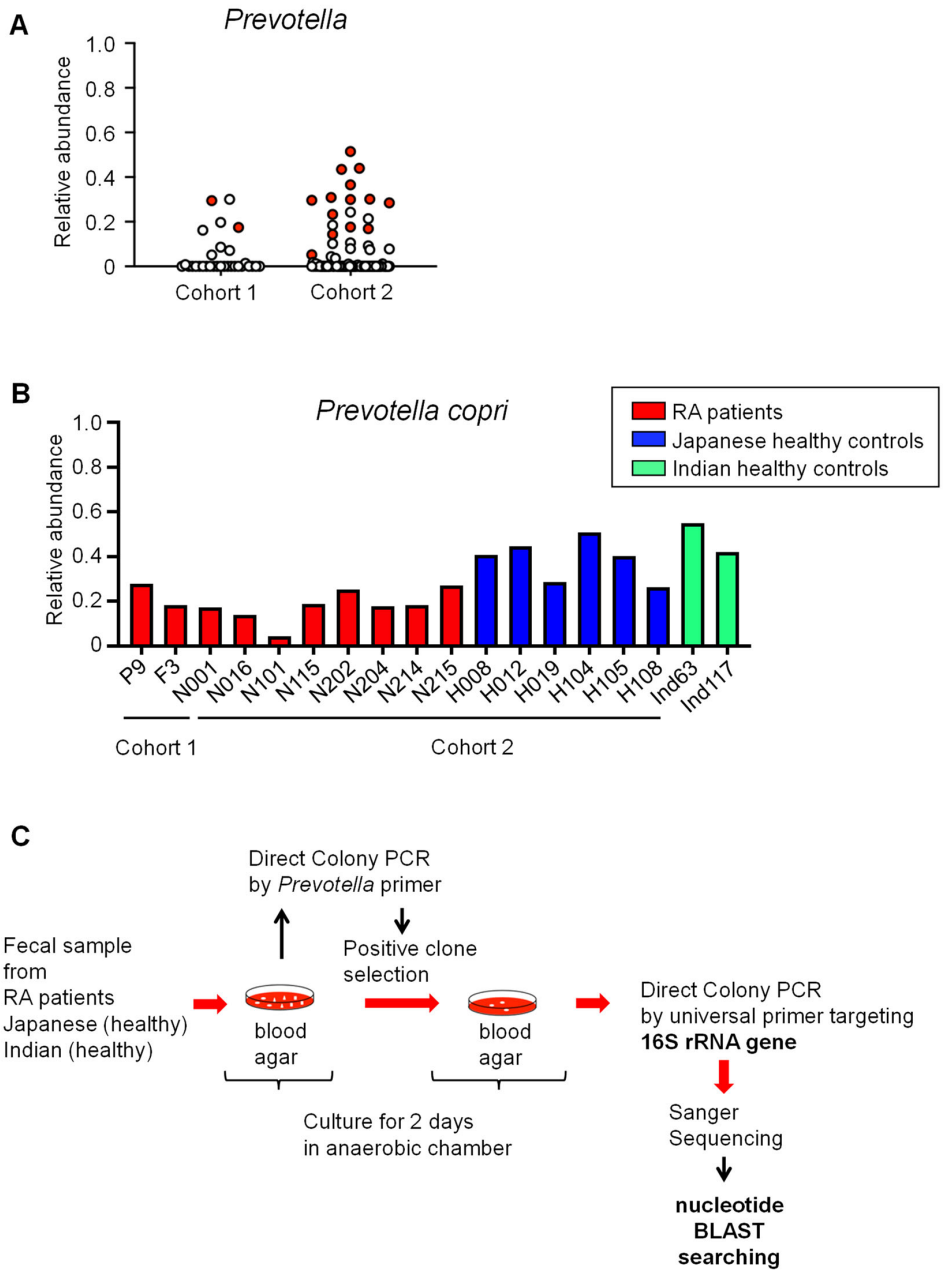


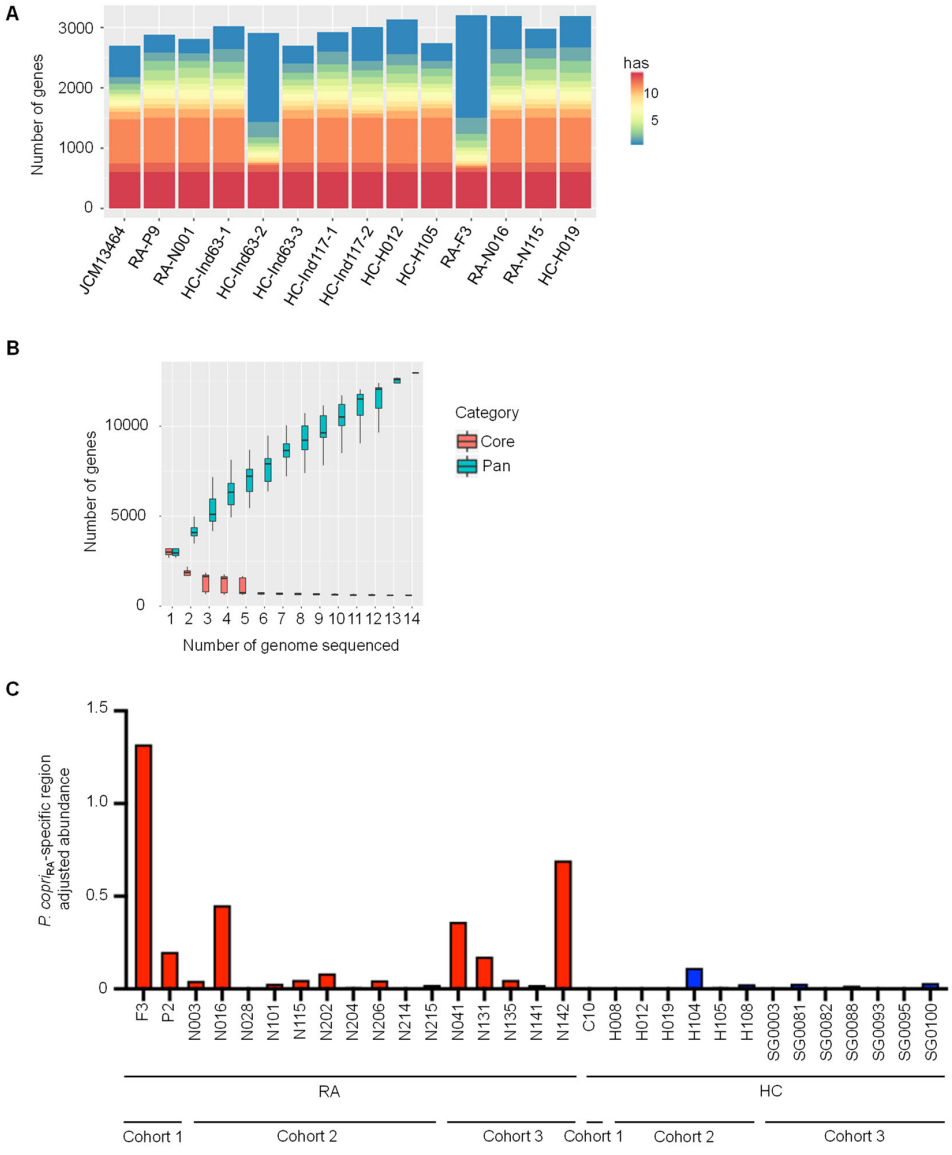
Supplementary Figure 1. Isolation of *P. copri* from RA patients

(A) Relative abundance of *Prevotella* at the genus level in the feces of RA patients (RA) and healthy controls (HC). Cohort 1: Fecal DNA samples from Maeda Y. et al. (1) (RA;

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

n=17, HC; n=14), Cohort 2: Fecal DNA samples from this study (RA; n=55, HC; n=28). Among them, samples with high abundance of *Prevotella* (highlighted by red color) were further analyzed for the presence of *P. copri*. Each vertical line represent the median. **(B)** Relative abundance of *P. copri* in the feces of RA patients and healthy controls with a high abundance of *Prevotella* in (A). **(C)** Schematic of the experimental protocol for isolation of *P. copri* from RA patients and healthy controls.

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

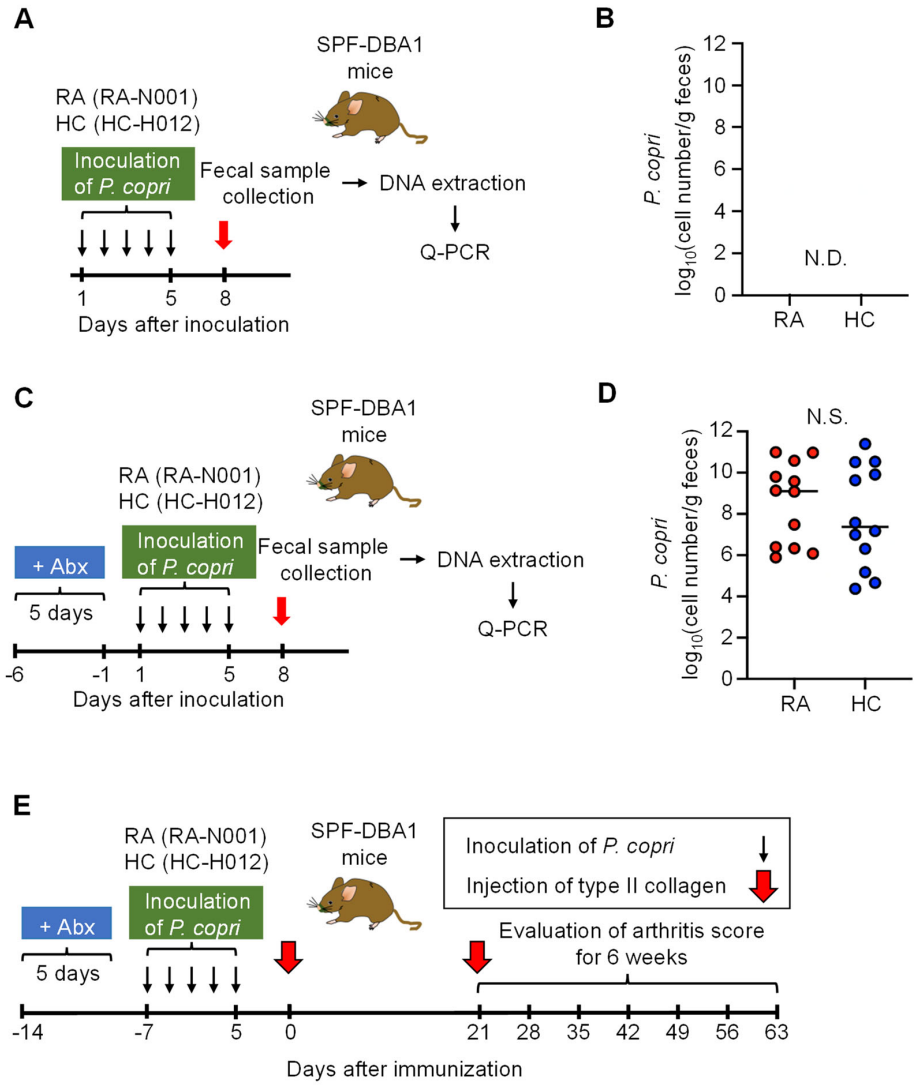


Supplementary Figure 2. Comparison of the genomic structures of *P. copri*$_{RA}$ and *P. copri*$_{HC}$

(A) The number of conserved genes in each strain. Bar color indicates the number of strains with each conserved gene. **(B)** The number of pan- and core-genes as a function of the number of assemblies that were randomly selected (n=1000) from 14 strains

including our 13 strains and the reference strain JCM13464. **(C)** Relative abundance of the *P. copri*$_{RA}$-specific genomic region in fecal DNA of RA patients (RA) and healthy controls (HC) with a high abundance of *P. copri* from cohort 1, 2 and 3.

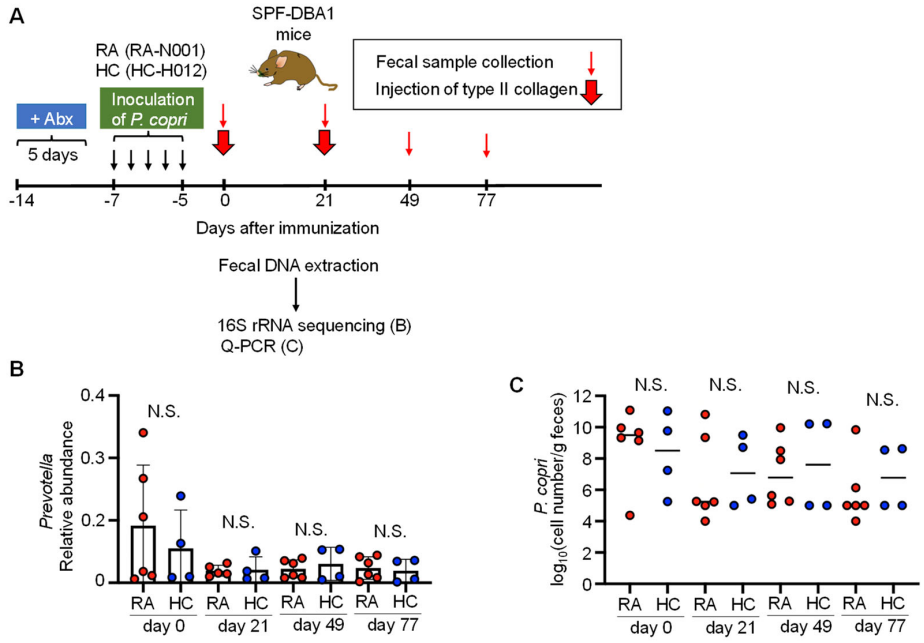
Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881



Supplementary Figure 3. Colonization of *P. copri* in DBA/1J mice

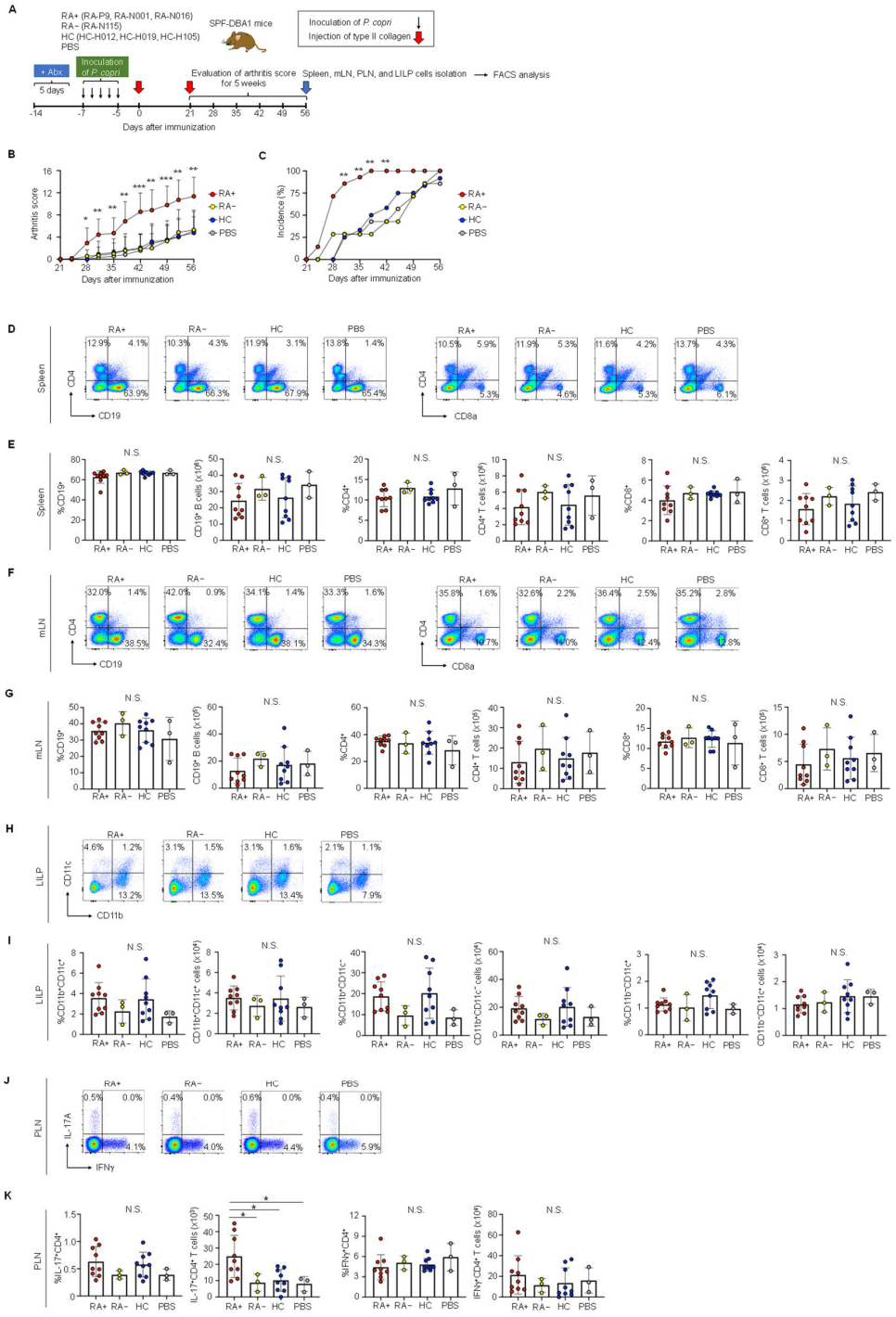
(A) Schematic of the colonization of *P. copri* in SPF DBA/1J mice. Fecal samples were collected from DBA/1J mice 3 days after 5 consecutive days of inoculation of *P. copri*$_{RA}$ (RA-N001: RA) or *P. copri*$_{HC}$ (HC-H012: HC). DNA was extracted from fecal samples and the numbers of *P. copri* were determined by Q-PCR. **(B)** The number of *P. copri* in the feces of *P. copri*$_{RA}$ (RA-N001: RA, n=5)) or *P. copri*$_{HC}$ (HC-H012: HC, n=5) -

inoculated mice without antibiotic treatment. **(C)** Schematic of the colonization of *P. copri* in DBA/1J mice treated with combinations of antibiotics (ampicillin, neomycin, metronidazole and vancomycin). SPF DBA/1J mice were orally treated with antibiotics for 5 days, and then inoculated with *P. copri*$_{RA}$ (RA-N001: RA) or *P. copri*$_{HC}$ (HC-H012: HC) for 5 consecutive days. Fecal samples were collected from DBA/1J mice 3 days after the final inoculation of *P. copri*$_{RA}$ or *P. copri*$_{HC}$. DNA was extracted from fecal samples and the numbers of *P. copri* were determined by Q-PCR. **(D)** The number of *P. copri* in the feces of *P. copri*$_{RA}$ (RA-N001: RA, n=12) or *P. copri*$_{HC}$ (HC-H012: HC, n=12) -inoculated mice with antibiotic treatment. **(E)** Schematic of the experimental protocol for type II collagen-induced arthritis in DBA1/J mice harboring *P. copri*. After colonization of *P. copri*, type II collagen was immunized at day 0 and day 21. The arthritis score and the incidence of arthritis was monitored for 6 weeks. Two-tailed Student's t-tests (B, D) were performed for statistical analysis. N.D. = not detected. N.S. = not significant.



Supplementary Figure 4. Presence of *P.copri* during the course of arthritis development

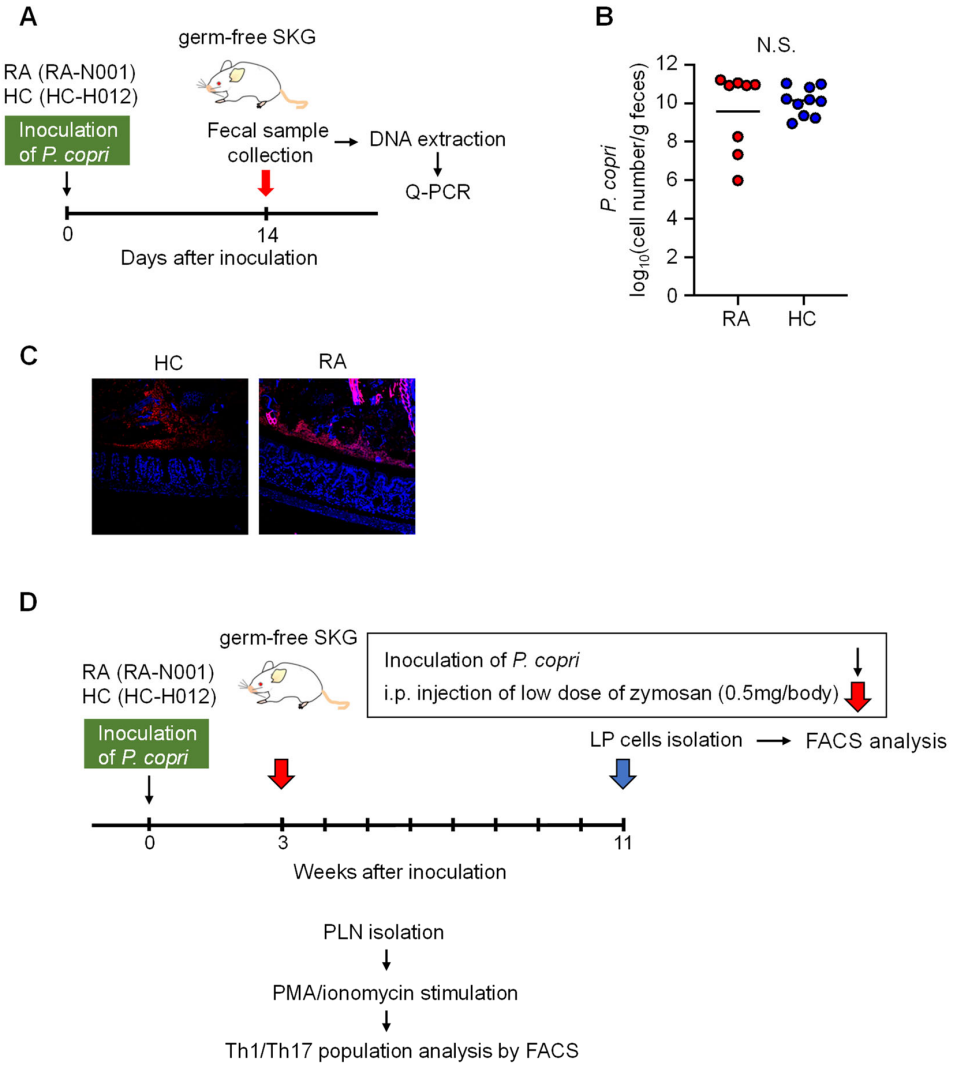
(A) Schematic of the experimental protocol for the analysis of *P. copri* during type II collagen-induced arthritis of DBA/1J mice. **(B)** Relative abundance of *Prevotella* genus in the feces at each time point was determined by 16S rRNA sequencing. **(C)** The numbers of *P. copri* in feces at each time point were determined by Q-PCR. (*P. copri*$_{RA}$ (RA-N001)-inoculated mice; n=6, *P. copri*$_{HC}$ (HC-H012)-inoculated mice; n=4). Two-tailed Student's t-tests (B, C) were performed for statistical analysis. N.S. = not significant.



Supplementary Figure 5. Colonization of *P. copri*$_{RA}$ with a specific genomic region

exacerbates collagen-induced arthritis in DBA/1J mice

(A) Schematic of the experimental protocol for type II collagen-induced arthritis in DBA1/J mice harboring *P. copri*. After colonization of *P. copri*, type II collagen was immunized at day 0 and day 21. The arthritis score and the incidence of arthritis was monitored for 5 weeks. **(B, C)** Arthritis scores (B) and incidence (C) of isolated *P. copri*-inoculated DBA/1J mice in CIA models under specific-pathogen-free (SPF) conditions. DBA/1J mice harboring *P. copri*$_{\mathrm{RA}}$ with CTnPc (RA+) (RA-P9, n=4; RA-N001, n=5; RA-N016, n=5), *P. copri*$_{\mathrm{RA}}$ without CTnPc (RA−) (RA-N115, n=7), *P. copri*$_{\mathrm{HC}}$ (HC) (HC-H012, n=3; HC-H019, n=5; HC-H105, n=4) or with PBS treatment (PBS) (n=7) were analyzed. Each symbol and vertical line represent the mean ± s.d. **(D-K)** Flow cytometric analysis of immune cells harvested from CIA model mice harboring *P. copri*$_{\mathrm{RA}}$ with CTnPc (RA+) (RA-P9, n=3; RA-N001, n=3; RA-N016, n=3), *P. copri*$_{\mathrm{RA}}$ without CTnPc (RA−) (RA-N115, n=3), *P. copri*$_{\mathrm{HC}}$ (HC) (HC-H012, n=3; HC-H019, n=3; HC-H105, n=3) or with PBS treatment (PBS, n=3). (D, F, H, J) Representative flow cytometry plots of splenic cells (D), mesenteric lymph node (mLN) cells (F), large intestinal lamina propria (LILP) cells (H), and popliteal lymph node (PLN) $CD4^+$ T cells (J). (E, G, I, K) Percentage and number of $CD19^+$ B cells and $CD4^+$ and $CD8^+$ T cells harvested from spleen (E) and mLN (G), $CD11b^+CD11c^+$, $CD11b^+CD11c^-$, and $CD11b^-CD11c^+$ cells harvested form LILP (I), and $IL17A^+$ and $IFN\gamma^+$ $CD4^+$ T cells harvested from PLN (K). (J, K) $IL17A^+$ and $IFN\gamma^+$ $CD4^+$ T cells of PLN were stimulated with PMA and ionomycin at day 56 after primary immunization. Bar graphs show the mean ± s.d of data in pool of RA+ (RA-N001, RA-P9, RA-N016), RA− (RA-N115), HC (HC-H012, HC-H105, HC-H019). One-way ANOVA followed by a Tukey's multiple comparison tests (B, E, G, I, K) and χ^2-test (C) were performed for statistical analysis. *P<0.05, **P<0.01, ***P<0.001. N.S. = not significant. All data were reproduced in another independent experiment with similar results.

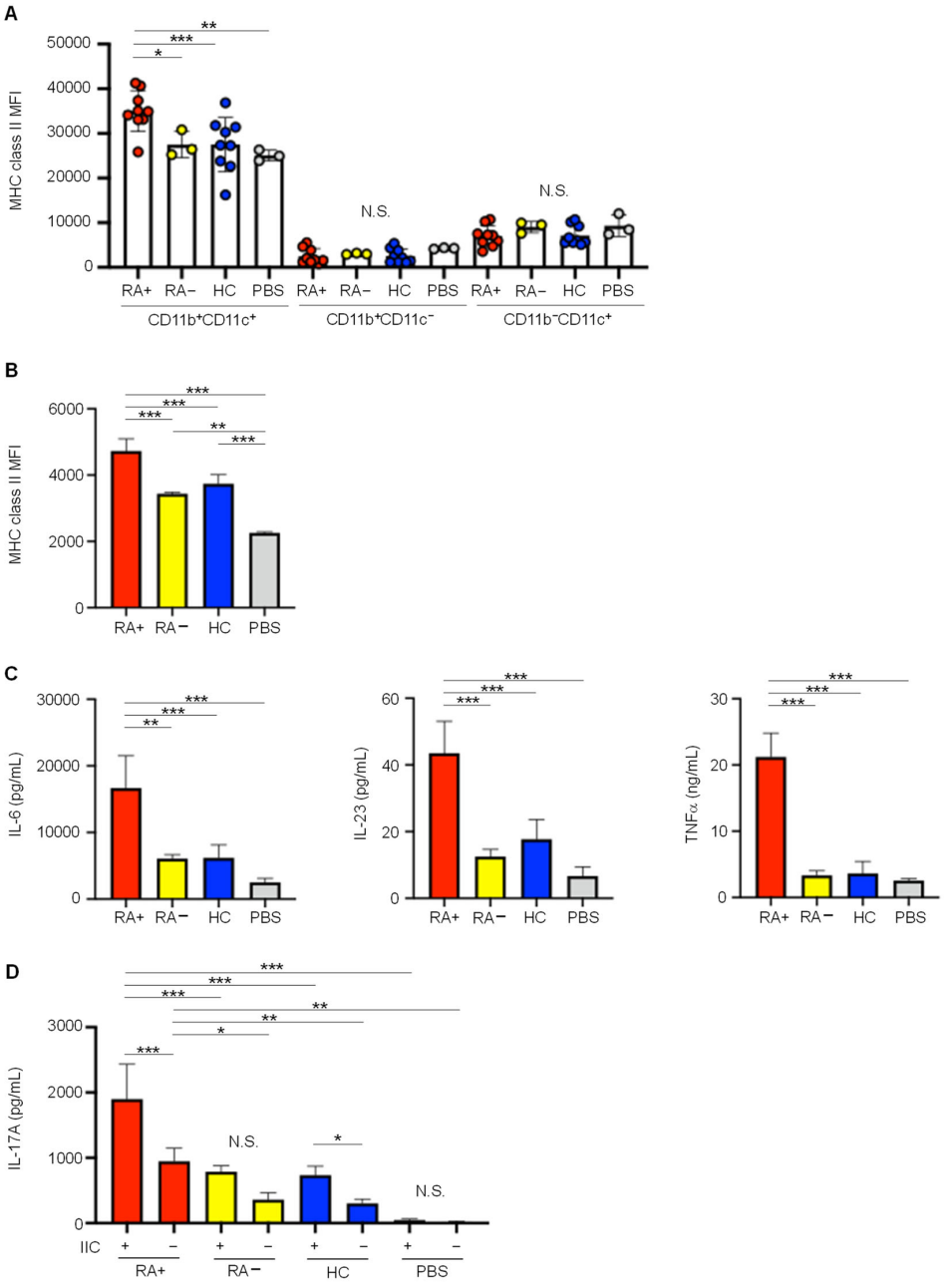


Supplementary Figure 6. Colonization of *P. copri* in germ free SKG mice

(A) Schematic of monocolonization of *P. copri* in germ free SKG mice. Female germ free SKG mice were inoculated with *P. copri*$_{RA}$ (RA-N001: RA) or *P. copri*$_{HC}$ (HC-H012: HC) and maintained in separate gnotobiotic isolators. Fecal samples were collected at 14 days after the inoculation. DNA was isolated from the fecal samples. The numbers of *P. copri* were determined by Q-PCR. **(B)** The number of *P. copri* in the feces of *P. copri*$_{RA}$

(RA-N001, n=8) or *P. copri*$_{HC}$ (HC-H012, n=10) -monocolonized SKG mice 14 days after the first inoculation of *P. copri*. **(C)** Detection of *P. copri*$_{RA}$ or *P. copri*$_{HC}$ in the lumen of the large intestine of *P. copri*-monocolonized SKG mice at 11 weeks after colonization using a eubacteria-specific FISH probe. Eubacteria-specific probe (red) and 4′, 6-diamidino-2-phenylindole (DAPI; blue) staining of Carnoy-fixed sections. **(D)** Schematic of the analysis of T cell populations in SKG mice monocolonized with *P. copri*. Female germ free SKG mice were inoculated with *P. copri*$_{RA}$ (RA-N001) or *P. copri*$_{HC}$ (HC-H012) and maintained in separate gnotobiotic isolators. Three weeks after inoculation, mice were intraperitoneally injected with 0.5 mg of zymosan and lymphocytes were isolated at 11 weeks after the injection. (B) Two-tailed Student's t-test was performed for statistical analysis. N.S. = not significant. All data were reproduced in another independent experiment with similar results.

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881



Supplementary Figure 7. *P. copri*$_{RA}$ with a specific genomic region induces MHC class II upregulation in innate immune cells and Th17-related immune responses

(A) MHC class II expression of $CD11b^{+}CD11c^{+}$, $CD11b^{+}CD11c^{-}$ and $CD11b^{-}CD11c^{+}$ cells harvested from LILP of CIA model mice harboring *P. copri*$_{RA}$ with CTnPc (RA+) (RA-P9, n=3; RA-N001, n=3; RA-N016, n=3), *P. copri*$_{RA}$ without CTnPc (RA−) (RA-N115, n=3), *P. copri*$_{HC}$ (HC) (HC-H012, n=3; HC-H019, n=3; HC-H019, n=3) or with PBS treatment (PBS, n=3). **(B)** Expression levels of MHC class II in bone marrow-derived dendritic cells (BMDCs). BMDCs were stimulated with each of heat-killed *P. copri* (*P. copri*$_{RA}$ with CTnPc (RA+): RA-P9, RA-N001, RA-N016; *P. copri*$_{RA}$ without CTnPc (RA−): RA-N115; *P. copri*$_{HC}$ (HC): HC-H012, HC-H019, HC-H105) or PBS for 72 h. (A, B) Expression of MHC class II was determined as the mean fluorescence intensity (MFI) by flow cytometric analysis. **(C)** Concentrations of IL-6, IL-23 and TNFα were measured by ELISA in the supernatants of BMDCs stimulated with the indicated heat-killed *P. copri*. **(D)** The concentration of IL-17A was measured by ELISA in the supernatants of naïve $CD4^{+}$ T cells co-cultured with BMDCs with the indicated heat-killed *P. copri*. Cells were stimulated with or without bovine type II collagen (IIC). Bar graphs show the mean ± s.d of data in pool of RA+ (RA-P9, RA-N001, RA-N016), RA− (RA-N115), HC (HC-H012, HC-H019, HC-H105). One-way ANOVA followed by a Tukey's multiple comparison tests were performed for all statistical analysis. **P*<0.05, ***P*<0.01, ****P*<0.001, N.S. = not significant. All data are representative of at another independent experiments.

	RA (n=50)	HC (n=28)	*P* value
Sex(%female)	37	24	0.2293
Age	62.2 ± 13.3	56.5 ± 8.4	0.0127
BMI	22.1 ± 2.9	21.6 ± 3.2	1.0000
Vegetable(d/w)	6.1 ± 1.9	5.8 ± 1.5	0.1568
Meat(d/w)	3.3 ± 1.8	3.2 ± 1.6	0.8420
Fish(d/w)	1.9 ± 1.7	1.9 ± 1.3	0.7166
Dairy products(d/w)	4.8 ± 2.8	4.2 ± 3.1	0.3585

Supplementary Table 1. The dietary preferences of RA patients and healthy controls

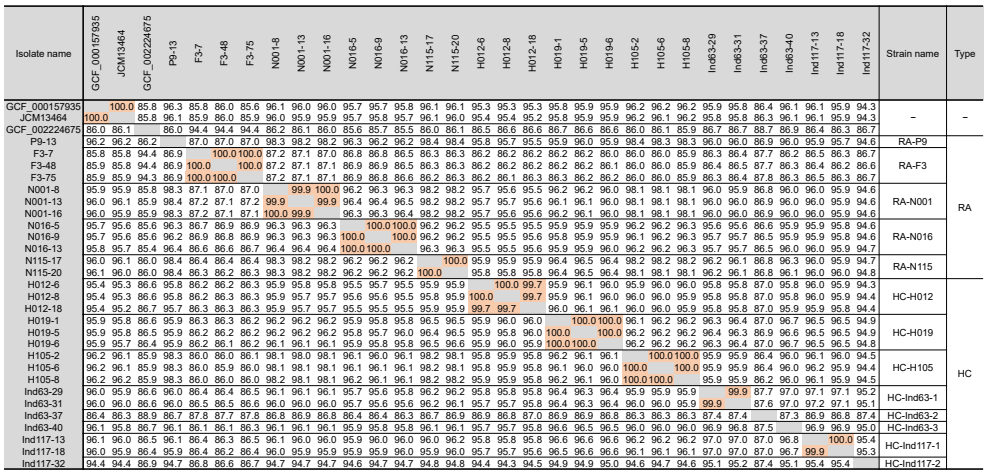
Questionnaires on dietary habits were administered among participants in cohort 2. Answers were received from 50 out of 55 rheumatoid arthritis patients (RA) and all 28 healthy controls (HC). Sex was shown as percentage of female and χ^2-test was performed for comparison between RA and HC group. Age, body mass index (BMI), frequency of consumption of vegetables, meat, fish, and dairy products (days/week consumed) were shown as mean ± s.d. and two-tailed Mann-Whitney U tests were performed.

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

Sample NO.	Status	Sex	Age	RF	ACPA	DAS-28CRP	Isolation	WGS	Duration of disease	Treatment
P9	Japanese RA patient	female	69	35	183	3.91	Yes	Yes	within 6 months	No
F3	Japanese RA patient	female	77	22	103	5.25	Yes	Yes	within 6 months	No
N001	Japanese RA patient	female	73	553	150	2.52	Yes	Yes	within 6 months	No
N016	Japanese RA patient	male	72	15	<0.6	2.82	Yes	Yes	within 6 months	No
N101	Japanese RA patient	female	63	15	142	3.02	No	No	within 6 months	No
N115	Japanese RA patient	female	75	100	886	2.33	Yes	Yes	within 6 months	No
N202	Japanese RA patient	female	53	<5	1.99	3.94	No	No	within 6 months	No
N204	Japanese RA patient	female	70	369	>132	3.94	No	No	within 6 months	No
N214	Japanese RA patient	female	48	13	32	4.48	No	No	within 6 months	No
N215	Japanese RA patient	male	77	97	45	5.00	No	No	within 6 months	No
H008	Japanese Healthy volunteer	female	55	<5	<0.6	–	No	No	–	–
H012	Japanese Healthy volunteer	female	55	<5	<0.6	–	Yes	Yes	–	–
H019	Japanese Healthy volunteer	female	65	<5	<0.6	–	Yes	Yes	–	–
H104	Japanese Healthy volunteer	female	64	<5	<0.6	–	No	No	–	–
H105	Japanese Healthy volunteer	female	37	<5	<0.6	–	Yes	Yes	–	–
H108	Japanese Healthy volunteer	male	62	<5	<0.6	–	No	No	–	–
Ind63	Indian Healthy volunteer	male	45	–	–	–	Yes	Yes	–	–
Ind117	Indian Healthy volunteer	male	25	–	–	–	Yes	Yes	–	–

Supplementary Table 2. Clinical features of RA patients and healthy controls from whom *P. copri* strains were isolated.

Age (year old); RF, Rheumatoid factor (IU/mL); ACPA, anti-cyclic citrullinated peptide (U/mL); DAS28-CRP, Evaluation methods of disease activity for RA; Isolation, isolated *P. copri*; WGS, whole genome sequencing; Duration of disease, duration between RA onset and until collecting the fecal samples; Treatment, treatments of anti-rheumatic drugs until the fecal samples were collected.



Isolate name	GCF_000157935	JCM13464	GCF_002224675	P9-13	F3-7	F3-48	F3-75	N001-8	N001-13	N001-16	N016-5	N016-9	N016-13	N115-17	N115-20	H012-6	H012-8	H012-18	H019-1	H019-5	H019-6	H105-2	H105-6	H105-8	Ind63-29	Ind63-31	Ind63-37	Ind63-40	Ind117-13	Ind117-18	Ind117-32	Strain name	Type
GCF_000157935		100.0	85.8	96.3	85.8	86.0	85.6	96.1	96.0	96.0	95.7	95.7	95.8	96.1	96.1	95.3	95.3	95.3	95.8	95.9	95.9	96.2	96.2	96.2	95.9	95.8	86.4	96.1	96.1	95.9	94.3	–	–
JCM13464	100.0		85.8	96.1	85.9	86.0	85.9	96.0	95.9	95.9	95.7	95.8	95.7	96.1	96.0	95.4	95.4	95.2	95.8	95.9	95.9	96.2	96.1	96.2	95.8	95.8	86.3	96.1	96.1	95.9	94.3		
GCF_002224675	86.0	86.1		86.0	94.4	94.4	94.4	86.2	86.1	86.0	85.6	85.7	85.5	86.0	86.1	86.5	86.6	86.6	86.7	86.6	86.6	86.0	86.1	85.9	86.7	86.7	88.7	86.9	86.4	86.3	86.7		
P9-13	96.2	96.2	86.2		87.0	87.0	87.0	98.3	98.2	98.2	96.3	96.2	96.2	98.4	98.4	95.8	95.7	95.5	95.9	96.0	95.9	98.4	98.3	98.3	96.0	96.0	86.9	96.0	95.9	95.7	94.6	RA-P9	
F3-7	85.8	85.8	94.4	86.9		100.0	100.0	87.2	87.1	87.0	86.8	86.8	86.5	86.3	86.3	86.2	86.2	86.2	86.2	86.2	86.2	86.0	86.0	86.0	85.9	86.3	86.4	87.7	86.2	86.5	86.3	86.7		
F3-48	85.9	85.8	94.4	86.9	100.0		100.0	87.2	87.1	87.1	86.9	86.9	86.5	86.3	86.3	86.2	86.2	86.2	86.2	86.2	86.1	86.0	86.0	85.9	86.4	86.5	87.7	86.3	86.4	86.2	86.6	RA-F3	
F3-75	85.9	85.9	94.3	86.9	100.0	100.0		87.2	87.1	87.1	86.9	86.8	86.6	86.2	86.3	86.2	86.1	86.3	86.3	86.2	86.2	86.0	86.0	85.9	86.3	86.4	87.8	86.3	86.5	86.3	86.7		
N001-8	95.9	95.9	85.8	98.3	87.1	87.0	87.0		99.9	100.0	96.2	96.3	96.3	98.2	98.2	95.7	95.6	95.5	96.2	96.2	96.0	98.1	98.1	98.1	96.0	95.9	86.8	96.0	96.0	95.9	94.6		
N001-13	96.0	96.1	85.9	98.4	87.2	87.1	87.2	99.9		99.9	96.4	96.4	96.5	98.2	98.2	95.7	95.7	95.6	96.1	96.1	96.0	98.1	98.1	98.1	96.0	96.0	86.9	96.0	96.0	95.9	94.6	RA-N001	RA
N001-16	96.0	95.9	85.9	98.3	87.2	87.1	87.1	100.0	99.9		96.3	96.3	96.4	98.2	98.2	95.7	95.6	95.6	96.2	96.1	96.0	98.1	98.1	98.1	96.0	96.0	86.9	96.0	96.0	95.9	94.6		
N016-5	95.7	95.6	85.6	96.3	86.7	86.9	86.9	96.3	96.3	96.3		100.0	100.0	96.2	96.2	95.5	95.5	95.5	95.9	95.9	95.9	96.2	96.2	96.3	95.6	95.6	86.6	95.9	95.9	95.8	94.6		
N016-9	95.7	95.6	85.6	96.2	86.9	86.8	86.9	96.3	96.3	96.3	100.0		100.0	96.2	96.2	95.5	95.5	95.6	95.8	95.9	95.9	96.1	96.2	96.3	95.7	95.7	86.5	95.9	95.9	95.8	94.6	RA-N016	
N016-13	95.8	95.7	85.4	96.4	86.6	86.6	86.7	96.4	96.4	96.4	100.0	100.0		96.3	96.3	95.5	95.5	95.6	95.9	95.9	96.0	96.2	96.2	96.3	95.7	95.7	86.5	96.0	96.0	95.9	94.7		
N115-17	96.0	96.1	86.0	98.4	86.4	86.4	86.4	98.3	98.2	98.2	96.2	96.2	96.2		100.0	95.9	95.9	95.9	96.4	96.5	96.4	98.2	98.2	98.2	96.2	96.1	86.8	96.3	96.0	95.9	94.7	RA-N115	
N115-20	96.1	96.0	86.0	98.4	86.3	86.2	86.3	98.3	98.2	98.2	96.2	96.2	96.2	100.0		95.8	95.8	95.8	96.4	96.5	96.4	98.1	98.1	98.1	96.2	96.1	86.8	96.1	96.0	96.0	94.8		
H012-6	95.4	95.3	86.6	95.8	86.2	86.2	86.3	95.9	95.8	95.8	95.5	95.7	95.5	95.9	95.9		100.0	99.7	95.9	96.1	96.0	95.9	96.0	96.0	95.8	95.8	87.0	95.8	96.0	95.9	94.3		
H012-8	95.4	95.3	86.6	95.8	86.2	86.3	86.3	95.9	95.7	95.7	95.6	95.6	95.5	95.8	95.9	100.0		99.7	95.9	96.1	96.0	96.0	96.0	96.0	95.9	95.8	95.8	87.0	95.8	96.0	95.9	94.4	HC-H012	
H012-18	95.4	95.2	86.7	95.7	86.3	86.3	86.3	95.9	95.7	95.7	95.5	95.5	95.5	95.9	95.9	99.7	99.7		96.0	96.1	96.1	96.0	96.0	96.0	95.9	95.8	95.8	87.0	95.9	95.9	95.8	94.4		
H019-1	95.9	95.8	86.6	95.9	86.3	86.3	86.2	96.2	96.2	96.2	95.9	95.8	95.8	96.5	96.5	95.9	96.0	96.0		100.0	100.0	96.1	96.2	96.2	96.3	96.4	87.0	96.7	96.5	96.5	94.9		
H019-5	95.9	95.8	86.5	95.9	86.2	86.2	86.2	96.2	96.2	96.2	95.8	95.7	96.0	96.4	96.5	95.9	95.8	96.0	100.0		100.0	96.2	96.2	96.2	96.4	96.3	86.9	96.6	96.5	96.5	94.9	HC-H019	
H019-6	95.9	95.7	86.4	95.9	86.2	86.1	86.2	96.1	96.1	96.1	95.9	95.8	95.8	96.5	96.6	95.9	96.0	95.9	100.0	100.0		96.2	96.2	96.2	96.3	96.4	87.0	96.7	96.5	96.5	94.8		
H105-2	96.2	96.1	85.9	98.3	86.0	86.0	86.1	98.1	98.0	98.1	96.1	96.0	96.1	98.2	98.1	95.8	95.9	95.8	96.2	96.1	96.1		100.0	100.0	95.9	95.9	86.4	96.0	96.1	96.0	94.5		
H105-6	96.2	96.1	85.9	98.3	86.0	85.9	86.0	98.1	98.1	98.1	96.1	96.1	96.1	98.2	98.1	95.8	95.9	95.8	96.1	96.0	96.0	100.0		100.0	95.9	95.9	86.4	96.0	96.2	95.9	94.4	HC-H105	HC
H105-8	96.2	96.2	85.9	98.3	86.0	86.0	86.0	98.2	98.1	98.1	96.2	96.1	96.1	98.2	98.2	95.9	95.9	95.8	96.2	96.1	96.0	100.0	100.0		95.9	95.9	86.2	96.0	96.1	95.9	94.5		
Ind63-29	96.0	95.9	86.6	96.0	86.4	86.4	86.5	96.1	96.1	96.1	95.7	95.6	95.8	96.2	96.2	95.8	95.8	95.8	96.4	96.3	96.4	95.9	95.9	95.9		99.9	87.7	97.0	97.1	97.1	95.2	HC-Ind63-1	
Ind63-31	96.0	96.0	86.6	96.0	86.5	86.5	86.6	96.0	96.0	96.0	95.7	95.6	95.6	96.2	96.1	95.7	95.7	95.8	96.4	96.3	96.4	96.0	96.0	95.9	99.9		87.6	97.0	97.2	97.1	95.1		
Ind63-37	86.4	86.3	88.9	86.7	87.8	87.7	87.8	86.8	86.9	86.8	86.4	86.4	86.3	86.7	86.9	86.9	86.8	87.0	86.9	86.9	86.8	86.3	86.3	86.3	87.4	87.4		87.3	86.9	86.8	87.4	HC-Ind63-2	
Ind63-40	96.1	95.8	86.7	96.1	86.1	86.1	86.3	96.1	96.1	96.1	95.9	95.8	95.8	96.1	96.1	95.7	95.7	95.8	96.6	96.5	96.5	96.0	96.0	96.0	96.9	96.8	87.5		96.9	96.9	95.0	HC-Ind63-3	
Ind117-13	96.1	96.0	86.5	96.1	86.4	86.3	86.5	96.1	96.0	96.0	95.9	96.0	96.0	96.0	96.2	95.8	95.8	95.8	96.6	96.6	96.6	96.2	96.2	96.2	97.0	97.0	87.0	96.8		100.0	95.4	HC-Ind117-1	
Ind117-18	96.0	95.9	86.4	95.9	86.4	86.2	86.4	96.0	95.9	95.9	95.9	95.9	96.0	95.9	96.0	95.7	95.7	95.6	96.5	96.6	96.6	96.1	96.1	96.1	97.0	97.0	87.0	96.7	99.9		95.3		
Ind117-32	94.4	94.4	86.9	94.7	86.8	86.6	86.7	94.7	94.7	94.7	94.6	94.7	94.7	94.8	94.8	94.4	94.3	94.5	94.9	94.9	95.0	94.6	94.7	94.6	95.1	95.2	87.4	95.1	95.4	95.4		HC-Ind117-2	

Supplementary Table 3. Average nucleotide identity analysis of 28 isolates and the reference strains

Average nucleotide identity analysis was performed among 28 isolates, JCM13464 and the reference strains (DSM 18205; GCF_000157935.1, Indica; GCF_002224675.1). Percentage of average nucleotide identity between indicating two isolates are shown. HC, isolate from healthy controls; RA, isolate from rheumatoid arthritis patient.

Strain name	Isolate name			Definitions of strain	Clade
RA-P9	P9-13			*P. copri* derived from RA patients (P9) living in Japan	A
RA-F3	F3-7	F3-48	F3-75	*P. copri* derived from RA patients (F3) living in Japan	C
RA-N001	N001-8	N001-13	N001-16	*P. copri* derived from RA patients (N001) living in Japan	A
RA-N016	N016-5	N016-9	N016-13	*P. copri* derived from RA patients (N016) living in Japan	A
RA-N115	N115-17	N115-20		*P. copri* derived from RA patients (N115) living in Japan	A
HC-H012	H012-6	H012-8	H012-18	*P. copri* derived from healthy control (H012) living in Japan	A
HC-H019	H019-1	H019-5	H019-6	*P. copri* derived from healthy control (H019) living in Japan	A
HC-H105	H105-2	H105-6	H105-8	*P. copri* derived from healthy control (H105) living in Japan	A
HC-Ind63-1	Ind63-29	Ind63-31		*P. copri* derived from healthy control (Ind63) living in India	A
HC-Ind63-2	Ind63-37			*P. copri* derived from healthy control (Ind63) living in India	D
HC-Ind63-3	Ind63-40			*P. copri* derived from healthy control (Ind63) living in India	A
HC-Ind117-1	Ind117-13	Ind117-18		*P. copri* derived from healthy control (Ind117) living in India	A
HC-Ind117-2	Ind117-32			*P. copri* derived from healthy control (Ind117) living in India	A

Supplementary Table 4. Definitions of isolated strains

Isolates with >99% homology by average nucleotide identity analysis were re-defined as identical strains. The table information includes the strain name, the corresponding isolates, host information, and belonging clade of each strain.

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

Epitope ID:606343	**Liner sequence: DYRGYWTMRYQFDSATVS**	
Strain name	**Gene possession**	**Gene name**
JCM13464	yes	hypothetical protein
RA-P9	yes	FIG00937602:hypothetical protein
RA-F3	no	−
RA-N001	yes	hypothetical protein
RA-N016	yes	hypothetical protein
RA-N115	yes	hypothetical protein
HC-H012	yes	hypothetical protein
HC-H019	yes	hypothetical protein
HC-H105	yes	hypothetical protein
HC-Ind63-1	yes	hypothetical protein
HC-Ind63-2	no	−
HC-Ind63-3	yes	hypothetical protein
HC-Ind117-1	yes	hypothetical protein
HC-Ind117-2	no	−

Supplementary Table 5. The presence of Pc-p27 epitope in *P. copri* strains

Almost all of our isolated *P. copri* strains, except for RA-F3, HC-Ind63-2, HC-Ind117-2, possessed Pc-p27 epitope sequences in their genome. The genes corresponding to the protein-coding sequences (CDSs), including the Pc-p27 epitope, were characterized by blastn searches. The linear sequence of the Pc-p27 epitope was available in the Immune Epitope Database (IEDB) (http://www.iedb.org/home_v3.php.)

NO.	Locus		Size(bp)	ID	Gene name
1	475072	475926	854	fig|165179.43.peg.373	hypothetical protein
2	475941	476294	353	fig|165179.43.peg.374	hypothetical protein
3	476525	476917	392	fig|165179.43.peg.375	hypothetical protein
4	476904	477440	536	fig|165179.43.peg.376	hypothetical protein
5	477520	479544	2024	fig|165179.43.peg.377	hypothetical protein
6	479631	480635	1004	fig|165179.43.peg.378	hypothetical protein
7	480632	481294	662	fig|165179.43.peg.379	hypothetical protein
8	481974	483662	1688	fig|165179.43.peg.380	ATP-dependent DNA helicase RecG-related protein
9	484286	484414	128	fig|165179.43.peg.381	hypothetical protein
10	485234	485482	248	fig|165179.43.peg.382	hypothetical protein
11	485724	485951	227	fig|165179.43.peg.383	hypothetical protein
12	485963	486505	542	fig|165179.43.peg.384	hypothetical protein
13	486692	486829	137	fig|165179.43.peg.385	hypothetical protein
14	487154	487429	275	fig|165179.43.peg.386	hypothetical protein
15	487438	487578	140	fig|165179.43.peg.387	hypothetical protein
16	487696	487821	125	fig|165179.43.peg.388	hypothetical protein
17	487773	488282	509	fig|165179.43.peg.389	hypothetical protein
18	488765	489076	311	fig|165179.43.peg.390	hypothetical protein
19	489122	489913	791	fig|165179.43.peg.391	Chromosome (plasmid) partitioning protein ParA
20	490074	490793	719	fig|165179.43.peg.392	hypothetical protein
21	490806	490967	161	fig|165179.43.peg.393	hypothetical protein
22	491300	491830	530	fig|165179.43.peg.394	hypothetical protein
23	492007	492654	647	fig|165179.43.peg.395	DNA primase (EC 2.7.7.-);Ontology_term=KEGG_ENZYME:2.7.7.-
24	492678	493718	1040	fig|165179.43.peg.396	Toprim domain protein
25	493831	494436	605	fig|165179.43.peg.397	hypothetical protein
26	494438	495775	1337	fig|165179.43.peg.398	hypothetical protein
27	495772	496071	299	fig|165179.43.peg.399	hypothetical protein
28	496083	496217	134	fig|165179.43.peg.400	hypothetical protein
29	496339	496644	305	fig|165179.43.peg.401	hypothetical protein
30	496956	497219	263	fig|165179.43.peg.402	hypothetical protein
31	497921	498568	647	fig|165179.43.peg.403	hypothetical protein
32	498776	499348	572	fig|165179.43.peg.404	Chromate transport protein
33	499371	499820	449	fig|165179.43.peg.405	Chromate transport protein
34	500072	501373	1301	fig|165179.43.peg.406	Nucleoside permease NupG
35	501761	502024	263	fig|165179.43.peg.407	hypothetical protein
36	502085	502840	755	fig|165179.43.peg.408	Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-); Ontology_term=KEGG_ENZYME:2.1.1.-
37	502879	504252	1373	fig|165179.43.peg.409	hypothetical protein
38	504271	504915	644	fig|165179.43.peg.410	ABC transporter ATP-binding protein
39	505005	506204	1199	fig|165179.43.peg.411	hypothetical protein
40	506574	509780	3206	fig|165179.43.peg.412	Oar protein
41	510074	511348	1274	fig|165179.43.peg.413	Beta-lactamase (EC 3.5.2.6);Ontology_term=KEGG_ENZYME:3.5.2.6
42	511583	512716	1133	fig|165179.43.peg.414	putative phosphoesterase

Supplementary Table 6. The list of *P. copri*$_{RA}$-specific gene in CTnPc

The genes corresponding to the protein-coding sequences (CDSs) were characterized by blastn searches. Locus indicates the sequence number of the reference strain N001-13.

MATERIALS AND METHODS

Patient participation

To collect the fecal samples with a high abundance of *P. copri*, we used the data from three cohort. Cohort 1 consists of fecal samples obtained from 17 rheumatoid arthritis (RA) patients and 14 healthy controls collected in a previously published study (1). Cohort 2 consists of fecal samples obtained from newly recruited 55 Japanese RA patients at Osaka University Hospital, the National Hospital Organization Osaka Minami Medical Center, and Daini Osaka Police Hospital. RA patients were diagnosed according to the 2010 Rheumatoid Arthritis Classification Criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative (2). We also recruited 28 healthy Japanese controls at Osaka University Hospital, the National Hospital Organization Osaka Minami Medical Center in this study, and added the fecal samples of two healthy Indian controls obtained in a previous study (3). We excluded participants using antibiotics for at least 3 months prior to sampling and known history of malignancy or serious disease of the heart, liver, or kidney. The study was approved by the ethical committees of Osaka University and related medical institutions. All subjects provided written informed consent before participation. Finally, cohort 3 which consists of fecal samples with a high abundance of *P. copri*, obtained in a previous study (4) was applied for metagenome shotgun sequencing analysis.

Fecal sample collection and DNA extraction

Feces were collected into collection tubes (Sarstedt) containing RNA*later* (Ambion, Austin, TX, USA) and stored at 4°C until DNA extraction. Fecal samples for isolation of *Prevotella copri* were immediately placed into a box with AnaeroPack (Mitsubishi Gas Chemical) to ensure anaerobic conditions. Bacterial DNA from human and murine fecal samples was extracted as previously described (1). Briefly, fecal samples collected in tubes containing RNA*later* were homogenized with a 10-fold amount (v/w) of RNA*later*.

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

Then, the homogenates were washed twice with 1 ml PBS and stored at -20°C until DNA extraction. The fecal suspension with 300 μl of sodium dodecyl sulfate–

Tris solution, 0.3 g glass beads (diameter 0.1 mm) (BioSpec Products) and 500 μl TE-saturated phenol was vortexed vigorously for 30 s at a power level of 5.0 using a FastPrep-24 (MP Biomedicals). The mixtures were centrifuged at 20,000 × g for 5 min at 4°C, and 400 μl of supernatant was collected. Subsequently, the supernatant mixed with phenol-chloroform-isoamylalcohol (25:24:1) was vortexed vigorously for 45 s at a power level of 4.0 using a FastPrep-24 and subjected to centrifugation (20,000 × g for 5 min at 4°C). The supernatant (250 μl) was collected and DNA was precipitated with isopropanol. Finally, the DNA was suspended in 200 μl of TE buffer and stored at -20°C.

16S rRNA gene sequencing

Fecal DNA was extracted from human or murine fecal samples by the method written in 'Fecal sample collection and DNA extraction'. Each library was prepared according to the Illumina 16S Metagenomic Sequencing Library Preparation Guide with primer set 27Fmod: 5′AGRGTTTGATCMTGGCTCAG-3′ and 338R: 5′-TGCTGCCTCCCGTAGGAGT-3′ targeting the V1−V2 region of 16S rRNA genes; 251-bp paired end sequencing of the amplicons was performed on a MiSeq system (Illumina) using a MiSeq Reagent v2 500 cycle kit. The paired end sequences obtained were merged, filtered, and denoised using DADA2. Taxonomic assignment was performed using QIIME2 feature-classifier plugin with the Greengenes 13_8 database. The QIIME2 pipeline, version 2020.2, was used as the bioinformatics environment for the processing of all relevant raw sequencing data (https://qiime2.org).

Quantitative PCR of bacterial 16S rRNA genes

Fecal DNA was screened for the presence of isolated *P. copri* using quantitative PCR (Q-PCR) primers designed against conserved regions of the *P. copri* genome from whole

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

genome sequencing data. Q-PCR was performed with the GoTaq qPCR master mix (Promega) on the StepOnePlus™ (Applied Biosystems), with the following cycle conditions: 95°C for 5 min, then 40 cycles of 95°C for 30 s, 62°C for 20 s and 72°C for 20 s. The standard curve obtained from DNA of *P. copri* JCM 13464 was used to determine the absolute quantity of *P. copri* in the feces. Absolute copy numbers per 1 g of feces were assessed with respect to standard curve values obtained for this bacterial analysis. A melting curve analysis was performed after amplification to distinguish the targeted PCR products from the non-targeted products. The melting curve was obtained by slow heating at temperatures from 60°C to 95°C with continuous fluorescence monitoring. The following primer set was used: *P. copri*-1Fw: 5′-GTCTGCCTCTGCTCTCTATGGT-3′ and *P. copri*-1Rv: 5′-ACCTAATCCCTTCTTGGCTTTC-3′.

Isolation of *Prevotella copri* from RA patients and healthy controls

To isolate *P. copri* strains from human fecal samples, a spoonful of feces was resuspended in phosphate-buffered saline (PBS), then 50 µl of serial dilutions were placed on Columbia agar with 5% sheep blood (Becton Dickinson). After 24–48 h culture in an anaerobic chamber, isolated colonies were picked, streaked onto fresh plates, and direct colony PCR was performed to identify *P. copri* candidates using the following specific primers against the *Prevotella* genus 16S rRNA gene: gPrev Fw: 5′-CACRGTAACGATGGATGCC-3′ and gPrev Rv: 5′-GGTCGGGTTGCAGACC-3′. Positive colonies were again subjected to direct colony PCR using primers 8F: 5′-AGAGTTTGATCMTGGTCAG-3′ and 15R: 5′-AAGGAGGTGATCCARCCGCA-3′ targeting the full-length 16S rRNA gene. The 16S rRNA gene sequences of the positive colonies were determined using Sanger sequencing methods. Sanger sequencing was performed using the BigDye Terminator v3.1 cycle sequencing kit on an Applied Biosystems 3730 DNA analyzer. *P. copri* was identified by homology searches using the

nucleotide blast (blastn) tool. Glycerol stocks of *P. copri* isolates were stored at -80°C.

Genome sequencing and core genome analysis

Genomic DNA was extracted from bacterial isolates using the DNeasy PowerSoil kit (QIAGEN, Hilden, Germany) and was then used to prepare short and long DNA libraries. Short DNA libraries were prepared as we described previously for the Illumina sequencer. Sequences were obtained as dual-indexed, 251-bp paired-end reads using the MiSeq Reagent Kit v2 (500 Cycles; Illumina) and the MiSeq sequencer (Illumina). Long DNA libraries were prepared as we described previously for the Oxford Nanopore Technologies (ONT) sequencer. To prepare each library, 1 μg of sheared DNA was used and the library was prepared using the Ligation Sequencing Kit (ONT, Oxford, UK) and the Native Barcoding Kit (ONT), according to the manufacturers' instructions. Sequencing runs were performed using a MinION (ONT). The reads were assembled using Unicycler v0.4.7 and annotated by the RAST annotation pipeline. Coding DNA sequences (CDSs) were obtained from the RAST annotations of 29 assemblies and were clustered with an amino acid sequence identity of ≥80% and a length ratio of ≤0.8. Core genomes were constructed from conserved CDS clusters among >80% of 29 assemblies, as we described previously (5). The size of pan- and core- genomes were calculated using randomly picked strains (n=1000).

Genomic regional analysis

A homology search was performed with blastn using each genomic assembly as the query and strain N001-13 as the reference to determine an *P. copri*$_{RA}$-specific region. The chromosome of strain N001-13 was separated by a 10-kb width of genome bins and was evaluated in the presence of the bin in the query assembly whether or not the search results entirely covered each bin. The core region of density was defined as the ratio of the number of strains that possessed each genomic region. The *P. copri*$_{RA}$ *and P. copri*$_{HC}$ core

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

regions of density were also defined for *P. copri*$_{RA}$ *and P. copri*$_{HC}$ strains, respectively. The *P. copri*$_{RA}$-specific region of density was calculated as the ratio of the *P. copri*$_{RA}$ core region of density to the *P. copri*$_{HC}$ core region of density.

Comparison of *P. copri*$_{RA}$ -specific genomic regions between healthy controls and RA patients in the fecal metagenome

To examine the difference in *P. copri*$_{RA}$-specific genomic regions between healthy controls and RA patients, we used the phylogenetic analysis data from the metagenomic shotgun sequencing in the reference article (4). Samples of which the relative abundance of *P. copri* was higher than one percent were selected for analysis. We calculated the content of *P. copri*$_{RA}$-specific genomic regions in each sample divided by the relative proportion of *P. copri* to estimate the relative proportion of strains with *P. copri*$_{RA}$-specific genomic regions among the total *P. copri* of each sample.

Mice

DBA/1J mice were obtained from Oriental Yeast in Japan, and were bred and maintained under specific-pathogen-free (SPF) conditions until inoculation with *P. copri*. SKG mice in a BALB/c background were also used, as described previously (1), and were maintained under SPF conditions until inoculated with *P. copri*. All animal experiments were performed in accordance with the Osaka University Animal Experimental Guidelines.

Colonization of isolated *P. copri* and the induction of arthritis

To generate *P. copri*-colonized DBA/1J mice, DBA/1J mice of 6–8 weeks of age, maintained under SPF conditions, were provided with autoclaved drinking water containing dissolved ampicillin sodium salt (500 mg/l; Nacalai Tesque), neomycin sulfate (500 mg/l; Nacalai Tesque), metronidazole (500 mg/l; Nacalai Tesque) and vancomycin

(250 mg/l; Duchefa Biochemie B.V.) for 5 days. Three days after antibiotic treatment, mice were inoculated daily with 200 μl of *P. copri*$_{RA}$ or *P. copri*$_{HC}$ (6.0×10^{10} *P. copri*/mouse, in PBS) for the following 5 days. Three days after the final inoculation, arthritis was induced in DBA/1J mice by injection of type II collagen, as previously described (6). Briefly, bovine type II collagen (Collagen Research Center) was dissolved in 0.05 M acetic acid to a concentration of 2 mg/ml, and mixed with an equal volume of Freund's complete adjuvant (263810: *Mycobacterium butyricum*; BD). On day 0 and 21, DBA/1J mice were immunized at the base of the tail with 100 μl of emulsion.

To generate *P. copri*-monocolonized SKG mice, SPF SKG mice between 5 and 6 weeks of age were orally inoculated with pure preparations of cultured 1.0×10^8 *P. copri*$_{RA}$ or *P. copri*$_{HC}$ and maintained in separate gnotobiotic isolators. The bacterial counts of *Prevotella* in the feces of inoculated mice were analyzed by Q-PCR of the bacterial 16S rRNA gene in each mouse at 2 weeks after colonization. For the induction of arthritis in SKG mice, the mice were intraperitoneally injected with zymosan (0.5 mg) at 3 weeks after bacterial inoculation.

Clinical assessment of arthritis

In mice with collagen-induced arthritis (CIA), disease severity was measured as the mean clinical score for each of the animal's four paws, where 0 = no swelling, 1 = mild swelling confined to the tarsals or ankle joint, 2 = mild swelling extending from the ankle to the tarsals, 3 = moderate swelling extending from the ankle to the metatarsal joints and 4 = severe swelling encompassing the ankle, foot and digits. The cumulative score for the four paws of each mouse (maximum score: 16) was used as the arthritis score to represent overall disease severity. Arthritis scores of CIA mice were evaluated twice per week.

In SKG mice, joint swelling was monitored by inspection and scored as follows: 0 = no joint swelling, 0.1 = swelling of one finger joint, 0.5 = mild swelling of the wrist or ankle and 1.0 = severe swelling of the wrist or ankle. The scores for all fingers of the fore paws

and hind paws, wrists and ankles were totaled for each mouse. Arthritis scores of SKG mice were evaluated once per week.

Incidence evaluation of arthritis

The incidence of arthritis was calculated by dividing the number of mice with an arthritis score of >1 point by the total number of mice in each experiment. The rate of incidence for each arthritis model was evaluated at each scoring time point.

Histological analysis

Tissues were fixed with 4% paraformaldehyde phosphate buffer solution. Paraffin-embedded sections mounted on glass slides were used for hematoxylin and eosin (H&E) staining. The lower ankle joints of DBA/1J mice and SKG mice after the induction of arthritis were used for the histological study. Images were taken under a microscope. The severity of inflammation in the lower ankle joints was scored as follows: 0 = none, 1 = mild, 2 = moderate and 3 = severe.

ELISA for anti-type II collagen antibody

The determination of anti-type II collagen antibodies was performed as follows. ELISA plates were

coated with bovine collagen II (Chondrex, Seattle, WA, USA) at 4°C. Following blocking with 3% bovine serum albumin (BSA), sera were applied at a 1:200 dilution for 2 h. Binding was detected using a goat anti-mouse-IgG-alkaline-phosphatase secondary antibody, diluted 1:5000 with the appropriate substrate. Data were read using an ELISA reader at 450 nm optical density (OD). A standard pool of anti-type II collagen antibodies was obtained by combining sera from several mice.

Isolation of lymphocytes

Intestinal lamina propria lymphocytes and myeloid cells were prepared as described previously (1). In brief, large intestines were opened longitudinally, washed to remove fecal content, and incubated in HBSS with 5 mM EDTA at 37°C for 20 min in a shaking water bath. After washing with PBS, tissues were cut into small pieces and incubated into RPMI 1640 medium containing 4% FBS, 1 mg/ml collagenase D (Roche Diagnostics), 0.5 mg/ml dispase (Gibco) and 40 μg/ml DNase I (Roche Diagnostics) for 50 min at 37°C in a shaking water bath. The digested tissues were resuspended in 5 ml of 40% Percoll (GE Healthcare) and then overlaid on 2.5 ml of 80% Percoll. Percoll gradient separation was performed by centrifugation at 2,000 rpm for 20 min at room temperature. The lamina propria lymphocytes were collected at the interface of the Percoll gradient and washed with PBS containing 2% FBS. To prepare lymphocytes from regional lymph nodes, tissues were removed from mice and mechanically mashed by slide glasses in RPMI 1640 medium. Collected cells were washed with PBS containing 2% FBS.

Flow cytometry

$CD4^+$ T cell intracellular IL-17A and IFN-γ expression was analyzed with a Cytofix/Cytoperm Plus kit with GolgiStop, according to the manufacturer's instructions. In brief, $CD4^+$ T cells were incubated with 50 ng/ml PMA, 5 μM calcium ionophore A23187 and GolgiStop in complete RPMI 1640 with 10% FBS at 37°C for 4 h. Surface staining was performed with anti-CD45 and anti-CD4 antibodies for 20 min at 4°C, and intracellular cytokine staining was performed with anti-IL-17A and anti-IFN-γ antibodies for 20 min. For the analysis of myeloid cells from the intestinal lamina propria, cells were pretreated with anti-CD16/32 monoclonal antibody to block nonspecific binding, and then incubated with the indicated monoclonal antibodies. $CD11b^+CD11c^+$ myeloid cells in the intestinal lamina propria were stained with anti-CD45, anti-CD11b, anti-CD11c and anti-MHC class II antibodies for 20 min at 4°C. Flow cytometric analysis was performed using a FACSCanto II flow cytometer (BD) with FlowJo software (Tree Star).

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

Antibodies and reagents

Fluorescein isothiocyanate (FITC) anti-IFNγ (XMG1.2), allophycocyanin (APC) anti-IL-17A (TC11- 18H10.1), PerCP/Cyanine5.5 anti-CD4 (GK1.5), Pacific Blue anti-CD8a (53-6.7), APC anti-CD19 (1D3/CD19), Pacific Blue anti-CD45 (30-F11), APC anti-CD11b (M1- 70) and PE/Cyanine7 anti-CD11c (N418) antibodies were purchased from BioLegend. Phycoerythrin (PE) anti-CD80 (16-10A1) antibodies were purchased from BD Biosciences. FITC anti-CD86 (GL-1) antibodies were purchased from ThermoFisher scientific. PE anti-MHC class II (M5/114.15.2) and anti-CD16/32 (2.4G2) antibodies were purchased from Tonbo Biosciences. The Ca2+ ionophore A23187, phorbol 12-myristate 13-acetate (PMA) and zymosan A were purchased from Sigma-Aldrich. Bovine type II collagen (K41) was purchased from Collagen Research Center. Freund's complete adjuvant (263810: *Mycobacterium butyricum*) and Cytofix/Cytoperm Plus Fixation/Permeabilization Solution Kit with BD GolgiPlug were purchased from BD Biosciences. The Mouse Naïve CD4+ T cell isolation kit was purchased from Miltenyi Biotec.

Bacterial culture

Isolated *P. copri and* JCM 13464 were cultured in Gifu anaerobic medium (GAM) broth (Nissui Pharmaceutical) under anaerobic conditions (10% CO2, 10% H2, 80% N2) at 37°C for 24 h. Bacterial suspensions were washed with PBS, and the OD at 600 nm was measured and Q-PCR analysis was performed to estimate bacterial numbers.

Stimulation of bone marrow-derived dendritic cells

Bone marrow was isolated from the femur and tibia of DBA/1J mice and differentiated in the presence of GM-CSF (FUJIFILM Wako Pure Chemical) for 7 days to generate bone marrow-derived dendritic cells (BMDCs). In the analysis of cytokine production,

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

BMDCs (1×10^5) were stimulated with heat- killed *P. copri*$_{RA}$ or *P. copri*$_{HC}$ for 12 h. Cytokine concentrations (i.e., IL-6, IL-23, and TNF-α) in the supernatants were analyzed by ELISA (R&D Systems). In the analysis of mRNA expression, BMDCs (1×10^6) were stimulated with heat-killed *P. copri*$_{RA}$ or *P. copri*$_{HC}$ for 6 h. Total RNA of stimulated BMDCs was isolated using the GenElute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich), and 1–2 μg of RNA were reverse transcribed using the ReverTra Ace qPCR RT Master Mix with gDNA Remover (Toyobo; Osaka, Japan). Real-time RT-PCR was performed on a Step One Plus Real-Time PCR System (Applied Biosystems) using GoTaq qPCR Master Mix (Promega). All values were normalized to the expression of the *Gapdh* gene, encoding glyceraldehyde-3-phosphate dehydrogenase, and the fold difference in expression relative to that of *Gapdh* is shown. The following primer sets were used: *Il6*, forward 5′-CTGCAAGAGACTTCCATCCAGTT-3′ and reverse 5′-AAGTAGGGAAGGCCGTGGTT-3′; *Il23a*, forward 5′-CCAGCAGCTCTCTCGGAATCT-3′ and reverse 5′-GGGTCACAACCATCTTCACAC-3′; *Gapdh*, forward 5′-TGGATATTGTTGCCATCAATG-3′ and reverse 5′-TGATGGGATTTCCATTGATGA-3′.

Activation of T cells by dendritic cells

BMDCs (2×10^5) prepared from DBA/1J mice were incubated with heat-killed *P. copri* for 12 h. Splenic naïve $CD4^+$ T cells from 8-week-old DBA/1J mice were sorted by magnetic-activated cell sorting. Then, 5×10^6 naïve $CD4^+$ T cells were co-cultured with 1×10^5 bacteria-stimulated dendritic cells for 5 days. T cells (5×10^5) were restimulated with BMDCs (1×10^5) from DBA/1J mice in the presence or absence of type II collagen. Two days later, the culture supernatant was collected for ELISA to evaluate cytokine production.

Nii T, *et al. Ann Rheum Dis* 2023; 82:621–629. doi: 10.1136/ard-2022-222881

Fluorescence *in situ* hybridization

The tissues of the large and small intestine were isolated from SKG mice at 11 weeks after colonization of *P. copri*$_{RA}$ or *P. copri*$_{HC}$ and were fixed in methanol-Carnoy's fixative solution (60% methanol, 30% chloroform and 10% acetic acid). Paraffin-embedded sections (5 μm) were then dewaxed and hydrated. For the detection of inoculated *P. copri*, probe Cy3-conjugated Eubacteria (5′-GCTGCCTCCCGTAGGAGT-3′) was used. The sections were incubated with 1 μg of the respective probes in 200 μl hybridization buffer (750 mM NaCl, 100 mM Tris-HCl [pH 7.4], 5 mM EDTA, 0.01% bovine serum albumin, 10% dextran sulfate) at 40° C for 16 h. After thorough rinsing with wash buffer (50 mM NaCl, 4 mM Tris-HCl [pH 7.4], 0.02 mM EDTA) at 45° C for 20 min, the sections were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) (Vector Laboratories). Then, the sections were observed using a confocal microscope (FV1000-D; Olympus).

Statistical analysis

Data are presented as the mean ± s.d., as indicated in the figure legends. Statistical significance to compare two groups were evaluated using two-tailed unpaired Student's *t*-test or two-tailed Mann-Whitney U test. Differences among categorical values were analyzed by performing a χ^2-test. For multiple groups comparisons, one-way analysis of variance (ANOVA) followed by a Tukey's multiple comparison test was performed. Differences of $P<0.05$ were considered statistically significant. Statistical analysis was performed using the GraphPad version 8.4.3 (GraphPad Software)

REFERENCES

1. Maeda Y, Kurakawa T, Umemoto E, Motooka D, Ito Y, Gotoh K, et al. Dysbiosis Contributes to Arthritis Development via Activation of Autoreactive T Cells in the Intestine. Arthritis & rheumatology (Hoboken, NJ). 2016;68(11):2646-61.

2. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis and rheumatism. 2010;62(9):2569-81.

3. Pareek S, Kurakawa T, Das B, Motooka D, Nakaya S, Rongsen-Chandola T, et al. Comparison of Japanese and Indian intestinal microbiota shows diet-dependent interaction between bacteria and fungi. NPJ biofilms and microbiomes. 2019;5(1):37.

4. Tomofuji Y, Kishikawa T, Maeda Y, Ogawa K, Nii T, Okuno T, et al. Whole gut virome analysis of 476 Japanese revealed a link between phage and autoimmune disease. Annals of the rheumatic diseases. 2022;81(2):278-88.

5. Matsumoto Y, Kinjo T, Motooka D, Nabeya D, Jung N, Uechi K, et al. Comprehensive subspecies identification of 175 nontuberculous mycobacteria species based on 7547 genomic profiles. Emerging microbes & infections. 2019;8(1):1043-53.

6. Brand DD, Latham KA, Rosloniec EF. Collagen-induced arthritis. Nature protocols. 2007;2(5):1269-75.