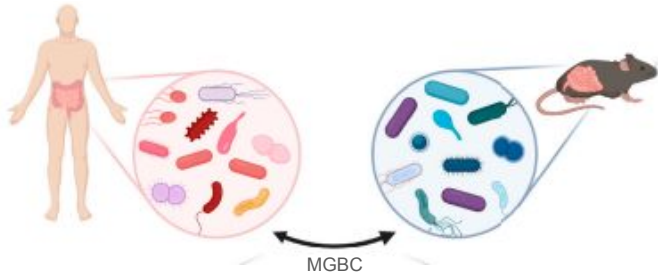


Translation of microbiome-phenotype association between hosts

Noa Mark, Supervised by Prof. Elhanan Borenstein, Aug 24



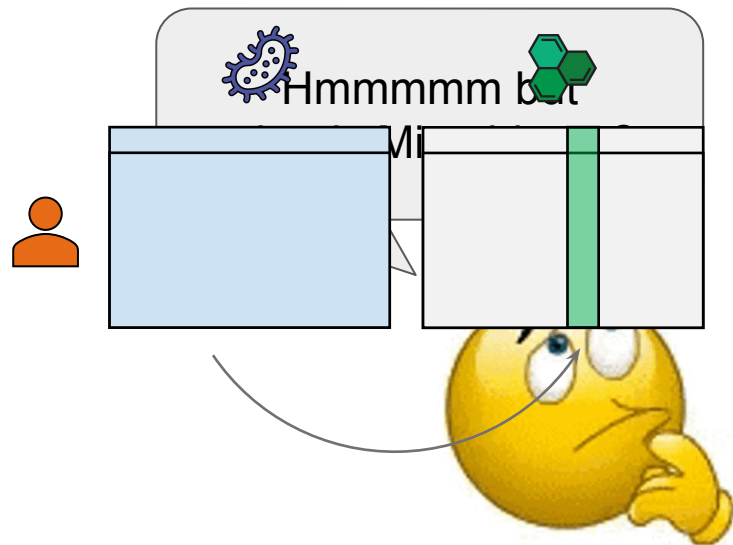
Borenstein Lab
Metagenomic Systems Biology



The research question

Can we apply domain adaptation / transfer learning between hosts to predict metabolite level from the microbiome community?

- Understand if, how and under which conditions: Microbiome-Metabolome association preserved between hosts?
- Can we predict metabolome levels from host Microbiome?





Human gut microbiome

A collection of microorganism that live within a host.

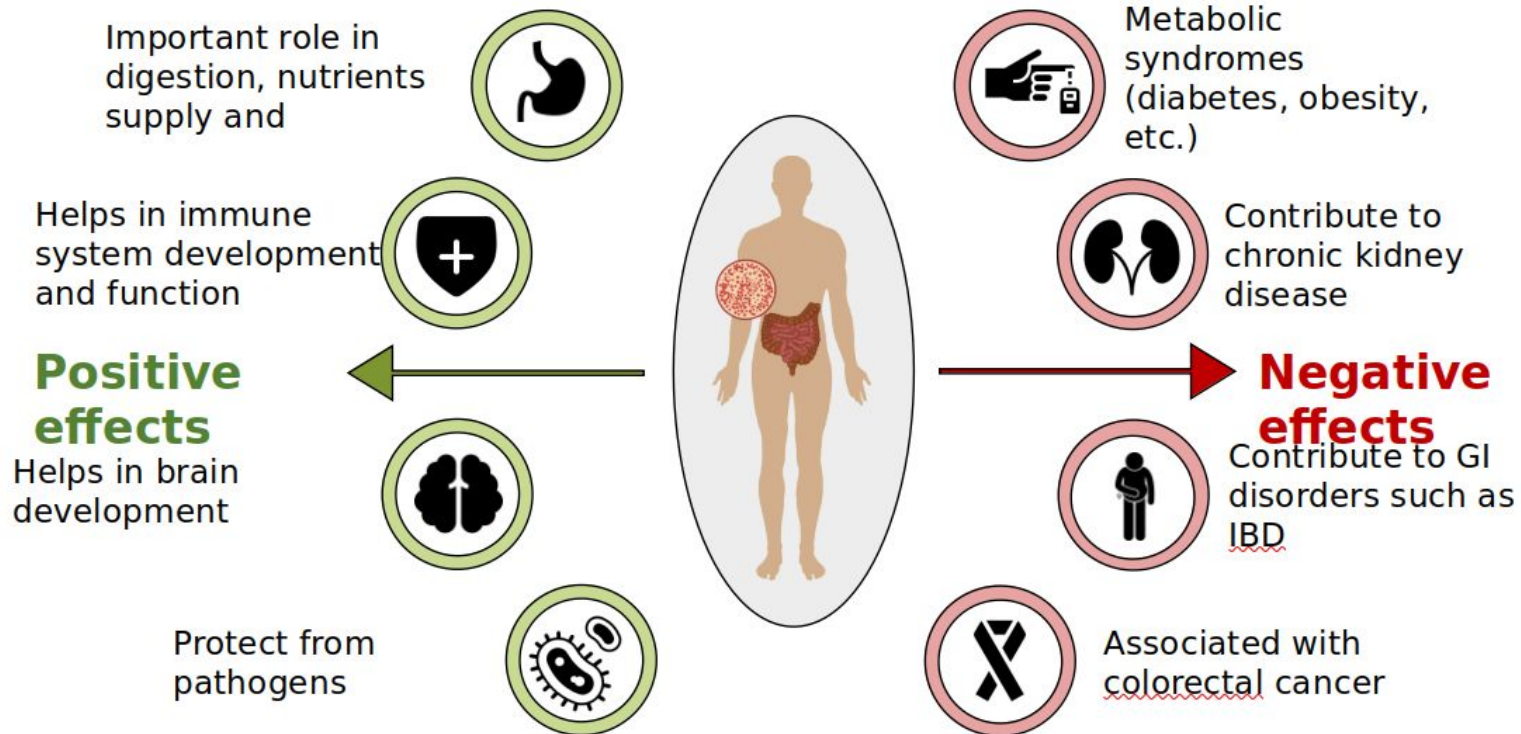
Micro (small) Biome (large community)

Including **bacteria**, archaea, viruses and fungus.

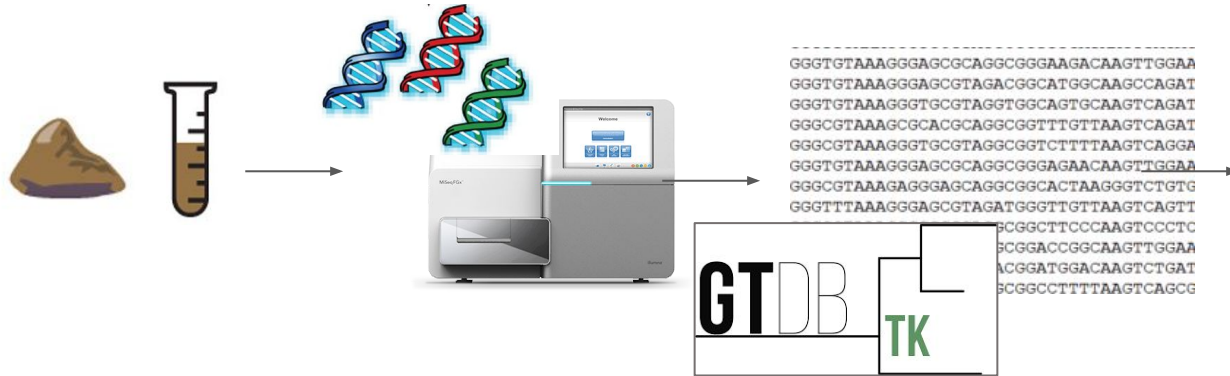
	Weight (kg)?	Number of cells?	Number of cell types?	Number of genes?
 Human	~70	$\sim 3 \times 10^{13}$	$\sim 10^0$	20,000
 Microbial	~0.2	$\sim 4 \times 10^{13}$	~1000	2,000,000



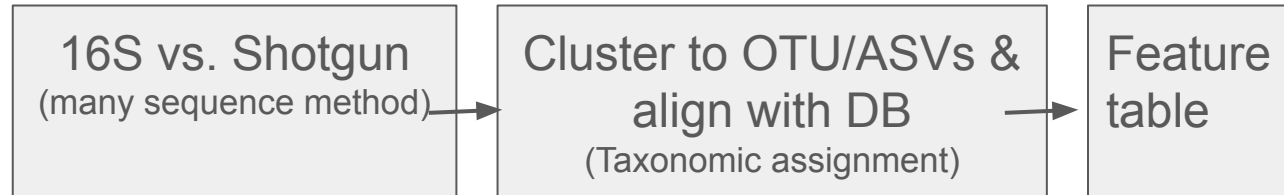
The gut microbiome's role in human health



Human gut microbiome - The data

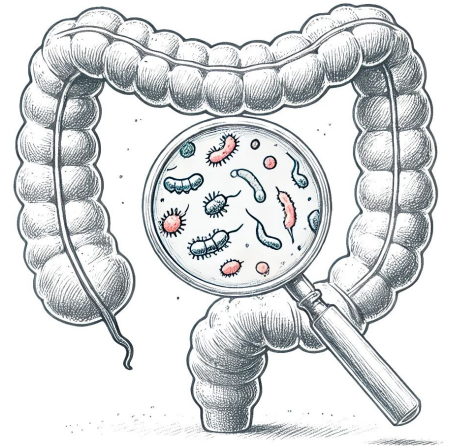


	Genus x	Genus y	...	Genus z
Sample 1	0.01	0.2	...	0.0008345
Sample 2 ...	0.00000034	0.142369	..	0.05
Sample n	0.12435	0	...	0



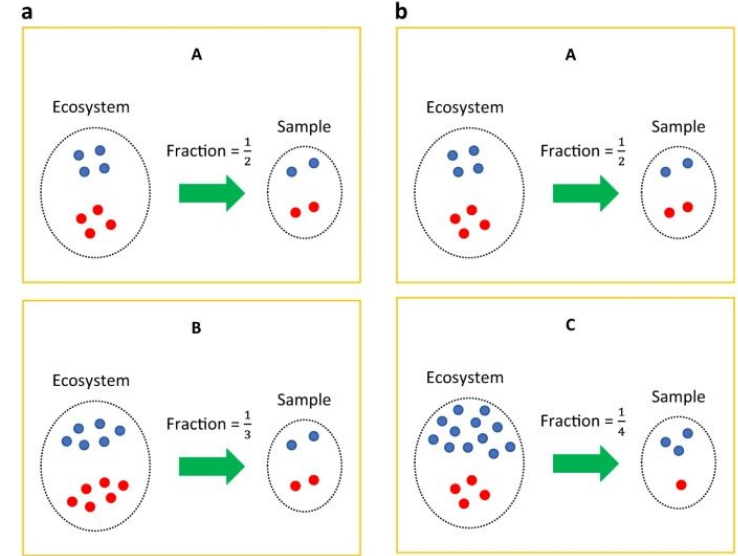
Human gut microbiome - The data

- DNA Sequences mapped to **taxonomic assignment** (Feature table, Tabular data)
 - *which (%) genus/ species in the sampled community*
- DNA Sequences mapped to **Functional annotation** (Feature table, Tabular data)
 - *which genes exists (potentially) in the sample*



Human gut microbiome - Challenges

- High dimensional
- Few samples
- Very noisy
- Strong covariates and batch effect
- Relative abundance only

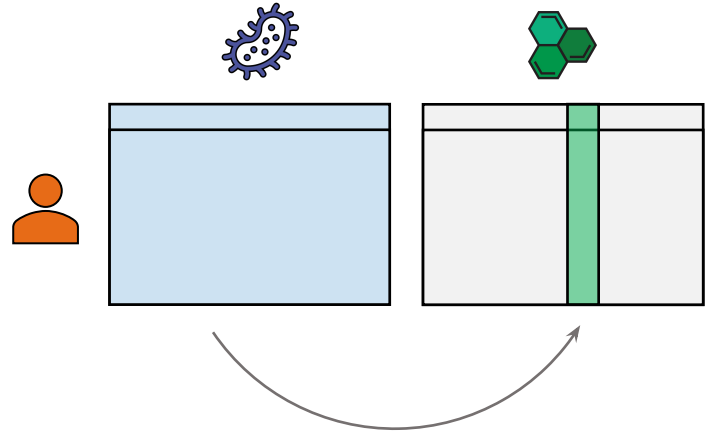


Can't cultivated most species in the lab Therefore, computation tools for analysing the hole community is really necessary!



The research question

Can we apply domain adaptation / transfer learning between hosts to predict metabolite level from the microbiome community?

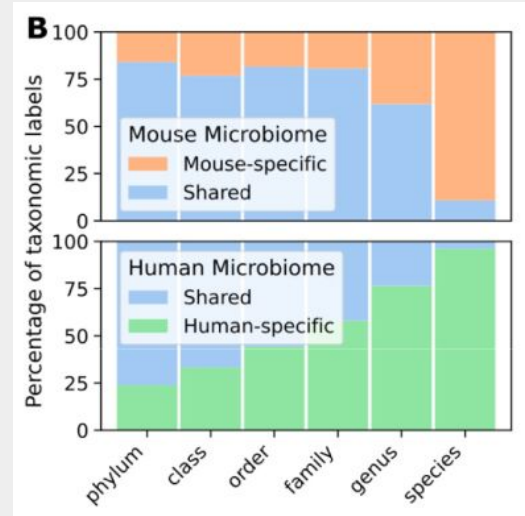


Background & motivation

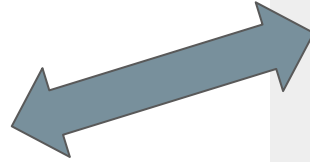
Disease markers obtained in a mouse model fail to carry over to humans.

Due experiment on mice are significant for the research community: allowing controlled experimental studies in a physiologically and genetically tractable system.

- *Mouse microbiome was largely unknown (<40% read-coverage).*
- *MGBC & CMMG contribute significantly to coverage (~80%).*
- Human and mouse taxa barely overlap (3-10% species level, 60% genus level)



Background & motivation



Improve our ability to utilize mouse microbiome experiments for human-relevant diagnosis!

by providing analysis tools to translate the finding from mice to humans!

The research question

Collecting data for Microbiome-disease across hosts is hard

→ Let's look at Metabolome as the phenotype.

As:

- One of the microbiome crucial and clinical metabolites.
- Causal relationship between metabolites and medical conditions or benefits.
- Microbiome-associated metabolites may be used to mitigate the adverse effects of microbiome.



Did you know?

Microbiome feces transplant (FMT) is already been shown to help in practice for specific disease! (*C. difficile* infection)

The research question

Mice dataset contain only a few subjects

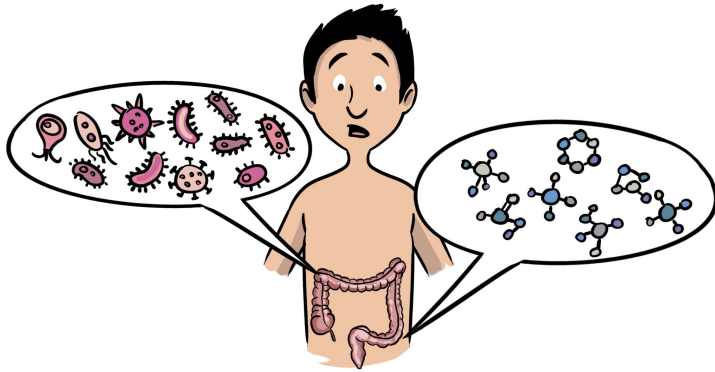
→ Let's do a **meta-analysis**

- We will need to verify we have enough **overlap** in
Our features & target
- Careful for **cage effect** & **batch effect**

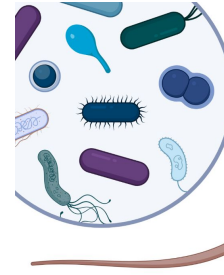


Related works

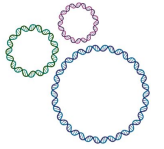
Microbiome-metabolome association in humans



Human-Mice translation



**Mouse
Gastrointestinal
Bacterial
Catalogue**



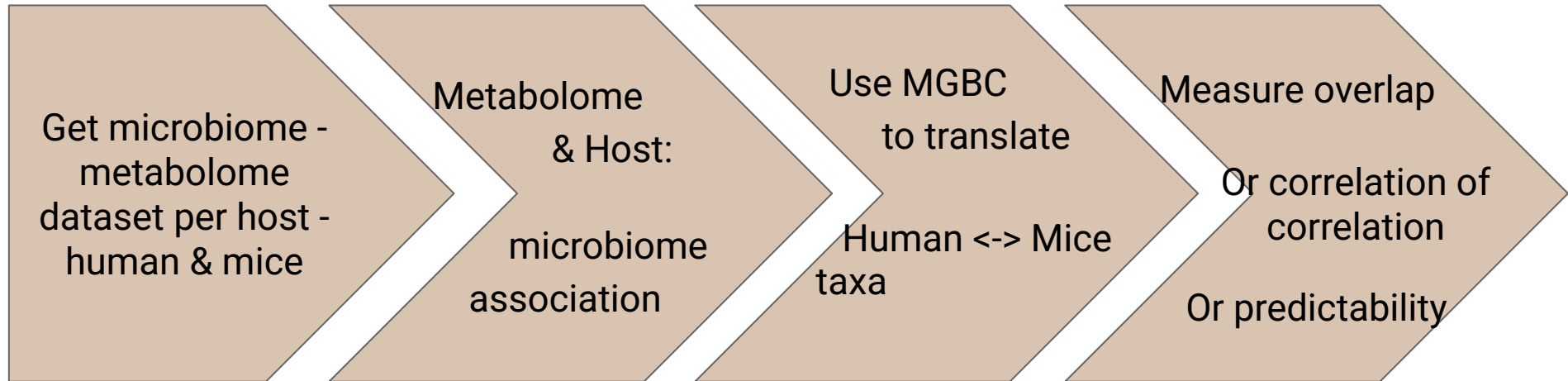
“A meta-analysis study of the robustness and universality of gut microbiome-metabolome associations”

Current work overview

- Collect meta-analysis cohort (literature overview)
- Data preparation and processing of microbiome & metabolome
- Data analysis:
 - Analysis correlation (note FDR corrections)
 - Predict metabolite levels from microbiome community

Current work overview

- Collect meta-analysis cohort (literature overview)
- Data preparation and processing of microbiome & metabolome
- Data analysis:
 - Analysis correlation (note FDR corrections)
 - Predict metabolite levels from microbiome community

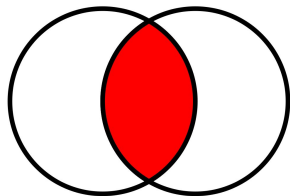


2 Approaches for Univariate microbiome-metabolome association analysis

Discrete

Per Metabolite

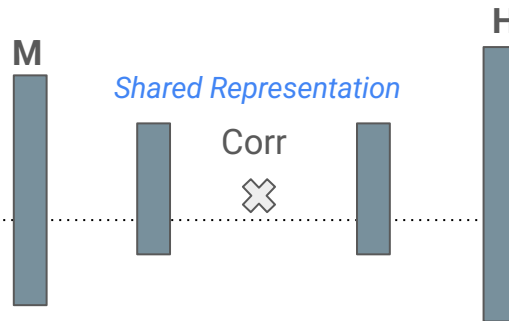
- Per host:
Find a set of significant bacteria.
- Translate mice bacteria to human.
- How much this set overall?
Or, What is the Recall?



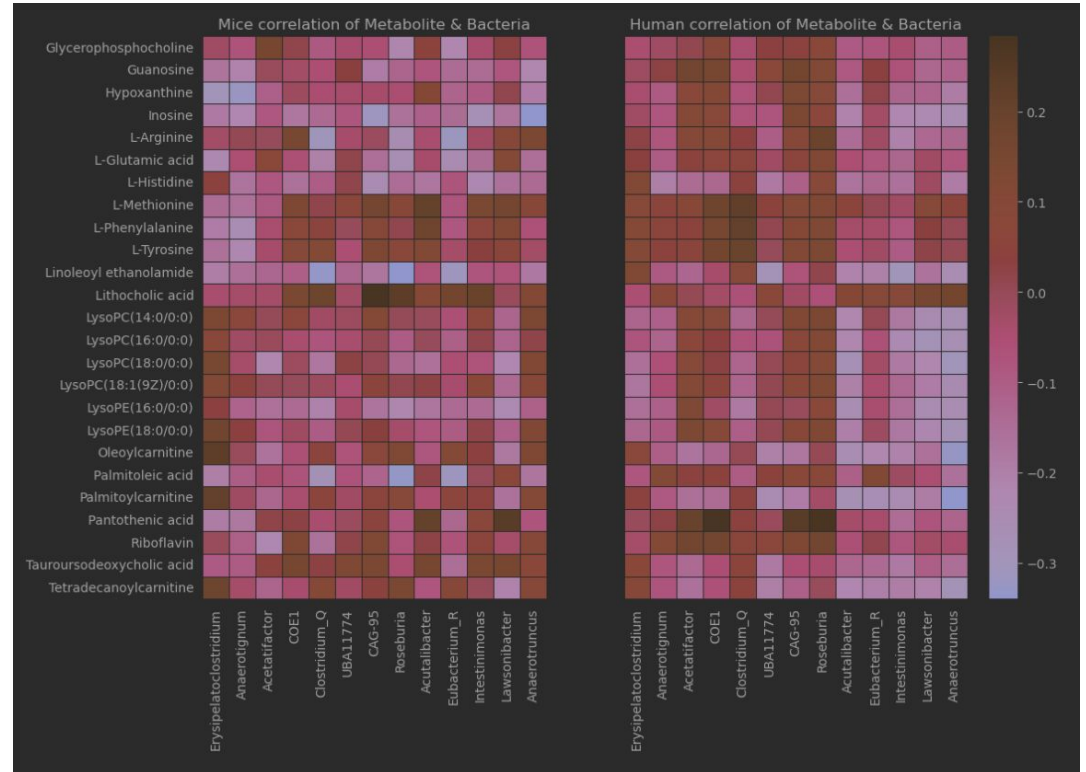
$$\frac{A \cap B}{A \cup B}$$

Continuous

- Per host:
Compute the microbiome coefficient vector to this metabolite.
- Align (map) the entries of this vector
- How much this vector correlates?



2 Approaches for Univariate microbiome-metabolome association analysis



Predict metabolite levels using the microbiome

- Train a Random Forest regression to predict metabolite levels based on genera relative abundance (evaluate using leave-one-subject-out cross validation)
- Define well-predicted metabolites:
Spearman correlation coefficient between the actual and predict metabolite level > 0.3 .*
- *Is well-predictability in mice indicates well-predictability in human?*
- *Does universal-well-predicted metabolites in human are indeed more predictable in mice?*

*same as in the article: "A meta-analysis study of the robustness and universality of gut microbiome-metabolome associations"

Availability	Verified it's a fecal metabolites	Metabolomics approach	Metagenomics approach	No. samples	No. subjects	Chorot description	Ref article
Not public. Appone request	V	Targeted, NMR	16S v3-v4	27-33	27-33	Mice prone genetically to dementia, with normal or fatty diet	Microbiome-metabolome signatures in mice genetically prone to develop dementia, fed a normal or fatty diet
Not public. Appone request	V	Targeted, GC-MS	16S	12	12	HFD-induced obese mice vs. control.	Gut Microbiome and Metabolome Profiles Associated with High-Fat Diet in Mice
Publicly available (raw metabolome only)	V	untargeted LC-MS/MS	16S	56 (control should be around 30)	27 overall (13 control)	Sleep disruption in mice.	Repeated sleep disruption in mice leads to persistent shifts in the fecal microbiome and metabolome

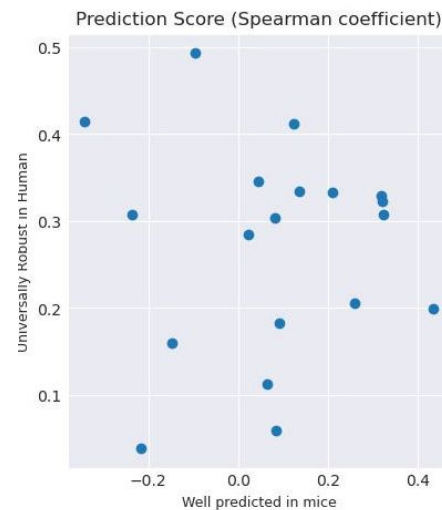
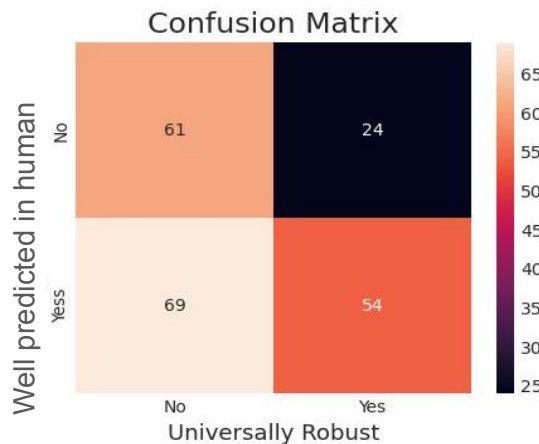
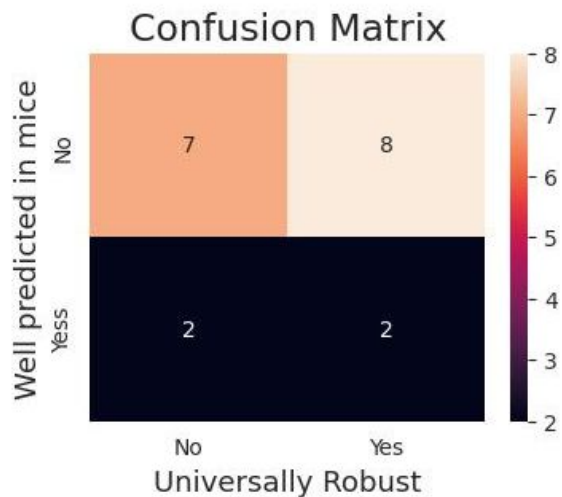
Results

- Low overlap between compared datasets. & Hard to find a signal.

Comparing the predictability of metabolites in mice - OSA and Human metabolites in Efrat's dataset:

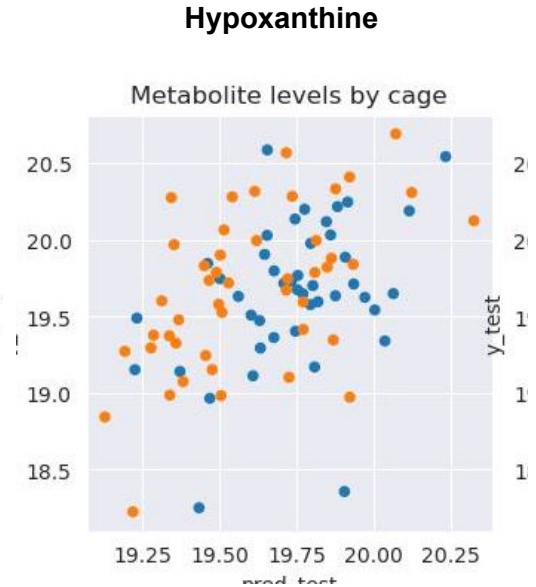
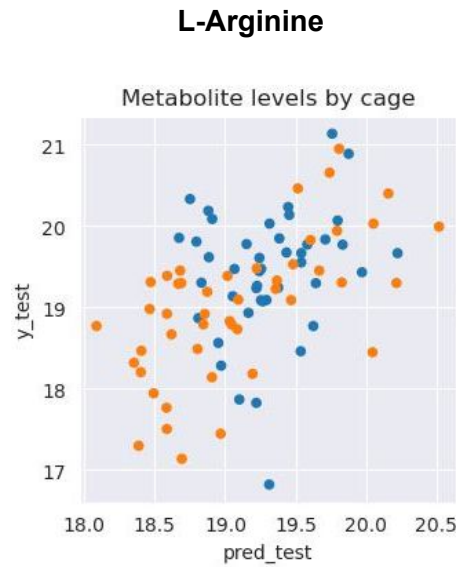
There are 19 HMDB in intersection between my OSA dataset metabolites and Human metabolites in Efrat's dataset.

Our of 61 metabolites in OSA and 273 in the Human meta-analysis



Results

- Very strong cage effect in mice.
- Taking the cage into account, Performance drop significantly

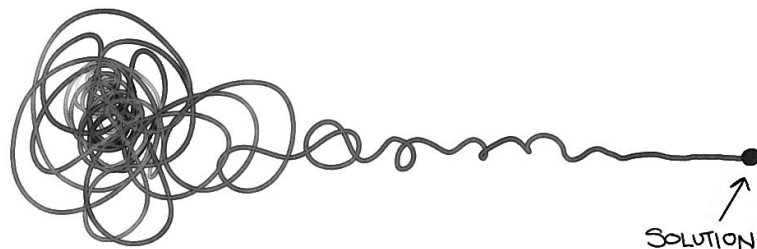


Change of plan

- Let's try to tackle the problem in different hosts.
- We have a large dataset of dogs (~900 subjects) and humans (~1k).
- 2 dataset only instead of meta-analysis.

Less conversions & batch effect.

lifelines



"Research is not a straight line"

Current ideas

To be continued...
Questions?

