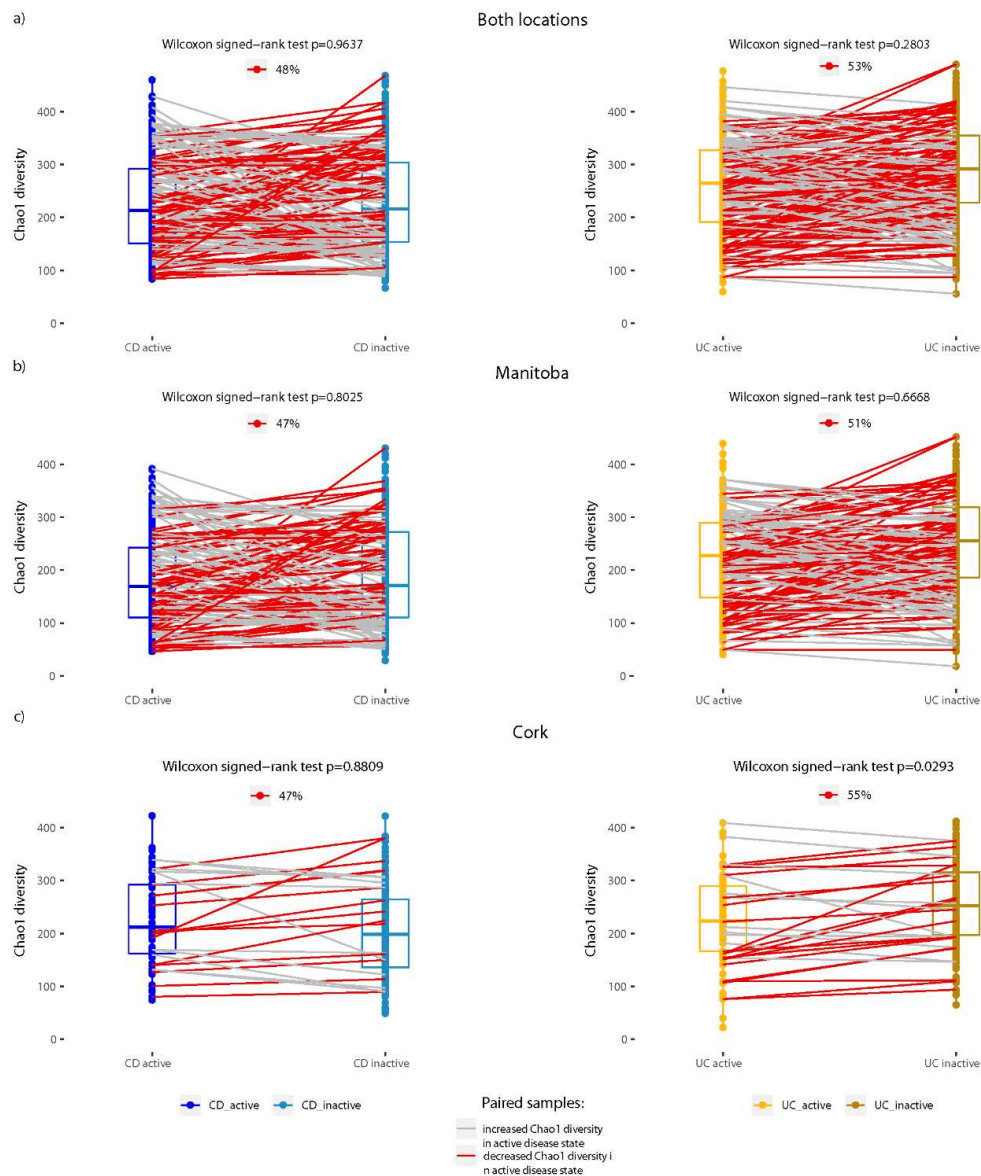
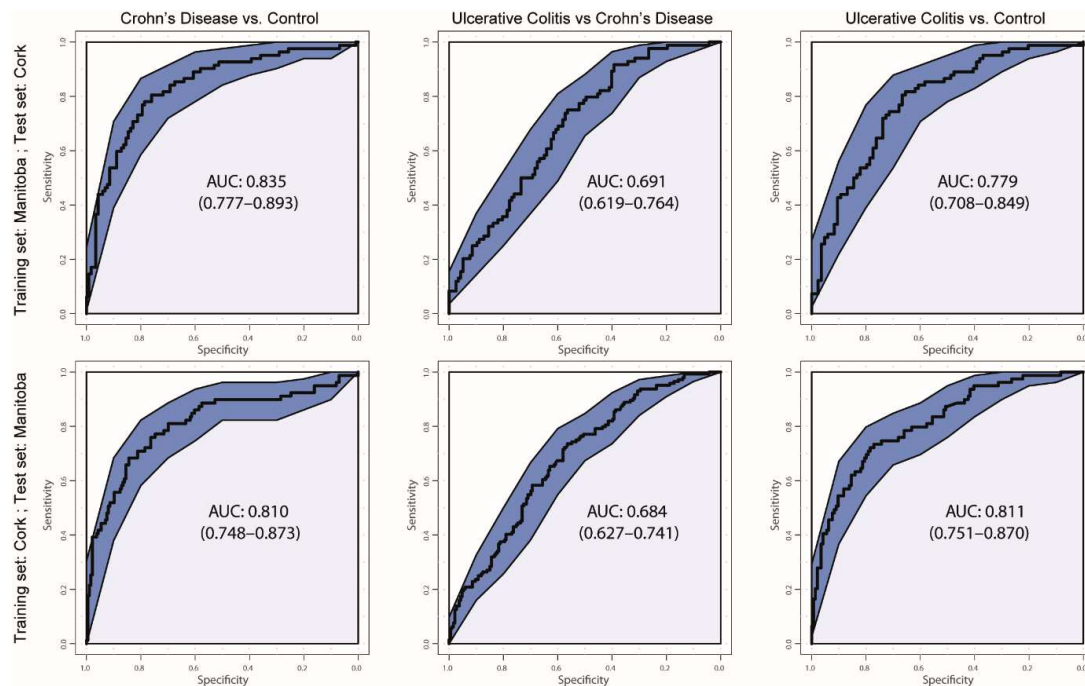


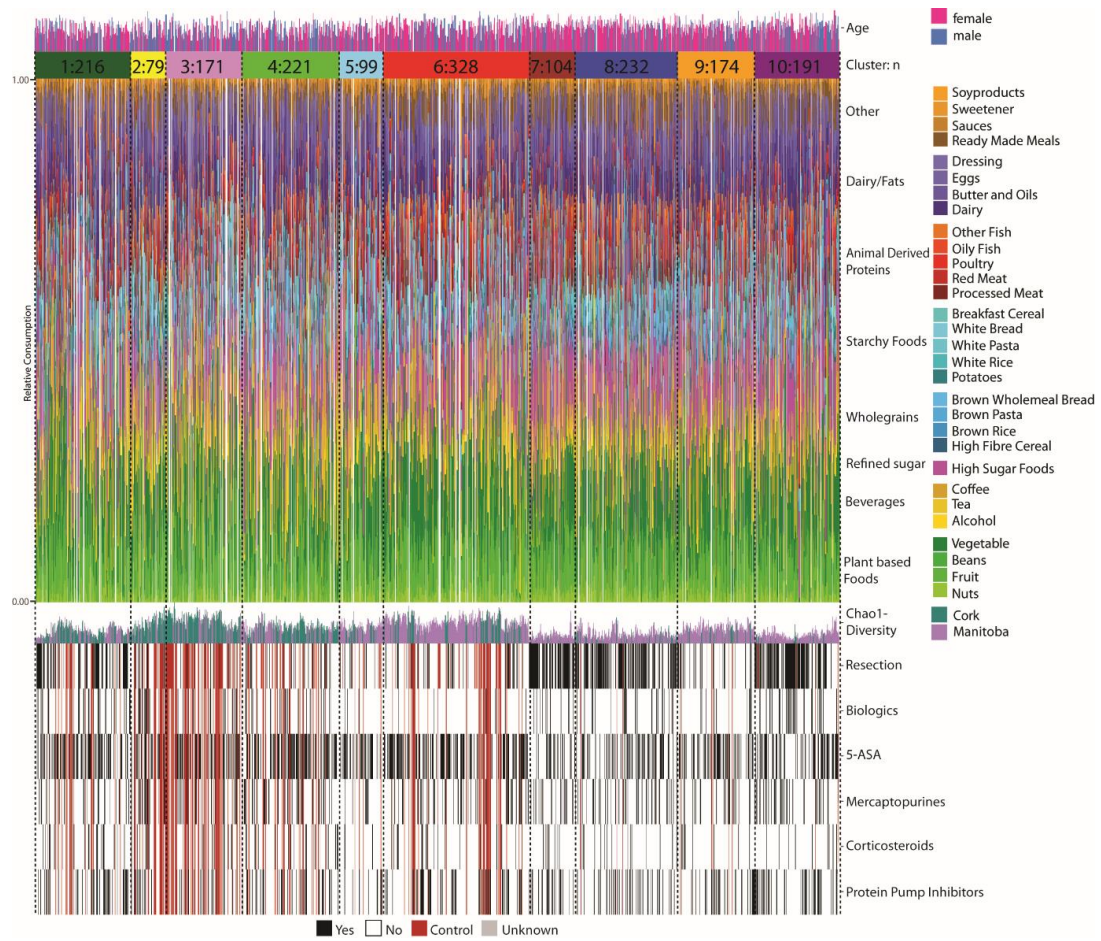
Supplementary Figure 1. Comparison of samples stored for various periods of time at room temperature. a) Chao1 diversity, b) samples clustered by time at RT, c) samples clustered by subjects.



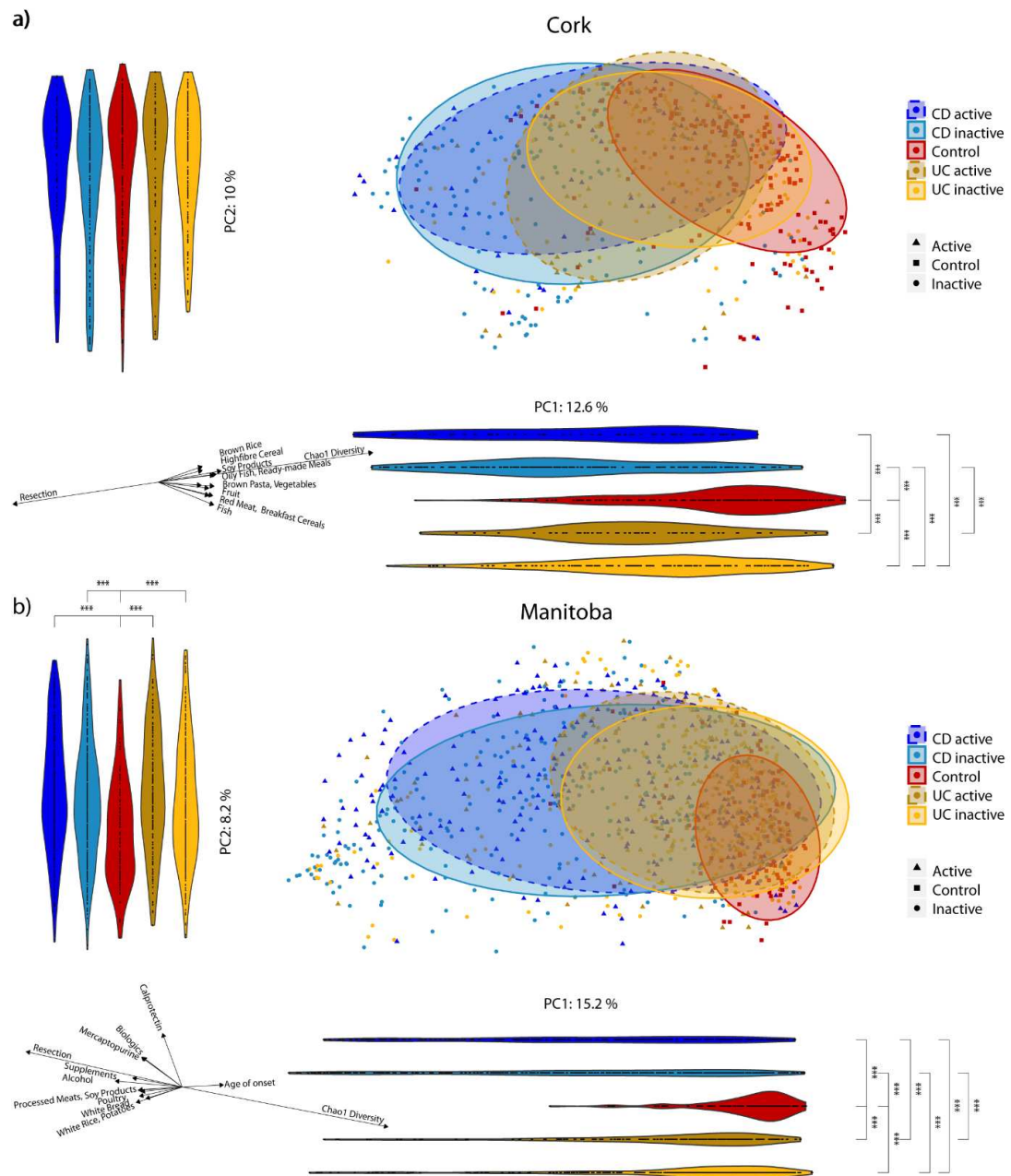
Supplementary Figure 2. Changes in alpha diversity (Chao1) between active and inactive disease state, lines connect samples of patients whose activity status change between time-points. a) Across both geographic locations separately for b) Manitoba and c) Cork. Red lines represent cases where alpha diversity was lower in active than in inactive disease, and grey lines where the opposite was true.



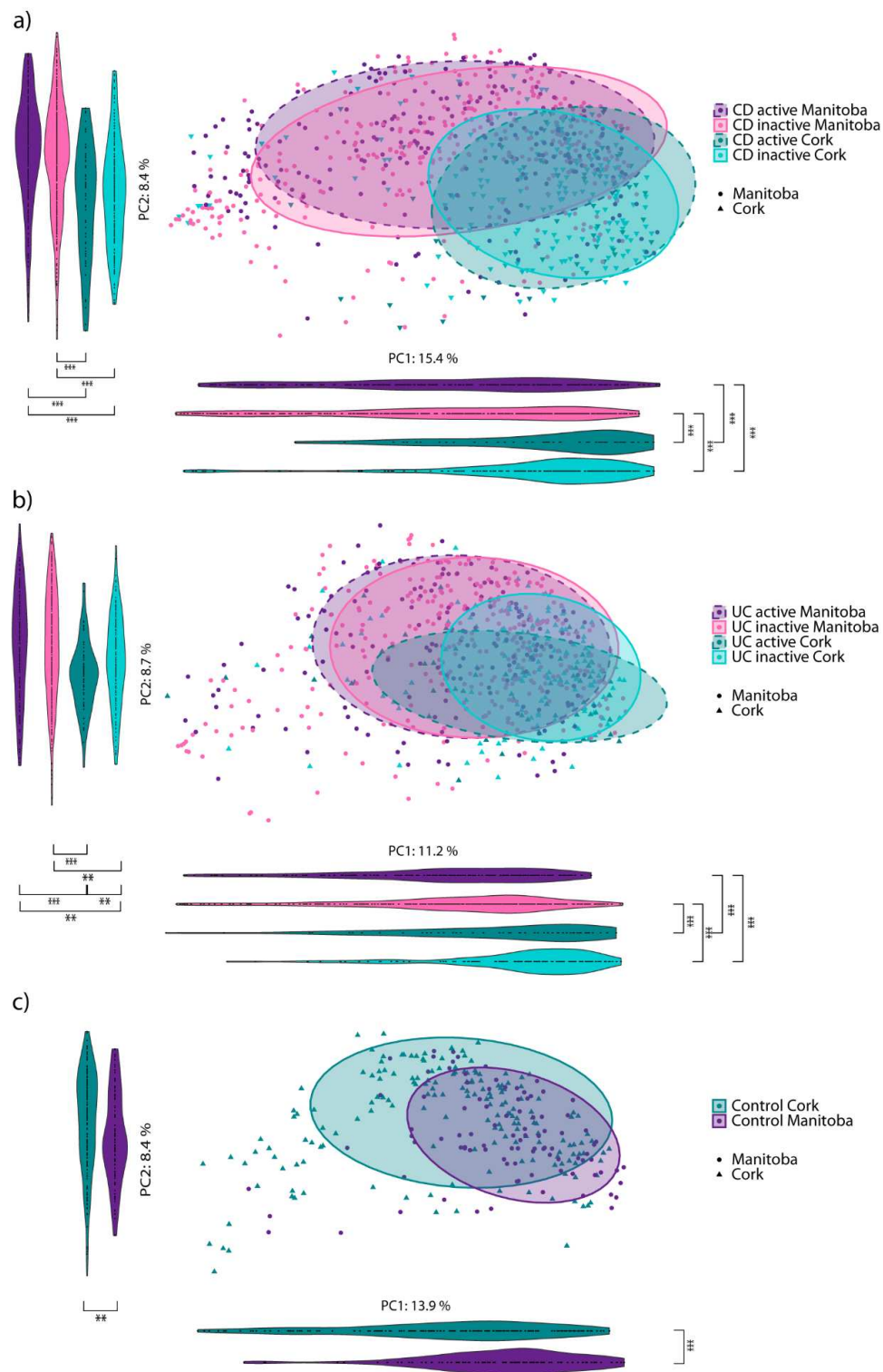
Supplementary Figure 3. Cross-geographic Machine Learning classification of the subject cohorts based on proportional normalized OTUs present in more than 5% of samples. Receiver operating characteristic curves (ROC) for the intercontinental validation of boosted tree classifiers between the same groups as in Figure 3. Here, the first row of ROC-curves used Manitoba subjects as training set and Cork subjects as test set, whereas the second row shows the results of the opposite case (Cork subjects as training model and Manitoba as test set). A model with an AUC of 0.5 has no discriminatory capacity, whereas an AUC of 1 indicates perfect separation of the response variables.



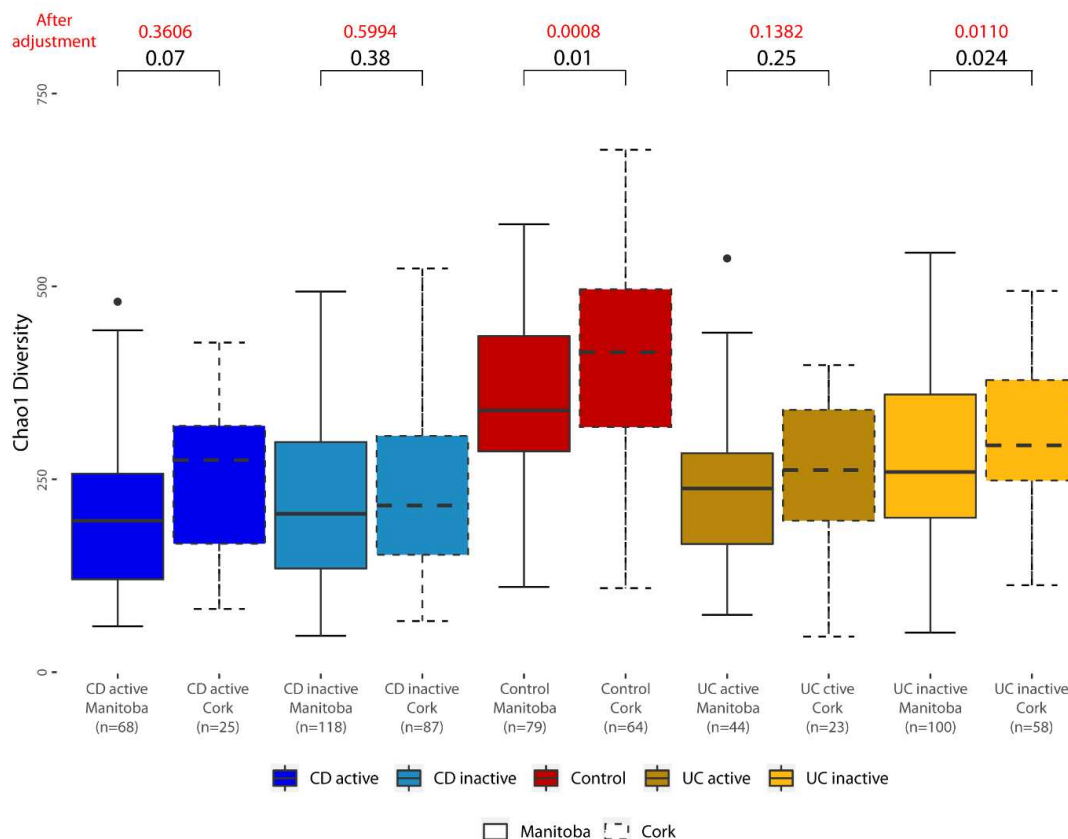
Supplementary Figure 4. Gender, medication status, resection status, dietary categories and Chao1 diversity with samples displayed in the same order as Figure 5.



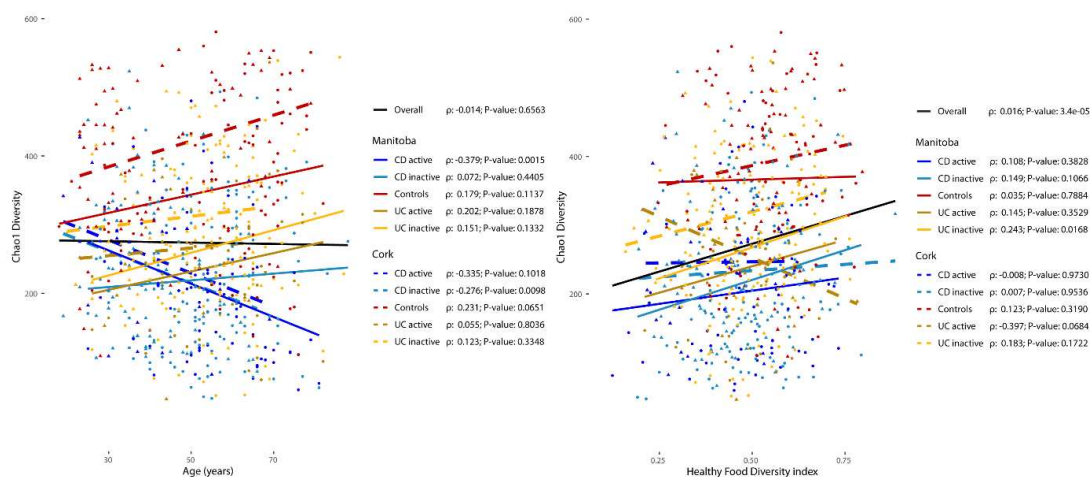
Supplementary Figure 5. Principle Coordinate Analysis based on Bray-Curtis distances with proportional normalization on all OTUs present in more than 5% of samples, with samples grouped by disease and activity state for a) Cork and b) Manitoba separately. Violin plots show projections of PCoA points onto PC1 and PC2. The arrows represent Spearman correlations between PC axes and food groups/metadata and alpha diversity. Only categories with significant correlations are represented.



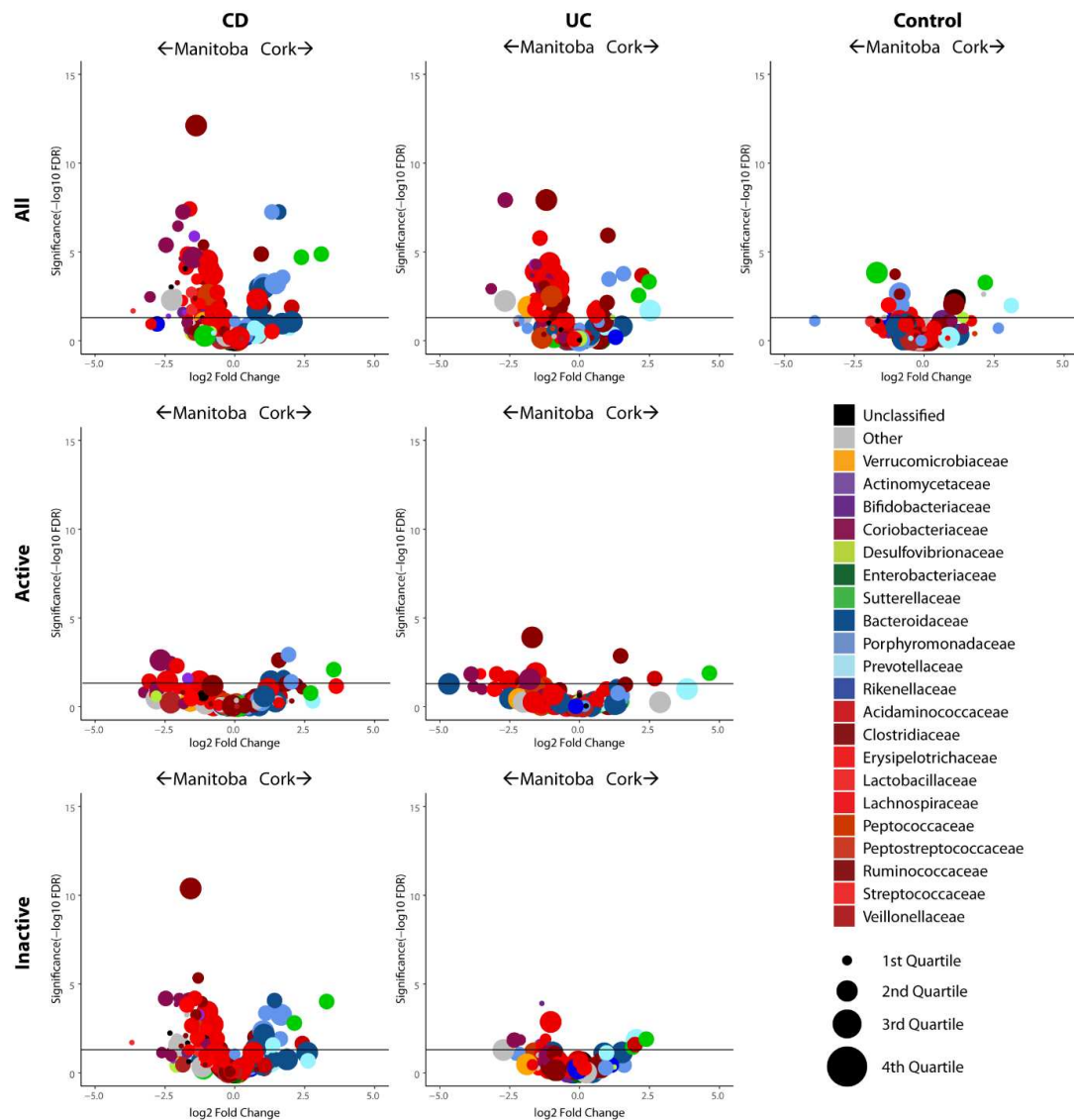
Supplementary Figure 6. Principle Coordinate Analysis based on Bray-Curtis distances with proportional normalization on all OTUs present in more than 5% of samples, grouped by disease state and geographic location for a) CD patients, b) UC patients, and c) controls. Violin plots show projections of PCoA points onto PC1 and PC2.



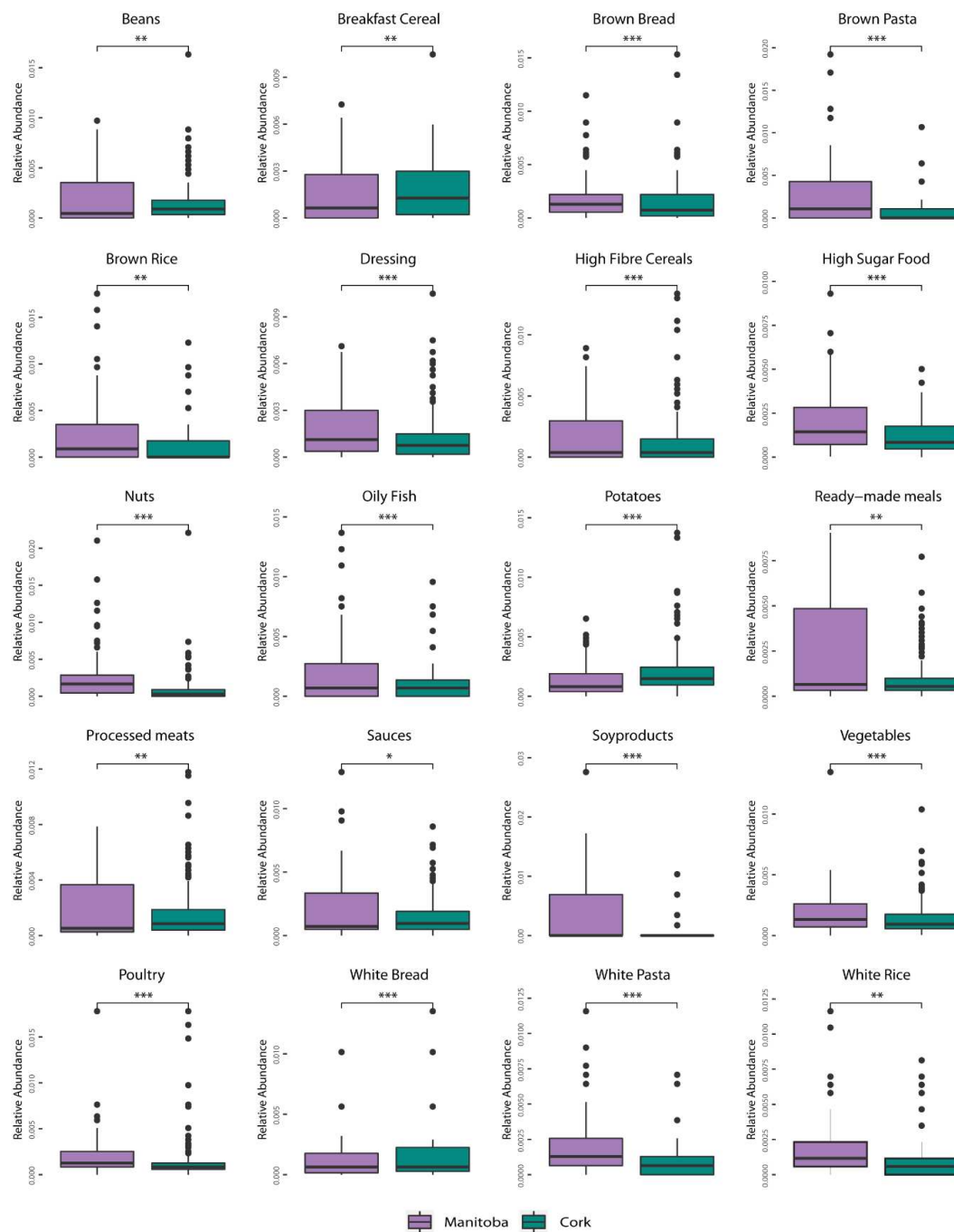
Supplementary Figure 7. Comparison of Alpha diversity across disease groups, disease activity and geographic location. P-values in black represent differences before adjustment for age while the red values show P-values after adjustment.



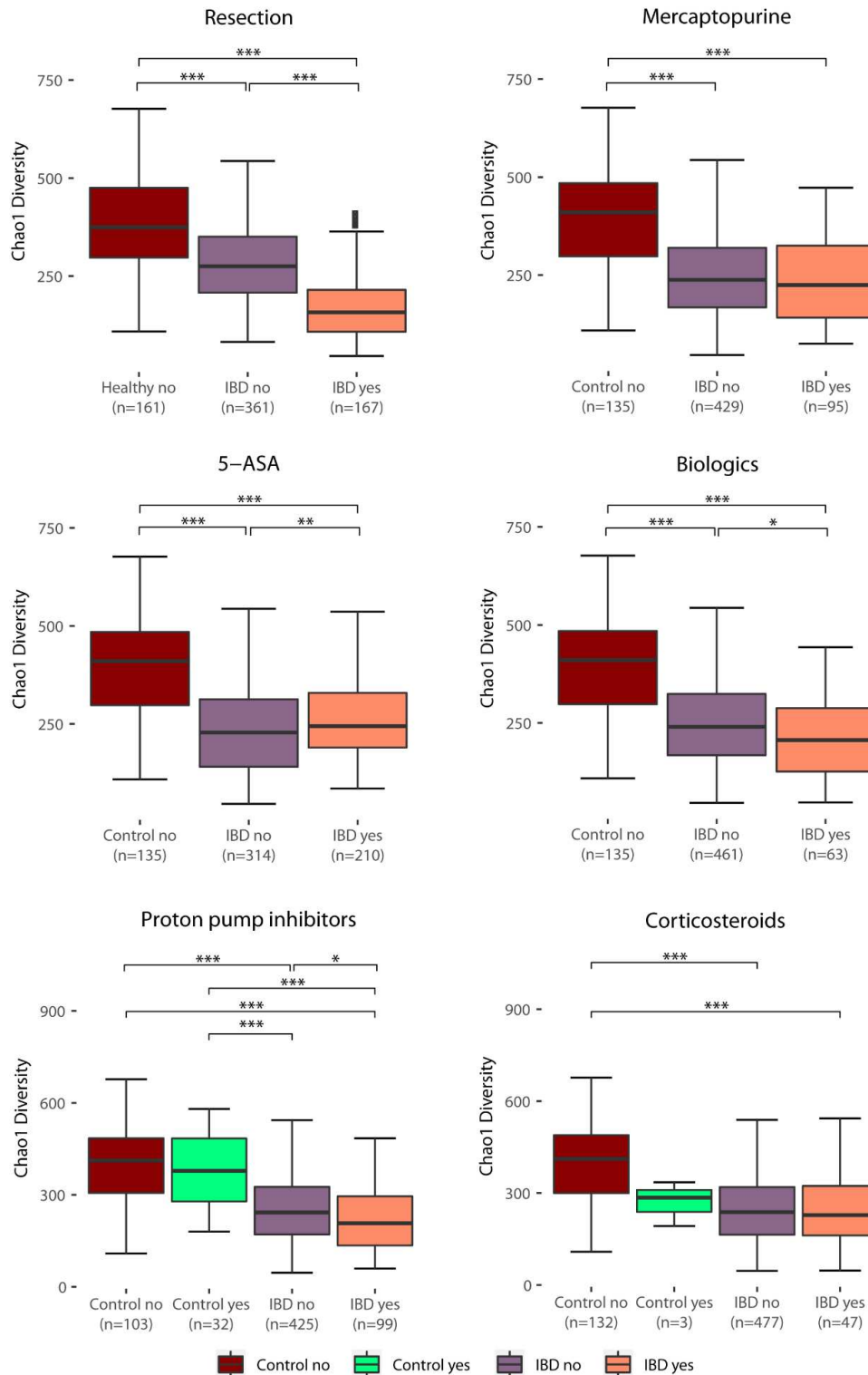
Supplementary Figure 8. Spearman correlations between alpha diversity (chao1) and a) age and b) Healthy Food Diversity (HFD) index.



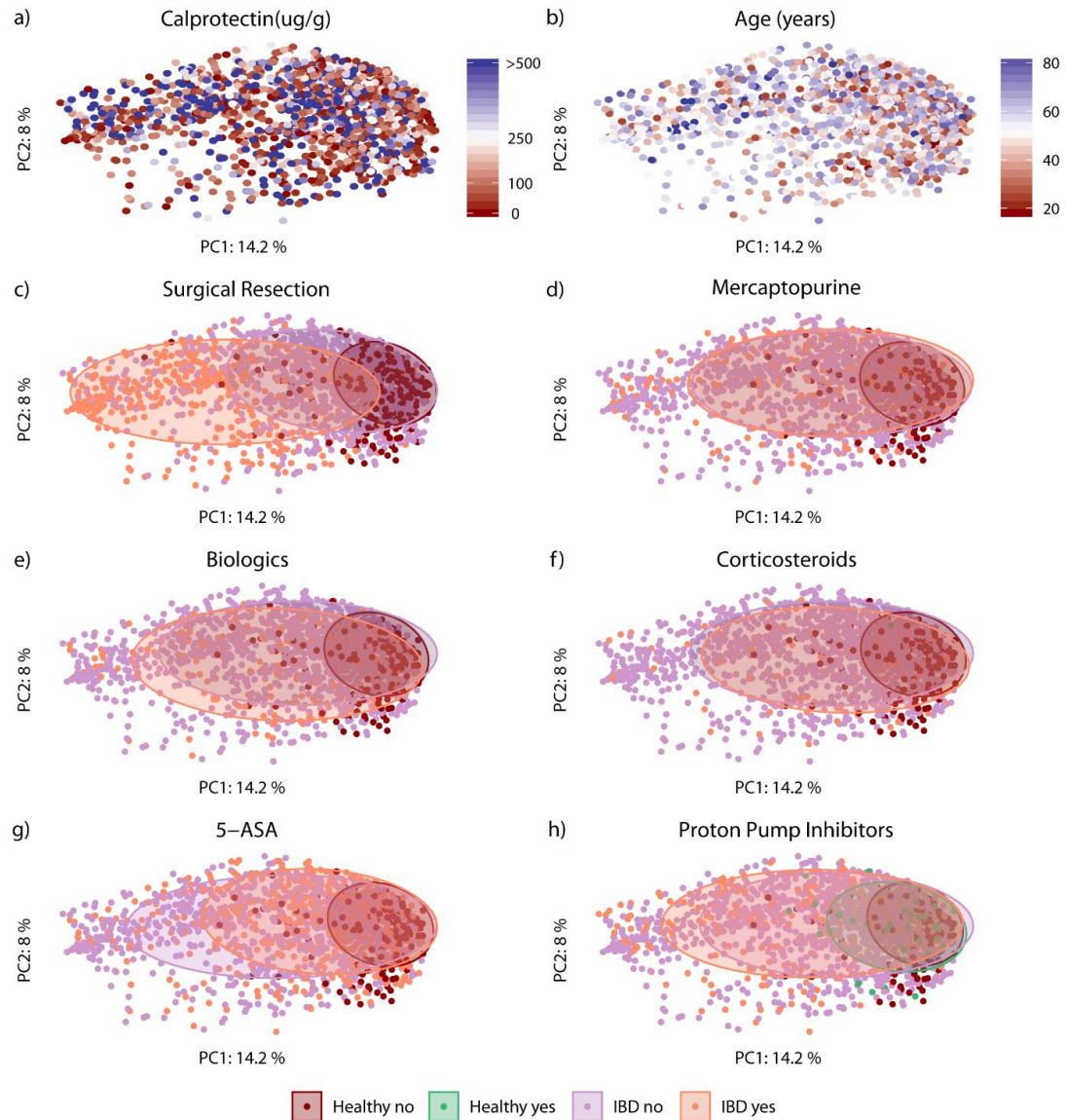
Supplementary Figure 9. Volcano plots showing differential abundant species between the Cork and Manitoba cohorts across disease and disease status. Points above the horizontal line are significant while the position on the x-axis of each point indicates the direction of fold change. The size of each point refers to the abundance of the species across the cohort while the color indicates the family rank.



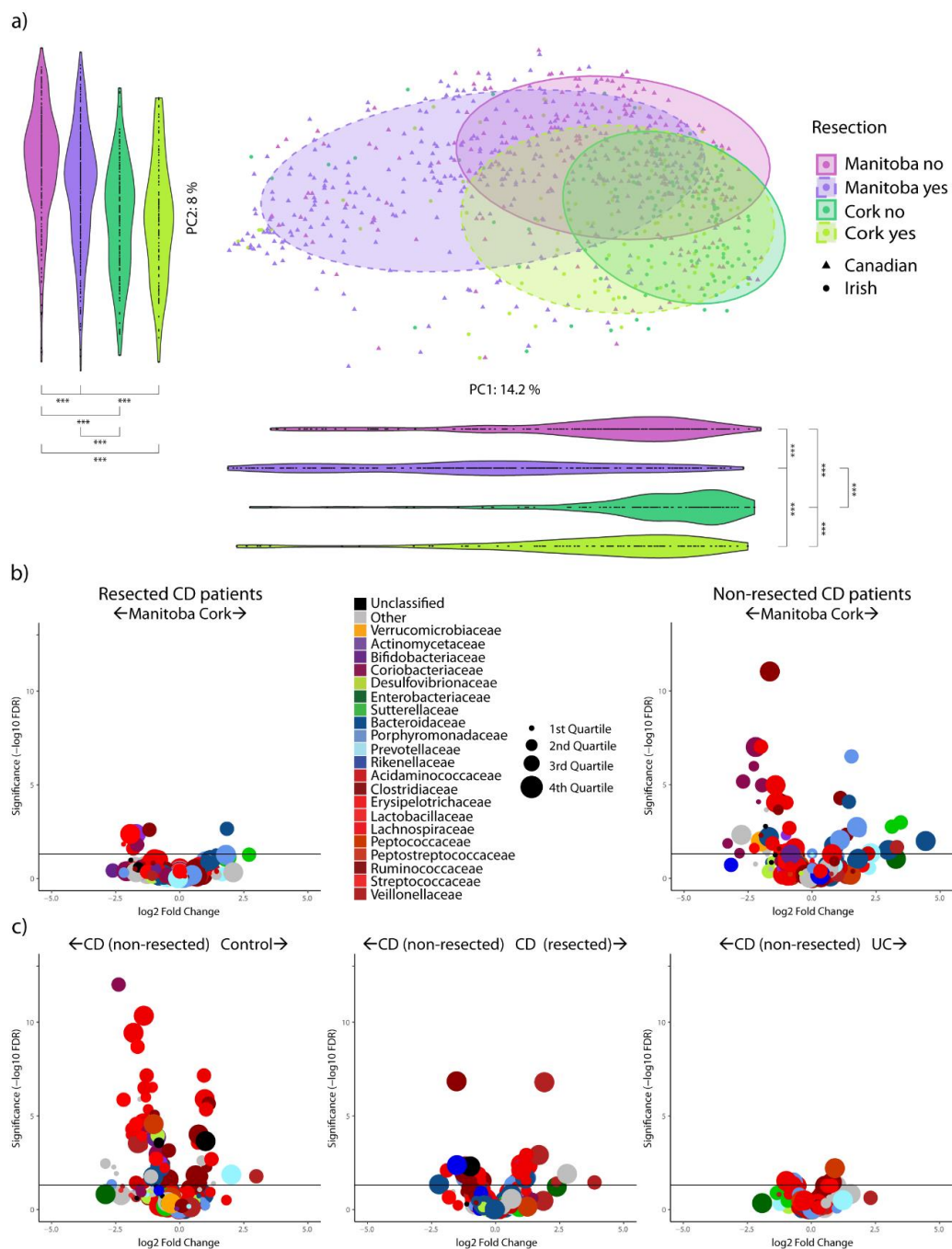
Supplementary Figure 10. Boxplots representing differential consumption of food groups between the Cork and Manitoba cohorts. Only food groups with a significant difference were plotted (20).



Supplementary Figure 11. Alpha diversity (Chao1) categorized by medication information and resection status.



Supplementary Figure 12. Principle Coordinate Analysis based on Bray-Curtis metric with proportional normalization on all OTUs present in greater than 5% of samples with samples grouped by a) fecal calprotectin, b) age, c) resection status, and d)-h) medication information.



Supplementary Figure 13. a) Principle Coordinate Analysis based on Bray-Curtis distances with proportional normalization on all OTUs present in more than 5% of samples, grouped by grouped resection status and geographic location. Volcano plots showing differential abundant species between b) Cork and Manitoba cohorts across resected and non-resected CD patients and c) between non-resected CD patient and Controls, resected CD patients, and UC patients, respectively. Points above the horizontal line are significant while the position on the x-axis of each point indicates the direction of fold change. The size of each point refers to the abundance of the species across the cohort while the color indicates the family rank.