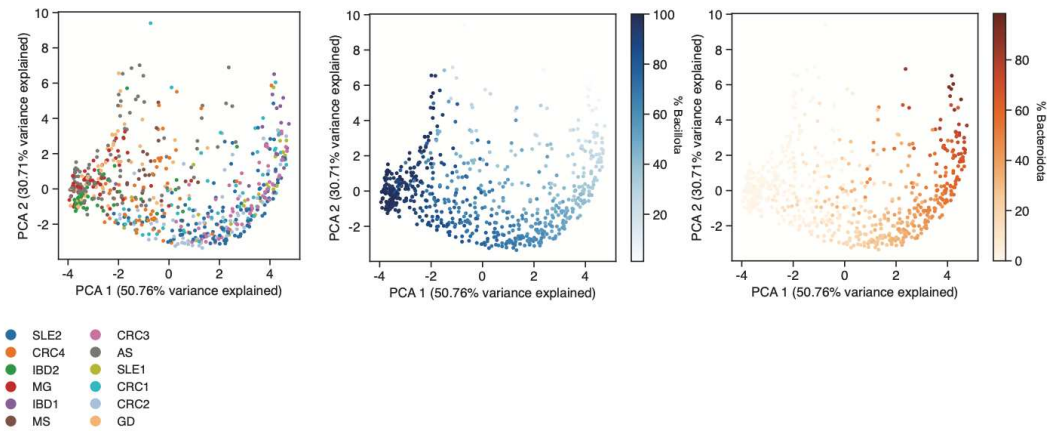
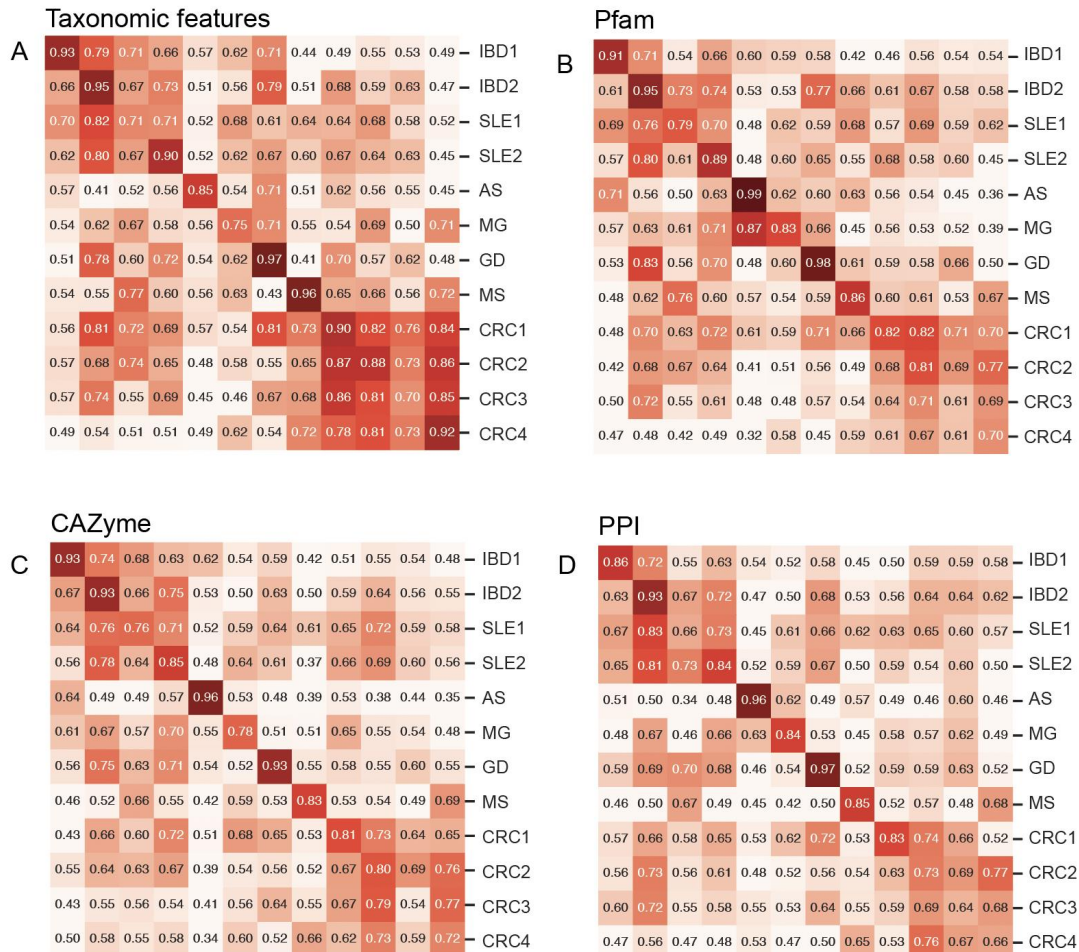


1 **Supplemental Figure 1: Principal Coordinate Analysis (PCoA) of taxonomic compositions in**
2 **healthy controls.** This figure presents a PCoA based on Bray-Curtis dissimilarity, illustrating the
3 taxonomic compositions from healthy control groups across various studies. The analysis is displayed in
4 three parts: the first segment (left) shows the distribution colored by study, the middle segment illustrates
5 the percentage of Bacillota, and the third segment (right) depicts the percentage of Bacteroidota.
6
7



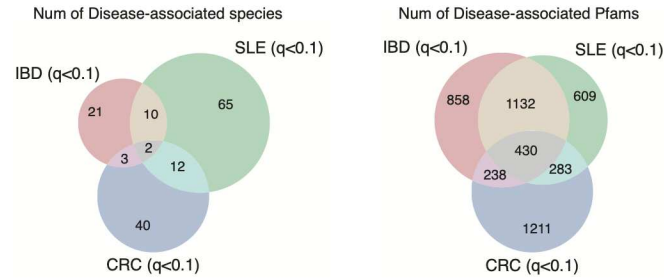
8
9

10 **Supplemental Figure 2.** Evaluation of area under the precision recall curve (AUPRC) for random forest
 11 models trained with diverse feature sets - taxonomic (A), Pfam (B), CAzyme (C), and PPI-relevant
 12 features (D) - for predicting health status (healthy/diseased) in the test cohort based on corresponding
 13 individual data. Median values were calculated from multiple evaluations, encompassing both cross-study
 14 testing and 5-fold cross-validation iterations.
 15



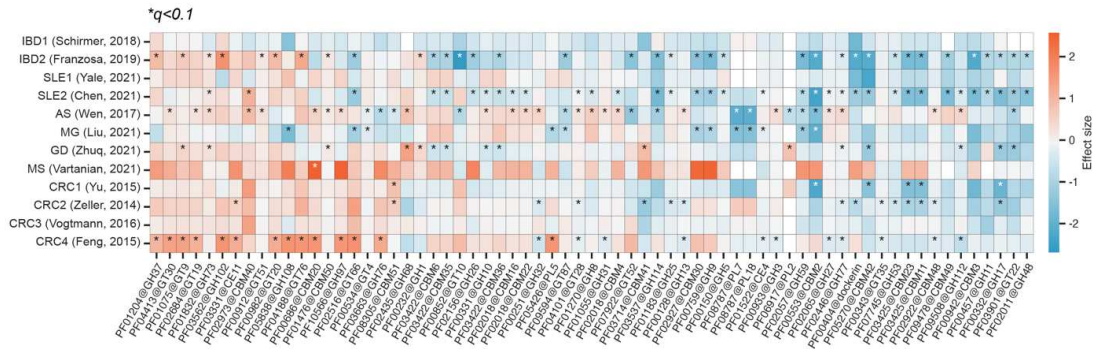
16
 17
 18

19 **Supplemental Figure 3. Numbers of significantly associated species and gene families in IBD, SLE**
20 **and CRC.** The analysis used adjusted p-values as a cutoff ($q < 0.1$). Species or Pfams included were
21 identified as significant in at least one study.



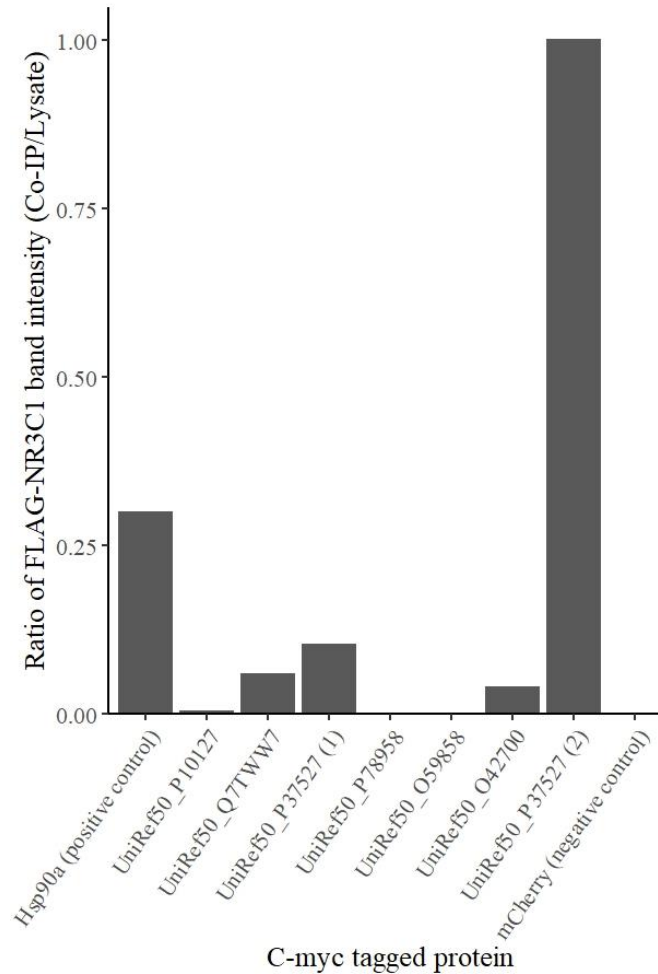
22
23
24
25

26 **Supplemental Figure 4.** CAZymes with q-values < 0.1 in two or more studies are plotted across all
 27 studies. Asterisks represent those studies in which the q-value < 0.1.



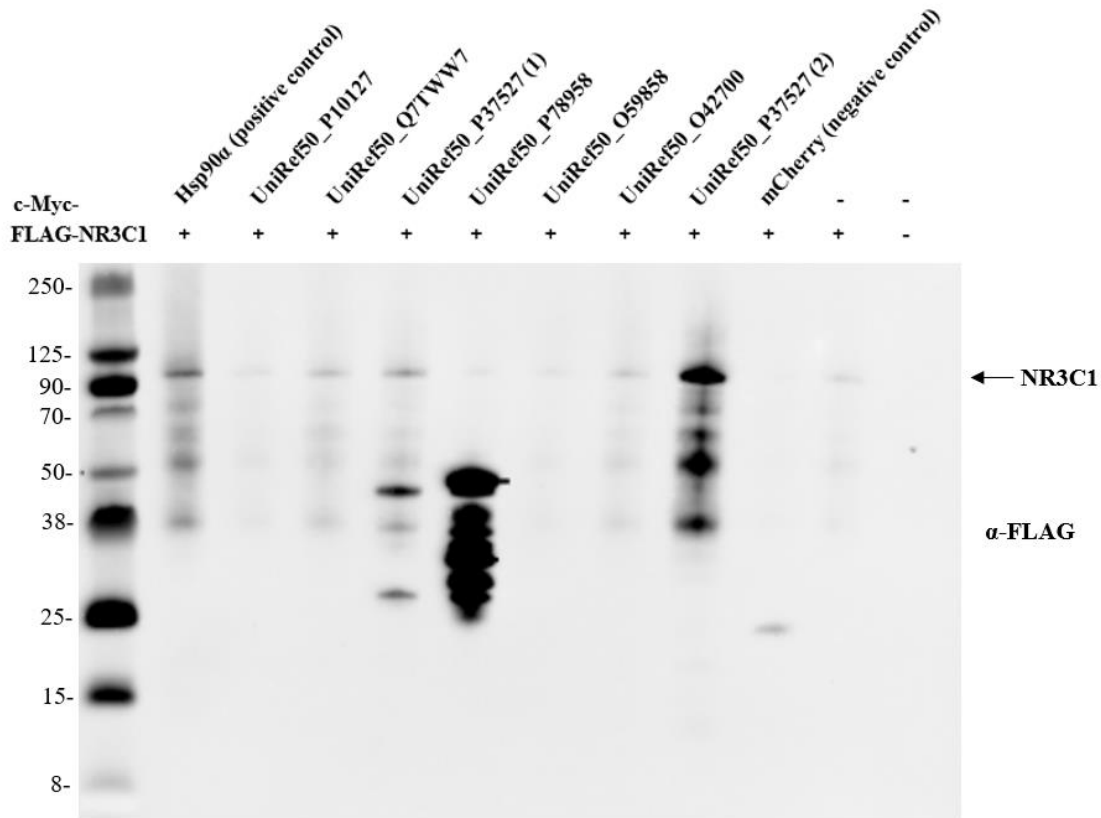
28
 29
 30
 31

32 **Supplementary Figure 5:** Quantification of Western blot. The bar graph represents the ratios between
33 the protein band intensities of co-immunoprecipitated FLAG-NR3C1 and FLAG-NR3C1 in the total cell
34 lysate. Ratios were normalized to the negative control with empty C-myc vector and FLAG-NR3C1.



35

36 **Supplemental Figure 6.** Uncropped anti-FLAG Western blots of whole cell lysates and anti-c-Myc Co-
37 IP samples from HEK293T cells co-transfected with FLAG-tagged NR3C1 and c-Myc-tagged bacterial
38 proteins from the predicted UniRef50 clusters.
39



40
41

- 42 **Supplemental Tables** (Also available at https://github.com/britolab/ppi_autoimmune/tree/main)
43
- 44 **Supplemental Table 1:** Metadata of cases and controls from all studies, detailing disease condition, age,
45 gender, and visit number.
46
- 47 **Supplemental Table 2:** Disease-associated species in each study, including p-values and adjusted p-
48 values.
49
- 50 **Supplemental Table 3:** Disease-associated Pfams in each study, including p-values and adjusted p-values.
51
- 52 **Supplemental Table 4:** Disease-associated SCFAs in each study, including p-values and adjusted p-
53 values.
54
- 55 **Supplemental Table 5:** Disease-associated CAZymes in each study, including p-values and adjusted p-
56 values.
57
- 58 **Supplemental Table 6:** Representative proteins from each UniRed50 cluster tested in Co-IP assay.
59