PATIENTS AND METHODS

RA patients and control subjects

The study, "Autoantigens in RA," was approved by the Human Investigations Committee at Massachusetts General Hospital (MGH); all subjects gave written informed consent. All RA patients met the American College of Rheumatology/ European League Against Rheumatism Collaborative Initiative criteria for that disease (24). All study patients with RA or other rheumatic diseases were recruited from the Rheumatology Clinic at Massachusetts General Hospital (MGH) or from suburban MGH clinics.

For isolation of HLA-DR-presented peptides, synovial tissue, SFMC or PBMC were obtained from 5 patients with RA. To test implicated peptides and their source proteins for immunoreactivity in larger numbers of patients, we used our cohort of NORA patients from whom systematic clinical information, PBMC and serum samples, and in some cases, SF were available. For comparison, PBMC and serum samples were available from patients with Lyme arthritis. In addition, we used our cohort of CRA patients in whom serum and sometimes SF samples were collected. Serum samples were obtained from patients with other types of arthritis or connective tissue diseases, and from healthy control subjects. HLA-DR typing was performed on blood samples from RA patients at the American Red Cross Laboratory in Dedham, MA. ACPA and RF determinations were made in the clinical laboratories at MGH.

Enzyme-linked immunospot (ELISpot) T cell assay

The detailed methods for isolation and identification of HLA-DR presented peptides are given in our previous publication (16). In this study, one microbial peptide (^{2}KRIILILTVLLAMLGQVAY^{20}) derived from a 27 kDa *P. copri* protein (WP_022121928.1)

was identified from the PBMC of one RA patient. This *P. copri* peptide and 2 additional peptides from the same protein, which were predicted to be promiscuous T cell epitopes (^{52}DYRGYWTMRYQFDSATVS^{69}, ^{118}EKINSLPTSSTGI^{130}), were synthesized and HPLC-purified in the Core Proteomics Laboratory at MGH. PBMC from RA patients were then stimulated with the peptides (1 μM) in duplicate along with positive (phytohemagglutinin) and negative (no antigen) controls, and incubated for 5 days in culture in a CO_2 incubator at 37^0C. Cells were then transferred to Dual Color ELISpot plates coated with IFN-γ/IL17 antibodies (Cellular Technology Limited), and incubated overnight at 37^0C. Images of wells were captured using ImmunoSpot series 3B analyzer, and spots were counted using ImmunoSpot software. A positive T cell response was defined as 3 standard deviations (SD) above the mean value of healthy subjects.

ELISA for serum IgG and IgA antibodies to the *P. copri* protein (*Pc*-p27)

ELISA plates were coated with 0.25 μg/ml of the *P. copri* protein *Pc*-p27 (GenScript USA Inc, Piscataway, NJ) overnight at 4°C. Afterwards, plates were incubated with blocking buffer (5% nonfat dry milk in PBST) for one hour. Each patient's serum sample (diluted 200-fold) was then added in duplicate wells for 1.5 hours, followed by horseradish-peroxidase (HRP)-conjugated goat anti-human IgG (sc-2453, Santa Cruz Biotech) or HRP-conjugated goat anti-human IgA (Bio-Rad, Hercules, CA) and then TMB substrate (BD, San Diego, CA). For inter-plate standardization, 2 control samples were included on each plate. In addition, using Heteroblock, we tested serum samples from 15 patients who had a range of OD values by ELISA, and confirmed that RF did not alter ELISA results. Therefore, Heteroblock was not used in subsequent antibody determinations.

ELISA for serum IgG and IgA antibodies to microbial organisms

The *P. copri* type strain (DSM 18205) was obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). *B. fragilis* (ATCC 25285), *E. coli* (ATCC 25922), and *P. gingivalis* (ATCC 33277) isolates were obtained from the American Type Culture Collection (ATCC, Manassas, VA). The bacterial cultures were inactivated in 1% formalin for 24 hours and washed twice in PBS and diluted in PBS at a final concentration of 10^6 cells/ml.

IgG and IgA antibody responses to whole-cell *P. copri*, *B. fragilis* or *E. coli* were determined by ELISA. The plates were coated overnight at 4˚C with suspensions of inactivated bacterial cells (10^6 cells/ml). Afterwards, plates were incubated with blocking buffer for one hour. Patients' serum samples (diluted 1:100) were added in duplicate wells for 1.5 hours, followed by HRP-conjugated goat anti-human IgG or HRP-conjugated goat anti-human IgA and then TMB substrate. For inter-plate standardization, 2 control samples were included on each plate. The ELISA for *P. gingivalis* was performed as previously described (28).

Cytokine and chemokine determinations

The levels of 14 cytokines and chemokines associated with innate immune responses (IFN-α, TNF, MIP-1α, MIP-1β), and Th1 (IFNγ, IL-12, CXCL9, and CXCL10) or Th17 (IL-1β, IL-17A, IL-17E, IL-17F, IL-22, and IL-23) adaptive immune responses were determined in serum or SF samples from RA patients. Samples were diluted 1:5 in PBS and incubated with Heteroblock (Omega Biologicals) at a concentration of 150 μg/ml to limit the possible confounding effects of rheumatoid factor. Protein levels of all 14 inflammatory mediators in serum or SF were assessed

in one complete experiment using bead-based Luminex assays (EMD-Millipore) coupled with the Luminex-200 System Analyzer (Luminex, Austin, TX). Data were assessed using the xPONENT 3.1 software.

Detection of *P.copri and B. fragilis* 16S rDNA in patients' samples by PCR

DNA was isolated from 200 µl of serum or SF using the QIAamp DNA mini (QIAGEN). Nested PCR primers were designed to detect DNA for *P. copri* or *B. fragilis* 16S rDNA using Primer3 software; the sequences are given in Table S1. DNA was amplified for outer PCR using the species-specific forward and reverse primers; 1 µl of the amplified DNA (1:10 in sterile distilled water) was then used for the nested PCR reaction. All reactions were carried out using 2.5 U of HotStarTaq DNA polymerase (QIAGEN). The amplification program included 40 cycles with denaturation 94°C for 30 s, annealing 59°C for 30 s, extension 72°C for 50 s, and a final extension of 10 min. For both outer and nested PCR reactions, a positive control (*P. copri* or *B. fragilis* DNA) and a negative control (sterile distilled water) were included. When enough DNA was available, samples were tested in duplicates. Amplified products (10 µl) were visualized by electrophoresis in a 2% agarose gel. Correct identity of PCR products were validated by direct DNA sequencing carried out in the DNA CORE facility at MGH. The sequenced product was aligned with all human and known microbial genomes using Genomic Blast Sequence.

Statistical analyses

Categorical data were analyzed by Fisher's exact test, and quantitative data were analyzed using unpaired t test with Welch correction. Correlations were sought using Pearson's correlation test.

All analyses were performed using GraphPad Prism 6. All *P* values were two-tailed. P values $\leq$ 0.05 were considered statistically significant.

Supplementary Table 1: Clinical information and laboratory results of the 18 RA patients tested[a]

Pt ID	Disease Duration (yrs)	DAS28	ACPA	RF	*P. copri* PCR serum	*P. copri* PCR SF	*B. fragilis* PCR SF
IgG *P. copri* antibodies (N=5)							
RA1	4.5	3.5	18	<30	-	-	-
	5	3.1	18	<30	-	-	-
	7	1.5	173	<30	-	-	-
	9	2.9	120	<30	-	**POS**	-
RA2	0.75	7.5	3	<30	-	**POS**	-
	1	5.1	3	<30	-	**POS**	-
RA3	13	3.6	7	<30	-	-	-
RA4	2	7.3	6	<30	-	-	-
RA5	1	6.6	268	3280	-	**POS**	-
Subtotal No. pos/total			**2/5**	**1/5**	**0/5**	**3/5**	**0/5**
IgA *P. copri* antibodies (N=2)							
RA6	0.25	8	152	730	-	-	-
RA7	3	4.2	183	155	-	-	-
No P. copri antibodies (N=13)							
RA8	3.7	2.6	235	<30	-	-	-
RA9	0.5	4	174	284	-	-	-
RA10	24	2.9	31	1410	-	-	-
RA11	0.75	6.6	8	<30	-	-	-
RA12	>50	NA	133	116	-	-	-
RA13	7.4	4.5	153	510	-	-	-
RA14	0.1	6.6	204	115	-	-	-
RA15	0.17	6.3	202	37	-	-	**POS**
RA16	1	3.8	19	782	-	-	-
RA17	3	2.1	7	<30	-	-	-
RA18	1	3	71	<30	-	-	-
Subtotal No. pos/total			**10/13**	**9/13**	**0/13**	**0/13**	**1/13**

[a]RA, rheumatoid arthritis; DAS28, disease activity score; ACPA, anti-citrullinated protein antibodies; RF, rheumatoid factor; POS, positive.