_060827077.1; WP_014223521.1.	HNISVEAGSDK	99.70%	98.40%	1
SusC/RagA family TonB-linked outer membrane protein	WP_046825759.1; WP_014224771.1.	AVPSGQGTSLHR	99.70%	100.00%	12
		DPSSGYTSQISNLGDVR	99.70%	100.00%	23
		FTPGAPNDGSSTANAGR	99.70%	100.00%	23

		IIENEQQVKPYVFLK	99.70%	100.00%	8
		VAGVDISTAPGPGATQNVIIR	99.70%	100.00%	12
		VQINHDAMATYK	99.70%	100.00%	8
		YVAATIQNPGYPK	99.70%	100.00%	19
SusC/RagA family TonB-linked outer membrane protein	WP_060830766.1.	AAGVVIGSTGSPGSK	98.80%	100.00%	6
		ASPWVPVYDIQGHFAGSK	99.70%	100.00%	6
		ATNFISTYSTNPSR	99.70%	100.00%	6
SusC/RagA family TonB-linked outer membrane protein	WP_060830766.1; WP_046825920.1; WP_041590573.1.	TPSASATQQLQGR	99.70%	100.00%	6
	WP_060830987.1.	ADLDALDPEVIDAFSVLK	99.60%	100.00%	5
SusC/RagA family TonB-linked outer membrane protein	WP_060830987.1; WP_060827529.1.	NQVLEYDEAPGLRPALR	99.70%	100.00%	5
		NVTDVLDIIFSGTGLAYR	99.70%	100.00%	8
		QVSESGSPSKPNLNAGSGVSGDR	99.70%	100.00%	15
		SLKPIVSNFK	99.50%	100.00%	5
SusC/RagA family TonB-linked outer membrane protein	WP_060830987.1; WP_060827529.1; WP_014224392.1.	ILSIPLSEDAQALGEVVVSAFNTGQK	99.70%	100.00%	8
SusC/RagA family TonB-linked outer membrane protein	WP_060831252.1; WP_046824538.1; WP_014225420.1; WP_060827965.1.	ATPVVDDNNTYQPYSK	99.70%	99.80%	3
		FGFINSLR	99.00%	100.00%	8
SusC/RagA family TonB-linked outer membrane protein	WP_060831791.1; WP_060828570.1; WP_052449072.1; WP_014224606.1.	IIDEGIVNSGNNTLDDHGDLKR	99.70%	98.60%	1
Susd and RagB outer membrane lipoprotein domain protein	WP_060828127.1; WP_060831380.1.	IFVLPFHTDNPYKK	99.70%	98.90%	2
SusD family protein [hypothetical protein]	WP_060830986.1; WP_014224393.1.	GAANPEVLFADTR	99.70%	98.40%	1
		LVYPWPVIR	99.30%	96.60%	1
T9SS C-terminal target domain-containing protein	WP_041590835.1.	LGANSAAENQGLNR	99.60%	97.50%	1

Tetratricopeptide repeat protein [hypothetical protein]	WP_060830863.1; WP_046825760.1; WP_014224770.1.	VWWNVKPEIVVK	99.70%	98.40%	1
TonB-dependent receptor	WP_060831832.1; WP_046825715.1; WP_060828609.1; WP_014225392.1.	IAGEGALDNVDFSR	99.70%	100.00%	3
		KANRPFVGTVSTR	99.20%	100.00%	3
Type I glyceraldehyde-3-phosphate dehydrogenase	WP_046825854.1; WP_014223728.1.	VPTLDVSVVDLTANLEKPATK	99.70%	100.00%	9

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1D - *Treponema denticola*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
AI-2E family transporter	WP_002681303.1; WP_002676680.1; WP_002672876.1; WP_010688795.1; WP_002687333.1; WP_002684957.1; WP_010698459.1; WP_002692331.1.	LFLPYASVLLWSAVIYVLVRPLYNKILSR	99.70%	98.80%	1
Cobalt transporter	WP_002668171.1.	IKNLKSVLGVLPFVIMVLLPFVIR	99.00%	95.40%	1
Elongation factor Tu	WP_002689871.1.	VDLVDDPELVELVEEEVRETLK	99.70%	100.00%	7
Flagellar filament core protein	WP_010697276.1; WP_002678135.1.	NQILSNTGVAMLAQANNNSQLVMSLLR	99.70%	98.40%	3
Flagellar filament outer layer protein	[AFD04804.1]	EMIWNNPSYVSHVK	99.60%	100.00%	4
		MAQVDDQGNIQDPTDEDKASGK	99.70%	100.00%	7
Flagellar protein	WP_002670927.1; WP_002669397.1.	MAQVDDQGNLQEPTAEDK	99.70%	100.00%	13
		MATEEGFTPSEGSEK	99.70%	100.00%	5
Flagellin	WP_002668952.1.	NQILSNTGIAMLAQANNNSQLVMSLLR	99.70%	100.00%	6
Flagellin	WP_002670454.1	GLNQASANAQNGISFIQVAEAFLQETTDVIQR	99.70%	100.00%	12
Hypothetical protein	WP_002676180.1.	KDKFPLLAK	99.50%	97.90%	2
Hypothetical protein	WP_002677442.1.	VSFELLYSAPNHTER	99.70%	98.50%	1
Hypothetical protein	WP_002678577.1.	MKKNIILLSIFIFSLLVLYAKR	99.70%	98.60%	3
Hypothetical protein	WP_002680619.1; WP_002676079.1;	IKIALGISIILIFFLLGLVIGKYKER	99.70%	98.60%	1

	WP_002673630.1; WP_002684338.1; WP_010698124.1; WP_010689495.1.				
Hypothetical protein	WP_002686836.1.	TWRIIALASLSFLFLSLLLLLYAIKLPK	99.60%	98.10%	1
Hypothetical protein	WP_002687397.1; WP_010698426.1; WP_002692266.1.	TIQARPLVETLVLIAGLLLINILAR	99.10%	96.80%	1
Hypothetical protein	WP_002689432.1; WP_010697066.1.	FIILLVKNLIAKILNKNLNIR	99.70%	98.60%	1
Hypothetical protein	WP_002693057.1.	NKLMIVFGLLILAVGFTLALLAVRTAR	99.20%	97.10%	1
Hypothetical protein	WP_002693225.1.	VLFIVLTTILVCIMVILIMLGIGYLIR	99.50%	97.70%	1
Hypothetical protein	WP_002693627.1.	LVNSVLPLFIYGLILSVLSLGIIVKSIEVR	99.70%	98.60%	1
Hypothetical protein	WP_010692308.1.	MAKKSRLVLFVVFAVIIAIIIR	99.70%	98.60%	1
Hypothetical protein	WP_010692309.1.	KNIVGLLIYLIIVALFFSACSKR	98.60%	95.20%	1
Hypothetical protein	WP_010692370.1.	SELIKKLVVFVKPLSGVMTITVILR	99.70%	98.60%	2
Hypothetical protein	WP_010693799.1.	FSLRNKLMIVFGLLILTVGFILSVLAIR	99.60%	98.00%	1
Hypothetical protein	WP_010693826.1.	AKKIVIVGGVAGGASVAARVR	98.80%	99.60%	2
Hypothetical protein	WP_010694677.1.	IKLRKLLSNILISIGMGSVFALFLIFVSPR	99.60%	98.10%	1
Hypothetical protein	WP_010697464.1.	IIAAFLLIVIIIVVLFTSVLNKKHDR	99.30%	97.70%	1
Hypothetical protein	WP_010698465.1.	KIKMPKFIAILFCVLIVLAVFYVLR	99.60%	98.10%	1
Hypothetical protein	WP_010699674.1.	NALKLENIK	99.70%	97.40%	1
Hypothetical protein HMPREF9353_00988	[EGC78141.1]	FGGGIFVK	99.00%	96.00%	1
Hypothetical protein HMPREF9353_01145	[EGC78298.1]	LVIAVYVIFTMITVSILLVKENKKR	99.20%	97.30%	1
Long-chain-fatty-acid-CoA ligase [hypothetical protein]	WP_010692313.1; WP_010698915.1; WP_002693047.1.	AGKITMLLGVPLLFNKLLAGIFRGIR	98.90%	95.10%	1
Major outer sheath protein [hypothetical protein]	WP_002681434.1; WP_002676761.1; WP_002672745.1;	FGSNGSWK	99.00%	100.00%	7

	WP_002685068.1; WP_010688700.1; WP_002687235.1; WP_010693658.1.				
Major surface protein [hypothetical protein]	WP_002687235.1.	LGSKPVDGLALTLAMDALTNVGTDSK	99.70%	100.00%	13
Major surface protein/ major outer sheath protein [hypothetical protein]	WP_002681434.1; WP_002676761.1; WP_002672745.1; WP_002685068.1; WP_010688700.1; WP_002687235.1.	LNHLLSAVPTGDK	99.70%	100.00%	12
		VVIPLYMGTLNSK	99.70%	100.00%	12
Major surface protein/ major outer sheath protein [hypothetical protein]	WP_002681434.1; WP_002676761.1; WP_002672745.1; WP_002685068.1; WP_010688700.1; WP_002687235.1; WP_010693658.1.	LGYASDDLAGTGLK	99.70%	100.00%	7
MATE efflux family protein	[EGC77779.1]	IPLMLGILRIWLLRYIFILATER	98.90%	95.00%	1
Methyl-accepting chemotaxis protein	WP_002681835.1.	IMIALISVLSGFLILTSIVLVKQLSGTSKR	99.50%	97.90%	1
Oligopeptide/dipeptide ABC transporter, ATP-binding protein	[EMB22321.1]	VILQNGLER	98.90%	95.50%	1
Peptide ABC transporter permease	WP_002670594.1.	TLLLFTTAFVINTLIGLALGLKKAQKPGSR	98.50%	98.90%	2
Peptide ABC transporter substrate-binding protein	WP_002670584.1.	VVPSETELDQLLQGEVDMLTSQGQAEK	99.70%	99.80%	4
Signal peptidase II	WP_010692516.1; WP_002691525.1.	IMLVFLPFLLLIALTVAYLKSVELTR	99.50%	97.70%	1
Signal peptidase II	WP_010698126.1; WP_002676082.1; WP_002673628.1; WP_002684345.1; WP_002688035.1.	ILLVFLPFLLLIALTVAYLKSAELTR	99.60%	98.00%	1
Two-component sensor histidine kinase	WP_002687124.1.	LALRFSLLLAAIIVVLSSGIILLLRMNIR	99.60%	98.00%	2
Two-component system response regulator	WP_002668975.1.	QIGQILNSEGYEVVATAVDGFEGVEK	99.70%	98.80%	2
Two-component system response regulator	WP_002693807.1.	QIGQILNSEGYEVIATAVDGFEGVEK	99.70%	98.50%	2

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1E - *Parvimonas micra*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
ATPase/histidine kinase/DNA gyrase B/HSP90 domain protein	[EDP23616.1]	SRKKIILSIMGSIIILFAVTLSVILLASFR	99.60%	98.00%	1
Hypothetical protein	WP_004832038.1.	GNKKFFKIISIILAVICIVFLSIFLR	99.70%	98.60%	1
Hypothetical protein	WP_004832400.1.	NGKLPLTSNLIKNIHKILLNNVR	99.50%	97.70%	1
Hypothetical protein	WP_004833662.1.	TSVLIVLLAIFVAMVVTYLLSR	99.50%	97.50%	1
Hypothetical protein	WP_029948939.1	IVNIVLSVINIIAGVCVHLIPNKIIKISVGMGK DGAPR	99.60%	98.10%	1
Hypothetical protein	WP_029948950.1; WP_004833498.1.	KLIFSSIGISLIGLIILLFFVLRGNFFR	99.00%	96.60%	1
Hypothetical protein	WP_041953664.1; WP_004832074.1.	TYLKVGFLVFLIMILILALDNLNLIIKNSIKNK	99.60%	97.90%	1
Hypothetical protein	WP_041953891.1.	SITNSNKIIDNIINKKIQDINVK	99.70%	98.40%	1

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1F - *Prevotella intermedia*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
Conserved hypothetical protein with SusD domain	WP_045167938.1.	APITEEEATDVKK	99.70%	97.90%	2
Flagellar protein FliS	WP_044047581.1.	LLGLIIIGVICIIITALAGYIRQR	99.30%	97.70%	2
Hypothetical protein	WP_014709881.1.	GLVKAVLLLLFGALLVGGVFSFNRYR	98.50%	95.00%	1
Hypothetical protein	WP_051129504.1; WP_050955385.1; WP_045168207.1.	RKAIAIALLLLSVLTVSAQEEIVR	99.30%	97.60%	1
N-acetyltransferase B complex (NatB) non catalytic subunit domain protein	WP_014709131.1; WP_045166621.1.	LAELIADKQLKIALLTRAIATQR	99.50%	97.70%	1
SusC/RagA family TonB-linked outer membrane protein	WP_045167937.1.	VQSQDQYTTAASDR	99.70%	98.60%	2

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1G - *Fusobacterium nucleatum ssp.†/ Fusobacterium periodonticum*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
3-ketoacyl-CoA thiolase	WP_008797581.1; WP_008799181.1; WP_008802473.1; WP_005914415.1; WP_005911138.1; WP_029600142.1; WP_005903642.1.	NILEETKVDPANIDEVIVGNVLSAGQAQGVGR	99.70%	100.00%	3
ABC transporter	WP_059223149.1; WP_011016023.1; WP_005895008.1; WP_023040760.1; WP_032847995.1.	LITSVIVIVILWIQKQKDKR	99.30%	96.90%	1
ABC transporter permease	WP_008700375.1.	IMLNIFLVLLVITLIQLIFLDRSSDIKMIILR	99.60%	100.00%	2
		QKKALKIMLNIFLVLLVITLIQLIFLDR	99.60%	100.00%	2
Acetyl-CoA acetyltransferase	WP_029758534.1; WP_005911138.1.	SAIGSFLGSLAPLKPGDLGAQIVK	99.70%	100.00%	6
Acetyl-CoA--acetoacetyl-CoA transferase subunit alpha	WP_008798487.1; WP_005909971.1; WP_005906021.1;	NFNPLMATAADLVIVEALEVVPAGSLSPEHLDISR	99.70%	100.00%	12
	WP_022069468.1; WP_029597783.1; WP_032838956.1.	VIASHIGTNAETGRR	99.30%	100.00%	8
Acyl-CoA dehydrogenase	WP_008796823.1; WP_005903799.1; WP_023039674.1; WP_005910603.1; WP_008693325.1; WP_005887888.1; WP_032838542.1.	AIEQGYQALEVPEK	99.70%	100.00%	8

Amino acid transporter LysE	WP_005889891.1; WP_005911766.1.	LIKIIKLANGIIFILALFVTVYSTRQIIIN	99.60%	100.00%	2
Cation diffusion facilitator transporter	WP_008797291.1; WP_005890180.1; WP_005913304.1; WP_008799415.1; WP_008802633.1; WP_023036054.1.	VEYSNITFIILVVSIILKLLLGK	99.60%	97.60%	1
Conjugal transfer protein TrbI	WP_008799789.1.	KPISKIKTKNIILIACGIAAVLMVALFR	99.40%	96.90%	1
Diguanylate phosphodiesterase	WP_005896091.1; WP_005902564.1; WP_011016360.1; WP_032837436.1; WP_008692845.1; WP_008700001.1; WP_020789089.1.	SLDPHASNDNPSSNVR	99.70%	99.00%	6
DNA starvation/stationary phase protection protein	WP_005896619.1; WP_060496102.1; WP_032888860.1; WP_008794449.1; WP_005898917.1.	DFTIPEVVASIKEDMELMLADAR	99.70%	98.10%	2
DNA starvation/stationary phase protection protein	WP_008802732.1; WP_008797196.1.	DFTIPEVIASIKEDMELMLADAK	99.70%	98.10%	2
		DFTIPEVIASIKEDMELMLADAKK	99.70%	98.50%	2
DNA-binding protein	WP_029758046.1; WP_005909242.1.	LVNAFLETIEDALLK	99.70%	99.00%	6
Electron transfer flavoprotein subunit alpha	WP_008795695.1; WP_005906228.1.	ANFQLLEDLAAEIGGIVSASR	99.70%	100.00%	12
Electron transfer flavoprotein subunit alpha	WP_029757636.1; WP_008797426.1; WP_005903460.1; WP_005890490.1; WP_008795695.1; WP_005913360.1;	DGILQNVGLELLGK	99.70%	100.00%	14

	WP_005906228.1; WP_008802554.1.				
Glutamate dehydrogenase	WP_008691370.1.	LTGEQNIGVFTGKPLTYGGSK	99.70%	100.00%	13
Glutamate dehydrogenase	WP_060676006.1; WP_029758541.1; WP_008691370.1; WP_032838361.1.	SKETLNPLESGQQQVK	99.70%	100.00%	7
Hemolysin D	WP_005903417.1; WP_029597148.1.	LKFIILLILIVLGLIYYFTHR	99.30%	97.70%	1
		NIFKGKLKFIILLILIVLGLIYYFTHR	98.90%	95.40%	1
Hemolysin D / efflux transporter periplasmic adaptor subunit	WP_060798428.1; WP_029598966.1; WP_008798349.1; WP_023041952.1; WP_016339940.1; WP_008691562.1; WP_008795405.1; WP_008701034.1; WP_023039441.1; WP_005908714.1.	GKLKFIILLILIILGLIYYFTHR	99.70%	98.60%	1
Hypothetical protein	[EGN66372.1]	LKKLFLSIFSLILIGYASLETLKISR	99.50%	97.60%	1
Hypothetical protein	[EGQ80340.1]	NSKTLVRVILVLVVIVIGFYLIKR	99.10%	97.10%	1
Hypothetical protein	[EGQ80481.1]	LEEKIKFALEYLVVKVLITMVLLN	98.70%	95.50%	1
Hypothetical protein	[EPC07777.1]	QLLLIVIVIISYTILINIQISMGK	99.40%	96.80%	1
Hypothetical protein	WP_005903627.1; WP_011016056.1.	MKKLLLGLFLLASALSFAAGR	98.80%	95.20%	1
Hypothetical protein	WP_005904230.1.	SKMESLIKLSGIIIILIFIILYK	99.30%	97.70%	1
Hypothetical protein	WP_008691669.1.	VAPAIFLEATK	99.70%	100.00%	4

Hypothetical protein	WP_008700013.1; WP_060496225.1; WP_029598725.1.	NSLKIIEKIR	99.70%	98.70%	1
Hypothetical protein	WP_008797448.1; WP_005889917.1.	WIFLILAIIIILIFGIIKSCQR	99.50%	97.80%	1
Hypothetical protein	WP_008797542.1; WP_008799214.1; WP_008802512.1; WP_005912895.1.	MKRIFILLIVLLGLLVVSCGK	99.30%	96.30%	1
Hypothetical protein	WP_008799008.1; WP_023041739.1; WP_023038381.1; WP_023039693.1; WP_016340063.1; WP_005910564.1; WP_008795010.1.	MNSKTLIRIILILIVIVIGFYLIKR	99.60%	98.00%	1
Hypothetical protein	WP_008799068.1; WP_008701272.1.	LKTIVLVSPELSSAYFIKIILEKR	99.50%	97.80%	1
Hypothetical protein	WP_011016730.1.	GIILLYILIFLFIVKNGILKGDSGTIFR	98.70%	95.50%	1
Hypothetical protein	WP_016339736.1.	YSKIFLAIIVVVSIIIKIIRGIR	99.70%	98.60%	1
Hypothetical protein	WP_020789205.1; WP_008796996.1.	SIIIFIVIFLGLGVIMGLITGVITSK	99.30%	96.00%	1
Hypothetical protein	WP_023036497.1.	QGDAFVAFK	99.50%	97.50%	1
Hypothetical protein	WP_023036933.1.	DKILILSIVGIILLIIVIFLFMKNR	99.70%	99.60%	2
		MDKILILSIVGIILLIIVIFLFMKNR	99.10%	99.30%	2
Hypothetical protein	WP_023039392.1.	TYKYLKLLKILDVIIIGVIFFSNINK	98.60%	95.10%	1
Hypothetical protein	WP_029599273.1; WP_032841138.1; WP_020789395.1.	KIILILLFLLINIGVFSVHSKKNLVR	99.70%	98.60%	1

Hypothetical protein	WP_029758616.1; WP_005897113.1.	GHDWLKEDEYIAK	95.20%	99.10%	1
Hypothetical protein	WP_059223007.1; WP_005903860.1.	ENIFLILAFFILIFVAAIALKIRQR	99.50%	98.20%	1
Hypothetical protein	WP_060496462.1; WP_005909975.1.	FGLRPGYAHNWSGHSNGGEGR	99.70%	100.00%	53
Hypothetical protein	WP_060496462.1; WP_032834443.1; WP_005909975.1.	ASYELYMLPTFQVSYKPTDFVK	99.70%	100.00%	45
		ASYSVYMLPTFQVAYKPTDFVK	99.70%	98.60%	1
		LQTTANVNFTK	99.70%	100.00%	20
Hypothetical protein	WP_060676618.1; WP_060676862.1; WP_041461237.1; WP_032840052.1; WP_032839640.1; WP_029757663.1; WP_008691612.1.	GEEEPIADNATTEGR	99.70%	99.00%	2
Hypothetical protein	WP_060676711.1; WP_023041864.1; WP_023039782.1; WP_005910878.1; WP_060675827.1; WP_032839248.1.	EAVQGTTAAPVANENK	99.70%	100.00%	14
Hypothetical protein	WP_060798423.1; WP_005900053.1; WP_032837031.1.	ENIFLILAFFILISLAAIALKIRHR	99.70%	98.40%	1
Hypothetical protein	WP_060797678.1; WP_005909469.1.	EEEKPAEQPAATEQTTTAEAPK	99.70%	100.00%	10
Hypothetical protein	WP_060798261.1; WP_005897607.1; WP_005910180.1.	EEEKPAEQPAAEATAPATEAPATEAAAEAK	99.70%	100.00%	23

LOS biosynthesis enzyme LBGB	WP_005890447.1.	KVYPYIKIGVVARGAAIDIIR	99.70%	98.40%	1
Membrane antigen	WP_060797678.1; WP_005905592.1; WP_005909469.1.	EGGVAAAEEYFK	99.70%	100.00%	10
Membrane protein	WP_008694349.1.	NYQTLNTNKEEKDPITGK	99.70%	100.00%	8
Membrane protein	WP_008797071.1; WP_023037361.1; WP_023039195.1; WP_005896265.1; WP_008802885.1.	DGGVAAAEEYFK	99.70%	100.00%	12
Membrane protein	WP_008800240.1.	STDKGDSNADDDQVR	99.70%	98.60%	2
Membrane protein	WP_011015959.1; WP_005903305.1; WP_023040702.1.	KKIAFLFIFVLVLLIPLELIKNLIDDR	99.70%	98.60%	1
Membrane protein	WP_011017017.1; WP_008797071.1; WP_008802885.1.	KEEEKPAEQPAATTEAPATAEAPK	99.70%	100.00%	4
Membrane protein	WP_029599066.1; WP_008797223.1; WP_005890252.1; WP_008799469.1; WP_005912656.1; WP_020788906.1.	TNKIIFYILLILVVISLIQLPFLDR	99.00%	95.40%	1
Membrane protein	WP_029599475.1; WP_008800240.1;	DHNKYEVYMLPTFQVAYKPTDFVK	99.70%	100.00%	3
	WP_008796594.1.	WAGHDSGVVGQPFSK	99.70%	100.00%	8
Membrane protein	WP_029758694.1; WP_008694349.1; WP_060496462.1.	VIDNGSGYER	99.70%	100.00%	14
Membrane protein	WP_005909975.1.	EYVTFDGYNGGTNNK	99.70%	100.00%	50
		SEDTDKDWANSK	99.70%	100.00%	15

Membrane protein [hypothetical protein]	WP_011015959.1; WP_005903305.1; WP_023040702.1.	KIAFLFIFVLVLLIPLELIKNLIDDR	99.60%	98.00%	1
Molecular chaperone DnaK	WP_005906171.1.	NAQAQQQAGAQANASSDENK	99.70%	98.40%	1
Nickel transporter	WP_005911130.1; WP_029598541.1; WP_005903634.1; WP_032838361.1; WP_008691370.1; WP_011016425.1.	DSGFTFEELEAAK	99.70%	100.00%	16
Nickel transporter	WP_060675608.1; WP_005903852.1; WP_032890213.1; WP_008800088.1; WP_005910197.1; WP_008803296.1; WP_005889025.1; WP_032837038.1; WP_032848328.1.	VAPGHNEIVPLVNPVNAWK	99.70%	100.00%	12
Outer membrane autotransporter barrel domain-containing protein	WP_005910328.1; WP_020992560.1.	AMTTSGFAGLENAR	99.70%	100.00%	3
Peptide ABC transporter permease	WP_008693282.1.	FIFKWLGIIFILSVITFLIVR	99.10%	96.70%	1
Phenylalanine--tRNA ligase subunit alpha	WP_011016059.1; WP_041461262.1; WP_032841095.1; WP_029599218.1; WP_005903621.1; WP_005890222.1; WP_020788884.1.	INSLLDERNK	98.80%	95.30%	1
Serine/threonine protein kinase	WP_005902075.1.	KLAKIILNIVLILAIIKLGSNIFQR	99.30%	100.00%	2
Transposase	WP_011017157.1; WP_054175325.1; WP_011015897.1; WP_005903793.1.	KNGETFSSLAK	99.70%	98.90%	8

Transposase	WP_060676577.1; WP_008798819.1; WP_008795215.1; WP_023038270.1; WP_023039981.1.	YDFQKIVIIR	99.60%	98.30%	3
tRNA (guanine-N(7)-)-methyltransferase	WP_008694462.1.	NYVLMIVPRHLERLPKIENLIK	99.70%	98.60%	1

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software); † *Fusobacterium nucleatum ssp.*: *Fusobacterium nucleatum subsp. nucleatum, Fusobacterium nucleatum subsp. polymorphum, Fusobacterium nucleatum subsp. vincentii, Fusobacterium nucleatum subsp. animalis, sp. CM1, sp. CM21, sp. CM22.*

S1G - *Streptococcus intermedius*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	WP_003075464.1; WP_003077406.1.	ENIPNALTLLRIALIPIFILVLSLGHSIR	99.70%	98.60%	1
Hypothetical protein	WP_003077480.1.	VSIINLLLLIVGTVEIFHYFIVLTGLKDR	99.10%	96.00%	1
Hypothetical protein	[EKU16991.1]	IVVALVIGVILLGIAGKPQIHKMAR	99.70%	100.00%	5
		SLFKKIVVALVIGVILLGIAGKPQIHKMAR	99.60%	100.00%	5
Putative membrane protein	[EID83471.1; BAM23756.1; AGU75869.1]	IRLLILLLSIALGFLLSSFFLNLYTLGR	99.60%	98.00%	1
YSIRK signal domain/LPXTG anchor domain surface protein	WP_003071367.1.	SSLGLTVLGLLGLSSAAILALIGMKRR	99.20%	97.20%	1

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1H - *Streptococcus oralis*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
Damage-inducible protein CinA	WP_001103720.1.	LVVLIPFLVLLLLTGIKEDVLDGGSRRR	99.00%	95.40%	1
Glucan-binding protein C, partial	[EJP21944.1]	LDPTATQGAETGK	99.70%	100.00%	6
		SEVGTAYDTTDQR	99.70%	100.00%	4
Gram-positive signal peptide protein, YSIRK family, partial	[EIC78459.1]	KEIINKNLKK	99.50%	97.90%	2
Hypothetical protein	WP_000859476.1.	MKTVIALLVLILIGFWAINILKPRK	98.40%	99.40%	2
Hypothetical protein	WP_001262749.1.	AALLQLLSEFLRELIKKFISWLK	99.20%	95.60%	1
Hypothetical protein	WP_002880152.1.	DQFKWLNEK	99.40%	98.00%	2
Hypothetical protein	WP_042902462.1.	SEIQQWLGK	99.60%	97.40%	2
Protease	WP_049500151.1; WP_000479681.1.	FLQILLVLLFLVVIASVIMVSSVYVVR	99.00%	95.40%	1

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

S1I - *Abiotrophia defectiva*

Protein Name	NCBI Ref. Sequence	Peptide Sequence	Best Peptide Probability	Best Protein Probability	Exclusive Unique Peptides
CAAX amino protease	WP_035364307.1.	KAIFLYLLGSLGQIWIISMIVFILR	99.60%	98.30%	1
Cell envelope-like function transcriptional attenuator common domain protein [hypothetical protein]	WP_035364821.1.	ILAGRGYQFWVLGGLGAVLLLLILLNGLKR	99.60%	98.10%	1
DUF368 domain-containing protein	WP_023391108.1.	QLDLDVLIPIGLGGLATIILLSKLIER	99.30%	97.70%	1
Hypothetical protein	WP_023391613.1.	ALKKVLLASLALTCLAGPMVQAAPRIINIR	99.60%	97.90%	1
Hypothetical protein	WP_023391813.1.	IKKSLLILLCIPPLVLFSINLYR	95.40%	95.40%	1
Hypothetical protein	WP_023391906.1.	RFDLSKLKLNISLFVLVLLAVLIGISQR	98.90%	95.00%	1
Nrdi protein	[ESK66283.1]	MLKNLLKEVLKLLVVYLSITGQTR	99.60%	100.00%	3
Twin arginine-targeting protein translocase TatA/E family	WP_023390673.1.	MGLLRDIGAPGLILIVLGALLIFGPKR	99.10%	97.00%	1

NCBI Ref. Sequence, National Center for Biotechnology Information Reference Sequence [NCBI GenBank accession numbers are given when no ref. sequence is available]; Protein/Peptide Probability as determined by Scaffold (Proteome Software)

Table S2. Citrullinated peptides identified in untreated PMNs (*provided as an Excel file*).

Table S3. Citrullinated peptides identified in LtxA-treated PMNs (*provided as an Excel file*).

Table S4. Characteristics of healthy controls without periodontitis and patients with chronic periodontitis.

	No PD (*n* = 100)†	**PD (*n* = 109)†**	***P***
Age, years (± SD)‡	35 ± 13	49 ± 11	<0.001
Male gender, *n* (%)‡	35 (35)	64 (59)	<0.001
Caucasian, *n* (%)‡	74 (73)	52 (52)	
African American, *n* (%)‡	13 (13)	41 (41)	
Other, *n* (%)‡	12 (12)	7 (7)	<0.001
Bleeding on probing, % of sites (± SD)§	16 ± 19	38 ± 22	<0.001
Pocket depth in mm, mean (± SD)§	2.0 ± 0.3	2.7 ± 0.5	<0.001
Clinical attachment level, mean (± SD)§	1.5 ± 0.7	2.5 ± 0.7	<0.001
Anti-LtxA positivity, *n* (%)	11 (11)	68 (62)	
Anti-LtxA units, median (IQR)	2.3 (0.0-2.8)	24 (2.5-201)	<0.001

†Smoking and diabetes mellitus were exclusion criteria in this study cohort.
‡Demographic information was available for 100/109 periodontitis patients (PD) and 99/100 controls (No PD).
§Clinical indices were available for 98/100 controls without periodontitis.

Multiple regression model for log anti-LtxA

	Beta (StdErr)	**Linear *P* value**	**Rank *P* value**
PD Case*	1.12 (0.12)	<0.0001	<0.0001

*Adjusted for race (3 options), age, gender, study

Table S5. Characteristics of RA patients by anti-LtxA antibody status.

	No anti-LtxA (*n* = 112)	**Anti-LtxA +ve (*n* = 84)**	***P***
Age, years	59 ± 8	60 ± 9	0.13
Male gender, *n* (%)	44 (39)	35 (42)	0.74
Caucasian, *n* (%)	103 (92)	65 (77)	0.004
African-American, *n* (%)	5 (5)	13 (15)	
Asian	3 (3)	4 (5)	
Hispanic	1 (1)	2 (2)	
RA duration, years	10 (4-17)	8 (5-20)	0.90
RF seropositivity, *n* (%)	64 (57)	65 (77)	0.003
Anti-CCP seropositivity, *n* (%)	82 (73)	69 (82)	0.14
Anti-CCP units among pos., median (IQR)	144 (93-182)	145 (104-182)	0.73
ACPA seropositivity, *n* (%)	75 (67)	69 (83)	0.011
DAS28, median	3.5 (2.9-4.3)	3.8 (2.9-4.3)	0.59
Swollen joint count, median	7 (3-10)	7 (4-10)	0.71
Tender joint count, median	6 (2-12)	6 (3-13)	0.70
Total SvdH score, median†	7 (0-32)	12 (1-65)	0.12
Total erosion score, median†	2 (0-8)	4 (0-22)	0.077
Total JSN score, median†	4 (0-20)	6 (0-38)	0.15
Pain (100mm VAS), median	18 (7-32)	24 (8-47)	0.042
Non-biologic DMARDs, *n* (%)	97 (87)	68 (82)	0.37
Biologic DMARDs, *n* (%)	57 (51)	32 (39)	0.087
Glucocorticoids, *n* (%)	44 (39)	31 (37)	0.73

†150 had interpretable radiographs at follow-up

Anti-LtxA, anti-Leukotoxin A antibodies, cut-off level for positivity >95th percentile of controls; RF, rheumatoid factor, cut-off for positivity >40 units; anti-CCP, anti-cyclic citrullinated peptide antibody, cut-off for positivity >20 units; ACPA, anti-citrullinated protein antibodies, any positivity by multiplex assay (ACPA fine specificities); DAS28, Disease Activity Score 28; DMARDs, disease-modifying antirheumatic drugs; JSN score, joint space narrowing score; SvdH score, Sharp/van der Heijde score; VAS, Visual Analogue Scale.

Table S6. The association of shared epitope alleles with ACPAs and RF according to anti-*P. gingivalis* (*Pg*) antibody status.

	Anti-*Pg* negative RA (*n* = 123)				Anti-*Pg* positive RA (*n* = 71)				
	SE negative (*n* = 36)	SE positive (*n* = 87)	OR	*P*	SE negative (*n* = 22)	SE positive (*n* = 49)	OR	*P*	*P-interaction*
Anti-CCP positivity, %									
Crude prevalence (95% CI)	56 (39, 71)	85 (76, 91)	4.55	0.001	64 (42, 81)	86 (73, 93)	3.43	0.041	0.71
Adjusted* prevalence (95% CI)	58 (40, 74)	89 (80, 94)	5.71	<0.001	65 (40, 84)	87 (73, 94)	3.45	0.071	0.55
ACPA positivity, %†									
Crude prevalence (95% CI)	61 (45, 75)	77 (67, 84)	2.10	0.082	68 (47, 84)	82 (68, 90)	2.07	0.22	0.99
Adjusted* prevalence (95% CI)	64 (46, 78)	79 (69, 87)	2.18	0.081	64 (41, 83)	81 (66, 90)	2.31	0.18	0.94
RF positivity, %									
Crude prevalence (95% CI)	64 (47, 78)	79 (70, 87)	2.17	0.076	73 (51, 87)	88 (75, 94)	2.69	0.13	0.78
Adjusted* prevalence (95% CI)	68 (51, 82)	83 (73, 90)	2.26	0.082	77 (52, 91)	88 (75, 95)	2.31	0.25	0.98

†Any ACPA fine specificity (see Table 1)

Anti-*Pg*, anti-*P. gingivalis* antibodies as determined by ELISA, cut-off for positivity >95th percentile of healthy controls without periodontitis; anti-CCP, anti-cyclic citrullinated peptide antibody, cut-off for positivity >20 units; ACPA, anti-citrullinated protein antibodies, any positivity by multiplex assay as previously defined (*X*); RF, rheumatoid factor, cut-off for positivity >40 units; SE, HLA-DRB1 shared epitope allele

*Adjusted for age, Caucasian race, current smoking, IL-6, pain severity, and biologic use