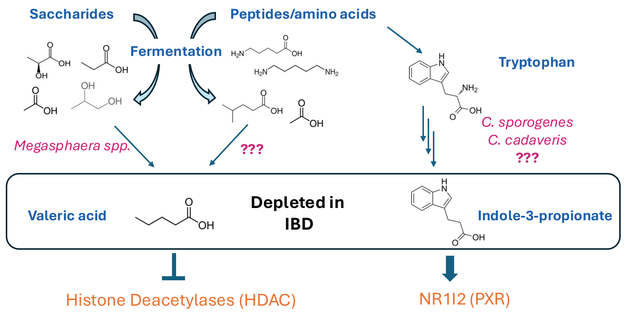
Identification of key bacterial taxa driving bioactive metabolite production in the gut microbiome using a reverse-translational approach

Marin Vulić, Nathan D. Hicks, Nicholas Beauchemin, Jessica Bryant, Christopher Ford, David I. Lichter, Lisa von Moltke, Matthew R. Henn

SERES THERAPEUTICS™

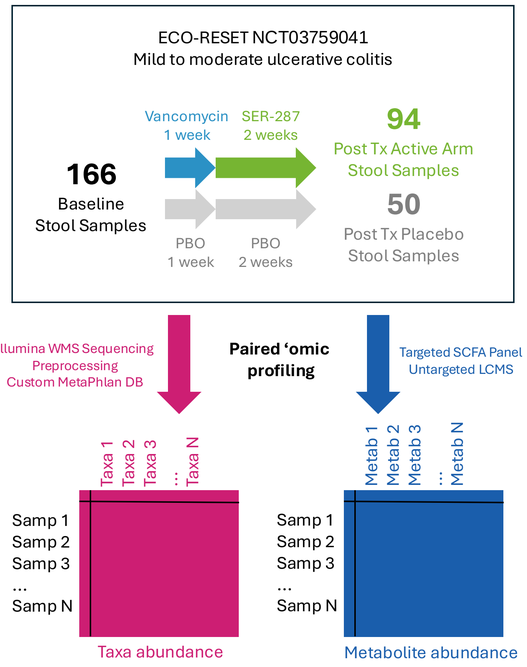
Introduction



The human gut microbiota collectively produce a diverse array of bioactive metabolites including short chain fatty acids (SCFAs) and amino acid derivatives. Altered concentrations of these metabolites are often associated with disease states. In cases where a metabolite has been demonstrated to directly interact with the host, modulating its production and/or consumption may represent a target for therapeutic intervention. However, for most metabolites, our current understanding of their production is incomplete and often restricted to bacterial taxa that are rarely present in the human gut microbiome, in model systems that may not accurately represent the conditions found within actual human gut communities.

Here we leverage paired metabolomic and metagenomic data from a microbiome therapeutic clinical trial to link bacterial taxa with metabolic transformations they perform in their human host.

Predicting stool metabolite concentrations using bacterial taxonomic profiles



Constructing interpretable metabolite prediction models with nested lasso regression

Model Type 1: Predict metabolite abundance by taxonomic profiles ***across*** subjects at baseline

$\log(\text{metab}) \sim \log(\text{taxa 1}) + \log(\text{taxa 2}) + \ldots + \log(\text{taxa n})$

Model Type 2: Predict ***change*** in metabolite abundance by ***change*** in taxonomic profiles ***within subjects over time***

Step 1: Calculate per-subject ***change*** in 'omics

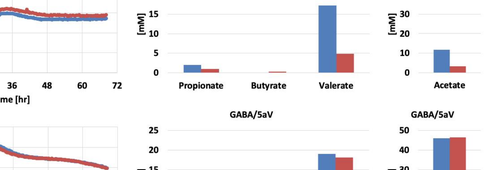
$\Delta\text{metab} = \log(\text{Post Tx metab} / \text{Baseline metab})$

$\Delta\text{taxa} = \log(\text{Post Tx taxa} / \text{Baseline taxa})$

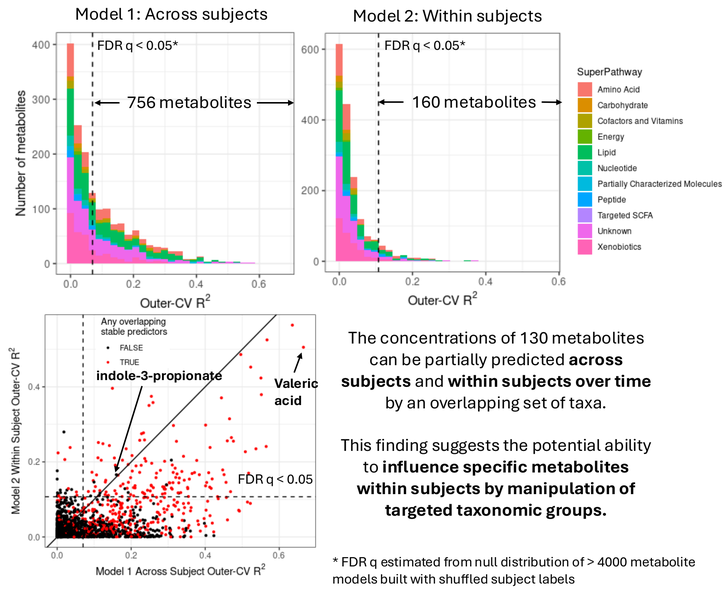
Step 2: Model ***change*** in metabolites by ***change*** in taxa

$\Delta\text{metab} \sim \Delta\text{taxa 1} + \Delta\text{taxa 2} + \ldots + \Delta\text{taxa n}$

10-fold nested lasso regression on each of 1767 metabolites



Hundreds of metabolites can be predicted by genus level taxonomic profiles



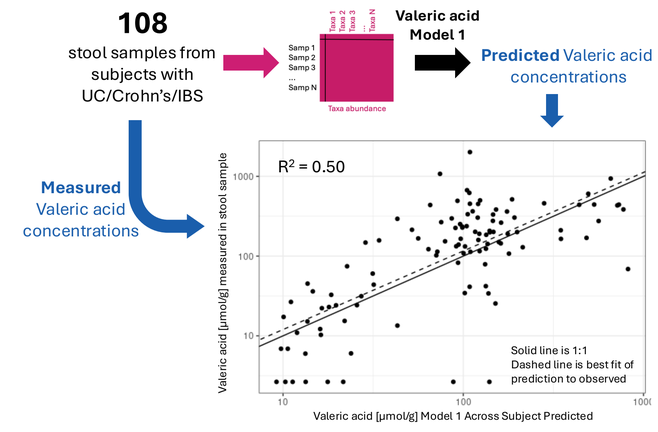
The concentrations of 130 metabolites can be partially predicted **across subjects** and **within subjects over time** by an overlapping set of taxa.

This finding suggests the potential ability to **influence specific metabolites within subjects by manipulation of targeted taxonomic groups.**

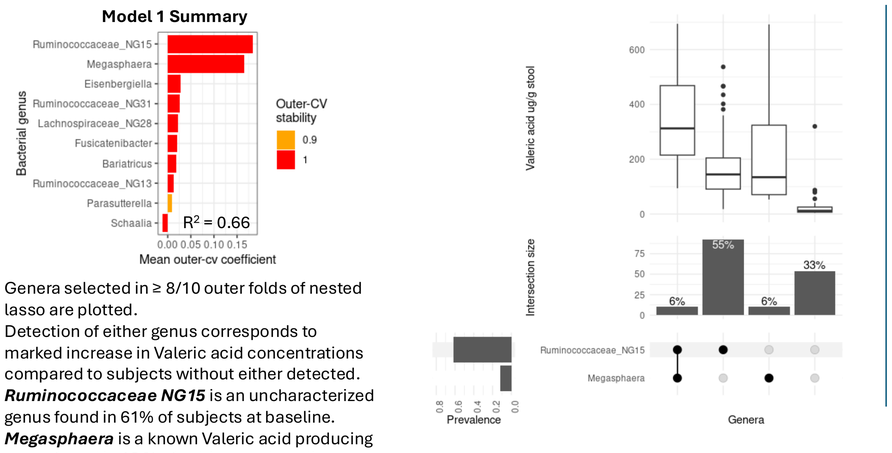
* FDR q estimated from null distribution of > 4000 metabolite models built with shuffled subject labels

Assessing model performance and taxonomic predictors for the SCFA Valeric acid

Valeric acid concentrations **across subjects** are well-predicted in a second dataset

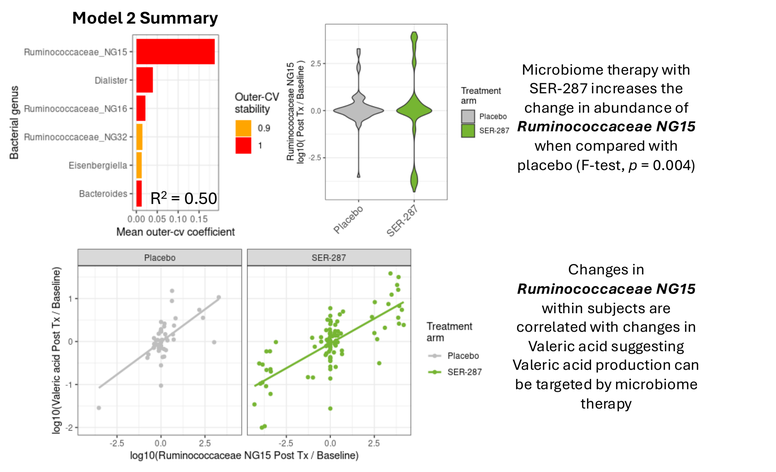


Genera ***Ruminococcaceae NG15*** and ***Megasphaera*** are strong and stable predictors of Valeric acid production ***across subjects at baseline***



- Genera selected in ≥ 8/10 outer folds of nested lasso are plotted.
- Detection of either genus corresponds to marked increase in Valeric acid concentrations compared to subjects without either detected.
- ***Ruminococcaceae NG15*** is an uncharacterized genus found in 61% of subjects at baseline.
- ***Megasphaera*** is a known Valeric acid producing genus found in 12 % of subjects at baseline.

Ruminococcaceae NG15 alone is a strong and stable predictor of Valeric acid production ***within subjects over time***



Microbiome therapy with SER-287 increases the change in abundance of ***Ruminococcaceae NG15*** when compared with placebo (F-test, $p = 0.004$)

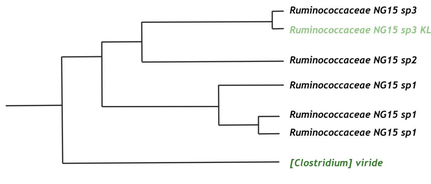
Changes in ***Ruminococcaceae NG15*** within subjects are correlated with changes in Valeric acid suggesting Valeric acid production can be targeted by microbiome therapy

Activity guided purification identifies novel Valeric acid producer

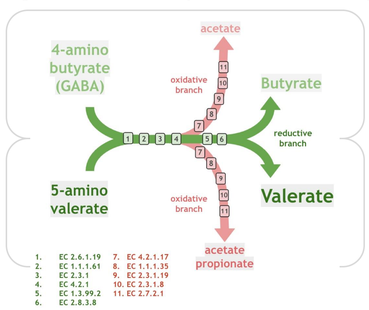
Megasphaera produces Valeric acid from fermentation products lactate, acetate, propionate through a poorly characterized mechanism.

2 described species belonging to ***Ruminococcaceae NG15*** genus are able to metabolize C4 and C5 amino-carboxylic acids.

- *[Clostridium] viride*, not present in human gut microbiome, is capable of growth on C5 **5-aminovaleric** acid.
- *Ruminococcaceae NG15* species (KLE1738), present in human gut is capable of growth on C4 **4-aminobutyric** acid (C5 5-aminovaleric acid was not tested as growth substrate).



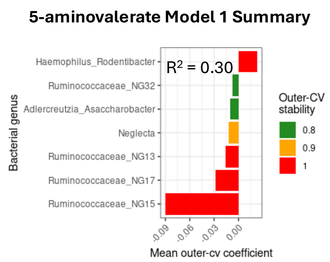
5-aminovalerate and GABA metabolic pathways of these two ***Ruminococcaceae NG15*** species consist of oxidative and reductive branch in which **5-aminovalerate** is oxidized to acetate and propionate and reduced to **valerate**, whereas GABA is oxidized to acetate and reduced to butyrate. At least some enzymes in these pathways are known to use both analogous C4 and C5 substrates, suggesting they might be identical and likely shared across the entire genus.



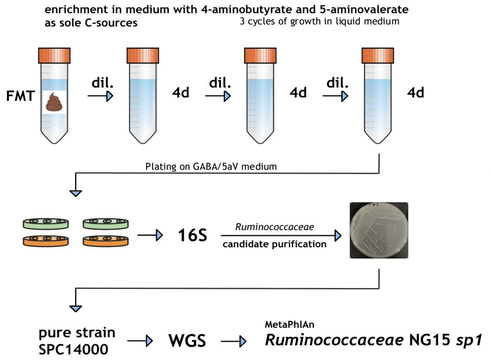
R. NG15 sp3

[C.] viride

Ruminococcaceae NG15 genus is negatively associated with **5-aminovalerate**, indicating that it may metabolize 5-aminovalerate to produce valeric acid *in vivo*.



By selecting for growth on 5-aminovalerate and GABA we were able to isolate a novel ***Ruminococcaceae NG15*** species from stool of a healthy donor.



Ruminococcaceae NG15 sp1 produces Valeric acid from 5-aminovalerate

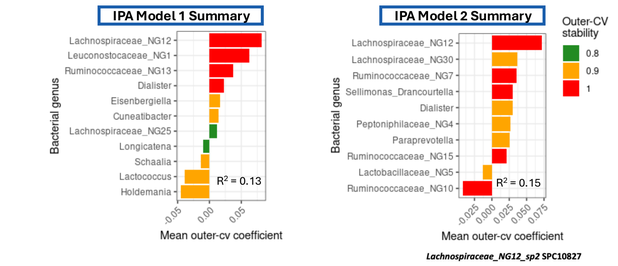
Ruminococcaceae NG15 sp1 grown in peptone medium with **A**) GABA, **B**) 5-aminovalerate and **C**) equimolar mix of GABA and 5-aminovalerate.

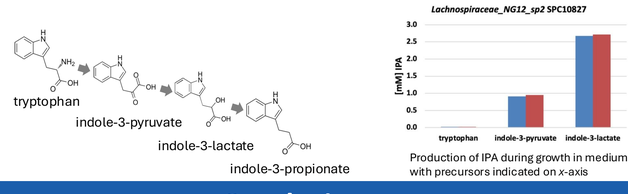
5-aminovalerate and GABA are carbon and energy sources for ***Ruminococcaceae NG15 sp1*** as they support growth as sole C-sources. Fermentation products show that 5-aminovalerate and GABA can both be oxidized or reduced.

Optimal growth is achieved when **5-aminovalerate** is reduced to **valeric acid** and GABA is oxidized to butyric acid. This, combined with ***Ruminococcaceae NG15 sp1***'s inability to utilize any other tested substrate, suggests a reliance on amino acid metabolism by other microbiome members (Arginine → Ornithine → Proline → 5-aminovalerate, Lysine → 5-aminovalerate, Glutamate → GABA, Lysine → GABA).

Indole-3-propionate models identify novel taxa

Immuno-modulatory metabolite **3-IPA** is produced by microbiome from amino acid tryptophan. Tryptophan to 3-IPA pathway has been described in *Clostridium sporogenes*, which is rarely part of human gut microbiome. We identify novel, highly prevalent genus from *Lachnospiraceae* family that is producing **3-IPA** from indole-3-pyruvate and indole-3-lactate.





Production of IPA during growth in medium with precursors indicated on x-axis

Conclusions

- Designing therapeutics to modulate microbiome-derived metabolite concentration in the gut requires a detailed understanding of the microbial taxa that can drive changes in metabolite production *within people over time*.
- Unexpectedly, relatively simple lasso models predicting a metabolite's concentration from microbiome composition at the genus level perform significantly better than expected by chance for hundreds of metabolites.
- For two common, immuno-modulatory microbially-derived metabolites (valeric acid and **i**ndole-3-**p**ropionic **a**cid), we demonstrate that these models identify taxa capable of producing them. This enabled the isolation of these taxa, validation of their valerate and 3-IPA production capabilities, and identification of the precursors and pathways they utilize to produce these metabolites.
- These taxa are highly prevalent when compared with model species that share metabolic capacities and are likely the key drivers of metabolite production *in vivo*.