# Gut microbiome structure and metabolic activity in IBD

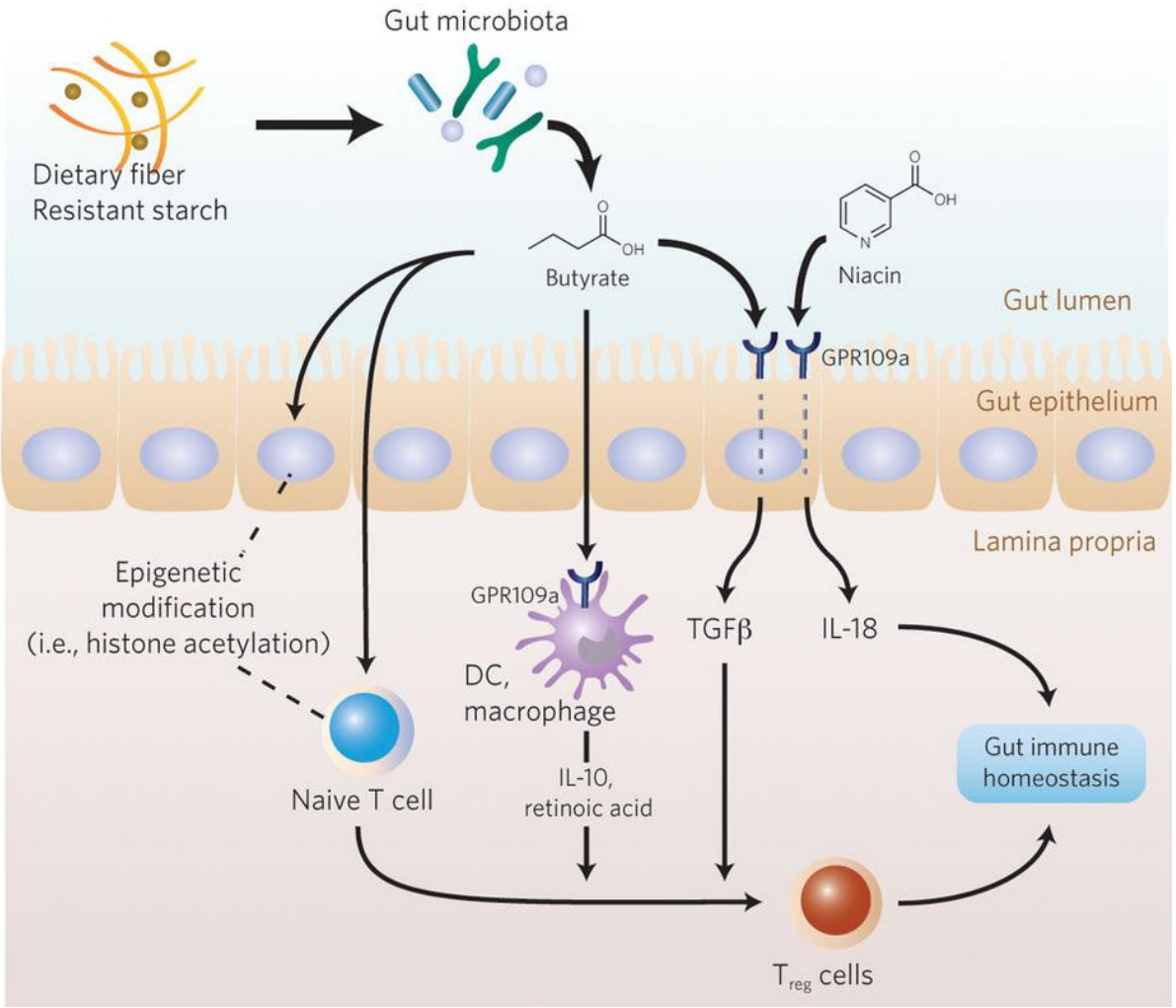
Eric Franzosa

Research Scientist, Huttenhower Lab

BST 273

15 October 2018

# The gut metabolome as a host-microbiome interface




Adapted from Won-Jae Lee & Koji Hase, *Nature Chemical Biology* 10, 416–424 (2014)

- Dietary compounds may directly promote/hinder the growth of certain microbes
- Microbes produce compounds with pro- and anti-inflammatory effects
- Most of what we know comes from targeted metabolomic surveys
- Potentially many microbially-derived, uncharacterized metabolites remain to be found

# Profiling metabolomes & metagenomes of the IBD gut

| | Non-IBD Control (control) | Crohn's Disease (CD) | Ulcerative Colitis (UC) |
|---|---|---|---|
| **Discovery cohort** (PRISM) | 34 | 68 | 53 |
| **Validation cohort** (NLIBD/LLDeep) | 22 | 20 | 23 |

**Multi'omic screening of stool samples**

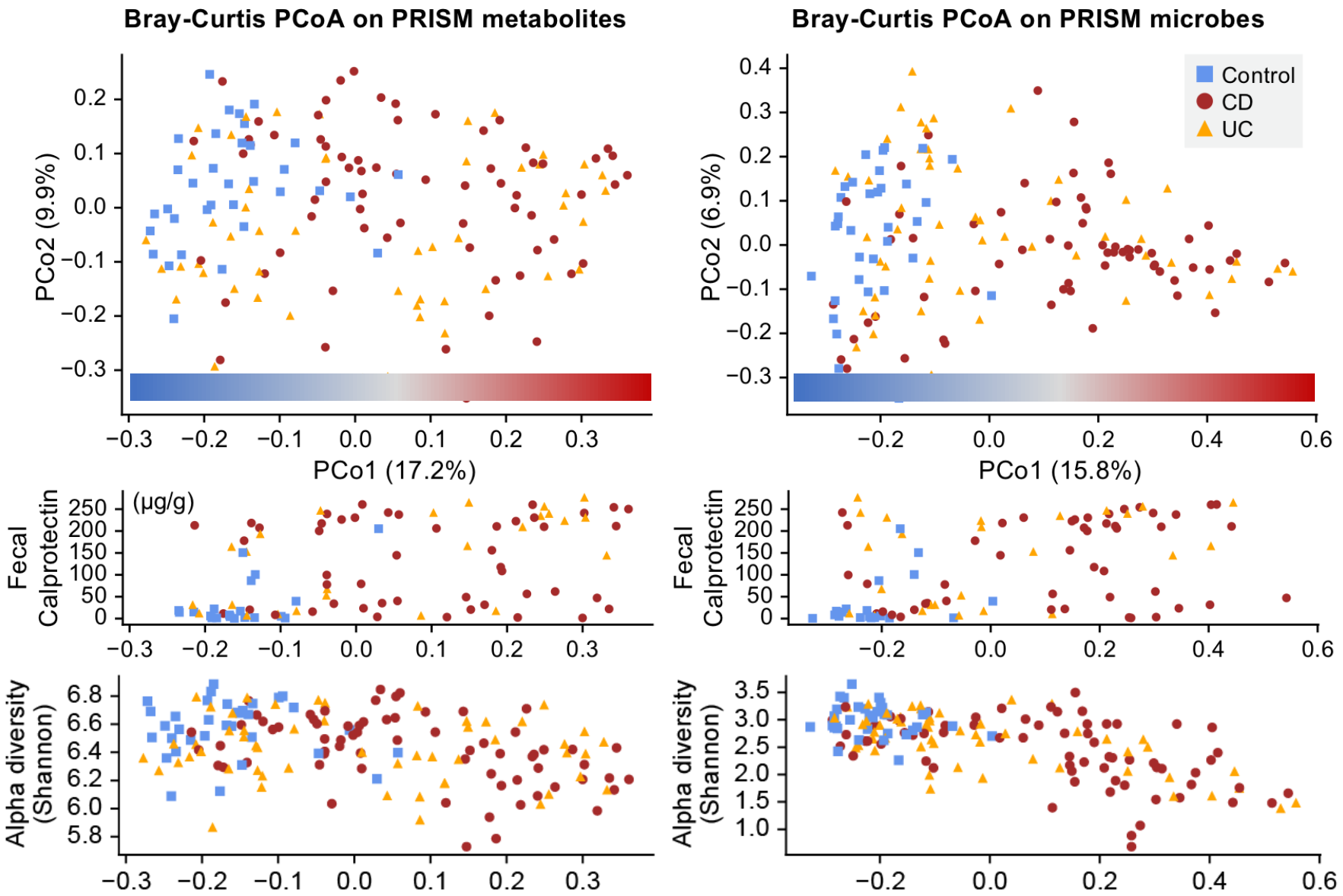
**Metagenomic shotgun sequencing**
Microbial taxa, genes, and pathways

+

**Four LC-MS metabolomics methods**
Lipids, polar metabolites, free fatty acids, & bile acids

- Metagenomes QC'ed by kneadData and profiled by MetaPhlAn2 + HUMAnN2
- ~8K measured metabolites from four untargeted metabolomics methods
- ~50% had approximate matches to HMDB (focus on molecular subclass)
- ~5% confidently matched standards from in-house compound library

# Broad ‘omics shifts correlate with host inflammation




Dominated by CD/control separation (UC is heterogeneous)

The two PCo1s were strongly correlated ($r_s$=0.66, $p$<$10^{-20}$)

PCo1s correlate with gut inflammation ($r_s$=0.49 and 0.44)

PCo1s anti-correlate with Shannon diversity ($r_s$=-0.32 and -0.57)

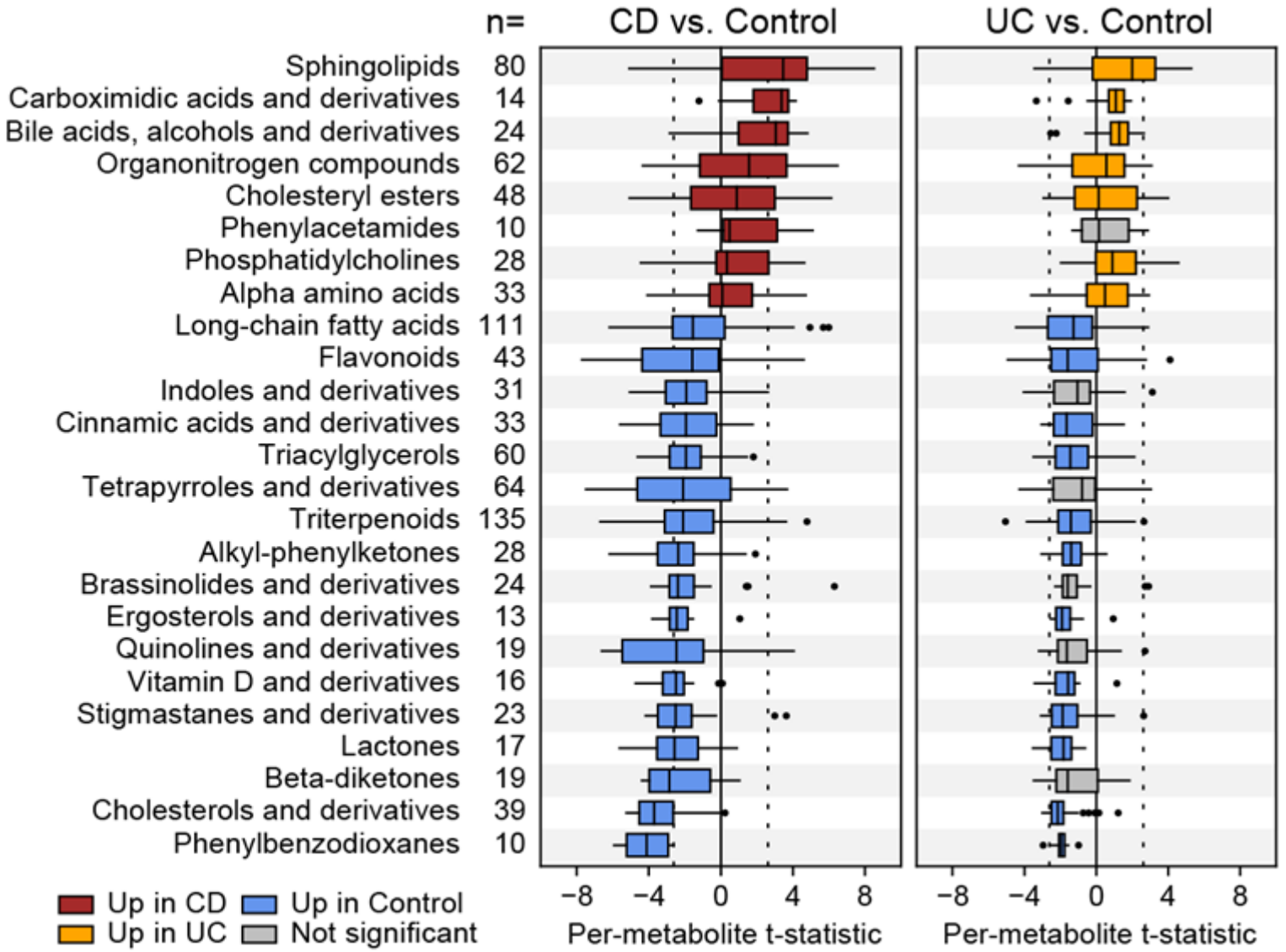
# Finding differentially abundant (DA) features in IBD

- Model (same for all features, sum-normalized within method):

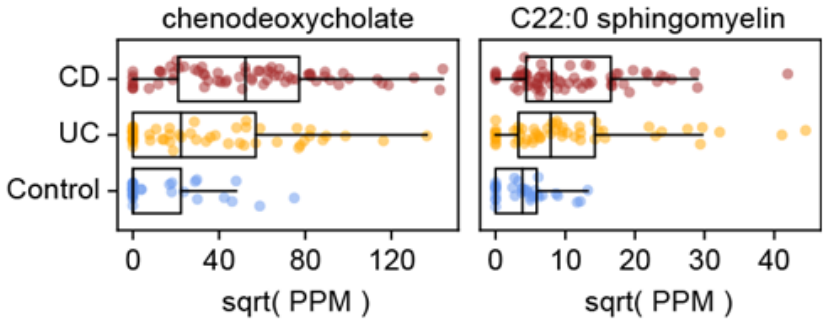
$$\log(\textbf{feature} + \epsilon) \sim \textbf{diagnosis} + \text{age} + \text{abx} + \text{mesalamine} + \text{immunosuppressants} + \text{steroids}$$

- Isolate features where diagnosis:CD or diagnosis:UC coefficient had Benjamini-Hochberg (FDR) $q$<0.05
- Save model residuals for downstream applications
- Metagenomic trends (species + enzymes) consistent with previous findings
  - We'll come back to these later…
- Many, many DA metabolites (~2.7K of 8K)
  - Used "GSEA" and clustering to simplify interpretation

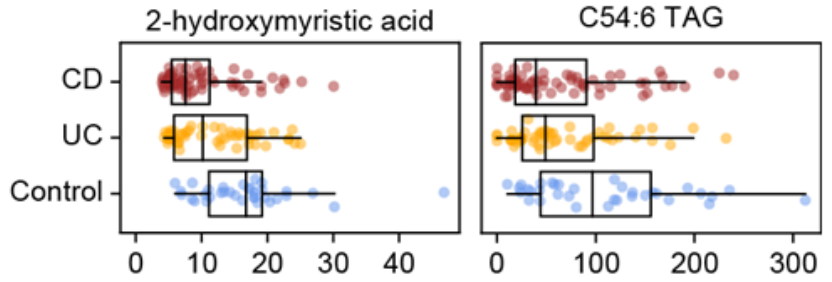
# Metabolite set enrichment analysis on model results



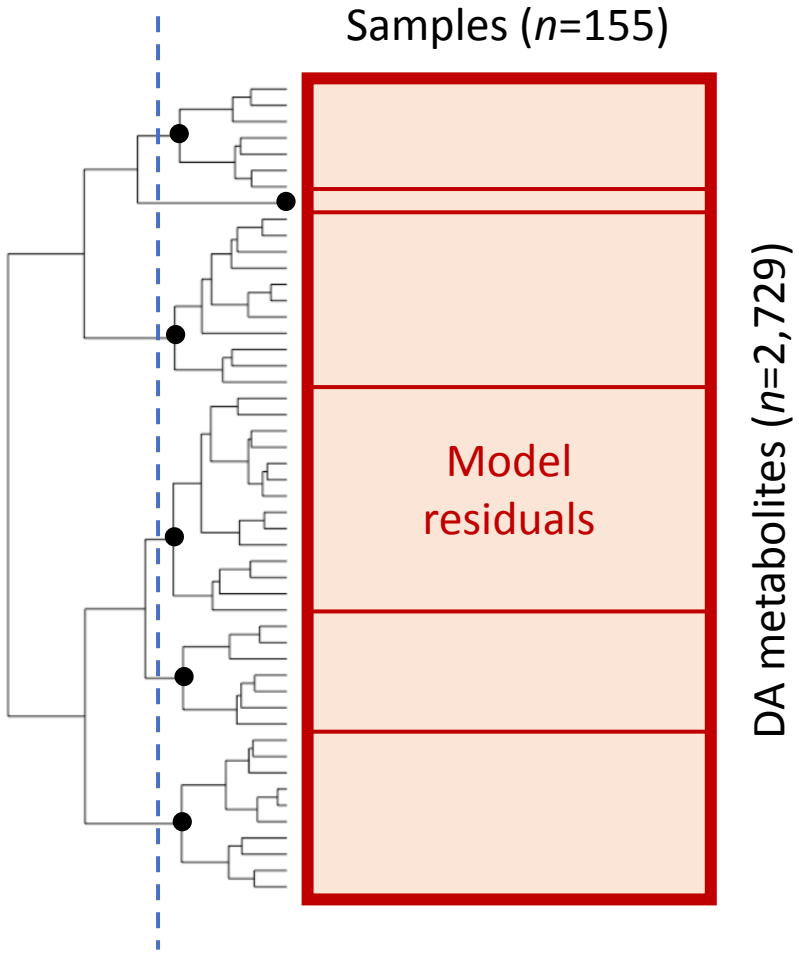

- UC trends are a weakened subset of CD trends
- Sphingolipids, bile acids up



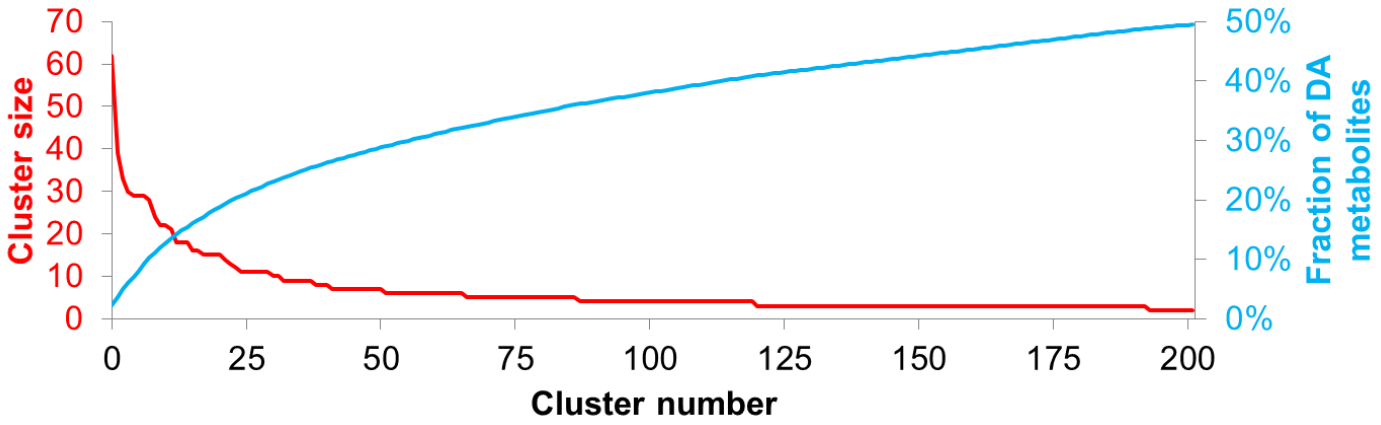

- Fatty acid subclasses down



# Finding clusters of co-varying, DA metabolites

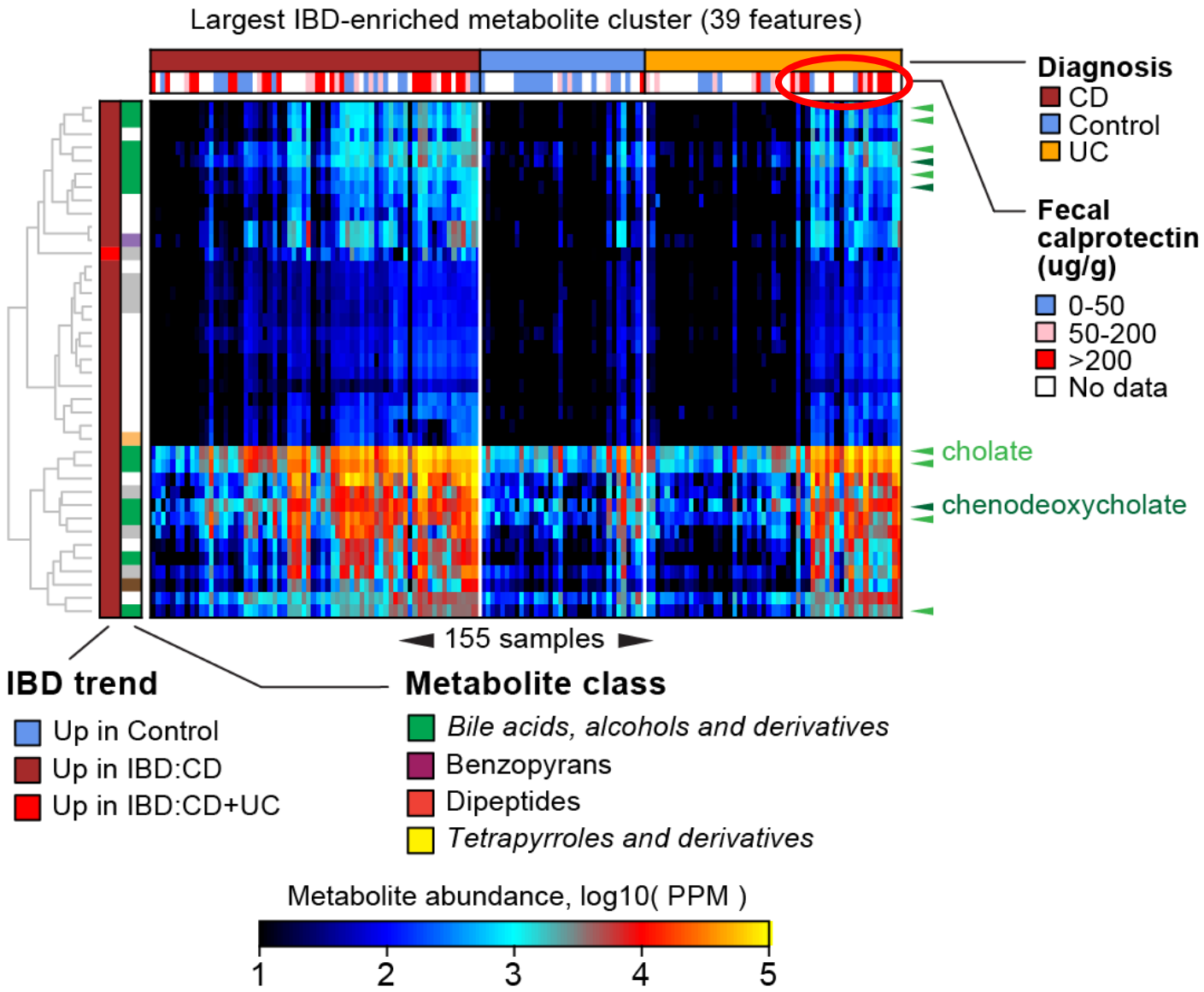



- Cluster DA metabolites based on residual abundance (Spearman as similarity measure)
- "Cut" hierarchy to define clusters with mean similarity $r_s$ = 0.7 (strongly correlated)



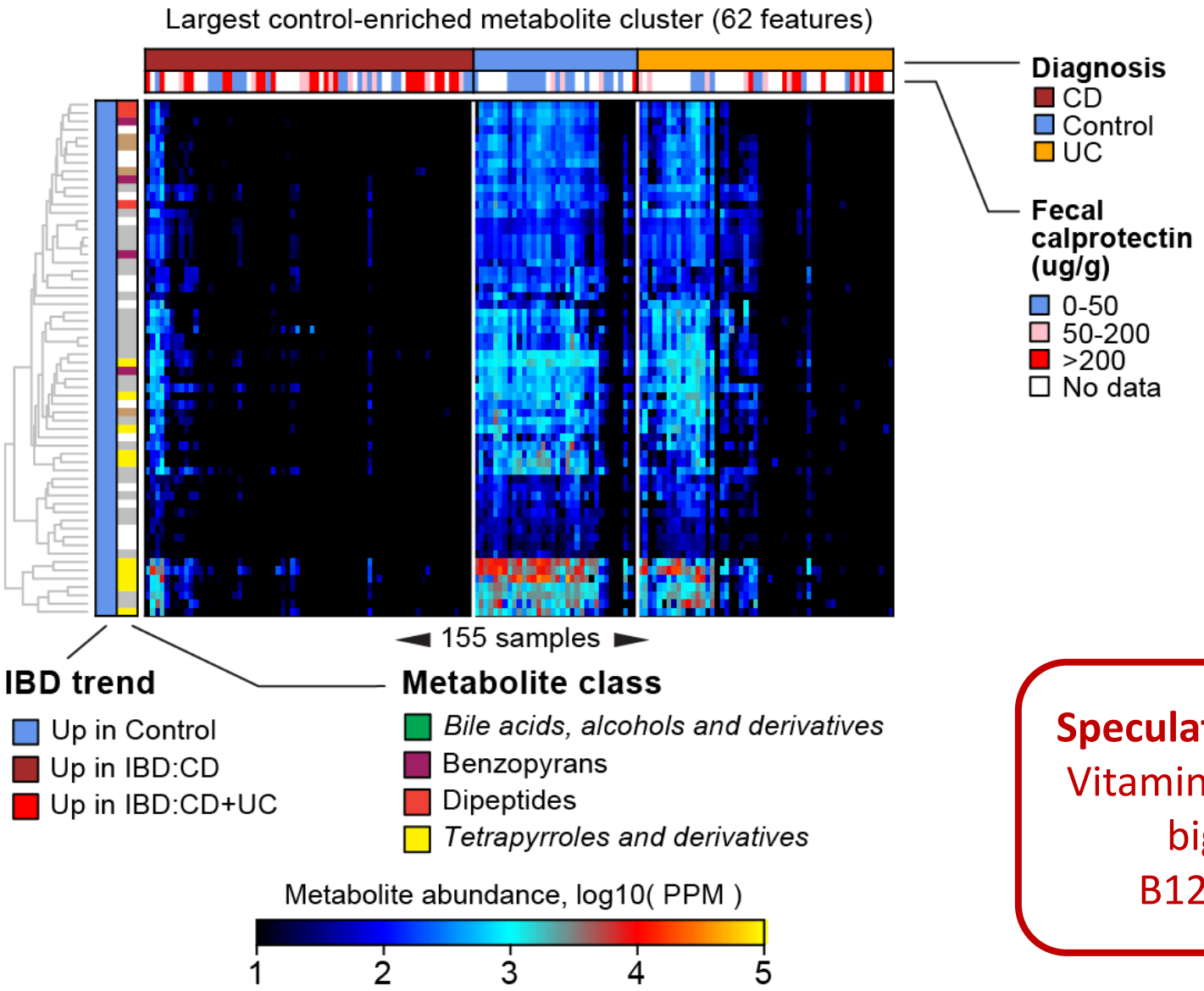

- Cluster members are **physicochemically similar**:
  - → 2.7x more similar retention times
  - → 3.0x more similar mass/charge ratios
  - → 15x more likely to be in same HMDB subclass

# The largest IBD-enriched cluster (mostly bile acids)




- Here, 39 bile-acid related metabolites co-vary strongly across samples
- All are elevated in CD
- Note substructure among UC patients, coinciding with high/low fecal calprotectin

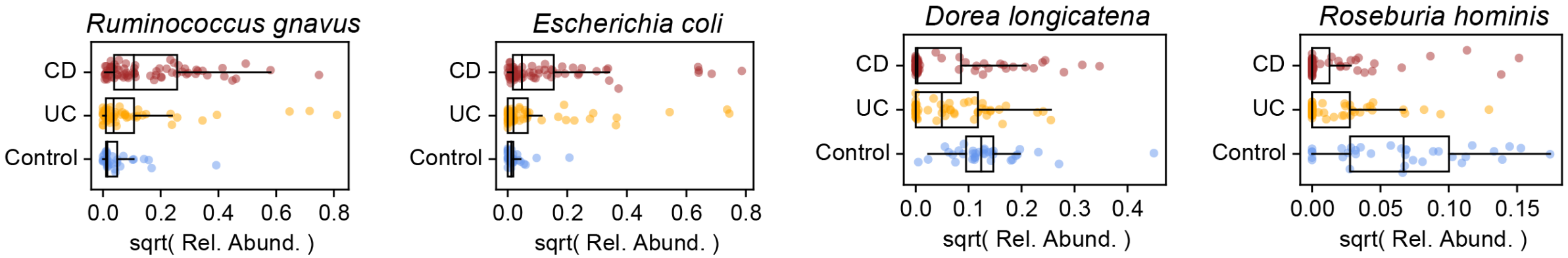
# The largest control-enriched cluster (mostly unknown)



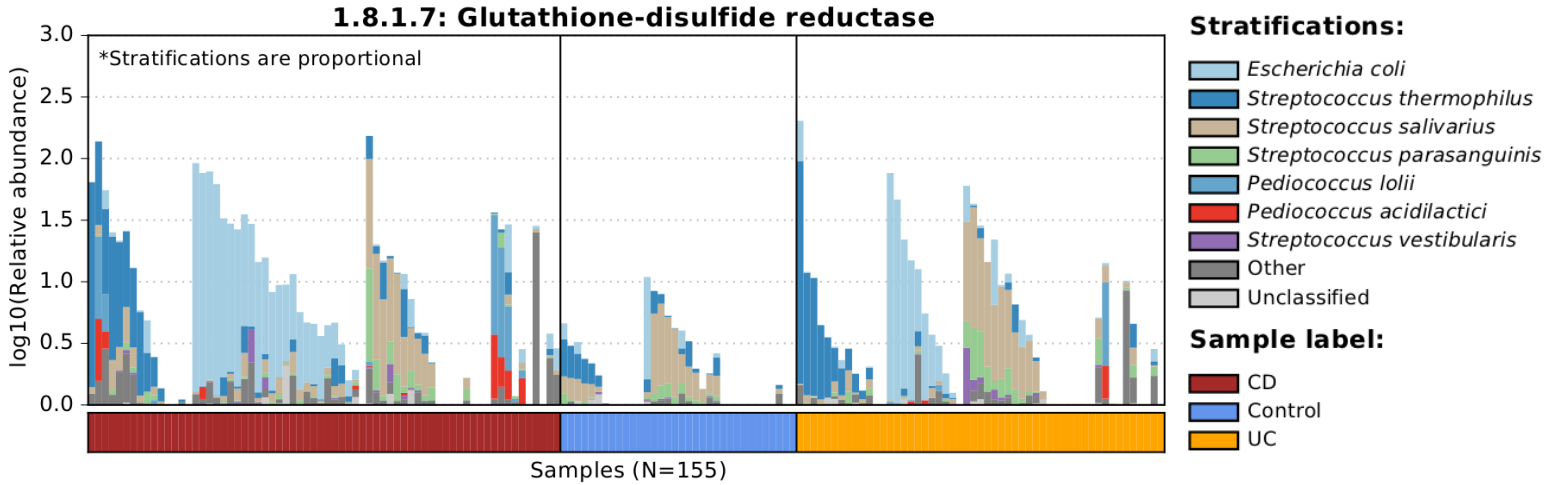

- This cluster (enriched in control subjects) contains no standard metabolites
- Could be enriched for uncharacterized microbially derived metabolites

**Speculation**: Lots of tetrapyrrole derivatives here. Vitamin B12 is a tetrapyrrole compound (but too big to be seen by these MS methods). B12-producing bugs are depleted in IBD.

# Metagenomic trends mostly follow previous findings



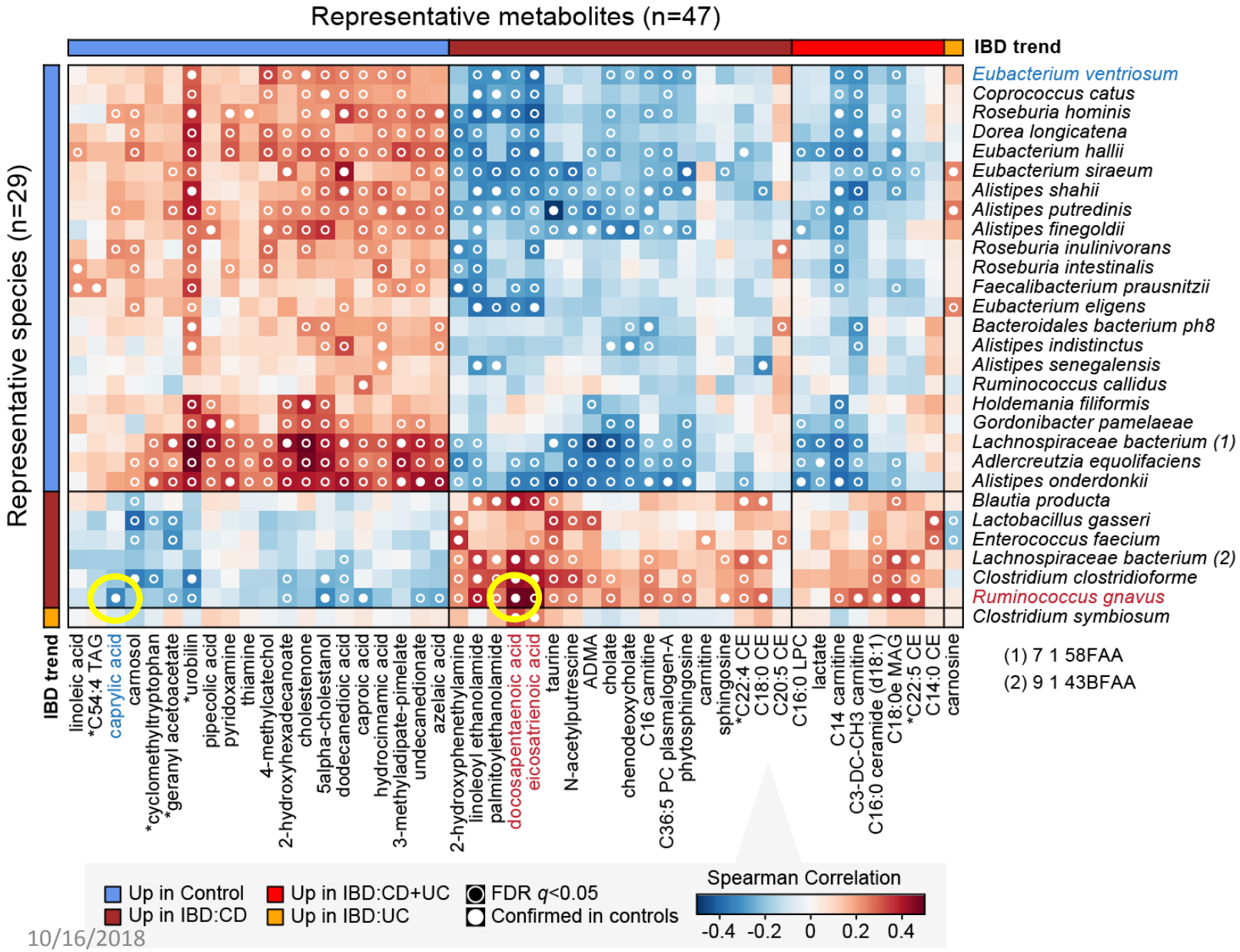

50 DA species, of which 35 were up in controls (reflects general loss of diversity in IBD)

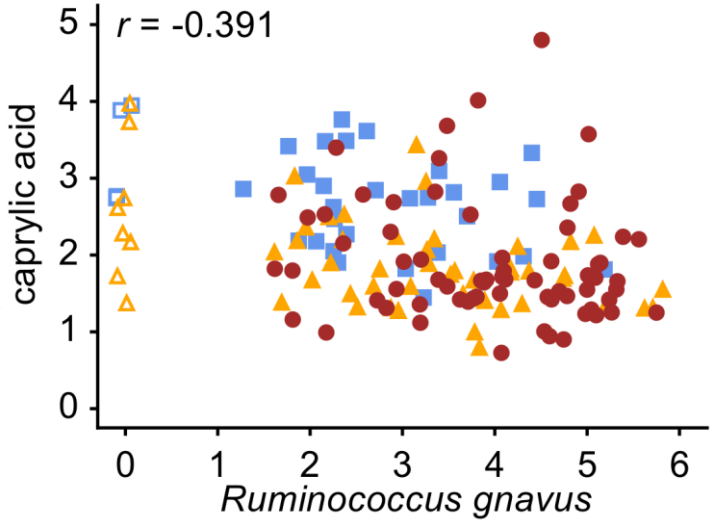


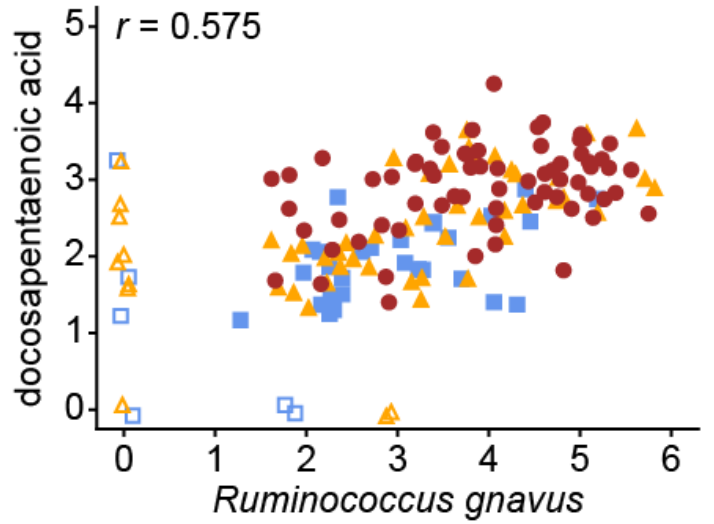

DA enzymes reflect community-level responses to a more aerobic environment in IBD (e.g. response to ox. stress)

***We focused on associating DA metagenomic features with DA metabolites!***

# Robust associations between metabolites & microbes

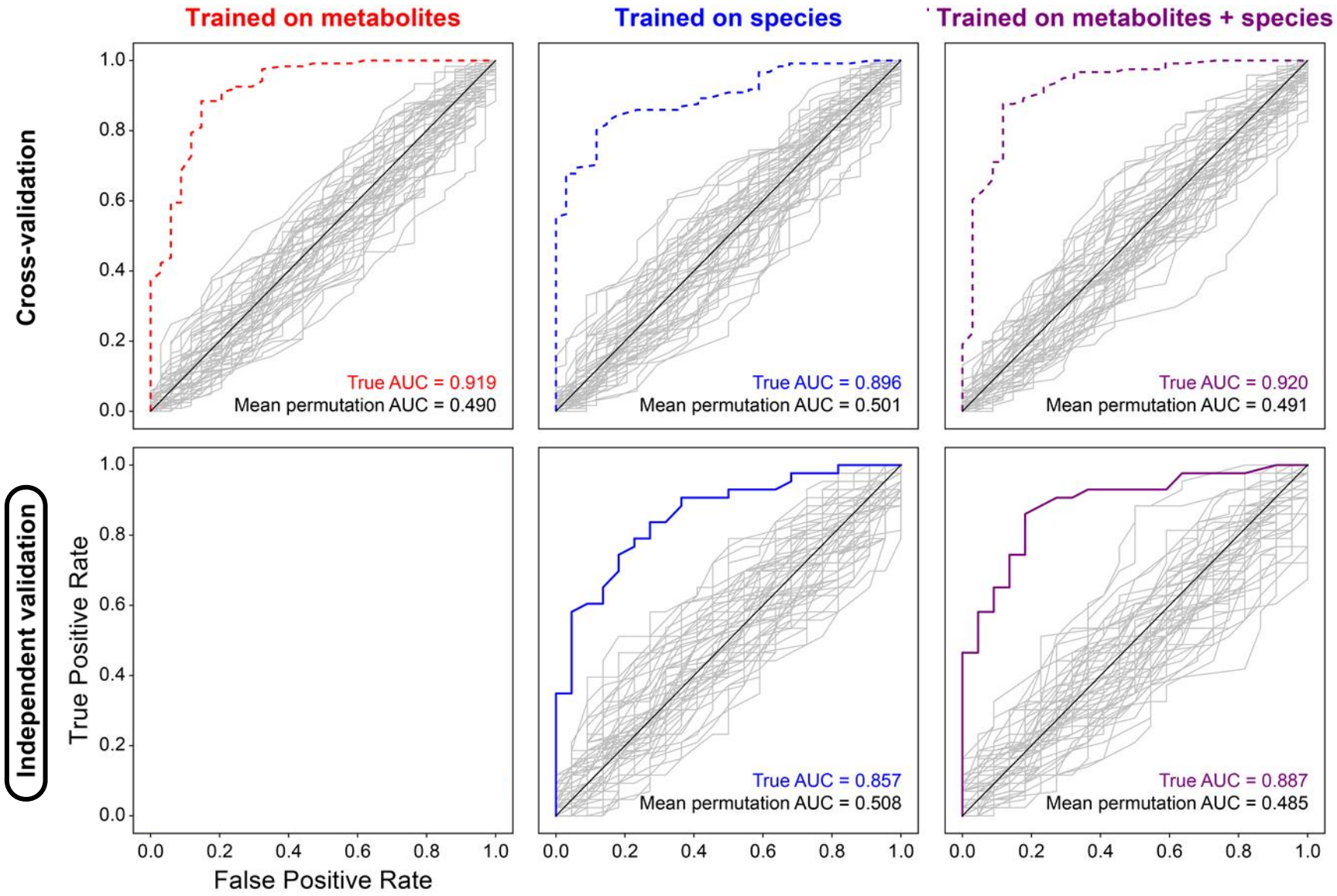



log10(PPM) units

# Random forest prediction of IBD status



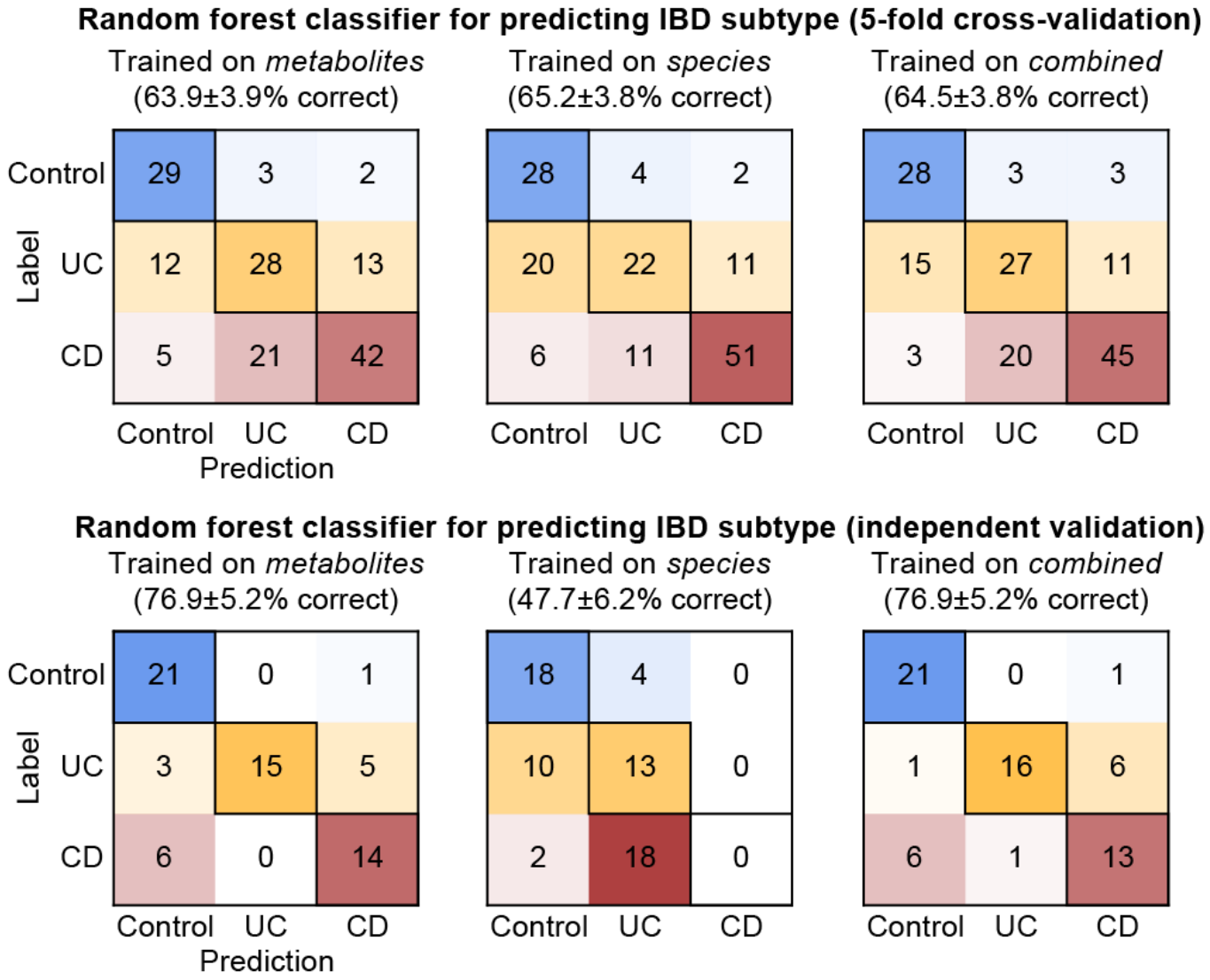

Metabolites predict IBD status well (AUC>0.9)

Predictor generalizes to new samples

Species also perform well, despite drawing on far fewer features

No marked boost in accuracy from combining metabolites and species

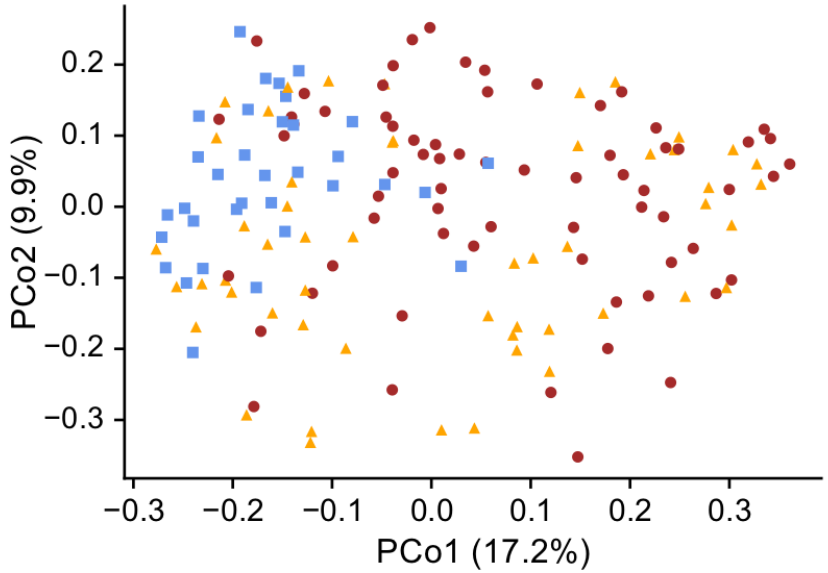
# Random forest prediction of IBD type: CD vs. UC



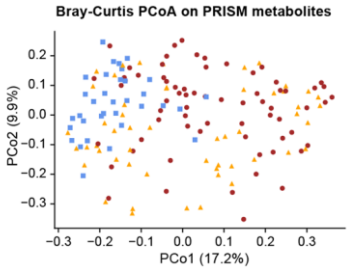

**Non-IBD** controls rarely classified as **IBD:CD** or **IBD:UC**

**IBD:CD** rarely classified as **Non-IBD** control, but sometimes as **UC**

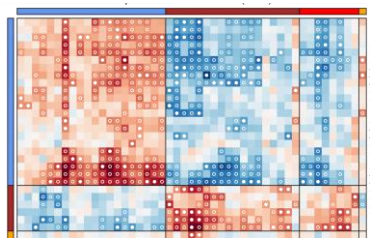
**IBD:UC** hard to get right



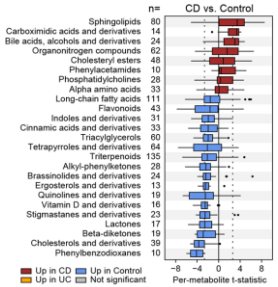
# Summary



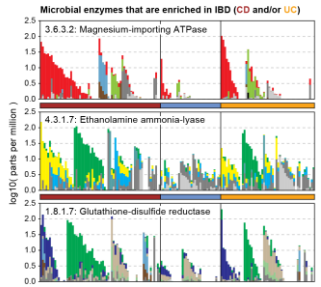

Metabolomes broadly stratified by subject inflammation level



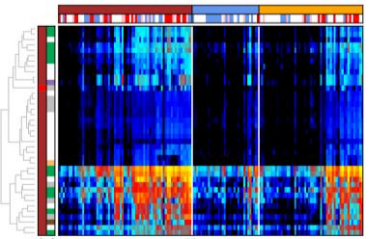
Putatively mechanistic microbe-metabolite associations



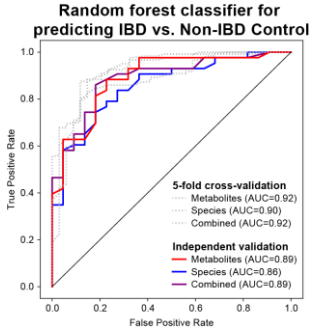

Many molecular classes are differentially abundant in IBD




Microbial functional adaptations to the IBD gut agree with earlier work



Some classes (clusters) are well understood, while others are largely uncharacterized




Metabolites and microbes can classify IBD status reasonably well, though subtype is harder

# Acknowledgements

We thank the members of the PRISM, NLIBD, and LLDeep cohorts for generously providing sample materials for use in the study.

HARVARD T.H. CHAN SCHOOL OF PUBLIC HEALTH

Curtis Huttenhower
Himel Mallick
Lauren McIver

BROAD INSTITUTE

Clary Clish
Julian Avila-Pacheco
Justin Scott
Kerry Peirce
Amy Deik
Kevin Bullock

MASSACHUSETTS GENERAL HOSPITAL

Ramnik Xavier
Jenny Sauk
Robin Wilson
Betsy Stevens

Hera Vlamakis
Alexandra Sirota-Madi
Nadine Fornelos
Tommi Vatanen
Brantley Hall

NOVARTIS

Henry Haiser
Jeffrey Porter

university of groningen

Rinse Weersma
Cisca Wijmenga
Jingyuan Fu
Floris Imhann
Alexandra Zhernakova

The Boston research team was funded by: the U.S. National Science Foundation, the U.S. National Institutes of Health, the Crohn's and Colitis Foundation of America, and the Center for Microbiome Informatics and Therapeutics.
The Netherlands research team was funded by: IN-CONTROL CVON, the MLDS Dutch Digestive Foundation, the Netherlands Organization for Scientific Research, a Spinoza Prize, and the European Research Council.