

Supplementary Table 1: Demographic and clinical characteristics of discovery and validation cohorts.

Discovery cohort			
Feature	Case	Control	P-value
Number	97	93	-
% Sulfasalazine	25.77	0	-
% TNFi	58.76	0	-
Age	35.66±11.32	34.46±11.65	0.47
BMI	22.61±2.84	23.18±2.55	0.15
% Male	80.41	84.95	0.41
% B27 Pos	96.91	3.63 (38 missed)	0.00
Disease Duration (yr)	10.87±9.23	0	-
% Smoker	43.30	40.86	0.74
BASDAI	3.33±2.22	0	-

Validation cohort			
Feature	Case	Control	P-value
Number	30	30	-
% Sulfasalazine	26.67	0	-
% TNFi	33.33	0	-
Age	41.67±17.22	23.8±0.89	0.00
BMI	21.27±2.18	22.69±2.37	0.02
% Male	86.67	90	0.69
% B27 Pos	96.67	0	0.00
Disease Duration (yr)	13.37±11.51	0	-
% Smoker	30.00	0	0.00
BASDAI	3.18±2.14	0	-

Supplementary Table 2: Differentially abundant taxa in cases and controls in discovery and validation cohorts, identified by linear modelling.

Discovery cohort			
Feature	Coefficient	P-value	Q-value
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Peptostreptococcaceae	-0.0099	0.00022	0.048
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Peptostreptococcaceae g_Peptostreptococcaceae_noname	-0.0098	0.00024	0.048
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Peptostreptococcaceae g_Peptostreptococcaceae_noname s_Peptostreptococcaceae_noname_unclassified	-0.0085	0.0010	0.13
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Roseburia s_Roseburia_inulinivorans	0.040	0.0027	0.22
k_Bacteria p_Firmicutes c_Erysipelotrichia o_Erysipelotrichales f_Erysipelotrichaceae g_Holdemania	-0.0071	0.0034	0.23
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Lachnospiraceae_noname s_Lachnospiraceae_bacterium_5_1_63FAA	0.0037	0.0063	0.28
k_Bacteria p_Actinobacteria c_Actinobacteria o_Bifidobacteriales f_Bifidobacteriaceae g_Bifidobacterium s_Bifidobacterium_adolescentis	0.012	0.0069	0.28
k_Bacteria p_Firmicutes c_Erysipelotrichia o_Erysipelotrichales f_Erysipelotrichaceae g_Holdemania s_Holdemania_filiformis	-0.0050	0.013	0.32
k_Bacteria p_Proteobacteria c_Betaproteobacteria o_Burkholderiales f_Sutterellaceae g_Parasutterella	0.0062	0.014	0.32
k_Bacteria p_Proteobacteria c_Betaproteobacteria o_Burkholderiales f_Sutterellaceae g_Parasutterella s_Parasutterella_excrementihominis	0.0062	0.014	0.32
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiaceae g_Clostridium s_Clostridium_hathewayi	-0.0052	0.016	0.32
k_Bacteria p_Proteobacteria c_Betaproteobacteria o_Burkholderiales f_Burkholderiales_noname	0.0060	0.016	0.32
k_Bacteria p_Proteobacteria c_Betaproteobacteria o_Burkholderiales f_Burkholderiales_noname g_Burkholderiales_noname	0.0060	0.016	0.32
k_Bacteria p_Proteobacteria c_Betaproteobacteria o_Burkholderiales f_Burkholderiales_noname g_Burkholderiales_noname s_Burkholderiales_bacterium_1_1_47	0.0060	0.016	0.32
k_Bacteria p_Proteobacteria c_Gammaproteobacteria o_Enterobacteriales f_Enterobacteriaceae g_Enterobacter s_Enterobacter_cloacae	0.013	0.017	0.32
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiales_noname	-0.0050	0.025	0.39
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Coproccoccus s_Coproccoccus_comes	0.013	0.025	0.39
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiales_noname g_Clostridiales_noname	-0.0036	0.030	0.40
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiales_noname g_Clostridiales_noname s_Clostridiales_bacterium_1_7_47FAA	-0.0036	0.030	0.40
k_Bacteria p_Bacteroidetes c_Bacteroidia o_Bacteroidales f_Porphyromonadaceae g_Odoribacter s_Odoribacter_splanchnicus	-0.0095	0.030	0.40
k_Bacteria p_Firmicutes c_Negativicutes o_Selenomonadales f_Veillonellaceae g_Megamonas s_Megamonas_rupellensis	-0.0023	0.035	0.43
k_Bacteria p_Firmicutes c_Negativicutes o_Selenomonadales f_Veillonellaceae g_Dialister s_Dialister_invisus	-0.022	0.043	0.49
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Blautia s_Ruminococcus_obeum	0.0042	0.046	0.50
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Coproccoccus	0.018	0.047	0.50
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiaceae g_Clostridium s_Clostridium_bolteae	-0.0075	0.050	0.51

Validation cohort			
Feature	Coefficient	P-value	Q-value
k_Bacteria p_Actinobacteria c_Actinobacteria o_Bifidobacteriales f_Bifidobacteriaceae g_Bifidobacterium s_Bifidobacterium_adolescentis	0.055	0.00059	0.070
k_Bacteria p_Bacteroidetes c_Bacteroidia o_Bacteroidales f_Bacteroidaceae g_Bacteroides s_Bacteroides_thetaiotaomicron	-0.044	0.00071	0.070
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiaceae g_Clostridium s_Clostridium_hathewayi	-0.0074	0.0040	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Dorea	0.036	0.0042	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Lachnospiraceae_noname s_Lachnospiraceae_bacterium_5_1_63FAA	0.011	0.0069	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Dorea s_Dorea_longicatena	0.021	0.0071	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Butyrvivrio	0.011	0.010	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Butyrvivrio s_Butyrvivrio_unclassified	0.011	0.010	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Coproccoccus s_Coproccoccus_comes	0.029	0.011	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Ruminococcaceae g_Anaerotruncus s_Anaerotruncus_colihominis	-0.0082	0.013	0.26
k_Bacteria p_Firmicutes	0.15	0.013	0.26
k_Bacteria p_Actinobacteria c_Actinobacteria o_Actinomycetales f_Micrococcaceae	-0.0077	0.015	0.26
k_Bacteria p_Actinobacteria c_Actinobacteria o_Actinomycetales f_Micrococcaceae g_Rothia	-0.0077	0.015	0.26
k_Bacteria p_Actinobacteria c_Actinobacteria o_Actinomycetales	-0.0084	0.016	0.26
k_Bacteria p_Actinobacteria c_Actinobacteria o_Actinomycetales f_Micrococcaceae g_Rothia s_Rothia_mucilaginos	-0.0073	0.017	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiales_noname g_Clostridiales_noname	-0.0082	0.018	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiales_noname g_Clostridiales_noname s_Clostridiales_bacterium_1_7_47FAA	-0.0082	0.018	0.26
k_Bacteria p_Actinobacteria c_Actinobacteria o_Bifidobacteriales f_Bifidobacteriaceae g_Bifidobacterium s_Bifidobacterium_pseudocatenulatum	0.022	0.018	0.26
k_Bacteria p_Bacteroidetes c_Bacteroidia o_Bacteroidales f_Bacteroidaceae	-0.13	0.019	0.26
k_Bacteria p_Bacteroidetes c_Bacteroidia o_Bacteroidales f_Bacteroidaceae g_Bacteroides	-0.13	0.019	0.26
k_Bacteria p_Firmicutes c_Clostridia	0.14	0.020	0.26
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales	0.14	0.020	0.26
k_Bacteria p_Firmicutes c_Negativicutes o_Selenomonadales f_Veillonellaceae g_Megasphaera s_Megasphaera_unclassified	0.023	0.033	0.39
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Dorea s_Dorea_formicigenerans	0.025	0.035	0.39
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Lachnospiraceae g_Roseburia s_Roseburia_inulinivorans	0.070	0.038	0.39
k_Bacteria p_Firmicutes c_Clostridia o_Clostridiales f_Clostridiaceae g_Clostridium s_Clostridium_bolteae	-0.0071	0.038	0.39

k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae g__Paraprevotella s__Paraprevotella_unclassified	0.013	0.041	0.39
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae g__Prevotella s__Prevotella_copri	0.0046	0.041	0.39
k__Bacteria p__Proteobacteria c__Betaproteobacteria o__Burkholderiales f__Burkholderiales_noname	0.014	0.047	0.39
k__Bacteria p__Bacteroidetes c__Bacteroidia o__Bacteroidales f__Prevotellaceae g__Paraprevotella s__Paraprevotella_clara	0.012	0.049	0.39

Supplementary Table 3: KEGG Orthogroup comparisons in cases and controls in discovery and validation cohorts, identified by linear modelling.

Discovery cohort			
Feature	Coefficient	P-value	Q-value
TRIOSEPISOMERIZATION-RXN	-0.00057	0.0017	0.39
TRYPTOPHAN--TRNA-LIGASE-RXN	-0.00065	0.0025	0.39
3.1.22.4-RXN	-0.00057	0.0056	0.39
RXN-12587	0.00050	0.0056	0.39
RXN-12588	0.00050	0.0056	0.39
RXN-14384	0.00050	0.0056	0.39
RXN-14385	0.00050	0.0056	0.39
RXN-14386	0.00050	0.0056	0.39
RXN-15881	0.00050	0.0056	0.39
RXN-9787	0.00050	0.0056	0.39
RXN0-308	0.00050	0.0056	0.39
3.1.11.6-RXN	-0.0013	0.0079	0.44
3.5.1.80-RXN	-0.00065	0.0089	0.44
NAG6PDEACET-RXN	-0.00065	0.0089	0.44
ACETOLACTREDUCTOISOM-RXN	-0.00051	0.0093	0.44
ACETOHBUTREDUCTOISOM-RXN	-0.00051	0.0093	0.44
GLURS-RXN	-0.00033	0.0098	0.44
3.4.21.53-RXN	0.00081	0.012	0.48
RXN-11322	0.0014	0.012	0.48
ATPPHOSPHORIBOSYLTRANS-RXN	0.00039	0.015	0.55
SERINE-O-ACETTRAN-RXN	-0.00052	0.018	0.64
HOLO-ACP-SYNTH-RXN	-0.00076	0.022	0.65
RXN-10994	-0.00076	0.022	0.65
RXN-16759	-0.00076	0.022	0.65
4.2.99.18-RXN	-0.00036	0.023	0.65
RXN0-2601	-0.00036	0.023	0.65
5.1.3.9-RXN	-0.00062	0.024	0.65
NANE-RXN	-0.00062	0.024	0.65
HOMOSERKIN-RXN	-0.00058	0.026	0.67
ENTDB-RXN	-0.00075	0.029	0.71
RXN-15889	-0.00075	0.029	0.71
HISTAMINOTRANS-RXN	0.00045	0.031	0.73
SUPEROX-DISMUT-RXN	-0.00080	0.032	0.74
DEOXYRIBOSE-P-ALD-RXN	-0.00063	0.035	0.79
POLYPHOSPHATE-KINASE-RXN	0.00056	0.037	0.81
SHIKIMATE-KINASE-RXN	0.00038	0.047	0.90
DTMPKI-RXN	-0.00054	0.048	0.90

Validation cohort			
Feature	Coefficient	P-value	Q-value
2.3.1.179-RXN	0.0025	8.56E-06	0.00037
RXN-8391	0.0025	8.56E-06	0.00037
RXN1G-1015	0.0025	8.56E-06	0.00037
RXN1G-1212	0.0025	8.56E-06	0.00037
RXN1G-172	0.0025	8.56E-06	0.00037

RXN1G-202	0.0025	8.56E-06	0.00037
RXN1G-32	0.0025	8.56E-06	0.00037
RXN1G-334	0.0025	8.56E-06	0.00037
RXN1G-374	0.0025	8.56E-06	0.00037
RXN1G-424	0.0025	8.56E-06	0.00037
RXN1G-557	0.0025	8.56E-06	0.00037
RXN1G-580	0.0025	8.56E-06	0.00037
RXN1G-582	0.0025	8.56E-06	0.00037
RXN1G-660	0.0025	8.56E-06	0.00037
RXN1G-79	0.0025	8.56E-06	0.00037
RXN1G-818	0.0025	8.56E-06	0.00037
RXN1G-883	0.0025	8.56E-06	0.00037
RXN1G-89	0.0025	8.56E-06	0.00037
RXN1G-94	0.0025	8.56E-06	0.00037
GLYC3PDEHYDROGBIOSYN-RXN	0.0010	2.78E-05	0.00114
RXN-11637	0.0011	9.14E-05	0.00356
RXN-11135	0.0020	0.00017	0.00650
DNA-DIRECTED-DNA-POLYMERASE-RXN	0.0024	0.00022	0.00715
GLYCOGEN-BRANCH-RXN	0.0028	0.00024	0.00715
RXN-14371	0.0028	0.00024	0.00715
RXN-14372	0.0028	0.00024	0.00715
RXN-7669	0.0028	0.00024	0.00715
RXN-7710	0.0028	0.00024	0.00715
PANTOTHENATE-KIN-RXN	0.0010	0.00026	0.00735
SUPEROX-DISMUT-RXN	-0.0026	0.00030	0.00830
3.4.21.53-RXN	0.0025	0.00037	0.00977
TRNA-GUANINE-N7--METHYLTRANSFERASE-RXN	0.0011	0.00047	0.012
HOMSUCTRAN-RXN	0.0009	0.00060	0.015
RXN-11322	0.0038	0.00069	0.016
RXN-12458	0.0011	0.00069	0.016
RXN-14517	0.0011	0.00069	0.016
2.7.1.148-RXN	0.0011	0.00076	0.017
N-ACETYLGLUTPREDUCT-RXN	0.0010	0.00086	0.018
RXN-15006	0.0010	0.00086	0.018
MALONYL-COA-ACP-TRANSACYL-RXN	0.00089	0.00093	0.019
RXN-8850	0.0022	0.00095	0.019
RXN-11638	0.0012	0.00098	0.019
3-DEHYDROQUINATE-DEHYDRATASE-RXN	0.0011	0.00100	0.019
POLYPHOSPHATE-KINASE-RXN	0.0022	0.0011	0.021
ATPSYN-RXN	0.0025	0.0012	0.022
DIHYDROXYISOVALDEHYDRAT-RXN	0.0014	0.0013	0.022
DIHYDROXYMETVALDEHYDRAT-RXN	0.0014	0.0013	0.022
METHIONYL-TRNA-FORMYLTRANSFERASE-RXN	0.0009	0.0013	0.022
AMINOCYL-TRNA-HYDROLASE-RXN	0.0010	0.0014	0.022
RXN-12460	0.0010	0.0014	0.022
RXN-16637	0.0010	0.0014	0.022
ATPPHOSPHORIBOSYLTRANS-RXN	0.0011	0.0014	0.022
ADENOSYLHOMOCYSTEINE-NUCLEOSIDASE-RXN	0.0012	0.0015	0.023
TRNA-PSEUDOURIDINE-SYNTHASE-I-RXN	0.0018	0.0015	0.023
ABC-27-RXN	0.0011	0.0015	0.023

RXN-13997	0.0009	0.0017	0.025
RXN-11633	0.0010	0.0017	0.025
6.3.5.7-RXN	0.0017	0.0018	0.025
TAGAKIN-RXN	0.0020	0.0021	0.029
HISTAMINOTRANS-RXN	0.0013	0.0021	0.029
PGLUCISOM-RXN	0.00080	0.0022	0.029
RXN-13720	0.00080	0.0022	0.029
RXN-6182	0.00080	0.0022	0.029
ADENYL-KIN-RXN	0.0011	0.0023	0.029
NICONUCADENYLYLTRAN-RXN	0.00093	0.0026	0.033
RXN-11592	0.0012	0.0027	0.033
1.5.1.20-RXN	0.00088	0.0028	0.034
3.1.26.3-RXN	0.00051	0.0031	0.035
ORNCARBAMTRANSFER-RXN	0.0017	0.0032	0.035
RXN-13482	0.0017	0.0032	0.035
DIHYDRODIPICSYN-RXN	0.0012	0.0032	0.035
L-LACTATE-DEHYDROGENASE-RXN	0.0019	0.0033	0.035
DMBPPRIBOSYLTRANS-RXN	0.0010	0.0034	0.035
1.8.1.4-RXN	-0.0021	0.0035	0.035
RXN-7716	-0.0021	0.0035	0.035
RXN-7719	-0.0021	0.0035	0.035
RXN-8629	-0.0021	0.0035	0.035
RXN0-1132	-0.0021	0.0035	0.035
RXN-8001	0.0011	0.0036	0.035
OROPRIBTRANS-RXN	0.00078	0.0036	0.035
6.3.2.10-RXN	0.00073	0.0037	0.035
UDP-NACMURALGLDAPAALIG-RXN	0.00073	0.0037	0.035
NACGLCTRANS-RXN	0.00078	0.0037	0.035
RXN-11029	0.00078	0.0037	0.035
RXN-11346	0.00078	0.0037	0.035
RXN-8976	0.00078	0.0037	0.035
UDPNACETYLGLUCOSAMENOLPYRTRANS-RXN	0.00073	0.0038	0.035
RIBOSYLHOMOCYSTEINASE-RXN	0.0013	0.0038	0.035
PRPPSYN-RXN	0.0012	0.0040	0.036
GLUC1PADENYLTRANS-RXN	0.0024	0.0040	0.036
GCVT-RXN	-0.0016	0.0041	0.036
HEMN-RXN	0.0010	0.0041	0.036
GLUTAMATE-SYNTHASE-NADH-RXN	0.0030	0.0043	0.038
DIHYDROOROT-RXN	0.00094	0.0043	0.038
GLYCOGENSYN-RXN	0.0015	0.0051	0.044
RXN-13163	0.0015	0.0058	0.049

Supplementary Table 4: MetaCyc metabolic pathway comparisons in cases and controls in discovery and validation cohorts, identified by linear modelling.

Discovery cohort			
Feature	Coefficient	P-value	Q-value
PWY_7234_inosine_5__phosphate_biosynthesis_III	-0.0015	0.001	0.26
PWY_7383_anaerobic_energy_metabolism_invertebrates_cytosol_	-0.0014	0.021	1.00
PWY_72824_amino_2_methyl_5_phosphomethyl_pyrimidine_biosynthesis_yeast_	-0.0012	0.036	1.00
GLYCOGENSYNTH_PWY_glycogen_biosynthesis_I_from_ADP_D_Glucose_	0.0013	0.036	1.00
HISTSYN_PWY_L_histidine_biosynthesis	-0.00074	0.039	1.00
PWY_4242_pantothenate_and_coenzymeA_biosynthesis_III	-0.00087	0.042	1.00
PWY_6700_queuosine_biosynthesis	0.0012	0.044	1.00
PWY66_399_gluconeogenesis_III	-0.00094	0.045	1.00

Validation cohort			
Feature	Coefficient	P-value	Q-value
ARO_PWY_chorismate_biosynthesis_I	0.0025	0.0011	0.074
PWY_6151S_adenosyl_L_methioninecycle_I	0.0027	0.0012	0.074
COMPLETE_ARO_PWY_superpathway_of_aromatic_amino_acid_biosynthesis	0.0023	0.0014	0.074
PWY_6163_chorismate_biosynthesis_from_3_dehydroquinone	0.0027	0.0014	0.074
PYRIDOSYN_PWY_pyridoxal_5__phosphatebiosynthesis_I	-0.0026	0.011	0.44
NONMEVIPP_PWY_methylerythritol_phosphate_pathway_I	0.0017	0.013	0.44
NONOXIPENT_PWY_pentose_phosphate_pathway_non_oxidative_branch_	0.0024	0.022	0.52
PWY_6386_UDP_N_acetylmuramoyl_pentapeptide_biosynthesis_II_lysine_containing_	0.0017	0.029	0.52
ARGININE_SYN4_PWY_L_ornithine_denovo_biosynthesis	-0.0027	0.031	0.52
PWY_6317_galactose_degradation_I_Le_loir_pathway_	0.0020	0.036	0.52
SER_GLYSYN_PWY_superpathway_of_L_serine_and_glycine_biosynthesis_I	0.0017	0.037	0.52
PWY66_422_D_galactose_degradation_V_Le_loir_pathway_	0.0019	0.044	0.52
PWY_7357_thiamin_formation_from_pyriothiamine_and_oxythiamine_yeast_	0.0017	0.045	0.52
PWY0_1296_purine_ribonucleosides_degradation	0.0017	0.047	0.52
PWY_6387UDP_N_acetylmuramoyl_pentapeptide_biosynthesis_I_meso_diaminopimelate_containing_	0.0014	0.049	0.52

Supplementary Table 5: Correlation between the indicator species (cases and controls) and the commonly-differentiated KEGG Orthogroups.

	EC 2.6.1.9: Histidinol-phosphate transaminase		EC 2.7.4.1: Polyphosphate kinase		EC 4.3.3.6: Pyridoxal 5'-phosphate synthase		EC 1.15.1.1: Superoxide proteinase		EC 3.4.21.53: ATP-dependent serine phosphatase		EC 2.4.2.17: ATP phosphoribosyltransferase	
	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value
Bifidobacterium adolescentis	0.0078	0.012	0.012	0.00036	0.00080	0.019	0.011	0.044	0.0092	2.63E-04	0.0056	0.0013
Clostridiales bacterium 1 7 47FAA	0.056	4.04E-06	0.054	1.49E-05	0.017	6.71E-04	0.027	6.68E-05	0.033	1.47E-05	0.028	5.72E-04
Clostridium bolteae	0.068	2.62E-06	0.099	2.90E-07	0.052	3.34E-07	0.131	2.12E-06	0.10	3.18E-09	0.048	3.25E-05
Clostridium hathewayi	0.054	5.91E-05	0.072	1.36E-06	0.042	9.83E-07	0.047	1.25E-06	0.063	1.08E-06	0.015	1.19E-04
Coprococcus comes	0.032	8.06E-04	0.076	2.31E-07	0.0090	0.020	0.061	1.75E-05	0.045	5.65E-06	0.019	1.20E-04
Lachnospiraceae bacterium 5 1 63FAA	-	NS	-	NS	-	NS	-	NS	-	NS	-	NS
Roseburia inulinivorans	0.022	0.010	0.034	4.01E-04	-	NS	0.040	1.66E-03	-	NS	0.065	6.96E-04

Supplementary Table 6: Differentially abundant bacterial species according to TNFi treatment, identified by linear modelling.

Feature	Value	Coefficient	P-value	Q-value
Bilophila unclassified	Untreated (relative to HC)	-0.0118	0.0218	0.3860
Bilophila unclassified	Untreated (relative to Treated)	-0.0146	0.0133	0.6714
Bilophila unclassified	Treated (relative to HC)	-	NS	NS
Clostridium symbiosum	Untreated (relative to HC)	0.004	0.0033	0.2168
Clostridium symbiosum	Untreated (relative to Treated)	0.0033	0.0366	0.7519
Clostridium symbiosum	Treated (relative to HC)	-	NS	NS
Eggerthella unclassified	Untreated (relative to HC)	0.0055	0.0022	0.1619
Eggerthella unclassified	Untreated (relative to Treated)	0.006	0.0034	0.5034
Eggerthella unclassified	Treated (relative to HC)	-	NS	NS
Eubacterium bifforme	Untreated (relative to HC)	-0.0129	0.0275	0.4325
Eubacterium bifforme	Untreated (relative to Treated)	-0.0152	0.0238	0.7258
Eubacterium bifforme	Treated (relative to HC)	-	NS	NS
Faecalibacterium prausnitzii	Untreated (relative to HC)	-0.0739	0.0052	0.3059
Faecalibacterium prausnitzii	Untreated (relative to Treated)	-0.0751	0.0129	0.6714
Faecalibacterium prausnitzii	Treated (relative to HC)	-	NS	NS
Klebsiella pneumoniae	Untreated (relative to HC)	-0.0101	0.0153	0.3777
Klebsiella pneumoniae	Untreated (relative to Treated)	-	NS	NS
Klebsiella pneumoniae	Treated (relative to HC)	-	NS	NS
Prevotella copri	Untreated (relative to HC)	-0.0518	0.0231	0.3860
Prevotella copri	Untreated (relative to Treated)	-	NS	NS
Prevotella copri	Treated (relative to HC)	-	NS	NS
Ruminococcus bromii	Untreated (relative to HC)	-0.0318	0.0193	0.3799
Ruminococcus bromii	Untreated (relative to Treated)	-0.0324	0.0377	0.7519
Ruminococcus bromii	Treated (relative to HC)	-	NS	NS

Supplementary Table 7: Differentially abundant MetaCyc metabolic pathways for TNFi treatment. *** P < 0.001, ** P < 0.01, * P < 0.05.

Class	Pathway	HC vs Treated	HC vs Untreated	Treated vs Untreated
Amino acid biosynthesis	BRANCHED-CHAIN-AA-SYN-PWY Superpathway of branched amino acid biosynthesis		*	**
Amino acid biosynthesis	COMPLETE-ARO-PWY Superpathway of aromatic amino acid biosynthesis		***	*
Amino acid biosynthesis	ILEUSYN-PWY L-isoleucine biosynthesis (from threonine)	*	*	**
Amino acid biosynthesis	PWY-5103 L-isoleucine biosynthesis (from glutamate)		*	**
Amino acid biosynthesis	PWY-5347 Superpathway of L-methionine biosynthesis		*	*
Amino acid biosynthesis	VALSYN-PWY L-valine biosynthesis	*	*	**
Aromatic compound biosynthesis	PWY-6163 Chorismate biosynthesis from 3-dehydroquinate		***	*
Carbohydrate metabolism	CALVIN-PWY Calvin Benson Bassham cycle		*	**
Carbohydrate metabolism	PWY-6317 Galactose degradation (Leloir pathway)		*	*
Carbohydrate metabolism	PWY-6737 Starch degradation		*	*
Carbohydrate metabolism	PEPTIDOGLYCANSYN-PWY Peptidoglycan biosynthesis		*	**
Cell structure	PWY-6386 UDP N-acetylmuramoyl pentapeptide biosynthesis (lysine containing)		*	*
Cell structure	NONOXIPENT-PWY Pentose phosphate pathway (non-oxidative branch)		*	**
Metabolites and energy	PWY-6700 Queuosine biosynthesis		**	*
Nucleic acid processing	PWY-5686 UMP biosynthesis	*	*	**
Nucleotide biosynthesis	PWY-6545 Pyrimidine deoxyribonucleotides denovo biosynthesis		*	**
Nucleotide biosynthesis	PWY-7219 Adenosine ribonucleotides denovo biosynthesis		*	**
Nucleotide biosynthesis	PWY-7221 Guanosine ribonucleotides denovo biosynthesis		*	**
Nucleotide biosynthesis	PWY-7234 Inosine 5-phosphate biosynthesis		***	*
Secondary metabolite biosyn.	NONMEVIPP-PWY Methylerythritol phosphate pathway		*	*

Pathway	Value	Coefficient	P-value	Q-value
BRANCHED-CHAIN-AA-SYN-PWY Superpathway of branched amino acid biosynthesis	Untreated (relative to HC)	-0.0014	0.0118	0.3002
BRANCHED-CHAIN-AA-SYN-PWY Superpathway of branched amino acid biosynthesis	Untreated (relative to Treated)	-0.0022	0.0014	0.1026
COMPLETE-ARO-PWY Superpathway of aromatic amino acid biosynthesis	Untreated (relative to HC)	-0.0015	0.0002	0.0637
COMPLETE-ARO-PWY Superpathway of aromatic amino acid biosynthesis	Untreated (relative to Treated)	-0.0011	0.0286	0.2980
ILEUSYN-PWY L-isoleucine biosynthesis (from threonine)	Untreated (relative to HC)	-0.0010	0.0470	0.5470
ILEUSYN-PWY L-isoleucine biosynthesis (from threonine)	Untreated (relative to Treated)	-0.0019	0.0010	0.0933
ILEUSYN-PWY L-isoleucine biosynthesis (from threonine)	HC (relative to Treated)	-0.0010	0.0525	0.3822
PWY-5103 L-isoleucine biosynthesis (from glutamate)	Untreated (relative to HC)	-0.0014	0.0161	0.3658
PWY-5103 L-isoleucine biosynthesis (from glutamate)	Untreated (relative to Treated)	-0.0022	0.0026	0.1238
PWY-5347 Superpathway of L-methionine biosynthesis	Untreated (relative to HC)	-0.0007	0.0480	0.5470
PWY-5347 Superpathway of L-methionine biosynthesis	Untreated (relative to Treated)	-0.0010	0.0290	0.2980
VALSYN-PWY L-valine biosynthesis	Untreated (relative to HC)	-0.0010	0.0470	0.5470
VALSYN-PWY L-valine biosynthesis	Untreated (relative to Treated)	-0.0019	0.0010	0.0933
VALSYN-PWY L-valine biosynthesis	HC (relative to Treated)	-0.0010	0.0525	0.3822
PWY-6163 Chorismate biosynthesis from 3-dehydroquinate	Untreated (relative to HC)	-0.0018	0.0002	0.0637
PWY-6163 Chorismate biosynthesis from 3-dehydroquinate	Untreated (relative to Treated)	-0.0012	0.0356	0.3298
CALVIN-PWY Calvin Benson Bassham cycle	Untreated (relative to HC)	-0.0011	0.0300	0.4674
CALVIN-PWY Calvin Benson Bassham cycle	Untreated (relative to Treated)	-0.0018	0.0028	0.1238
PEPTIDOGLYCANSYN-PWY Peptidoglycan biosynthesis	Untreated (relative to HC)	-0.0011	0.0030	0.1584
PEPTIDOGLYCANSYN-PWY Peptidoglycan biosynthesis	Untreated (relative to Treated)	-0.0014	0.0021	0.1238
PWY-6317 Galactose degradation (Leloir pathway)	Untreated (relative to HC)	-0.0012	0.0412	0.5337
PWY-6317 Galactose degradation (Leloir pathway)	Untreated (relative to Treated)	-0.0014	0.0486	0.3693
PWY-6737 Starch degradation	Untreated (relative to HC)	-0.0020	0.0033	0.1584
PWY-6737 Starch degradation	Untreated (relative to Treated)	-0.0023	0.0076	0.1751
NONOXIPENT-PWY Pentose phosphate pathway (non-oxidative branch)	Untreated (relative to HC)	-0.0017	0.0053	0.1933
NONOXIPENT-PWY Pentose phosphate pathway (non-oxidative branch)	Untreated (relative to Treated)	-0.0022	0.0031	0.1238
PWY-6386 UDP N-acetylmuramoyl pentapeptide biosynthesis (lysine containing)	Untreated (relative to HC)	-0.0012	0.0030	0.1584
PWY-6386 UDP N-acetylmuramoyl pentapeptide biosynthesis (lysine containing)	Untreated (relative to Treated)	-0.0015	0.0041	0.1345
PWY-6700 Queuosine biosynthesis	Untreated (relative to HC)	-0.0018	0.0031	0.1584
PWY-6700 Queuosine biosynthesis	Untreated (relative to Treated)	-0.0017	0.0190	0.2519
PWY-5686 UMP biosynthesis	Untreated (relative to HC)	-0.0014	0.0043	0.1795
PWY-5686 UMP biosynthesis	Untreated (relative to Treated)	-0.0022	0.0002	0.0831
PWY-5686 UMP biosynthesis	HC (relative to Treated)	-0.0008	0.1187	0.6171
PWY-6545 Pyrimidine deoxyribonucleotides denovo biosynthesis	Untreated (relative to HC)	-0.0011	0.0135	0.3318

PWY-6545 Pyrimidine deoxyribonucleotides denovo biosynthesis	Untreated (relative to Treated)	-0.0016	0.0030	0.1238
PWY-7219 Adenosine ribonucleotides denovo biosynthesis	Untreated (relative to HC)	-0.0019	0.0015	0.1584
PWY-7219 Adenosine ribonucleotides denovo biosynthesis	Untreated (relative to Treated)	-0.0023	0.0016	0.1035
PWY-7221 Guanosine ribonucleotides denovo biosynthesis	Untreated (relative to HC)	-0.0013	0.0009	0.1425
PWY-7221 Guanosine ribonucleotides denovo biosynthesis	Untreated (relative to Treated)	-0.0018	0.0002	0.0831
PWY-7234 Inosine 5-phosphate biosynthesis	Untreated (relative to HC)	0.0017	0.0003	0.0637
PWY-7234 Inosine 5-phosphate biosynthesis	Untreated (relative to Treated)	0.0013	0.0242	0.2685
NONMEVIPP-PWY Methylethritol phosphate pathway	Untreated (relative to HC)	-0.0011	0.0033	0.1584
NONMEVIPP-PWY Methylethritol phosphate pathway	Untreated (relative to Treated)	-0.0013	0.0048	0.1355

Supplementary Table 8: Correlation between differentially abundant bacterial species and MetaCyc metabolic pathways for TNFi treatment.

	BRANCHED-CHAIN-AA-SYN-PWY: Superpathway of branched amino acid biosynthesis		COMPLETE-ARO-PWY: Superpathway of aromatic amino acid biosynthesis		ILEUSYN-PWY: L-isoleucine biosynthesis (from threonine)		PWY-5103: L-isoleucine biosynthesis (from glutamate)		PWY-5347: Superpathway of L-methionine biosynthesis (transsulfuration)		VALSYN-PWY: L-valine biosynthesis		PWY-6163: Chorismate biosynthesis from 3-dehydroquinate	
	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value
<i>Prevotella copri</i>	-	NS	-	NS	-	NS	-	NS	-	NS	-	NS	-	NS
<i>Faecalibacterium prausnitzii</i>	0.387	6.18E-28	0.027	1.33E-03	0.292	2.19E-21	0.3870	3.46E-27	-	NS	0.2922	2.19E-21	0.028	1.52E-03
<i>Bilophila unclassified</i>	-	NS	-0.030	6.22E-03	-	NS	-	NS	-	NS	-	NS	-0.031	5.00E-03
<i>Klebsiella pneumoniae</i>	-	NS	-0.025	3.29E-02	-0.034	5.05E-03	-	NS	-	NS	-0.034	5.05E-03	-	NS
<i>Ruminococcus bromii</i>	-	NS	0.031	5.79E-03	0.049	1.13E-03	-	NS	0.218	6.55E-05	0.049	1.13E-03	0.035	2.55E-03
<i>Eubacterium bifforme</i>	-	NS	0.002	1.17E-02	-	NS	-	NS	0.021	4.51E-02	-	NS	0.003	8.83E-03
<i>Clostridium symbiosum</i>	0.014	2.43E-05	0.015	7.62E-05	0.004	3.01E-02	0.011	2.59E-05	-	NS	0.004	3.01E-02	0.018	8.00E-05
<i>Eggerthella unclassified</i>	-	NS	0.018	2.91E-03	-	NS	-	NS	-	NS	-	NS	0.022	1.24E-03

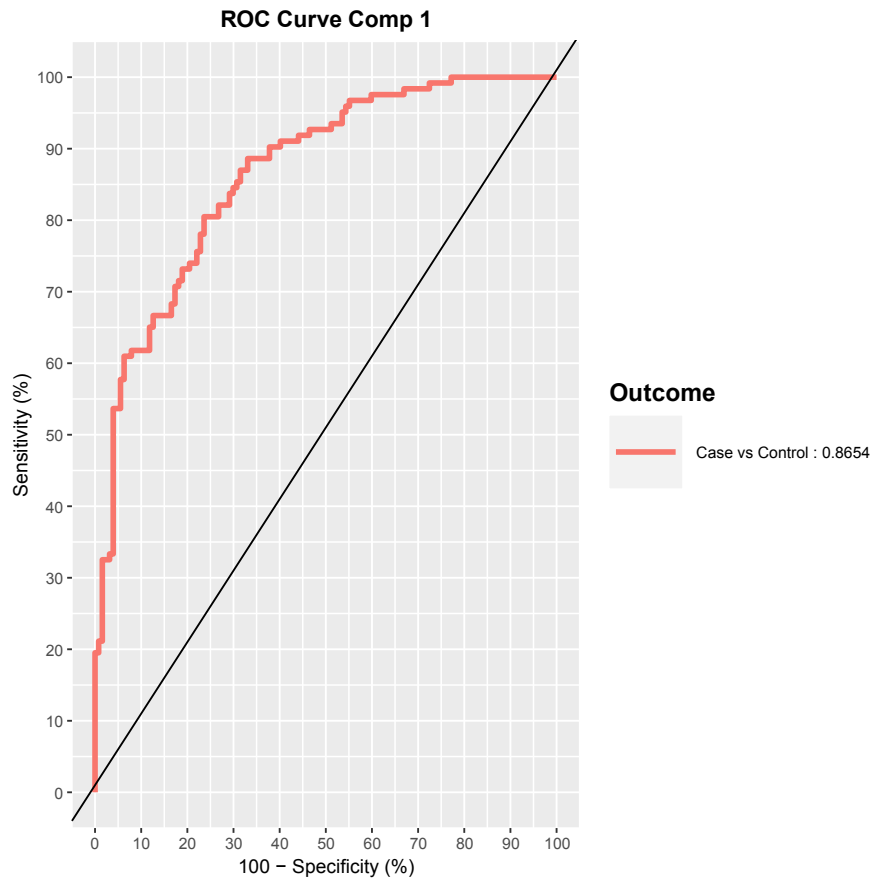
	PEPTIDOLYCATS-YN-PWY: Peptidoglycan biosynthesis (meso-diaminopimelate containing)		PWY-6386: UDP-N-acetylmuramoyl pentapeptide biosynthesis (lysine containing)		NONOXIPENT-PWY: Pentose phosphate pathway (non-oxidative branch)		PWY-6700: Queuosine biosynthesis		PWY-5686: UMP biosynthesis		PWY-6545: Pyrimidine deoxyribonucleotides denovo biosynthesis		PWY-7219: Adenosine ribonucleotides de novo biosynthesis	
	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value
<i>Prevotella copri</i>	0.104	3.97E-03	0.055	1.32E-02	-	NS	0.3243	5.40E-03	0.115	9.18E-03	-	NS	0.126	2.53E-03
<i>Faecalibacterium prausnitzii</i>	0.067	4.27E-07	0.104	1.13E-09	0.089	3.43E-09	0.030	7.50E-13	0.107	6.00E-12	-	NS	0.070	4.65E-09
<i>Bilophila unclassified</i>	-	NS	-	NS	-0.017	1.57E-02	-	NS	-	NS	-	NS	-	NS
<i>Klebsiella pneumoniae</i>	-0.045	3.70E-03	-0.044	2.00E-03	-	NS	-0.003	4.89E-03	-0.043	9.18E-03	-	NS	-0.032	2.61E-02
<i>Ruminococcus bromii</i>	-	NS	0.033	4.23E-02	0.053	4.35E-03	0.008	4.27E-02	0.025	4.88E-02	-	NS	-	NS
<i>Eubacterium bifforme</i>	-	NS	-	NS	-	NS	0.000	5.74E-03	0.010	9.18E-03	-	NS	0.016	7.61E-04
<i>Clostridium symbiosum</i>	0.011	1.18E-02	0.016	2.10E-03	0.009	8.51E-03	0.010	1.34E-03	0.015	5.36E-03	-	NS	0.036	5.82E-05
<i>Eggerthella unclassified</i>	-	NS	0.012	1.98E-02	-	NS	0.013	7.32E-03	0.011	9.76E-03	-	NS	0.017	8.20E-04

	CALVIN-PWY: Calvin Benson Bassham cycle		PWY-6317: Galactose degradation (Leloir pathway)		PWY-6737: Starch degradation		PWY-7221: Guanosine ribonucleotides de novo biosynthesis		PWY-7234: Inosine-5- phosphate biosynthesis		NONMEVPP- PWY: Methylerythritol phosphate pathway	
	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value	r ²	P-value
Prevotella copri	-	NS	-	NS	-	NS	0.2923	1.88E-02	0.039	7.68E-04	-	NS
Faecalibacterium prausnitzii	0.044	3.60E-08	0.076	1.78E-11	0.312	2.38E-24	0.042	1.06E-09	0.012	2.89E-02	0.040	6.18E-28
Bilophila unclassified	-	NS	-0.029	3.60E-03	-	NS	-	NS	-	NS	-	NS
Klebsiella pneumoniae	-0.014	2.94E-02	-	NS	-0.043	1.41E-02	-	NS	-	NS	-	NS
Ruminococcus bromii	0.032	1.19E-02	0.029	1.68E-04	0.034	1.35E-04	0.031	3.15E-03	-	NS	-	NS
Eubacterium bifforme	-	NS	-	NS	-	NS	-	NS	0.016	1.07E-03	-	NS
Clostridium symbiosum	0.007	1.50E-02	0.013	2.34E-04	0.034	5.86E-06	-	NS	0.057	3.13E-03	0.023	2.43E-05
Eggerthella unclassified	-	NS	-	NS	-	NS	-	NS	0.018	2.75E-03	-	NS

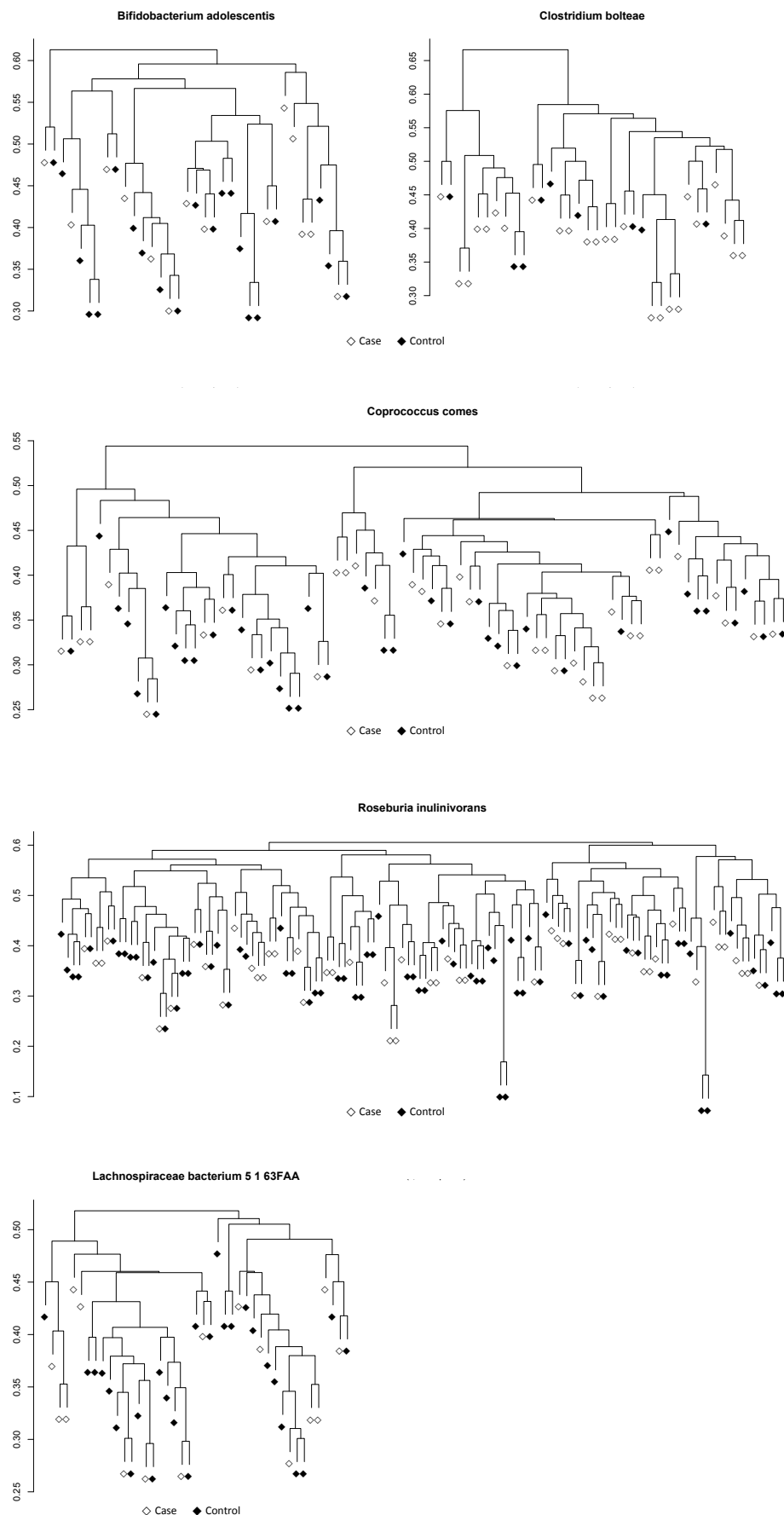
Supplementary Table 9: Bacterial epitopes homologous to HLA-B27-presented epitopes which were commonly enriched in the discovery and validation cohorts for AS cases. All peptides were found to be 100% homologous to bacterial proteins from multiple bacterial species according to NCBI nr. The abundance of each peptide in cases vs controls was evaluated by Wilcoxon rank-sum testing (FDR corrected) and the presence/absence of the peptides in cases vs controls was evaluated via Fisher's exact test (FDR corrected).

IEDB ID	HLA-B27 Subtype	Sequence	Protein	Wilcoxon (Q-value)	Fisher's (Q-value)
434944	B*2702; B*2705; B*2707; B*2709	ARTPHWALF	tRNA wybutosine-synthesizing protein 1 homolog	2.42x10 ⁻²	1.30x10 ⁻³
445935	B*2702; B*2703; B*2707; B*2708; B*2709	KRNTFVGTPPFWM	Serine/threonine-protein kinase 24	4.60x10 ⁻³	1.13x10 ⁻⁸
447192	B*2702	QRGLWGGEGW	39S ribosomal protein L28, mitochondrial	1.20x10 ⁻³	1.56x10 ⁻³⁰
490238	B*2702; B*2704	ARFKSNVTKTMKGFEY	PRELI domain-containing protein 1, mitochondrial	2.60x10 ⁻³	1.00x10 ⁻⁵
490741	B*2702	ATTAALLLEAQAAATGFLVDPVR	Plectin	1.20x10 ⁻³	2.03x10 ⁻¹⁰
490897	B*2702	ERIATEFNQLQF	Conserved oligomeric Golgi complex subunit 2	3.93x10 ⁻²	7.00x10 ⁻⁴
491073	B*2702	FQWMSSRVSPNTLW	Death ligand signal enhancer	2.63x10 ⁻²	5.18x10 ⁻¹³
491482	B*2702; B*2703; B*2704; B*2705; B*2706; B*2707; B*2708	GRFTIKSDVWSF	Tyrosine-protein kinase Fyn	9.30x10 ⁻³	4.73x10 ⁻⁶
491582	B*2702	GRINPNSGGTNY	Ig heavy chain V-I region V35	1.88x10 ⁻²	7.00x10 ⁻⁴
491796	B*2705	GRLYPWGVEVENPEHNDFL	Septin-2	5.69x10 ⁻²	2.29x10 ⁻⁹
492689	B*2703; B*2707; B*2709	KRFDDKYTLKLT	Signal peptidase complex subunit 2	2.42x10 ⁻²	1.30x10 ⁻³
492710	B*2703	KRFFFDVGSNKY	Transcriptional activator protein Pur-alpha	3.93x10 ⁻²	2.00x10 ⁻⁴
492876	B*2702	KRNEDEDSPNKLY	60S ribosomal protein L31	2.42x10 ⁻²	4.20x10 ⁻⁵
492935	B*2703	KRWESERVLSF	AP-3 complex subunit mu-1	2.98x10 ⁻²	2.50x10 ⁻³
492970	B*2702	KRYYDEVEAEGY	Ubiquitin carboxyl-terminal hydrolase 11	3.93x10 ⁻²	2.48x10 ⁻⁷
493378	B*2704; B*2705; B*2709	MRLPAQLLGLLM	Ig kappa chain V-II region RPMI 6410	9.30x10 ⁻³	2.00x10 ⁻⁴
493505	B*2702	NRHYTFYVW	Dol-P-Glc:Glc(2)Man(9)GlcNAc(2)-PP-Dol alpha-1,2-glucosyltransferase	9.30x10 ⁻³	2.00x10 ⁻⁴
493564	B*2702	NRQIVSGSRDKTIKLW	Guanine nucleotide-binding protein subunit beta-2-like 1	3.40x10 ⁻³	1.03x10 ⁻⁹
493651	B*2702	QRAIQVDPNYAY	Cell division cycle protein 27 homolog	2.42x10 ⁻²	2.28x10 ⁻²⁶
494083	B*2702	QTTFLVDNKKVFGTHL	N-alpha-acetyltransferase 35, NatC auxiliary subunit	1.88x10 ⁻²	7.00x10 ⁻⁴
494092	B*2702	RKFQPYKPFYY	Histone-lysine N-methyltransferase SETDB1	2.88x10 ⁻²	4.00x10 ⁻⁴
494108	B*2702	RQIMTGFGELSY	E3 ubiquitin-protein ligase RLIM	1.88x10 ⁻²	7.00x10 ⁻⁴
494342	B*2707; B*2709	RRIMRPTDVPDQGL	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	1.88x10 ⁻²	4.73x10 ⁻⁶
495095	B*2708	SRLEQGEEPWWL	Zinc finger protein 7	2.20x10 ⁻³	2.48x10 ⁻⁷

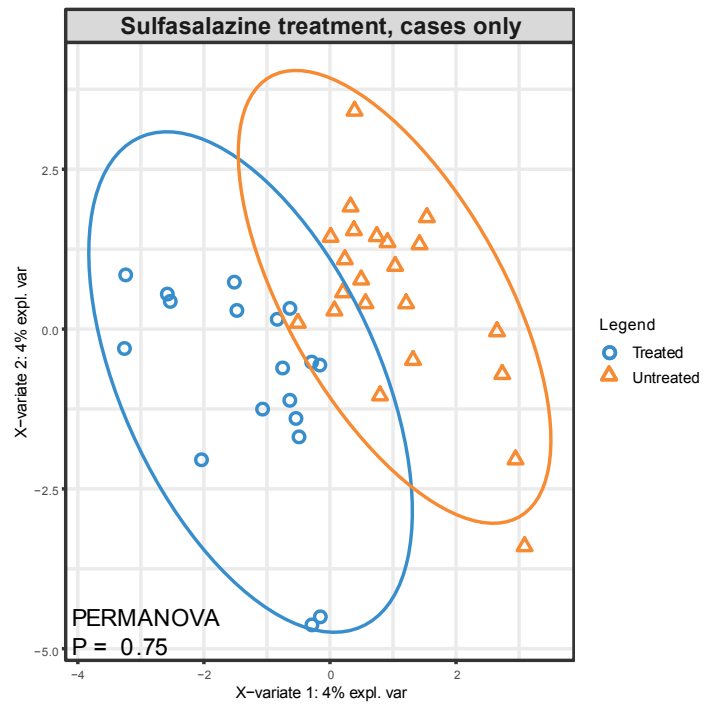
Supplementary Figure 1: Receiver operating characteristic curve revealing substantial differentiation between AS cases and controls within the cohort (250 individuals).



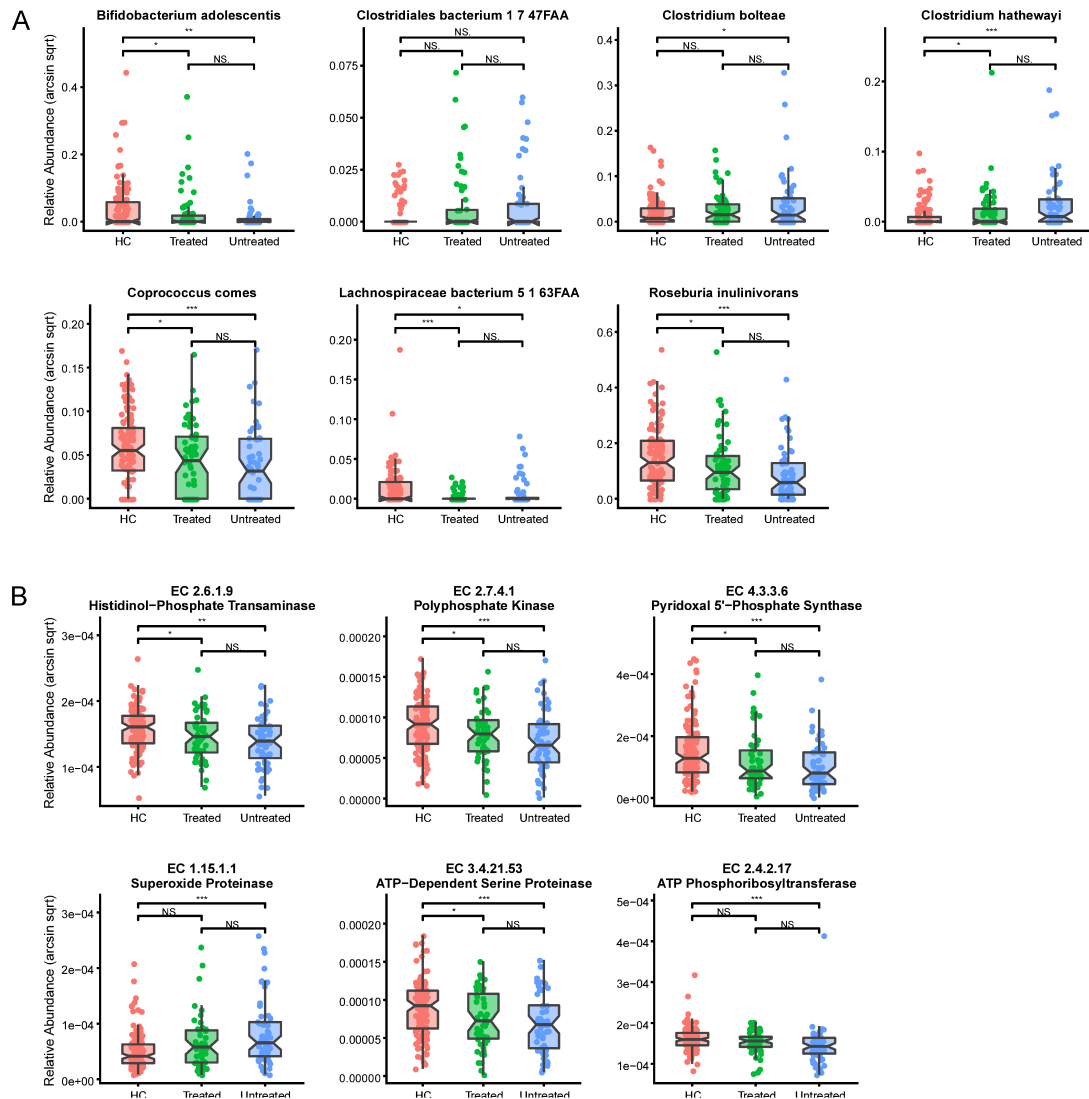
Supplementary Figure 2: Genetic-relatedness dendrograms of the strain population for each of the dysbiotic bacterial species in AS cases, as identified in Figure 1B. Samples without sufficient read coverage for these species were excluded. Each data point represents an individual sample, which may consist of multiple strains. AS status was unable to distinctively separate the strains present in AS cases compared to healthy controls, indicating that gut dysbiosis may primarily be attributed to differential species abundance with no additional strain-level effects.



Supplementary Figure 3: sPLSDA and PERMANOVA results measuring the effect of sulfasalazine treatment upon the microbiome of AS cases who have not undergone TNFi therapy, revealing a non-significant effect, potentially due to power or sample size constraints.



Supplementary Figure 4: Effect of TNFi therapy upon the dysbiotic A. bacterial species and B. KEGG Orthogroups which were commonly differentiated between cases vs controls in the discovery and validation cohorts, as shown in Figure 1. TNFi therapy appears to partially normalise dysbiosis, however statistical testing revealed non-significant differences between treated and untreated cases.



Supplementary Figure 5: Genetic-relatedness dendrograms of the strain population for each of the bacterial species modulated by TNFi therapy, as identified in Figure 2B. Samples without sufficient read coverage as well as species annotated as "unclassified" were excluded. Each data point represents an individual sample, which may consist of multiple strains. No distinctive separation of strains present was noted, indicating that TNFi therapy may primarily affect species abundance with no additional strain-level effects.

