AUTOIMMUNE DISEASES

Oral mucosal breaks trigger anti-citrullinated bacterial and human protein antibody responses in rheumatoid arthritis

R. Camille Brewer[1,2], Tobias V. Lanz[1,2,3], Caryn R. Hale[4], Gregory D. Sepich-Poore[5]†, Cameron Martino[6,7,8], Austin D. Swafford[7], Thomas S. Carroll[9], Sarah Kongpachith[1,2], Lisa K. Blum[1,2], Serra E. Elliott[1,2], Nathalie E. Blachere[4,10], Salina Parveen[4], John Fak[4], Vicky Yao[11,12], Olga Troyanskaya[12,13,14], Mayu O. Frank[4], Michelle S. Bloom[1,2], Shaghayegh Jahanbani[1,2], Alejandro M. Gomez[1,2], Radhika Iyer[1,2], Nitya S. Ramadoss[1,2], Orr Sharpe[1,2], Sangeetha Chandrasekaran[15], Lindsay B. Kelmenson[15], Qian Wang[1,2], Heidi Wong[1,2], Holly L. Torres[2], Mark Wiesen[2], Dana T. Graves[16], Kevin D. Deane[15], V. Michael Holers[15], Rob Knight[5,6,7,17], Robert B. Darnell[4,10], William H. Robinson[1,2]*‡, Dana E. Orange[4,18]*‡

Periodontal disease is more common in individuals with rheumatoid arthritis (RA) who have detectable anti-citrullinated protein antibodies (ACPAs), implicating oral mucosal inflammation in RA pathogenesis. Here, we performed paired analysis of human and bacterial transcriptomics in longitudinal blood samples from RA patients. We found that patients with RA and periodontal disease experienced repeated oral bacteremias associated with transcriptional signatures of ISG15^{+}HLADR^{hi} and CD48^{high}S100A2^{pos} monocytes, recently identified in inflamed RA synovia and blood of those with RA flares. The oral bacteria observed transiently in blood were broadly citrullinated in the mouth, and their in situ citrullinated epitopes were targeted by extensively somatically hypermutated ACPAs encoded by RA blood plasmablasts. Together, these results suggest that (i) periodontal disease results in repeated breaches of the oral mucosa that release citrullinated oral bacteria into circulation, which (ii) activate inflammatory monocyte subsets that are observed in inflamed RA synovia and blood of RA patients with flares and (iii) activate ACPA B cells, thereby promoting affinity maturation and epitope spreading to citrullinated human antigens.

INTRODUCTION

Periodontal disease (PD) is more common in patients with rheumatoid arthritis (RA), particularly RA with anti-citrullinated protein antibodies (ACPAs) (*1*). PD is a common disease that affects up to 47% of the adult population (*2*) and results in gingival bleeding with translocation of oral bacteria to blood (*3*). RA patients with ongoing PD have increased disease activity (*4*–*6*) and are more likely to have treatment-refractory disease (*7*), suggesting that PD may trigger systemic inflammatory pathways that are relevant for ongoing joint inflammation. However, the mechanisms by which PD and oral inflammation may contribute to ACPA development and persistence in RA are unclear.

ACPAs recognize an array of human citrullinated proteins, generated through posttranslational modifications of arginine to citrulline by peptidylarginine deiminase enzymes (PADs), including filaggrin, fibrinogen, α-enolase, histones, and vimentin (*8*), and are clinically useful for the classification of RA (*9*). In established RA, synovial ACPAs positively correlate with disease activity (*10*), and RA patients who are seropositive are more likely to have flares after discontinuation of conventional treatments (*11*–*13*). ACPAs often precede the onset of arthritis by years, suggesting that loss of tolerance to citrullinated human antigens is an early event in RA pathogenesis (*8*, *14*, *15*). In addition, before onset of RA, there is an increase in N-linked glycosylation sites in the variable regions of ACPA-expressing B cells (*16*–*18*), and these N-linked glycans may enhance binding to bacterial lectins (*19*, *20*). Whereas epitope spreading, development of high-titer ACPAs, and accumulation of N-linked glycans in variable regions of ACPAs are associated with the progression to established RA, the mechanisms underlying the induction and reactivation of ACPA-expressing B cell responses are not well defined.

[1]Division of Immunology and Rheumatology, Stanford University, Stanford, CA 94305, USA. [2]VA Palo Alto Health Care System, Palo Alto, CA 94304, USA. [3]Department of Neurology, Medical Faculty Mannheim, University of Heidelberg, Mannheim 68167, Germany. [4]Rockefeller University, New York City, NY 10065, USA. [5]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA. [6]Department of Pediatrics, University of California San Diego, La Jolla, CA 92093, USA. [7]Center for Microbiome Innovation, University of California San Diego, La Jolla, CA 92093, USA. [8]Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA 92093, USA. [9]Bioinformatics Resource Center, Rockefeller University, 1230 York Ave., New York, NY 10065, USA. [10]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA. [11]Department of Computer Science, Rice University, Houston, TX 77005, USA. [12]Department of Computer Science, Princeton University, Princeton, NJ 08544, USA. [13]Lewis-Sigler Institute of Integrative Genomics, Princeton University, Princeton, NJ 08544, USA. [14]Flatiron Institute, Simons Foundation, New York, NY 10010, USA. [15]Division of Rheumatology, University of Colorado Denver, Aurora, CO 80045, USA. [16]Department of Periodontics, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [17]Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA 92093, USA. [18]Hospital for Special Surgery, New York City, NY 10075, USA.
*Corresponding author. Email: w.robinson@stanford.edu (W.H.R.); dorange@mail.rockefeller.edu (D.E.O.)
†Present address: Micronoma, San Diego, CA 92121, USA.
‡These authors contributed equally to this work.

Here, we report our observation that RA patients with PD experience frequent bouts of oral bacteremias, that is, episodes of increased oral bacterial RNA in their blood. These oral bacteremias coincided with inflammatory monocyte transcriptional signatures. In RA patients with PD, flares were enriched with the same inflammatory monocyte signatures and antibody effector function pathways. This observation prompted an investigation of the antibody response to oral bacteria, and we found that oral bacteria that are repeatedly detected in RA blood are broadly citrullinated in the mouth and recognized by extensively somatic hypermutated ACPAs encoded by RA blood plasmablasts. Our findings indicate that damage of the oral mucosal barrier mediated by PD results in repeated, spontaneous translocation of citrullinated oral bacteria to the blood, which triggers innate and adaptive immune responses in RA associated with systemic disease flares.

RESULTS

Oral mucosal breaks trigger systemic inflammatory responses

To determine the role of the microbiome in RA patients with PD, we performed bulk RNA sequencing (RNA-seq) analysis on blood samples from RA patients with and without PD, obtained by weekly finger sticks over the course of 1 to 4 years (Rockefeller University longitudinal cohort, five patients, mean $n = 67$ time points per patient) (table S1). RNA-seq transcripts were first aligned to the human genome (hg38) for differential gene expression analysis. The human-depleted reads were then aligned to a microbial metagenomic database [Web of Life (WoL)] (*21*), followed by an in silico decontamination pipeline (Fig. 1, A and B, and fig. S1) (*22*). To ascertain the mucosal source of the bacteria in the blood, we inferred the relative abundances of bacteria from three oral mucosal sites (buccal mucosa, supragingival plaque, and tongue dorsum) and five other body sites (stool, vaginal fornix, anterior nares, left and right retroauricular creases) using SourceTracker2 (*23*) and matched body site samples from the Human Microbiome Project (HMP) (*24*). Because the HMP dataset did not use metatranscriptomics, we validated the approach by applying this pipeline to a publicly available metatranscriptomic dataset of human stool samples (*25*). We determined that nearly all predicted body site attributions were stool (fig. S2).

The inferred relative abundances of oral bacteria in blood were higher in the time points from RA patients with PD as compared with RA patients without PD (Fig. 1, C and D, and fig. S3, A to E). Conversely, there were no differences in the inferred relative abundances of bacteria from other body sites (Fig. 1C). The most abundant oral bacteria detected in blood by RNA-seq were *Streptococcus* species, mirroring the most abundant species detected in oral swabs by 16*S* sequencing (Fig. 1E and fig. S3F). Together, these data demonstrate that bacteremias from the oral cavity, but not other body sites, were more frequent in the blood of RA patients with PD compared with the blood of RA patients without PD.

We next compared human gene expression with the inferred relative abundances of bacteria from various body habitats. Because there were very few source contributions or variances in the levels of fornix- and stool-derived bacteria, these two sites were excluded (figs. S1 and S4A). We found that the inferred relative abundances of tongue dorsum, buccal mucosa, and supragingival plaque, but not any other body sites, were associated with human differentially expressed genes (DEGs) (Fig. 1F and fig. S4B). Moreover, these DEGs were enriched for overlapping innate immune, defense, and interferon pathways (Fig. 1G). Congruently, we observed a positive correlation between the percentage of monocytes, a key innate immune cell, and the inferred relative abundances of oral bacteria in the blood of RA patients with PD (Fig. 1H). We validated the robustness of these SourceTracker2-inferred body site contribution estimates by comparing each SourceTracker2-inferred body site contribution to matched log ratios of the RA data comprising the microbe counts determined to be differentially abundant [ANCOM-BC (*26*)] in the respective HMP body sites over all other decontaminated microbial read counts (fig. S4, C and D; see Materials and Methods). Given that these three oral sites are proximal to each other, breaches in the oral mucosa likely result in synchronous entry of bacteria from the oral cavity into the blood, inducing activation of innate immune cells, including monocytes, in the blood.

Given that this observation was correlative, we tested whether oral bacteria directly trigger similar gene expression responses in vitro. Given that both RA and healthy oral flora are composed of mostly *Streptococcus* species, we incubated pooled oral bacteria isolated directly from the oral cavities of healthy individuals with whole blood from healthy individuals in vitro, and we measured relative expression of genes concomitant with oral bacteremias in RA patients (MARCO, TNFAIP6, ERAP2, and ISG15) (fig. S5, A and B). We detected robust induction of these genes in whole blood after 6 and 20 hours of stimulation (Fig. 1I). Thus, oral bacteria can induce systemic inflammatory responses if oral bacteria translocate to blood as a result of an impaired oral barrier, as seen in PD, and this response is not specific to patients with RA.

The observed response to the oral bacteria was likely mediated by granulocytes and monocytes, because purified granulocytes and monocytes up-regulated expression of TNFAIP6 and ISG15 after 6 hours of stimulation with oral bacteria (Fig. 1J). In addition, monocytes stimulated with oral bacteria up-regulated ISG15 protein expression (Fig. 1K), an effect mediated, in part, by the activation of FcγR2a caused by immunoglobulin G (IgG) coating of oral bacteria (fig. S5C) (*27*). Together, antibody-coated oral bacteria in vitro induce a similar inflammatory immune gene signature as observed in vivo in RA patients, with PD likely causing direct activation of blood innate immune cells by oral bacteria.

Inflammatory synovial monocyte genes are enriched in blood during RA flares and in response to oral bacteremia

We next sought to determine whether the oral bacteremias in RA patients with PD contribute to RA flares. In our prior study of RA flares (*28*), we identified enrichment of myeloid-related pathways in blood during flares as compared with baseline. Using the same longitudinal cohort from Fig. 1, we compared gene expression during flares to baseline and identified DEGs for flares from RA patients with and without PD (Fig. 2A). We classified genes with consistent fold change direction in both PD-associated and non–PD-associated flares as "common flare genes." Conversely, we classified genes with discordant fold change direction in PD-associated and non–PD-associated flares as "periodontal disease–specific genes" (Fig. 2B). PD-specific flare genes, but not the common flare genes, were enriched for myeloid leukocyte migration, immune complex clearance, and Ig binding (Fig. 2C). Moreover, only the PD-specific flare genes, and not the common flare genes, were

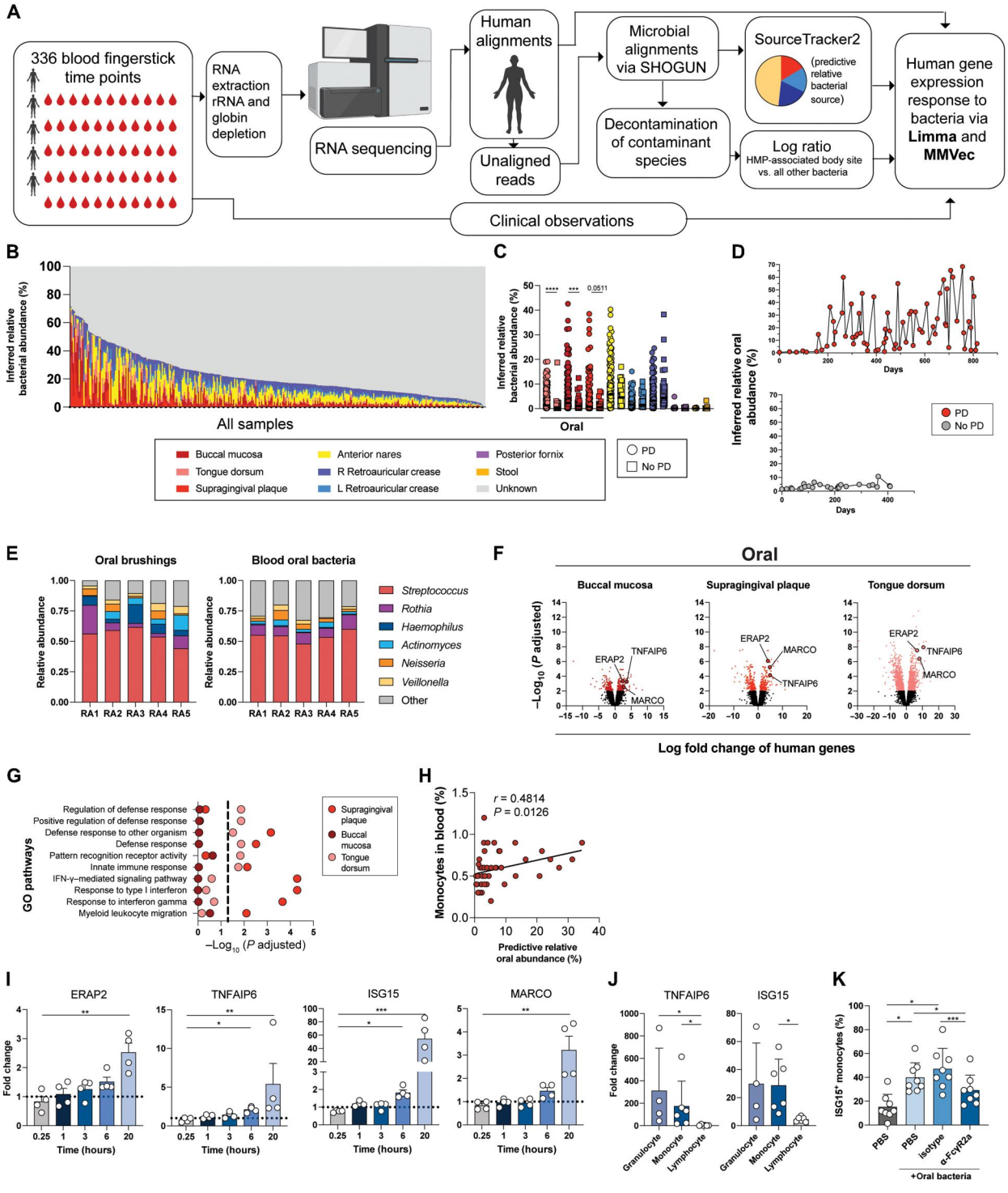

Fig. 1. Oral mucosal breaks trigger systemic inflammatory responses. (**A**) Experimental workflow. (**B**) Inferred relative bacterial abundances from eight HMP body sites (n = 336). (**C**) Body site–inferred relative abundances for time points from RA patients with and without periodontal disease, median. (**D**) Inferred relative oral abundances in blood for one RA donor with and one without periodontal disease. (**E**) Relative abundances of bacteria genera from oral brushings (left) and blood (right). (**F**) Log_{10}-adjusted P values versus log fold changes of human gene expression relative to bacterial abundances of the three oral body sites. (**G**) Enriched GO pathways in differentially expressed human genes from (F) (adjusted P values). IFN-γ, interferon-γ. (**H**) Percentage of monocytes in blood cell counts compared with inferred relative abundances of oral bacteria, Pearson's correlation. (**I** and **J**) RT-qPCRs of mRNA of inflammatory genes in (I) whole blood, (J) granulocytes, monocytes, and lymphocytes (n = 4 to 6) stimulated with oral bacteria versus unstimulated control. (**K**) Flow cytometry data showing proportion of $ISG15^+$ monocytes in $CD14^+$ monocytes incubated with PBS, oral bacteria, and anti-FcγR2a or isotype (n = 8). (B, H, and I) Two-tailed Kruskal-Wallis test, Dunnett-corrected multiple comparisons. (C) Mann-Whitney U test. (K) Within-subjects analysis of variance (ANOVA), Tukey-corrected multiple comparisons. Mean ± SD. *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.001 or exact value shown.

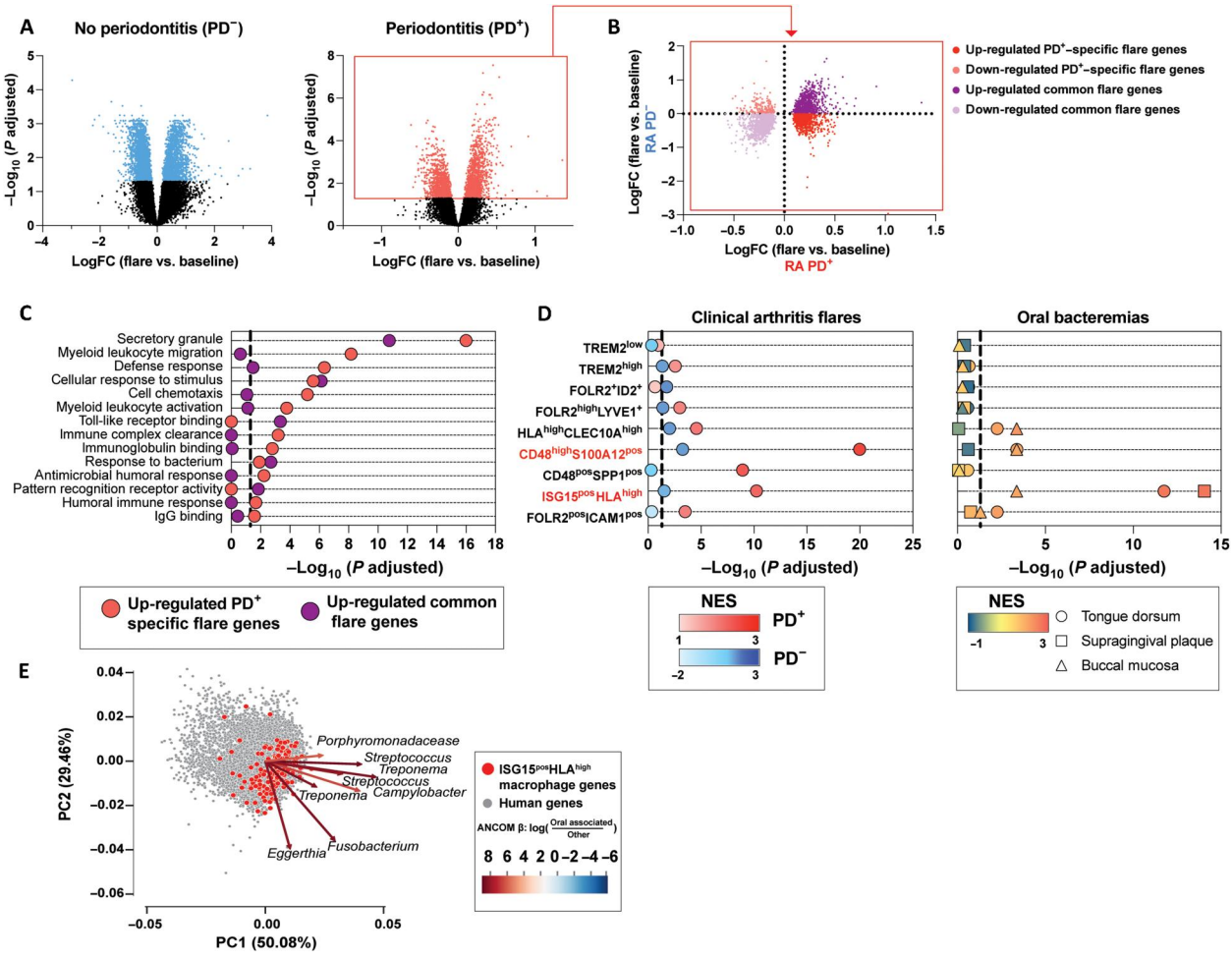

Fig. 2. Inflammatory synovial monocyte genes are enriched in RA blood during flares and in response to oral bacteremia. (**A**) Volcano plot of $-\log_{10}$-adjusted *P* value versus log fold change of blood samples in flare versus baseline for RA patients without periodontal disease (left) and RA patients with periodontal disease (right). Colored dots indicate false discovery rate (FDR) < 0.05. (**B**) For differentially expressed genes (DEGs) during flares in RA patients with periodontal disease, a scatterplot shows the direction and value of those DEGs during flares of RA patients with and without periodontal disease. (**C**) GO pathway enrichment analysis of up-regulated PD^+ specific flare genes and up-regulated common flare genes identified in (B). (**D**) Gene set enrichment analysis results of synovial macrophage clusters for flares in RA patients with periodontal disease (left, red) and flares in RA patients without periodontal disease (left, blue) and oral-derived bacteremias (right). Dashed line indicates FDR of 0.05. (**E**) Principal components analysis (PCA) biplot of bacteria-human gene expression with co-occurrence probabilities estimated from MMvec. Distances between points quantify the probability of co-occurrence strength between human genes (points). Distances between arrow tips quantify co-occurrence strength between microbes (arrows). Arrow color indicates strength of association of bacteria with the oral cavity body site.

enriched for antibody effector functions against microbes and myeloid immune responses (Fig. 2C). Together, these data suggest that (i) immune responses to repeated oral bacteremias contribute to RA flares; (ii) oral bacteremias promote an antibody effector response in RA patients with PD; and (iii) immune signatures in response to RA flares are variable, potentially due to distinct triggers of flare.

Given the specific enrichment of the myeloid leukocyte pathway during flares in RA patients with PD, we compared the host transcriptomic response to oral bacteremias and flares with a recently published single-cell RNA-seq analysis of RA synovial macrophages (*29*). We found that oral bacteremias and flares only in RA patients with PD, but not those without PD, were enriched for genes expressed by monocyte subsets analogous to $ISG15^+HLA^{hi}$, $CD48^{pos}SPP1^{pos}$, and $CD48^{high}S100A12^{pos}$ inflammatory synovial macrophage subsets (Fig. 2D). Similarly, the gene signature of $MerTK^-HLA^+ISG15^+$ synovial macrophages had a high probability of co-occurrence with species that dominate the oral cavity, including *Streptococcus* spp. (Fig. 2E), based on MMvec (*30*), suggesting that oral bacteria promote activation of the inflammatory monocyte subsets that are analogous to macrophages in the inflamed RA synovium.

Highly mutated RA antibodies bind citrullinated human and bacterial antigens in vitro

We reasoned that if flares in RA patients with PD were associated with antibody effector functions, RA patients might harbor additional more specific responses to oral bacteria relative to individuals without RA. Given that PD is specifically associated with $ACPA^+$ RA (*31*), we hypothesized that ACPAs may cross-react with oral

bacteria. We therefore screened monoclonal antibodies (mAbs) derived from human RA blood plasmablasts (tables S2 and S3) against a panel of human proteins and bacterial lysates characteristic of the oral mucosa, as well as other mucosal sites. The RA mAbs that bound the human citrullinated autoantigens (ACPA mAbs) cross-reacted with a broad array of in vitro citrullinated bacterial lysates, but not native bacterial lysates or the PAD enzyme used to citrullinate in vitro (Fig. 3A and fig. S6), indicating that they were also anti-citrullinated bacterial antibodies (ACBAs). RA antibodies without specificity against citrullinated human antigens (unreactive mAbs) showed little reactivity to any bacterial lysate (Fig. 3A and fig. S6). There was no difference in B cell receptor (BCR) characteristics, such as CDR3 length or hydrophobicity (GRAVY score), between ACPA and unreactive mAbs (fig. S7, A and B). In contrast, the ACPA mAbs harbored more somatic hypermutations (SHMs) in the V-region of the heavy and light chain (V_H and V_L) genes compared with the unreactive mAbs (Fig. 3B and fig. S7C), which had similar percentages of SHM as antibody responses in other clinical scenarios (*32*, *33*). In addition, we investigated the frequency of N-linked glycan motifs in the antigen-binding fragment (Fab) region,

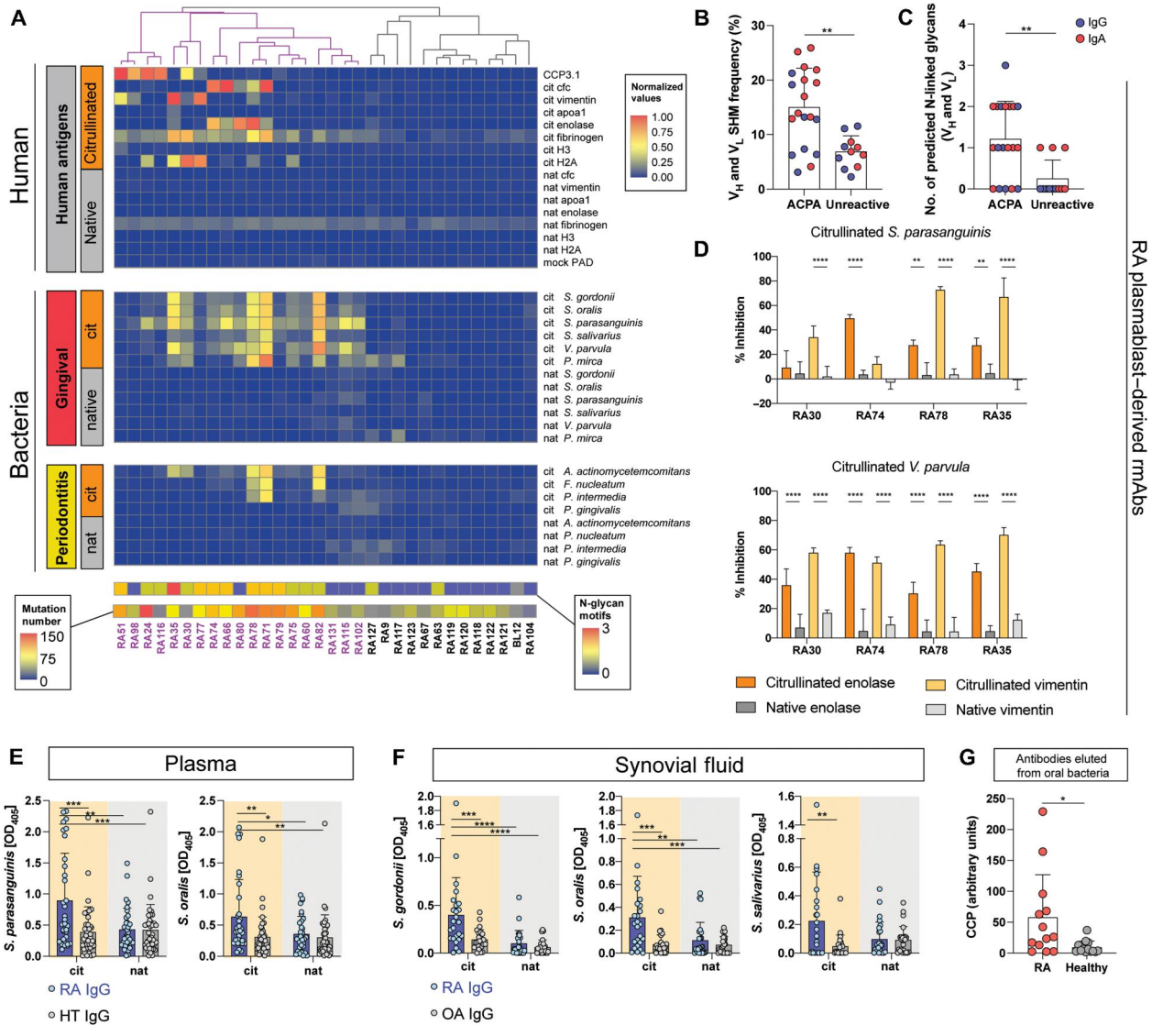

Fig. 3. RA antibodies from plasma, synovium, and gingiva bind both citrullinated human and oral bacteria antigens. (**A**) Heatmap of normalized mAb reactivity to citrullinated (cit) and native (nat) bacterial lysates and human recombinant proteins as measured by ELISA. Bars below indicate the mutation load in the V-region and predicted N-glycan motifs. (**B**) SHM frequency in the V-region in ACPA and unreactive mAbs from (A). (**C**) Quantification of predicted N-linked glycan motifs in ACPA mAbs and unreactive mAbs from (A). (**D**) Percentage inhibition of mAb binding to citrullinated *S. parasanguinis* and *Veillonella parvula* bacterial lysates after preincubation with citrullinated or native human proteins. (**E**) Reactivity of plasma IgG from RA (n = 33) and healthy controls (n = 40) against citrullinated and native bacterial lysates. Mean ± SD. (**F**) Reactivity of synovial fluid IgG from RA (n = 25) and OA (n = 25) against citrullinated and native bacterial lysates. Mean ± SD. (**G**) Cyclic citrullinated peptide (CCP) reactivity of IgG and IgA eluted from oral bacteria from RA patients (n = 13) and healthy controls (n = 13). Mean ± SD. For (B), two-tailed Student's t test was performed. For (C) and (G), two-tailed Mann-Whitney test was performed. For (D to F), two-way ANOVA followed by Dunnett's multiple comparisons test was performed. *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001.

characterized by the consensus sequence N-X-S/T (*34*), which are potential sites for glycosylation that facilitate binding to bacterial lectins (*19*). ACPA mAbs contained increased numbers of predicted N-linked glycan motifs (Fig. 3C and fig. S7D), consistent with prior reports (*16*, *35*).

To confirm cross-reactivity of ACPA mAbs between citrullinated oral bacterial and citrullinated human antigens, we preincubated mAbs with either a citrullinated or native human antigen and then tested reactivity of the mAbs to citrullinated oral bacterial lysates. The citrullinated, but not the native, forms of human enolase and human vimentin inhibited binding to the citrullinated bacterial lysates (Fig. 3D and fig. S8, A and B). Together, these data demonstrate that antibodies encoded by ACPA-expressing plasmablasts cross-react against both in vitro citrullinated oral bacterial and human antigens.

The data presented above represent studies of individual B cell clones. To determine whether ACPAs in plasma and synovial fluid of RA patients cross-react to citrullinated bacterial antigens, we compared the binding of IgG antibodies derived from plasma (n = 37 RA versus n = 40 healthy; table S4) and synovial fluid [n = 25 RA versus 25 osteoarthritis (OA)] with native and in vitro citrullinated streptococcal species (Fig. 3, E and F). IgG isolated from RA plasma, but not healthy donor plasma, bound citrullinated but not native *Streptococcus oralis* and *Streptococcus parasanguinis* (Fig. 3E). Similarly, RA synovial IgG, but not OA synovial fluid IgG, bound citrullinated but not native *Streptococcus gordonii*, *S. oralis*, and *Streptococcus salivarius* (Fig. 3F). We conclude that RA patients harbor antibodies in both their plasma and synovial fluid that cross-react with in vitro citrullinated oral bacterial and human antigens.

To determine whether oral bacteria are targeted by ACPAs in vivo, we tested whether antibodies bound to oral bacteria, when eluted, were reactive for cyclic citrullinated peptide (CCP) using a clinical anti-CCP antibody assay (*36*). Antibodies bound to oral bacteria in RA patients had higher CCP reactivity in comparison with those in healthy donors (Fig. 3G), demonstrating that ACPAs are bound to oral bacteria in RA patients. Collectively, these data show that ACPAs cross-react with citrullinated human and oral bacterial antigens.

Oral bacteria are highly citrullinated in vivo

We next tested to what extent bacteria from various mucosal sites are citrullinated in vivo. Citrullination of bacteria from three sites (oral, fecal, and vaginal) was analyzed by flow cytometry using an anti-citrullinated peptide antibody (Fig. 4A). The proportion of citrullinated bacteria was highest in oral bacteria as compared with fecal and vaginal bacteria (Fig. 4, A and B).

Porphyromonas gingivalis encodes bacterial PAD and is enriched in the gingiva of patients with PD (*37*, *38*), whereas *Aggregatibacter actinomycetemcomitans* expresses leukotoxin A, which induces hypercitrullination of human proteins by forming pores in neutrophils (*39*). To determine whether specific bacteria are preferentially citrullinated, citrullinated bacteria and unbound bacteria were sorted, followed by 16*S* ribosomal RNA (rRNA) amplicon sequencing. No significant bacterial family enrichment was identified between the two fractions, suggesting broad and indiscriminate citrullination (Fig. 4, C and D, and fig. S9). The most abundant family of oral bacteria in the citrullinated fraction was the Streptococcaceae family, which does not express a PAD ortholog (*40*), suggesting that citrullination occurred from an exogenous source of PAD.

Because neutrophils represent >95% of immune cells in the oral cavity (*41*) and express robust levels of PAD2 and PAD4 (*42*), we hypothesized that human neutrophils may citrullinate a Streptococcaceae family member, *S. parasanguinis*. After 3 hours of culture with human neutrophils and physiologic levels of $CaCl_2$, the cofactor for human PAD2 and PAD4, there was an increase in the proportion of citrullinated live *S. parasanguinis*, as measured by flow cytometry (Fig. 4E). As demonstrated earlier, antibodies bound to oral bacteria contribute to ISG15 expression (Fig. 1K). To determine whether ACPA binding to citrullinated *S. parasanguinis* augmented ISG15 expression in blood monocytes, we stimulated blood monocytes with citrullinated, intact *S. parasanguinis* in the presence of ACPA IgG1 mAbs (Fig. 3A) or isotype control. Congruently, ACPA binding to the citrullinated *S. parasanguinis* increased ISG15 expression (Fig. 4F). In conclusion, the oral cavity represents a source of citrullinated commensal bacteria, which can be citrullinated by host neutrophils, thereby forming immune complexes with ACPAs to augment ISG15 expression in blood monocytes when translocating into circulation.

RA ACPAs bind in vivo citrullinated bacterial peptides detected by mass spectrometry

To identify the citrullinated epitopes of in vivo citrullinated bacteria, we analyzed a publicly available mass spectrometry dataset of 30 human saliva samples (*43*) from PD patients, healthy individuals, and individuals with dental caries and identified five bacterial-derived peptides with deiminated arginines (Fig. 5A, fig. S10, and table S5). We then tested our library of 21 ACPA mAbs for binding to these five citrullinated bacterial epitopes detected by mass spectrometry. Eight of the ACPA mAbs reacted strongly with one or more of the citrullinated bacterial epitopes, demonstrating that citrullinated bacterial epitopes present in human saliva are also recognized by ACPAs (Fig. 5B).

We next sought to assess whether the ACPA mAbs initially develop against citrullinated bacterial antigens or citrullinated human autoantigens. We reverted four of the ACPA mAbs that bound a citrullinated bacterial epitope and expressed their respective germline mAbs (Fig. 5, C and D, and fig. S11). In addition, we used IgTree to predict ancestral BCRs in B cell clonal families, and we expressed predicted ancestral mAbs, as well as the germline mAb, of RA78 and RA80 (Fig. 5E and fig. S11). For each ACPA mAb and the respective ancestral and germline mAbs, we measured the binding affinity to the citrullinated bacterial peptide and citrullinated human peptide that was most reactive via enzyme-linked immunosorbent assay (ELISA). For the ACPA mAbs that bound citrullinated molecular chaperone DnaK (87–106) derived from *A. actinomycetemcomitans*, the corresponding germline or ancestral antibody bound the citrullinated DnaK peptide (87–106) with low affinity and had no detectable binding to the citrullinated human peptide (Fig. 5, C to E, and fig. S11). Furthermore, these three ACPA mAbs had a higher or equivalent affinity for the citrullinated bacterial epitope as compared with the human epitope (Fig. 5, C to E, and fig. S11). There were also ACPA mAbs for which the ancestral germline mAb did not bind more strongly to the citrullinated bacterial epitope. Because the proteome of citrullinated antigens (both bacterial and human) is vast and not fully defined, only a subset of citrullinated antigens was tested in these experiments. Together, these results suggest that some ACPA B cells may originally develop against citrullinated bacterial antigens, followed by affinity

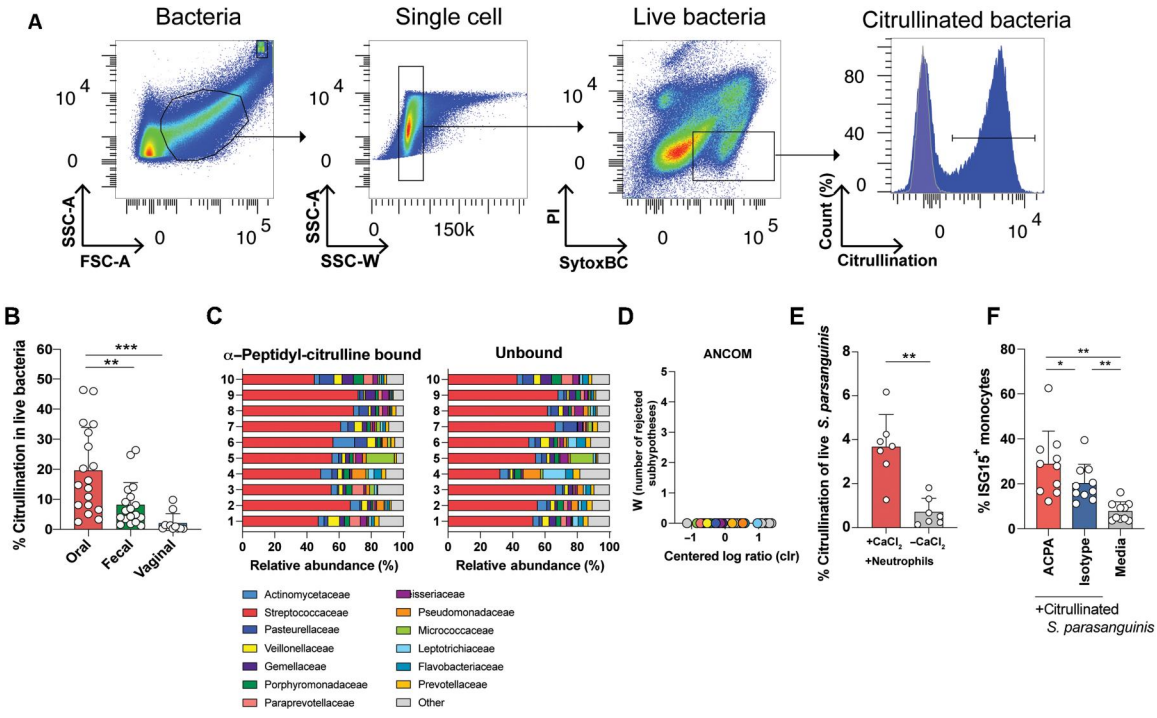

Fig. 4. Commensal bacteria from multiple mucosal sites are citrullinated in vivo. (**A**) Single-cell bacteria were analyzed by flow cytometry for citrullination and viability determined by SYTOX BC green, propidium iodide (PI), and anti–peptidyl-citrulline antibody staining. (**B**) Frequency of live citrullinated bacteria sampled from the mouth (n = 18), stool (n = 18), and vagina (n = 11). (**C**) Relative abundance of the 13 most abundant families in citrullinated (n = 10, left) and unbound (n = 10, right) oral sorted fractions. (**D**) Analysis of composition of microbiomes (ANCOM) of families in citrullinated and unbound oral sorted fractions. (**E**) Percentage of live citrullinated *S. parasanguinis* after culture with neutrophils (n = 7) with or without $CaCl_2$. (**F**) ISG15 expression by flow cytometry of blood monocytes stimulated with citrullinated *S. parasanguinis* and ACPA mAbs, citrullinated *S. parasanguinis* and hIgG1 isotype mAb, or PBS (n = 10). To produce citrullinated *S. parasanguinis*, *S. parasanguinis* was incubated in neutrophil conditioned medium with calcium and DTT. Results represent means ± SD. Statistical analysis for (D) was performed using ANCOM. One-way ANOVA followed by Dunnett's multiple comparisons test and paired t test was used for statistical analysis for (B) and (E), respectively. Repeated-measures one-way ANOVA followed by Tukey's multiple comparisons test was performed for (F). *P < 0.05, **P < 0.01, and ***P < 0.001.

maturation and epitope spreading, to bind citrullinated human antigens.

The most frequently detected citrullinated bacterial peptide in saliva in vivo was derived from *S. parasanguinis* amylase binding protein A (AbpA), which promotes bacterial binding to the tooth surface (*44*). To assess the generalizability of the citrullinated AbpA peptide as an antigenic target in other patients with RA, we compared IgG binding of plasma from RA (n = 46) and healthy patients (n = 89), as well as IgG binding of synovial fluid from RA (n = 31) and OA patients (n = 34) to citrullinated and native AbpA peptide. Plasma and synovial fluid from RA patients exhibited increased reactivity to citrullinated AbpA peptide when compared with those from non-RA controls (Fig. 5F). In addition, RA plasma absorbed against these three citrullinated bacterial peptides (citrullinated AbpA, citrullinated DnaK peptide, and citrullinated hypothetical protein) showed decreased reactivity to CCP, citrullinated vimentin, citrullinated enolase, and citrullinated histone 2A, demonstrating cross-reactivity (fig. S12). Furthermore, we measured the citrullinated AbpA antibody levels in healthy individuals with no or mild PD, patients with moderate to severe PD, and RA patients with moderate to severe PD. We found that RA patients with PD, but not those with only PD or healthy controls, harbored anti-citrullinated AbpA antibodies (Fig. 5G). Collectively, these data indicate that RA patients develop antibodies that bind citrullinated oral commensal bacterial proteins that are cross-reactive against known human citrullinated autoantigens.

DISCUSSION

Our data suggest that PD-associated oral mucosal breaks result in oral bacteremias that trigger innate and adaptive immune responses, which, when repeated over time in a susceptible host, likely contribute to the pathogenesis of RA. Oral bacteremias trigger activation of innate immune pathways, including inflammatory $ISG15^{+}$ $HLADR^{hi}$ monocytes, that are present in inflamed joints and the blood of patients with PD and RA flare. Whereas the innate interferon response to oral bacteremia is not specific to RA, antibodies to citrullinated oral bacteria (ACBA) are specific to RA. PD affects 47% of the population, whereas RA only affects 1%, indicating that additional host susceptibility factors must play a role (*2*, *45*). Our finding that some RA patients harbor B cells that bind citrullinated bacterial antigens and subsequently undergo affinity maturation and epitope spreading to bind citrullinated human antigen provides insight into the long-known association between PD and seropositive RA. Overall, the results presented here are consistent with a model where PD enables bacterial translocation that, in turn, stimulates innate and adaptive immune responses that contribute to RA pathogenesis.

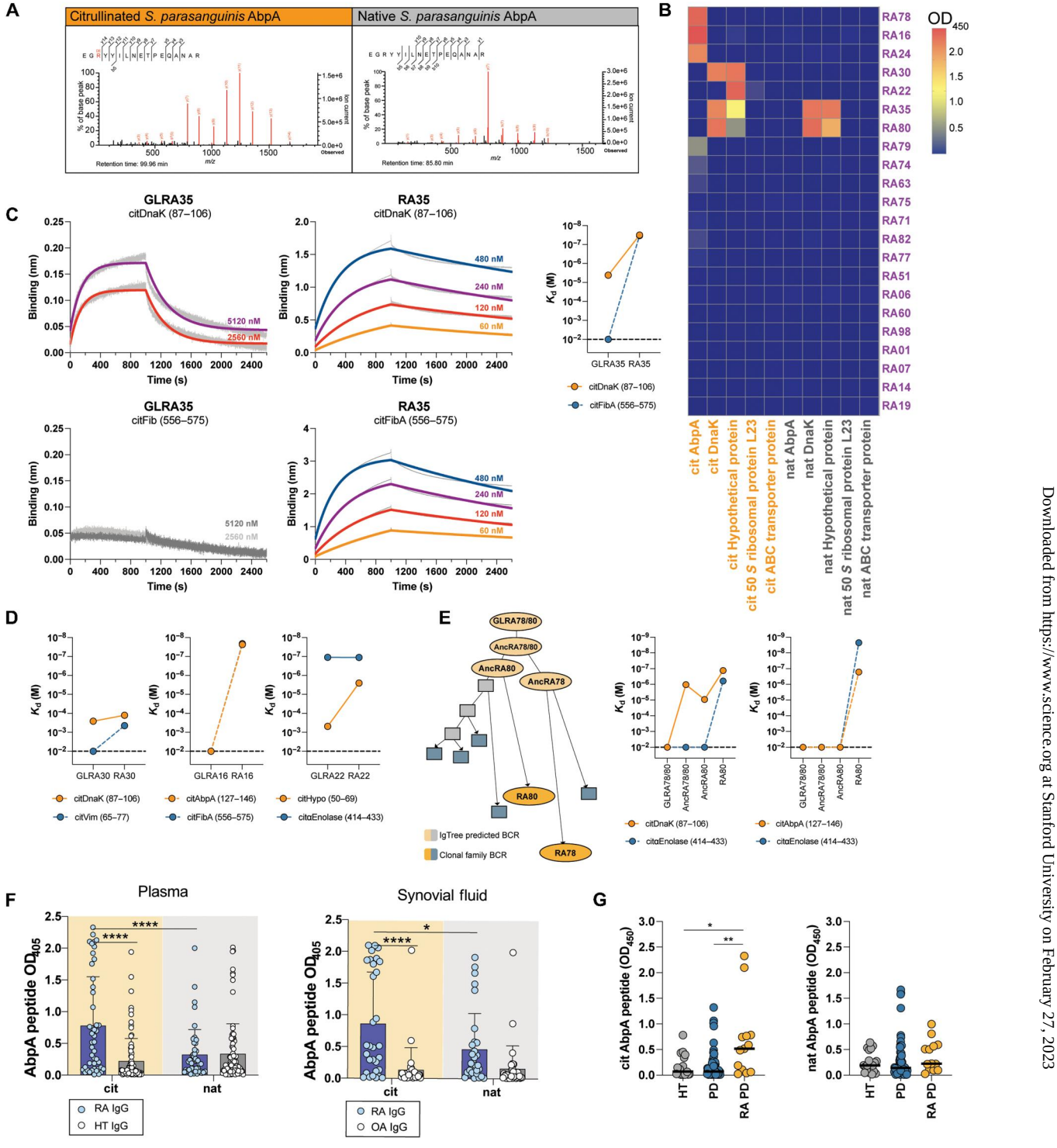

Fig. 5. Identification of citrullinated bacterial epitopes targeted by ACPA. (**A**) Mass spectrometry spectra and retention times of citrullinated and native *S. parasanguinis* amylase peptide from saliva. (**B**) Reactivity of RA mAbs to citrullinated and native bacterial peptides determined by ELISA. (**C**) Binding kinetics of RA35 and GLRA35 to citrullinated DnaK (87–106) and citrullinated fibrinogen (556–575). (**D**) K_d (dissociation constant) values derived from biolayer interferometry of germline (GL) and ACPA mAbs to citrullinated bacterial and human peptides. (**E**) Lineage tree analysis of ACPA clonal family as predicted by IgTree. K_d values derived from biolayer interferometry of corresponding mAbs to citrullinated bacterial and human peptides. (**F** and **G**) IgG reactivity of (F, left) plasma from RA patients ($n = 46$) and healthy patients ($n = 89$), (F, right), synovial fluid from RA patients ($n = 31$) and OA patients ($n = 34$) (right), and (G) plasma from individuals with no or mild periodontal disease (HT, $n = 20$), patients with moderate to severe periodontal disease (PD, $n = 60$), and RA patients with moderate to severe periodontal disease (RA PD, $n = 13$) against citrullinated and native AbpA (124–144) by ELISA. Mean ± SD. *$P < 0.05$, **$P < 0.01$, and ****$P < 0.0001$. For (F), two-way ANOVA followed by Dunnett's multiple comparisons test was performed. For (G), Kruskal-Wallis test followed by Dunnett's multiple comparisons test was performed.

Plasmablasts are a relatively rare B cell population in blood, and antigen-specific plasmablasts typically peak 6 or 7 days after vaccination and then drop to baseline within 2 to 3 weeks (*46*, *47*). We previously observed shared clonal lineages of IgA^+ and IgG^+ ACPA-expressing plasmablasts in RA blood for up to 12 months, far beyond the established duration of an acute immune response (*48*). Although we have specifically focused on IgG effector functions, given the numerous studies that implicate Fcγ receptor activation in RA (*49*, *50*), the presence of IgA isotype in the ACPA-expressing clonal families is suggestive of a mucosal drive. Nevertheless, it is possible that IgA and IgG ACPAs contribute to inflammation in RA by forming an immune complex with citrullinated bacterial antigens that activate their respective pro-inflammatory Ig receptors.

Our data show that RA ACPA-expressing plasmablasts exhibit extensive SHM that is greater than typical antibody responses to vaccines (*32*). Furthermore, we demonstrated that this extensive SHM gives rise to ACPAs that cross-bind both citrullinated oral bacterial antigens and citrullinated human antigens. Because B cells accumulate on average one nucleotide mutation per cycle of activation and affinity maturation in the germinal center (*51*), the extensive mutation burden of ACPA plasmablasts in RA suggests that they may have been repeatedly restimulated by the recurrent mucosal breaches of citrullinated oral bacteria. For the three mAbs that bound the bacterial citrullinated molecular chaperone DnaK (87–106), the inferred germline or ancestral antibodies that bound the bacterial citrullinated molecular chaperone DnaK (87–106) with low affinity had no detectable binding to citrullinated human antigens. Thus, we propose that PD-mediated repeated oral bacteremias stimulate recurrent rounds of ACPA-expressing B cell affinity maturation and SHM, eventually resulting in epitope spread of the ACPA response to target both citrullinated bacterial and human antigen(s).

We additionally found that non–PAD-expressing oral commensal bacteria are broadly citrullinated in vivo in individuals with PD. Culture of neutrophils with non–PAD-expressing *Streptococcus* species (*40*) led to bacterial citrullination, indicating that human-derived PADs released from NETosing neutrophils can mediate citrullination of oral bacteria. Our finding that ACPA-expressing plasmablasts target citrullinated commensal oral bacteria that do not necessarily express endogenous PAD(s) suggests that human neutrophil PADs mediate the generation of citrullinated bacterial epitopes, a result that we confirmed with experiments in vitro.

Entry of oral bacteria into blood was associated with innate immune and defense pathways, indicating a subclinical response to oral bacteremias in RA patients with PD. Several groups have demonstrated that oral dysbiosis in RA is characterized by enrichment of *Streptococcus* species (*52*) and that specific isolates of *S. parasanguinis* can induce arthritis in murine models (*53*). In PD, oral bacteria released into the blood may disseminate to the liver, spleen, joints, and gut (*54*–*56*). Congruently, recent studies have reported an enrichment of oral *Streptococcus* species in the guts of RA patients (*57*, *58*) that correlates with disease activity, further supporting hematogenous transmission of oral bacteria to distant body sites.

The temporal association of oral bacteremia with $ISG15^{+-}$ HLA^{high} monocyte gene expression suggests a mechanism by which these monocytes are activated by citrullinated oral bacteria alone or in ACPA-bound immune complexes. In RA patients with PD, the gene signatures of inflammatory $S100A12^+$ and $ISG15^+$ synovial macrophages were enriched during flares in RA patients with PD. Given that these gene signatures partially overlap, the $ISG15^+$ macrophage subset may represent a recently activated monocyte subset that trafficked from blood and further differentiated into a more inflammatory subset in the synovial tissue. The inflammatory monocyte analogous to the $ISG15^+HLA^{high}$ macrophage was only enriched in the flares of RA patients with PD, indicating that there is variability in the immune signature of flares, potentially due to distinct triggers. Given the heterogeneous nature of RA, oral bacteremias may trigger a subset of RA arthritis flares, particularly for RA patients with PD. Future efforts will focus on elucidating additional mechanisms of flare in patients without oral bacteremia or PD.

There are several limitations of our approach. It was not possible to know whether the bacterial reads that we identified in blood samples were intracellular, extracellular, or exosomal. The majority of the samples used in our study were from individuals with established RA on disease-modifying anti-rheumatic drugs and/or other biologics, and for some RA patients included in this study, their PD statuses were unknown. Further studies involving samples from at-risk individuals and medication-naïve individuals are needed to understand the impact of medication on RA progression related to oral bacterial translocation. In addition, given that ACPAs can be reactive against several citrullinated antigens and a vast number of citrullinated epitopes, we cannot precisely determine the driving epitope for each BCR with absolute certainty. Last, the proteomic analysis likely did not identify all bacterial citrullinated proteins present in the oral cavity given that >95% of proteins were human, and, thus, those identified only represent a fraction of citrullinated bacterial peptides present.

In summary, the findings described here indicate that PD, through repeated mucosal breaks, results in recurring innate and adaptive immune activation that may contribute to the pathogenesis of RA. Our findings suggest that future studies are needed to determine whether improved oral care may provide therapeutic benefit in the management of RA.

MATERIALS AND METHODS

Study design

The objectives of this study were to investigate the relationship between the oral microbiome and RA flares using finger-stick RNA-seq monitoring and B cell repertoire sequencing. The study used cohorts from Rockefeller University, University of Colorado, and Stanford University/VA Palo Alto.

Rockefeller University longitudinal cohort

Five female patients who met the ACR 2010 classification criteria (*9*) for RA and were seropositive for CCP^+ were followed for a minimum of 1 year with weekly finger stick samples and questionnaires to capture RAPID3 scores and changes in medication and dental work (table S1). They were also assessed with monthly clinic visits that included physical exams with assessment of tenderness and swelling in 28 joints. Additional samples were collected during self-reported flare events. This study was approved by the Rockefeller Investigational Review Board (IRB) (#DOR-0833).

Stanford University cohorts

Studies below were approved by the Stanford University IRB (#3780).

Plasmablasts. Plasmablasts were isolated from blood samples collected from 12 donors recruited at the VA Palo Alto who met 1987 American College of Rheumatology (ACR) criteria (*59*) and who were positive for anti-CCP antibodies (tables S2 and S3).

Plasma. Plasma samples were collected from individuals with RA who met the 1987 ACR criteria (*59*) and who were positive for anti-CCP antibodies, as well as from healthy donors (table S4).

Synovial fluid. Synovial fluids from 65 RA and OA patients as determined by board-certified rheumatologists were provided by the Stanford Rheumatic Diseases Biorepository.

University of Colorado PD cohort

Plasma samples from RA individuals who met ACR 2010 classification criteria for RA (*57*) or from normal donors were collected (table S6). The gingival health of all individuals was classified according to Centers for Disease Control and Prevention and American Academy of Periodontology case definitions (*60*). This study was approved by the Colorado Multiple Institute Review Board (#15-2288).

Clinical phenotype data and RNA-seq of fingerstick blood specimens from RA patients

The five female patients in the Rockefeller University longitudinal cohort self-collected finger stick blood samples. RNA-seq samples (N = 336) (mean n = 67 per patient) were available for the analysis comparing microbial and human gene expression (Fig. 1). RNA was purified using the Globin-Zero Kit (EpiCentre), cDNA libraries were prepared using the Illumina Truseq mRNA Stranded Library Kit, with 11 to 12 polymerase chain reaction (PCR) cycles for 5 to 8 nM input, and sequencing was performed on a HiSeq2500 system (Illumina) or a NovaSeq 6000 (Illumina) with 150–base pair paired-end reads. Transcript abundance was quantitated with salmon using an hg38 ensGene gene model.

Shotgun metatranscriptomic microbial assignments

Host-depleted and quality-controlled output fastq files were then uploaded to Qiita web server (*61*) (project ID 13456) for per-sample metatranscriptomic microbial classification. Qiita offers a graphical user interface that easily facilitates shotgun metatranscriptomic and/or metagenomic analysis based on Woltka v0.1.1 (*62*) and its associated "Web of Life" database containing 10,575 microbial genomes of bacteria and archaea (*21*) through direct genome alignments. A total of 371 blood-derived RA samples were aligned against the WoL reference genome database using SHOGUN v1.0.8 (*63*), with Bowtie2 v2.4.1 (*64*) as the back end. This process is equivalent to a Bowtie2 run with the following parameters: *--very-sensitive -k 16 --np 1 --mp "1,1" --rdg "0,1" --rfg "0,1" --score-min "L,0,-0.05."* The sequence alignment is treated as a mapping from queries (sequencing data) to subjects (microbial reference genomes). Reads mapped to a microbial reference genome are counted as hits such that the resultant feature table comprises samples (rows) by microbial genome IDs (columns) and concomitant abundances. These microbial genome IDs ["operational genomic units" (OGUs)] provide a shotgun metagenomic equivalent to amplicon sequence variants in 16*S* rRNA amplicon sequencing data (*62*). In the case that one sequence is mapped to multiple genomes by Bowtie2 (up to 16), each genome is counted $1/k$ times, where k is the number of genomes to which this sequence is mapped. The frequencies of individual genomes were then summed after the entire alignment was processed and rounded to the nearest even integer, thereby making the sum of OGU frequencies per sample nearly equal (considering rounding) to the number of aligned sequences in the dataset. The resultant count matrix was saved as a biom file for downstream analyses. In addition, 217 shotgun metagenomic samples from the HMP (*65*) that were previously used for source tracking against low biomass data (*66*) were processed through the same Woltka-WoL OGU pipeline as described above.

Microbial source tracking

Using the RA metatranscriptomic and HMP metagenomic data with annotated HMP body site data, microbial source tracking was performed using a Bayesian source tracking model, SourceTracker2 (https://github.com/biota/sourcetracker2) (*23*). Details of the Bayesian model have been previously described by our laboratory (*66*). Because metatranscriptomic data were also available for HMP fecal samples (*25*), we compared whether HMP fecal metatranscriptomic samples would be assigned as similar or equivalent using SourceTracker2 (fig. S2). On the basis of the near-perfect association between HMP metagenomic and metatranscriptomic data (fig. S3C), we proceeded with using the HMP metagenomic data for source tracking with the RA metatranscriptomic data. Using SourceTracker parlance, the HMP samples served as "sources," whereas the RA samples acted as "sinks," and the SourceTracker2 algorithm was used to calculate the proportion of each source attributable to each sink. In lay terms, we estimated the proportion of body site from HMP data attributable to each RA microbiome sample using the Bayesian model. After (i) intersecting the OGUs in our RA microbiome dataset with those in HMP and (ii) converting the data to BIOM table format (*67*), we applied the SourceTracking2 approach to all RA samples. SourceTracker2 default settings (alpha1 = 0.001, alpha2 = 0.1, beta = 10, restarts = 10, draws_ per_-restart = 1, burnin = 100, delay = 1) plus a sink rarefaction depth of 9000 were used. The outputs were calculated in terms of mean fractional contributions of each source to each sink.

Differential gene expression and pathway analysis for oral bacteremias

Differential gene expression was performed with limma (*68*). Genes with low expression were filtered out using the filterByExpr command in the edgeR R package (*60*). To remove batch effects and unwanted variation due to low RNA yields, the batch and starting RNA amounts were included in the linear model. Normalization factors were calculated using the "TMM" method, and voomWithQualityWeights was used before fitting a linear model with lmFit and eBayes. Genes with an adjusted P value (false discovery rate, using the Benjamini and Hochberg correction) less than 0.01 were considered significant. SourceTracker2-predicted relative contributions or log ratios (described below) from microbial read counts were used as a continuous variable. For Gene Ontology (GO) pathway analyses, goana function in edgeR was used (*69*). Gene set enrichment analysis was used for interpreting expression profiles (*70*).

RT-qPCR of stimulated whole blood and isolated immune cell populations

To isolate monocytes and lymphocytes, frozen peripheral blood mononuclear cells (PBMCs) were thawed and washed with 10% fetal bovine serum (FBS) (Corning) in RPMI 1640 (Thermo

Fisher Scientific) and resuspended in deoxyribonuclease in 10% FBS in RPMI 1640 to improve viability. PBMCs were washed and layered over the Percoll gradient (GE) to separate monocytes and lymphocytes. Monocyte-enriched buffy coat was spun down and resuspended in EasySep buffer (STEMCELL Technologies). Monocytes were then isolated using the EasySep Human Monocyte Isolation Kit (STEMCELL Technologies) as per the manufacturer's instructions. Neutrophils were isolated as described above. Isolated cells were checked by flow cytometry for >95% purity. The oral bacteria were isolated directly from the mouths of healthy individuals and frozen at −80°C before the experiment. For all of the experiments, a pool of oral bacteria from six healthy donors (combined and aliquoted before freezing) was used to stimulate the monocytes or whole blood. All monocytes, lymphocytes, or neutrophils were resuspended at 7×10^6 cells per ml and mixed with oral bacteria or medium. After incubation for 6 hours, total RNA was isolated using an RNeasy Plus Mini or Micro Kit according to the manufacturer's instructions. For experiments with whole blood, whole blood (2.5 ml) was incubated with oral bacteria (~200,000 bacteria cells) isolated from the oral cavity or mock-stimulated for 0.25 to 20 hours with orbital shaking. At designated time points, blood was added to a PAXgene Blood RNA tube (BD Biosciences) and incubated at room temperature for 2 to 24 hours. Total RNA was purified from the whole blood using the PAXgene Blood RNA Kit (Qiagen) as per the manufacturer's instructions. Total RNA was reverse-transcribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Reverse transcription quantitative PCR (RT-qPCR) was performed using TaqMan Gene Expression Master Mix and predesigned TaqMan qPCR probes (Thermo Fisher Scientific) on a QuantStudio 7 Flex Real-Time PCR Instrument (Thermo Fisher Scientific). Relative mRNA expression was normalized to the levels of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and expressed relative to the values for control cells using the ΔΔCt method.

ISG15 expression in blood monocytes

Blood monocytes were isolated from frozen PBMCs using the EasySep Human Monocyte Isolation Kit (STEMCELL Technologies) as per the manufacturer's instructions. Oral bacteria were isolated from the oral cavity as described above. Monocultured *S. parasanguinis* was grown and incubated with neutrophil-conditioned medium with $CaCl_2$ and dithiothreitol (DTT) for 3 hours before producing citrullinated *S. parasanguinis*. ACPA mAbs or isotype control mAb was incubated with citrullinated *S. parasanguinis* before incubation with monocytes. Monocytes were resuspended at a concentration of 4×10^6 cells per ml in 10% FBS (Corning), RPMI 1640 (Thermo Fisher Scientific), 25 mM Hepes (VWR), and 1× penicillin-streptomycin, and 100 μl of cells were mixed with 50 μl of oral bacteria or citrullinated *S. parasanguinis*. In certain experiments, anti-CD32A (STEMCELL Technologies) was added. Cells were incubated for 16 to 20 hours. Cells were washed and stained using CD14 allophycocyanin (APC)–Cy7 (BioLegend), AquaZombie (BioLegend), CD45 Brillant Violet 711 (BioLegend), and ISG15 Alexa 488 (R&D Systems) according to the manufacturer's instructions. Cells were analyzed on a Fortessa flow cytometer (BD Biosciences). Data analysis was performed using the FlowJo software.

Cell barcode–enabled sequencing of plasmablasts

Sequencing of Ig genes from individual plasmablasts was performed using cell barcodes as described previously (*48*, *71*–*73*). Briefly, $CD19^+CD3^-IgD^-CD14^-CD20^-CD27^+CD38^{++}$ plasmablasts were single cell–sorted using BD FACSAria3.1 or BD FACSAria3.2 (BD Biosciences) directly into lysis buffer in 96-well plates. RT of RNA with oligo-dT was carried out in separate wells, and unique well-ID barcodes were attached via template switching activity of Maxima Reverse Transcriptase (Thermo Fisher Scientific). Barcoded cDNA from each plate was pooled, and heavy chain (IgA, IgG, and IgM) and light chain (kappa and lambda) were amplified in three rounds of PCRs while attaching plate-specific barcodes and sequencing adapters. PCRs were carried out separating heavy chain of IgA, IgG, and IgM as well as kappa and lambda light chain, and separate libraries were generated from each, gel-purified, cleaned with AMPure XP beads (Beckman Coulter), and sequenced on Illumina MiSeq (Illumina) with 2 × 300 paired-end reads.

Bioinformatic analysis of Ig sequences

Sequencing data were processed as previously described (*48*, *71*–*73*). Briefly, FASTQ files were demultiplexed using the MiSeq FASTQ workflow. Poor-quality reads and bases were trimmed, and the paired reads were stitched together. Similar reads sharing the same plate and well IDs were clustered into operational taxonomic units (*74*). Heavy chain VDJ genes and light VJ chain were aligned to germline sequences using ImMunoGeneTics (IMGT) HighV-Quest (*75*). V-region SHMs were calculated using IMGT analysis based on comparison with germline V-region genes. SHM frequency was calculated by dividing the number of mutations by the length of the V-region of the corresponding chain. When comparing antigen-specific plasmablast sequences, the SHM frequency full V-region was compared. N-glycan sites in the heavy and light chain variable regions were predicted by NetNGlyc 1.0 (*76*).

Generation of recombinant mAbs

Previously published plasmablast antibody sequences from RA patients (*48*, *71*, *72*) were recombinantly produced in human IgG1 backbone to ensure consistency in the characterization assays. In-house production was done as described previously, using the Expi293 Expression System (Thermo Fisher Scientific) with Expi293F cells (*48*, *71*–*73*). Briefly, constructs including the heavy chain and light chain variable region sequences were synthesized as gBlock gene fragments (IDT) for cloning into pFUSE antibody plasmids (InvivoGen) using the Cold Fusion Cloning Kit (System Biosciences). We used pFUSEss-CHIghIg1 for gamma, pFUSE2ss-CLIg-hK for kappa, and pFUSE2ss-CLIg-hL2 for lambda. The Expi293 Expression System (Thermo Fisher Scientific) was used for transient transfections, and harvested culture supernatants were purified using Pierce Protein A Plus Agarose (Thermo Fisher Scientific).

Bacterial culture and preparation

Bacteria strains were grown to mid-logarithmic growth phase (0.4 to 0.9 optical density units), as determined spectrophotometrically at 600 nm, in American Type Culture Collection– recommended media and conditions (table S7). The bacteria were pelleted by spinning at 6000*g* for 10 min and then washed in phosphate-buffered saline (PBS) before storage at −20°C.

Bacterial flow cytometry and sorting

Oral "brushings" were collected using BBL culture swabs (BD Biosciences) to brush all teeth above the gingiva for about 2 min according to Stanford University IRB (#3780). Oral samples, healthy vaginal fluid (Lee BioSolutions), and healthy fecal swabs (Lee BioSolutions) were stored at −80°C. One oral swab, vaginal swab, or fecal swab was placed in staining buffer, PBS with 1% bovine serum albumin (BSA; Sigma-Aldrich) on ice for 10 min with vortexing. Each sample was spun down (8000*g*, 10 min, 4°C) and resuspended in 20% normal mouse serum (Jackson ImmunoResearch) in staining buffer for 20 min on ice before labeling with anti-citrulline mAb (clone 1D9, Cayman Chemical). The sample was washed after the primary staining and resuspended in phycoerythrin-conjugated F(ab′)2 donkey anti-mouse (H+L) (Jackson ImmunoResearch) and SYTO BC dye (Thermo Fisher Scientific) for 30 min at 4°C. The bacteria were washed twice and resuspended in 100 μl of staining buffer containing propidium iodide (Sigma-Aldrich) and 6-μm beads (Thermo Fisher Scientific). Bacteria were run on an LSRII or FACSAria cytometer and sorted directly into DNA/RNA Shield (Zymo Research). Samples with low bacteria concentrations were not included in the analysis.

Citrullination of bacteria via neutrophils

Neutrophils were isolated from freshly drawn peripheral blood from healthy donors as previously described (*77*). Briefly, neutrophils and red blood cells (RBCs) were isolated from PBMC via Ficoll gradient using Ficoll-Paque Plus (GE). The neutrophil and RBC pellet was diluted in 3% dextran in Hanks′ balanced salt solution (HBSS; Corning). After a 20-min incubation, the neutrophil-rich supernatant was spun down at 300*g* with no brake. Remaining RBCs were lysed using RBC lysis buffer (BioLegend). After washing, neutrophils were resuspended to a concentration of 2.5×10^7 cells per ml. *S. parasanguinis* was grown to mid-logarithmic growth phase and then washed three times in PBS. After the washes, *S. parasanguinis* was labeled with CellTrace Far Red (Thermo Fisher Scientific) and resuspended at a concentration of 2.5×10^8 cells per ml. Freshly isolated neutrophils were added to *S. parasanguinis* for a final concentration of 10^7 cells per ml of neutrophils and 10^8 cells per ml of *S. parasanguinis* in 1 mM DTT (Sigma-Aldrich), 2 mM $CaCl_2$ (specific conditions), and HBSS. After a 3-hour incubation, citrullination levels of *S. parasanguinis* were checked by flow as described above.

Bio-layer interferometry and thermostability assay

Antibody binding to peptide was assessed using ForteBio Octet KQ. The binding assays were performed at 30°C and with agitation set at 1000 rpm. Biotinylated peptides were loaded on High Precision Streptavidin Biosensors (ForteBio) at 100 nM concentration in 1× kinetic buffer (ForteBio). The sensors were then quenched using biocytin (Thermo Fisher Scientific). Antibody concentrations ranging from 10 to 8000 nM were measured for each peptide.

Recombinant peptide ELISA

For experiments with the biotinylated peptides, Nunc Immobilizer Streptavidin 384 plates (Thermo Fisher Scientific) were loaded with 1 μM of biotinylated peptide (table S8) for 1 hour at room temperature. Plates were blocked with 50% ChonBlock (Chondrex) in PBS (Chondrex) or 2% BSA in PBS for 1 hour at room temperature. After blocking, plasma or synovial fluid was diluted at 1:100 concentration or to 1 μg/ml for mAbs and was incubated with shaking at room temperature for 1.5 hours. Dilution series of a positive control serum was included on each plate to ensure consistency between plates. Antibody binding was detected using a horseradish peroxidase–conjugated goat anti-human IgG Fc antibody and Super AquaBlue ELISA substrate. Five washes with PBS–Tween 20 were performed between each step. Blank wells were subtracted for analysis.

Supplementary Materials

This PDF file includes:
Materials and Methods
Figs. S1 to S12
Tables S1 to S8
References (*78*–*86*)

Other Supplementary Material for this manuscript includes the following:
Data files S1 to S6
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. P. de Pablo, T. Dietrich, T. E. McAlindon, Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population. *J. Rheumatol.* **35**, 70–76 (2008).
2. P. I. Eke, B. A. Dye, L. Wei, G. O. Thornton-Evans, R. J. Genco; CDC Periodontal Disease Surveillance workgroup: James Beck (University of North Carolina, Chapel Hill, USA), Gordon Douglass (Past President, American Academy of Periodontology), Roy Page (University of Washington, Seattle, USA), Gary Slade (University of North Carolina, Chapel Hill, USA), George W. Taylor (University of Michigan, Ann Arbor, USA), Wenche Borgnakke (University of Michigan, Ann Arbor, USA), and representatives of the American Academy of Periodontology, Prevalence of periodontitis in adults in the United States: 2009 and 2010. *J. Dent. Res.* **91**, 914–920 (2012).
3. P. B. Lockhart, M. T. Brennan, M. Thornhill, B. S. Michalowicz, J. Noll, F. K. Bahrani-Mougeot, H. C. Sasser, Poor oral hygiene as a risk factor for infective endocarditis-related bacteremia. *J. Am. Dent. Assoc.* **140**, 1238–1244 (2009).
4. T. R. Mikuls, J. B. Payne, F. Yu, G. M. Thiele, R. J. Reynolds, G. W. Cannon, J. Markt, D. McGowan, G. S. Kerr, R. S. Redman, A. Reimold, G. Griffiths, M. Beatty, S. M. Gonzalez, D. A. Bergman, B. C. Hamilton 3rd, A. R. Erickson, J. Sokolove, W. H. Robinson, C. Walker, F. Chandad, J. R. O'Dell, Periodontitis and *Porphyromonas gingivalis* in patients with rheumatoid arthritis. *Arthritis Rheumatol.* **66**, 1090–1100 (2014).
5. C. Kaneko, T. Kobayashi, S. Ito, N. Sugita, A. Murasawa, K. Nakazono, H. Yoshie, Circulating levels of carbamylated protein and neutrophil extracellular traps are associated with periodontitis severity in patients with rheumatoid arthritis: A pilot case-control study. *PLOS ONE* **13**, e0192365 (2018).
6. D. A. González, M. L. Bianchi, P. A. Salgado, M. Armada, S. Seni, C. A. Isnardi, G. Citera, T. Ferrary, B. Orman, Disease activity and subcutaneous nodules are associated to severe periodontitis in patients with rheumatoid arthritis. *Rheumatol. Int.* **42**, 1331–1339 (2022).
7. B. Möller, P. Bender, S. Eick, S. Kuchen, A. Maldonado, J. Potempa, S. Reichenbach, A. Sculean, A. Schwenzer, P. M. Villiger, A. Wong, K. S. Midwood, Treatment of severe periodontitis may improve clinical disease activity in otherwise treatment-refractory rheumatoid arthritis patients. *Rheumatology* **59**, 243–245 (2020).
8. J. Sokolove, R. Bromberg, K. D. Deane, L. J. Lahey, L. A. Derber, P. E. Chandra, J. D. Edison, W. R. Gilliland, R. J. Tibshirani, J. M. Norris, V. M. Holers, W. H. Robinson, Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. *PLOS ONE* **7**, e35296 (2012).
9. D. Aletaha, T. Neogi, A. J. Silman, J. Funovits, D. T. Felson, C. O. Bingham III, N. S. Birnbaum, G. R. Burmester, V. P. Bykerk, M. D. Cohen, B. Combe, K. H. Costenbader, M. Dougados, P. Emery, G. Ferraccioli, J. M. W. Hazes, K. Hobbs, T. W. J. Huizinga, A. Kavanaugh, J. Kay, T. K. Kvien, T. Laing, P. Mease, H. A. Ménard, L. W. Moreland, R. L. Naden, T. Pincus, J. S. Smolen, E. Stanislawska-Biernat, D. Symmons, P. P. Tak, K. S. Upchurch, J. Vencovský, F. Wolfe, G. Hawker, 2010 rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum.* **62**, 2569–2581 (2010).

10. A. Sohrabian, L. Mathsson-Alm, M. Hansson, A. Knight, J. Lysholm, M. Cornillet, K. Skriner, G. Serre, A. Larsson, T. Weitoft, J. Rönnelid, Number of individual ACPA reactivities in synovial fluid immune complexes, but not serum anti-CCP2 levels, associate with inflammation and joint destruction in rheumatoid arthritis. *Ann. Rheum. Dis.* **77**, 1345–1353 (2018).
11. J. Haschka, M. Englbrecht, A. J. Hueber, B. Manger, A. Kleyer, M. Reiser, S. Finzel, H.-P. Tony, S. Kleinert, M. Feuchtenberger, M. Fleck, K. Manger, W. Ochs, M. Schmitt-Haendle, J. Wendler, F. Schuch, M. Ronneberger, H.-M. Lorenz, H. Nuesslein, R. Alten, W. Demary, J. Henes, G. Schett, J. Rech, Relapse rates in patients with rheumatoid arthritis in stable remission tapering or stopping antirheumatic therapy: Interim results from the prospective randomised controlled RETRO study. *Ann. Rheum. Dis.* **75**, 45–51 (2016).
12. S. Ajeganova, T. Huizinga, Sustained remission in rheumatoid arthritis: Latest evidence and clinical considerations. *Ther. Adv. Musculoskelet. Dis.* **9**, 249–262 (2017).
13. S. Khatri, J. Hansen, K. Astakhova, Antibodies to synthetic citrullinated peptide epitope correlate with disease activity and flares in rheumatoid arthritis. *PLOS ONE* **15**, e0232010 (2020).
14. S. Rantapää-Dahlqvist, B. A. W. de Jong, E. Berglin, G. Hallmans, G. Wadell, H. Stenlund, U. Sundin, W. J. van Venrooij, Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum.* **48**, 2741–2749 (2003).
15. H. Kokkonen, M. Mullazehi, E. Berglin, G. Hallmans, G. Wadell, J. Rönnelid, S. Rantapää-Dahlqvist, Antibodies of IgG, IgA and IgM isotypes against cyclic citrullinated peptide precede the development of rheumatoid arthritis. *Arthritis Res. Ther.* **13**, R13 (2011).
16. R. D. Vergroesen, L. M. Slot, L. Hafkenscheid, M. T. Koning, E. I. H. van der Voort, C. A. Grooff, G. Zervakis, H. Veelken, T. W. J. Huizinga, T. Rispens, H. U. Scherer, R. E. M. Toes, B-cell receptor sequencing of anti-citrullinated protein antibody (ACPA) IgG-expressing B cells indicates a selective advantage for the introduction of *N*-glycosylation sites during somatic hypermutation. *Ann. Rheum. Dis.* **77**, 956–958 (2018).
17. L. Hafkenscheid, E. Moel, I. Smolik, S. Tanner, X. Meng, B. C. Jansen, A. Bondt, M. Wuhrer, T. W. J. Huizinga, R. E. M. Toes, H. El-Gabalawy, H. U. Scherer, *N*-linked glycans in the variable domain of IgG anti–citrullinated protein antibodies predict the development of rheumatoid arthritis. *Arthritis Rheumatol.* **71**, 1626–1633 (2019).
18. T. Kissel, L. Hafkenscheid, T. J. Wesemael, M. Tamai, S.-Y. Kawashiri, A. Kawakami, H. S. El-Gabalawy, D. van Schaardenburg, S. Rantapää-Dahlqvist, M. Wuhrer, A. H. M. van der Helm-van Mil, C. F. Allaart, D. van der Woude, H. U. Scherer, R. E. M. Toes, T. W. J. Huizinga, IgG anti-citrullinated protein antibody variable domain glycosylation increases before the onset of rheumatoid arthritis and stabilizes thereafter: A cross-sectional study encompassing ~1,500 samples. *Arthritis Rheumatol.* **74**, 1147–1158 (2022).
19. D. Schneider, M. Dühren-von Minden, A. Alkhatib, C. Setz, C. A. M. van Bergen, M. Benkißer-Petersen, I. Wilhelm, S. Villringer, S. Krysov, G. Packham, K. Zirlik, W. Römer, C. Buske, F. K. Stevenson, H. Veelken, H. Jumaa, Lectins from opportunistic bacteria interact with acquired variable-region glycans of surface immunoglobulin in follicular lymphoma. *Blood* **125**, 3287–3296 (2015).
20. F. S. Ielasi, M. Alioscha-Perez, D. Donohue, S. Claes, H. Sahli, D. Schols, R. G. Willaert, Lectin-glycan interaction network-based identification of host receptors of microbial pathogenic adhesins. *MBio* **7**, (2016).
21. Q. Zhu, U. Mai, W. Pfeiffer, S. Janssen, F. Asnicar, J. G. Sanders, P. Belda-Ferre, G. A. Al-Ghalith, E. Kopylova, D. McDonald, T. Kosciolek, J. B. Yin, S. Huang, N. Salam, J.-Y. Jiao, Z. Wu, Z. Z. Xu, K. Cantrell, Y. Yang, E. Sayyari, M. Rabiee, J. T. Morton, S. Podell, D. Knights, W.-J. Li, C. Huttenhower, N. Segata, L. Smarr, S. Mirarab, R. Knight, Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. *Nat. Commun.* **10**, 5477 (2019).
22. N. M. Davis, D. M. Proctor, S. P. Holmes, D. A. Relman, B. J. Callahan, Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).
23. D. Knights, J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight, S. T. Kelley, Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763 (2011).
24. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
25. J. Lloyd-Price, C. Arze, A. N. Ananthakrishnan, M. Schirmer, J. Avila-Pacheco, T. W. Poon, E. Andrews, N. J. Ajami, K. S. Bonham, C. J. Brislawn, D. Casero, H. Courtney, A. Gonzalez, T. G. Graeber, A. B. Hall, K. Lake, C. J. Landers, H. Mallick, D. R. Plichta, M. Prasad, G. Rahnavard, J. Sauk, D. Shungin, Y. Vázquez-Baeza, R. A. White III; IBDMDB Investigators, J. Braun, L. A. Denson, J. K. Jansson, R. Knight, S. Kugathasan, D. P. B. McGovern, J. F. Petrosino, T. S. Stappenbeck, H. S. Winter, C. B. Clish, E. A. Franzosa, H. Vlamakis, R. J. Xavier, C. Huttenhower, Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019).
26. H. Lin, S. D. Peddada, Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* **11**, 3514 (2020).
27. Á. Simón-Soro, G. D'Auria, M. C. Collado, M. Džunková, S. Culshaw, A. Mira, Revealing microbial recognition by specific antibodies. *BMC Microbiol.* **15**, 132 (2015).
28. D. E. Orange, V. Yao, K. Sawicka, J. Fak, M. O. Frank, S. Parveen, N. E. Blachere, C. Hale, F. Zhang, S. Raychaudhuri, O. G. Troyanskaya, R. B. Darnell, RNA identification of PRIME cells predicting rheumatoid arthritis flares. *N. Engl. J. Med.* **383**, 218–228 (2020).
29. S. Alivernini, L. MacDonald, A. Elmesmari, S. Finlay, B. Tolusso, M. R. Gigante, L. Petricca, C. Di Mario, L. Bui, S. Perniola, M. Attar, M. Gessi, A. L. Fedele, S. Chilaka, D. Somma, S. N. Sansom, A. Filer, C. McSharry, N. L. Millar, K. Kirschner, A. Nerviani, M. J. Lewis, C. Pitzalis, A. R. Clark, G. Ferraccioli, I. Udalova, C. D. Buckley, E. Gremese, I. B. McInnes, T. D. Otto, M. Kurowska-Stolarska, Distinct synovial tissue macrophage subsets regulate inflammation and remission in rheumatoid arthritis. *Nat. Med.* **26**, 1295–1306 (2020).
30. J. T. Morton, A. A. Aksenov, L. F. Nothias, J. R. Foulds, R. A. Quinn, M. H. Badri, T. L. Swenson, M. W. Van Goethem, T. R. Northen, Y. Vazquez-Baeza, M. Wang, N. A. Bokulich, A. Watters, S. J. Song, R. Bonneau, P. C. Dorrestein, R. Knight, Learning representations of microbe-metabolite interactions. *Nat. Methods* **16**, 1306–1314 (2019).
31. K. Eriksson, L. Nise, L. Alfredsson, A. I. Catrina, J. Askling, K. Lundberg, L. Klareskog, T. Yucel-Lindberg, Seropositivity combined with smoking is associated with increased prevalence of periodontitis in patients with rheumatoid arthritis. *Ann. Rheum. Dis.* **77**, 1236–1238 (2018).
32. R. C. Brewer, N. S. Ramadoss, L. J. Lahey, S. Jahanbani, W. H. Robinson, T. V. Lanz, BNT162b2 vaccine induces divergent B cell responses to SARS-CoV-2 S1 and S2. *Nat. Immunol.* **23**, 33–39 (2022).
33. T. Cantaert, J.-N. Schickel, J. M. Bannock, Y.-S. Ng, C. Massad, F. R. Delmotte, N. Yamakawa, S. Glauzy, N. Chamberlain, T. Kinnunen, L. Menard, A. Lavoie, J. E. Walter, L. D. Notarangelo, J. Bruneau, W. Al-Herz, S. S. Kilic, H. D. Ochs, C. Cunningham-Rundles, M. van der Burg, T. W. Kuijpers, S. Kracker, H. Kaneko, Y. Sekinaka, S. Nonoyama, A. Durandy, E. Meffre, Decreased somatic hypermutation induces an impaired peripheral B cell tolerance checkpoint. *J. Clin. Invest.* **126**, 4289–4302 (2016).
34. F. S. van de Bovenkamp, L. Hafkenscheid, T. Rispens, Y. Rombouts, The emerging importance of IgG Fab glycosylation in immunity. *J. Immunol.* **196**, 1435–1441 (2016).
35. Y. Rombouts, A. Willemze, J. J. B. C. van Beers, J. Shi, P. F. Kerkman, L. van Toorn, G. M. C. Janssen, A. Zaldumbide, R. C. Hoeben, G. J. M. Pruijn, A. M. Deelder, G. Wolbink, T. Rispens, P. A. van Veelen, T. W. J. Huizinga, M. Wuhrer, L. A. Trouw, H. U. Scherer, R. E. M. Toes, Extensive glycosylation of ACPA-IgG variable domains modulates binding to citrullinated antigens in rheumatoid arthritis. *Ann. Rheum. Dis.* **75**, 578–585 (2016).
36. L. M. E. dos Anjos, I. A. Pereira, E. D'Orsi, A. P. Seaman, R. W. Burlingame, E. F. Morato, A comparative study of IgG second- and third-generation anti-cyclic citrullinated peptide (CCP) ELISAs and their combination with IgA third-generation CCP ELISA for the diagnosis of rheumatoid arthritis. *Clin. Rheumatol.* **28**, 153–158 (2009).
37. T. R. Mikuls, G. M. Thiele, K. D. Deane, J. B. Payne, J. R. O'Dell, F. Yu, H. Sayles, M. H. Weisman, P. K. Gregersen, J. H. Buckner, R. M. Keating, L. A. Derber, W. H. Robinson, V. M. Holers, J. M. Norris, *Porphyromonas gingivalis* and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis. *Arthritis Rheum.* **64**, 3522–3530 (2012).
38. T. Stobernack, C. Glasner, S. Junker, G. Gabarrini, M. de Smit, A. de Jong, A. Otto, D. Becher, A. J. van Winkelhoff, J. M. van Dijl, Extracellular proteome and citrullinome of the oral pathogen *Porphyromonas gingivalis*. *J. Proteome Res.* **15**, 4532–4543 (2016).
39. M. F. Konig, L. Abusleme, J. Reinholdt, R. J. Palmer, R. P. Teles, K. Sampson, A. Rosen, P. A. Nigrovic, J. Sokolove, J. T. Giles, N. M. Moutsopoulos, F. Andrade, Aggregatibacter actinomycetemcomitans-induced hypercitrullination links periodontal infection to autoimmunity in rheumatoid arthritis. *Sci. Transl. Med.* **8**, 369ra176 (2016).
40. N. Wegner, R. Wait, A. Sroka, S. Eick, K.-A. Nguyen, K. Lundberg, A. Kinloch, S. Culshaw, J. Potempa, P. J. Venables, Peptidylarginine deiminase from *Porphyromonas gingivalis* citrullinates human fibrinogen and α-enolase: Implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum.* **62**, 2662–2672 (2010).
41. C. Zenobia, K.-L. Herpoldt, M. Freire, Is the oral microbiome a source to enhance mucosal immunity against infectious diseases? *NPJ Vaccines* **6**, 80 (2021).
42. Y. Zhou, B. Chen, N. Mittereder, R. Chaerkady, M. Strain, L.-L. An, S. Rahman, W. Ma, C. P. Low, D. Chan, F. Neal, C. O. Bingham III, K. Sampson, E. Darrah, R. M. Siegel, S. Hasni, F. Andrade, K. A. Vousden, T. Mustelin, G. P. Sims, Spontaneous secretion of the citrullination enzyme PAD2 and cell surface exposure of PAD4 by neutrophils. *Front. Immunol.* **8**, 1200 (2017).
43. D. Belstrøm, R. R. Jersie-Christensen, D. Lyon, C. Damgaard, L. J. Jensen, P. Holmstrup, J. V. Olsen, Metaproteomics of saliva identifies human protein markers specific for individuals with periodontitis and dental caries compared to orally healthy controls. *PeerJ* **4**, e2433 (2016).
44. F. A. Scannapieco, L. Solomon, R. O. Wadenya, Emergence in human dental plaque and host distribution of amylase-binding streptococci. *J. Dent. Res.* **73**, 1627–1635 (1994).
45. Y. Xu, Q. Wu, Prevalence trend and disparities in rheumatoid arthritis among US adults, 2005-2018. *J. Clin. Med.* **10**, 3289 (2021).

46. T. A. Schwickert, R. L. Lindquist, G. Shakhar, G. Livshits, D. Skokos, M. H. Kosco-Vilbois, M. L. Dustin, M. C. Nussenzweig, In vivo imaging of germinal centres reveals a dynamic open structure. *Nature* **446**, 83–87 (2007).
47. J. Wrammert, K. Smith, J. Miller, W. A. Langley, K. Kokko, C. Larsen, N.-Y. Zheng, I. Mays, L. Garman, C. Helms, J. James, G. M. Air, J. D. Capra, R. Ahmed, P. C. Wilson, Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* **453**, 667–671 (2008).
48. S. E. Elliott, S. Kongpachith, N. Lingampalli, J. Z. Adamska, B. J. Cannon, R. Mao, L. K. Blum, W. H. Robinson, Affinity maturation drives epitope spreading and generation of proinflammatory anti–citrullinated protein antibodies in rheumatoid arthritis. *Arthritis Rheumatol.* **70**, 1946–1958 (2018).
49. Y. H. Lee, S.-C. Bae, G. G. Song, FCGR2A, FCGR3A, FCGR3B polymorphisms and susceptibility to rheumatoid arthritis: A meta-analysis. *Clin. Exp. Rheumatol.* **33**, 647–654 (2015).
50. C. Tan Sardjono, P. L. Mottram, N. C. van de Velde, M. S. Powell, D. Power, R. F. Slocombe, I. P. Wicks, I. K. Campbell, S. E. McKenzie, M. Brooks, A. W. Stevenson, P. M. Hogarth, Development of spontaneous multisystem autoimmune disease and hypersensitivity to antibody-induced inflammation in Fcgamma receptor IIa-transgenic mice. *Arthritis Rheum.* **52**, 3220–3229 (2005).
51. D. McKean, K. Huppi, M. Bell, L. Staudt, W. Gerhard, M. Weigert, Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3180–3184 (1984).
52. X. Zhang, D. Zhang, H. Jia, Q. Feng, D. Wang, D. Liang, X. Wu, J. Li, L. Tang, Y. Li, Z. Lan, B. Chen, Y. Li, H. Zhong, H. Xie, Z. Jie, W. Chen, S. Tang, X. Xu, X. Wang, X. Cai, S. Liu, Y. Xia, J. Li, X. Qiao, J. Y. Al-Aama, H. Chen, L. Wang, Q.-J. Wu, F. Zhang, W. Zheng, Y. Li, M. Zhang, G. Luo, W. Xue, L. Xiao, J. Li, W. Chen, X. Xu, Y. Yin, H. Yang, J. Wang, K. Kristiansen, L. Liu, T. Li, Q. Huang, Y. Li, J. Wang, The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat. Med.* **21**, 895–905 (2015).
53. R. Moentadj, Y. Wang, K. Bowerman, L. Rehaume, H. Nel, P. O. Cuiv, J. Stephens, A. Baharom, M. Maradana, V. Lakis, M. Morrison, T. Wells, P. Hugenholtz, H. Benham, K.-A. Le Cao, R. Thomas, *Streptococcus* species enriched in the oral cavity of patients with RA are a source of peptidoglycan-polysaccharide polymers that can induce arthritis in mice. *Ann. Rheum. Dis.* **80**, 573–581 (2021).
54. M. Tsukasaki, N. Komatsu, K. Nagashima, T. Nitta, W. Pluemsakunthai, C. Shukunami, Y. Iwakura, T. Nakashima, K. Okamoto, H. Takayanagi, Host defense against oral microbiota by bone-damaging T cells. *Nat. Commun.* **9**, 701 (2018).
55. S. Kitamoto, H. Nagao-Kitamoto, R. Hein, T. M. Schmidt, N. Kamada, The bacterial connection between the oral cavity and the gut diseases. *J. Dent. Res.* **99**, 1021–1029 (2020).
56. S. Chukkapalli, M. Rivera-Kweh, P. Gehlot, I. Velsko, I. Bhattacharyya, S. J. Calise, M. Satoh, E. K. L. Chan, J. Holoshitz, L. Kesavalu, Periodontal bacterial colonization in synovial tissues exacerbates collagen-induced arthritis in B10.RIII mice. *Arthritis Res. Ther.* **18**, 161 (2016).
57. T. S. Schmidt, M. R. Hayward, L. P. Coelho, S. S. Li, P. I. Costea, A. Y. Voigt, J. Wirbel, O. M. Maistrenko, R. J. Alves, E. Bergsten, C. de Beaufort, I. Sobhani, A. Heintz-Buschart, S. Sunagawa, G. Zeller, P. Wilmes, P. Bork, Extensive transmission of microbes along the gastrointestinal tract. *eLife* **8**, e42693 (2019).
58. R. Gacesa, A. Kurilshikov, A. Vich Vila, T. Sinha, M. A. Y. Klaassen, L. A. Bolte, S. Andreu-Sánchez, L. Chen, V. Collij, S. Hu, J. A. M. Dekens, V. C. Lenters, J. R. Björk, J. C. Swarte, M. A. Swertz, B. H. Jansen, J. Gelderloos-Arends, S. Jankipersadsing, M. Hofker, R. C. H. Vermeulen, S. Sanna, H. J. M. Harmsen, C. Wijmenga, J. Fu, A. Zhernakova, R. K. Weersma, Environmental factors shaping the gut microbiome in a Dutch population. *Nature* **604**, 732–739 (2022).
59. F. C. Arnett, S. M. Edworthy, D. A. Bloch, D. J. McShane, J. F. Fries, N. S. Cooper, L. A. Healey, S. R. Kaplan, M. H. Liang, H. S. Luthra, T. A. Medsger Jr., D. M. Mitchell, D. H. Neustadt, R. S. Pinals, J. G. Schaller, J. T. Sharp, R. L. Wilder, G. G. Hunder, The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* **31**, 315–324 (1988).
60. R. C. Page, P. I. Eke, Case definitions for use in population-based surveillance of periodontitis. *J. Periodontol.* **78**, 1387–1399 (2007).
61. A. Gonzalez, J. A. Navas-Molina, T. Kosciolek, D. McDonald, Y. Vázquez-Baeza, G. Ackermann, J. DeReus, S. Janssen, A. D. Swafford, S. B. Orchanian, J. G. Sanders, J. Shorenstein, H. Holste, S. Petrus, A. Robbins-Pianka, C. J. Brislawn, M. Wang, J. R. Rideout, E. Bolyen, M. Dillon, J. G. Caporaso, P. C. Dorrestein, R. Knight, Qiita: Rapid, web-enabled microbiome meta-analysis. *Nat. Methods* **15**, 796–798 (2018).
62. Q. Zhu, S. Huang, A. Gonzalez, I. McGrath, D. McDonald, N. Haiminen, G. Armstrong, Y. Vázquez-Baeza, J. Yu, J. Kuczynski, G. D. Sepich-Poore, A. D. Swafford, P. Das, J. P. Shaffer, F. Lejzerowicz, P. Belda-Ferre, A. S. Havulinna, G. Méric, T. Niiranen, L. Lahti, V. Salomaa, H.-C. Kim, M. Jain, M. Inouye, J. A. Gilbert, R. Knight, OGUs enable effective, phylogeny-aware analysis of even shallow metagenome community structures. bioRxiv 2021.04.04.438427 [Preprint]. 6 April 2021. https://doi.org/10.1101/2021.04.04.438427.
63. B. Hillmann, G. A. Al-Ghalith, R. R. Shields-Cutler, Q. Zhu, D. M. Gohl, K. B. Beckman, R. Knight, D. Knights, Evaluating the information content of shallow shotgun metagenomics. *mSystems* **3**, e00069-18 (2018).
64. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
65. J. Lloyd-Price, A. Mahurkar, G. Rahnavard, J. Crabtree, J. Orvis, A. B. Hall, A. Brady, H. H. Creasy, C. McCracken, M. G. Giglio, D. McDonald, E. A. Franzosa, R. Knight, O. White, C. Huttenhower, Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**, 61–66 (2017).
66. G. D. Poore, E. Kopylova, Q. Zhu, C. Carpenter, S. Fraraccio, S. Wandro, T. Kosciolek, S. Janssen, J. Metcalf, S. J. Song, J. Kanbar, S. Miller-Montgomery, R. Heaton, R. Mckay, S. P. Patel, A. D. Swafford, R. Knight, Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574 (2020).
67. D. McDonald, J. C. Clemente, J. Kuczynski, J. R. Rideout, J. Stombaugh, D. Wendel, A. Wilke, S. Huse, J. Hufnagle, F. Meyer, R. Knight, J. G. Caporaso, The biological observation matrix (BIOM) format or: How I learned to stop worrying and love the ome-ome. *Gigascience* **1**, 7 (2012).
68. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
69. M. D. Robinson, D. J. Mccarthy, G. K. Smyth, Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
70. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
71. S. E. Elliott, S. Kongpachith, N. Lingampalli, J. Z. Adamska, B. J. Cannon, L. K. Blum, M. S. Bloom, M. Henkel, M. J. McGeachy, L. W. Moreland, W. H. Robinson, B cells in rheumatoid arthritis synovial tissues encode focused antibody repertoires that include antibodies that stimulate macrophage TNF-α production. *Clin. Immunol.* **212**, 108360 (2020).
72. S. Kongpachith, N. Lingampalli, C.-H. Ju, L. K. Blum, D. R. Lu, S. E. Elliott, R. Mao, W. H. Robinson, Affinity maturation of the anti-citrullinated protein antibody paratope drives epitope spreading and polyreactivity in rheumatoid arthritis. *Arthritis Rheumatol.* **71**, 507–517 (2019).
73. L. K. Blum, J. Z. Adamska, D. S. Martin, A. W. Rebman, S. E. Elliott, R. R. L. Cao, M. E. Embers, J. N. Aucott, M. J. Soloski, W. H. Robinson, Robust B cell responses predict rapid resolution of Lyme disease. *Front. Immunol.* **9**, 1634 (2018).
74. R. C. Edgar, UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
75. E. Alamyar, P. Duroux, M.-P. Lefranc, V. Giudicelli, IMGT(®) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol. Biol.* **882**, 569–604 (2012).
76. K. Julenius, M. B. Johansen, Y. Zhang, S. Brunak, R. Gupta, *Bioinformatics for Glycobiology and Glycomics* (John Wiley & Sons Ltd., 2009), pp. 163–192.
77. H. Oh, B. Siano, S. Diamond, Neutrophil isolation protocol. *J. Vis. Exp.*, (2008).
78. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
79. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
80. M. W. Fedarko, C. Martino, J. T. Morton, A. González, G. Rahman, C. A. Marotz, J. J. Minich, E. E. Allen, R. Knight, Visualizing 'omic feature rankings and log-ratios using Qurro. *NAR Genom. Bioinform.* **2**, lqaa023 (2020).
81. T. P. Quinn, I. Erb, Amalgams: Data-driven amalgamation for the dimensionality reduction of compositional data. *NAR Genom. Bioinform.* **2**, lqaa076 (2020).
82. C. Martino, L. Shenhav, C. A. Marotz, G. Armstrong, D. McDonald, Y. Vázquez-Baeza, J. T. Morton, L. Jiang, M. G. Dominguez-Bello, A. D. Swafford, E. Halperin, R. Knight, Context-aware dimensionality reduction deconvolutes gut microbial community dynamics. *Nat. Biotechnol.* **39**, 165–168 (2021).
83. J.-C. J. Tsay, B. G. Wu, I. Sulaiman, K. Gershner, R. Schluger, Y. Li, T.-A. Yie, P. Meyn, E. Olsen, L. Perez, B. Franca, J. Carpenito, T. Iizumi, M. El-Ashmawy, M. Badri, J. T. Morton, N. Shen, L. He, G. Michaud, S. Rafeq, J. L. Bessich, R. L. Smith, H. Sauthoff, K. Felner, R. Pillai, A.-M. Zavitsanou, S. B. Koralov, V. Mezzano, C. A. Loomis, A. L. Moreira, W. Moore, A. Tsirigos, A. Heguy, W. N. Rom, D. H. Sterman, H. I. Pass, J. C. Clemente, H. Li, R. Bonneau, K.-K. Wong, T. Papagiannakopoulos, L. N. Segal, Lower airway dysbiosis affects lung cancer progression. *Cancer Discov.* **11**, 293–307 (2021).

84. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, S. P. Holmes, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
85. E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. Da Silva, C. Diener, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, S. M. Gibbons, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. D. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B. Rasmussen, A. Rivers, M. S. Robeson II, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
86. S. Mandal, W. Van Treuren, R. A. White, M. Eggesbø, R. Knight, S. D. Peddada, Analysis of composition of microbiomes: A novel method for studying microbial composition. *Microb. Ecol. Health Dis.* **26**, 27663 (2015).

Acknowledgments: We thank the VA Palo Alto FACS facility and Stanford University Mass Spectrometry Core as well as the staff of the Rockefeller University Hospital for assistance. We also thank L. Shenhav, S. Culshaw, and M. Kurowska-Stolarska for the helpful discussions and edits. **Funding:** The research reported in this publication was supported by National Institutes of Health (NIH) R01 AR063676 (to W.H.R.), NIH R01 AR078268 (to W.H.R. and D.E.O.), NIH U19 AI110491 (to W.H.R.), NIH U01 AI101981 (to V.M.H. and W.H.R.), NIH T32 HG003284 (to V.Y.), NIH R01 GM071966 (to O.T.), NIH T32 AI007290-35 (to R.C.B.), NIH P30 AR079369 (to V.M.H. and K.D.D.), NIH UC2 AR082186 (to D.E.O),
National Science Foundation (NSF) Graduate Research Fellowship (to R.C.B.), Robertson Foundation (to D.E.O.), Rockefeller University UL1 TR001866 (to D.E.O.), Bernard and Irene Schwartz Foundation (to D.E.O.), Iris and Junming Le Foundation (to D.E.O.), Rockefeller Clinical and Translational Science Award Program Pilot Award (to D.E.O.), Rheumatology Research Foundation (to L.B.K.), and National Cancer Institute (NCI) F30 CA243480 (to G.D.S.-P.). **Author contributions:** R.C.B., D.E.O., W.H.R., and R.B.D. conceptualized the study. R.C.B., T.V.L., C.R.H., G.D.S.-P., C.M., A.D.S., S.K., L.K.B., S.E.E., N.E.B., S.P., J.F., V.Y., O.T., M.W., R.K., Q.W., H.W., M.O.F., S.C., and L.B.K. developed the methodology for the study. R.C.B., T.V.L., C.R.H., G.D.S.-P., C.M., A.D.S., and D.E.O. conducted investigation. C.R.H., G.D.S.-P., C.M., A.D.S., R.C.B., and D.E.O. performed bioinformatic analyses. R.C.B. and T.V.L. performed and analyzed the in vitro experiments. R.C.B., D.E.O., and W.H.R. wrote the original draft. All authors reviewed and edited the manuscript. **Competing interests:** W.H.R. is a founder and member of the Board of Directors and a consultant to Atreca Inc. D.E.O. and R.B.D. have filed a provisional patent encompassing aspects of this technology. G.D.S.-P. is a cofounder of and reports stock interest in Micronoma. The other authors declare that they have no competing interests. **Data and materials availability:** All data associated with this study are present in the paper or the Supplementary Materials. Plasmablast sequences can be accessed at NCBI Sequence Read Archive (accession nos. PRJNA503739 and SRP150122) and GEO (accession number GSE114310). Raw RNA-seq data are available at DbGAP (phs003179.v1.p1). Data files S1 to S6 contain the raw data from figures. Further information and requests for resources and reagents should be directed to and will be fulfilled by W.H.R. (w.robinson@stanford.edu).

Submitted 6 May 2022
Resubmitted 19 August 2022
Accepted 2 February 2023
Published 22 February 2023
10.1126/scitranslmed.abq8476

Science Translational Medicine

Oral mucosal breaks trigger anti-citrullinated bacterial and human protein antibody responses in rheumatoid arthritis

R. Camille Brewer, Tobias V. Lanz, Caryn R. Hale, Gregory D. Sepich-Poore, Cameron Martino, Austin D. Swafford, Thomas S. Carroll, Sarah Kongpachith, Lisa K. Blum, Serra E. Elliott, Nathalie E. Blachere, Salina Parveen, John Fak, Vicky Yao, Olga Troyanskaya, Mayu O. Frank, Michelle S. Bloom, Shaghayegh Jahanbani, Alejandro M. Gomez, Radhika Iyer, Nitya S. Ramadoss, Orr Sharpe, Sangeetha Chandrasekaran, Lindsay B. Kelmenson, Qian Wang, Heidi Wong, Holly L. Torres, Mark Wiesen, Dana T. Graves, Kevin D. Deane, V. Michael Holers, Rob Knight, Robert B. Darnell, William H. Robinson, and Dana E. Orange

Sci. Transl. Med., **15** (684), eabq8476.
DOI: 10.1126/scitranslmed.abq8476