

---

# LITERATURE REPORT

**nature biotechnology**

## Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine

Published: 19 January 2023

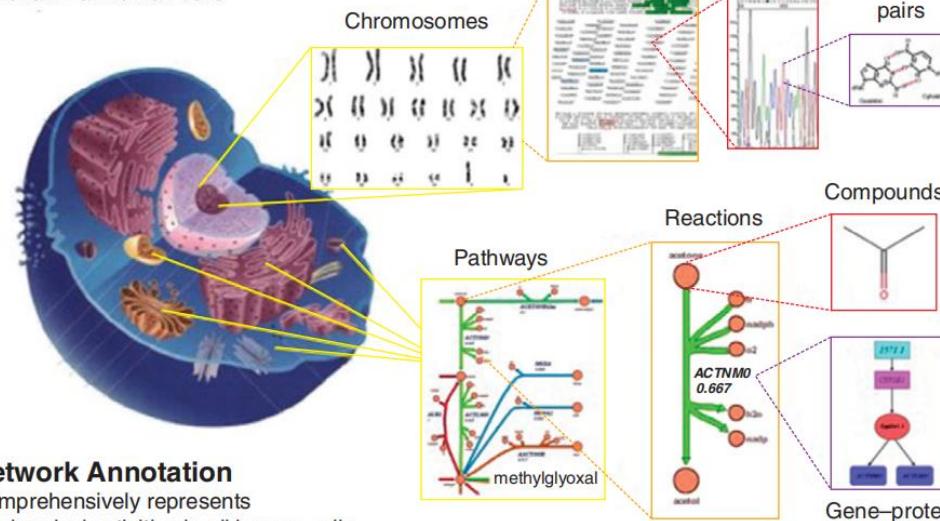
汇报人: 程科林  
日期: 2023. 04. 14

# Background

## Genome-scale metabolic reconstruction

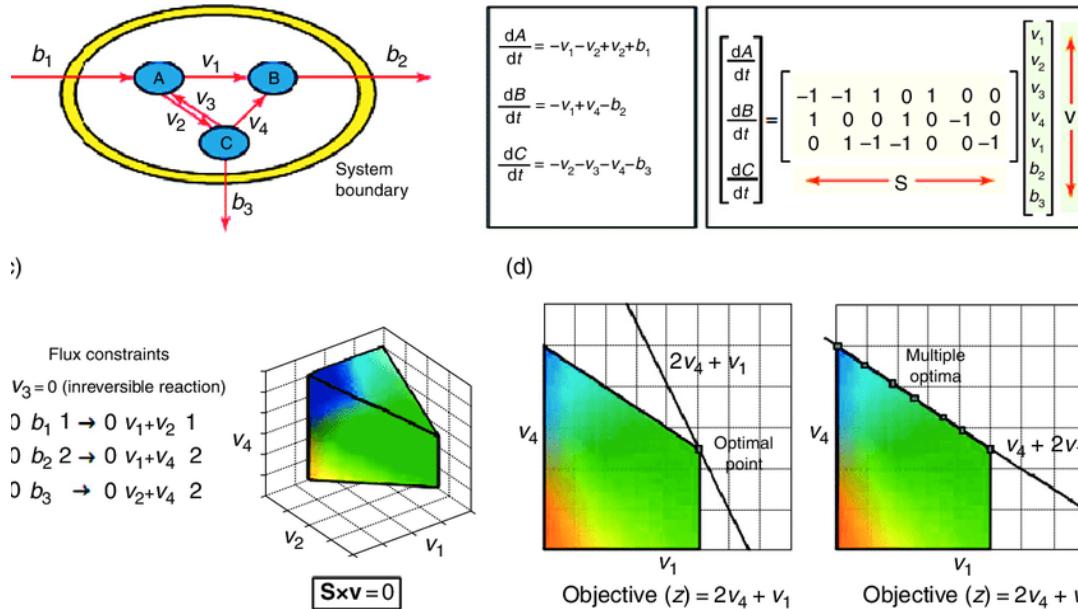
### Genome Assembly

Comprehensively represents genetic material in all human cells

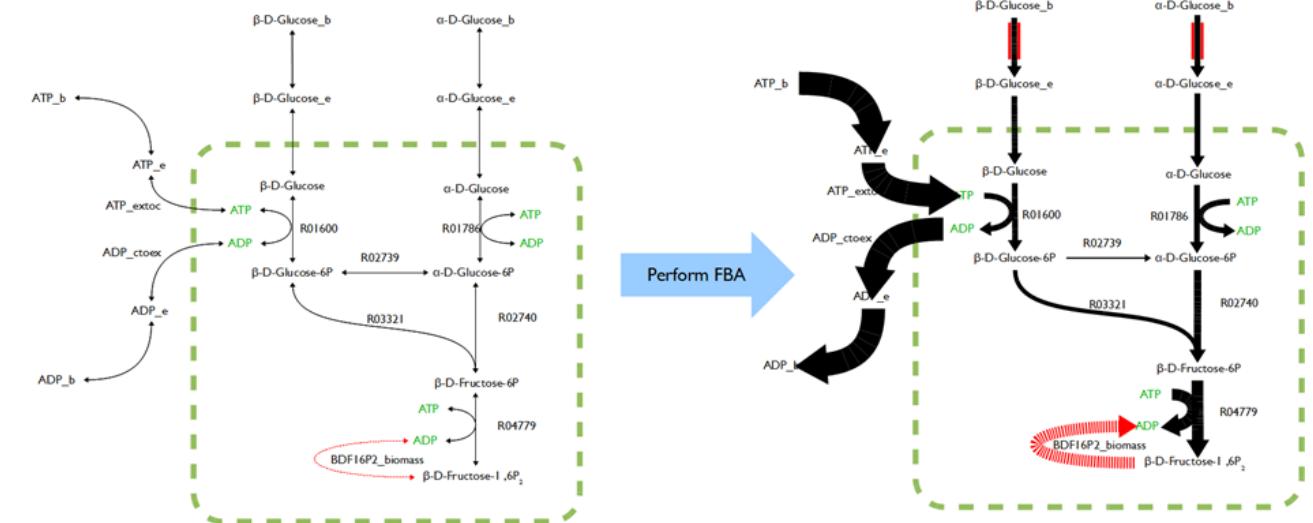
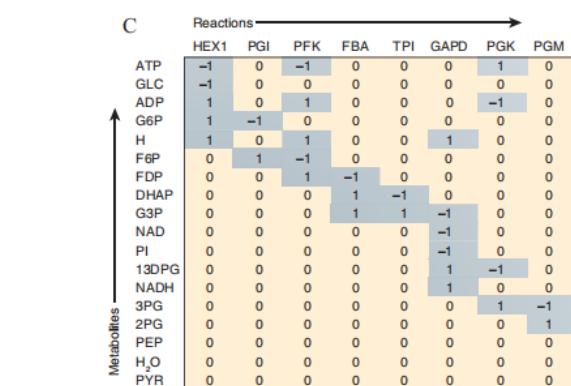
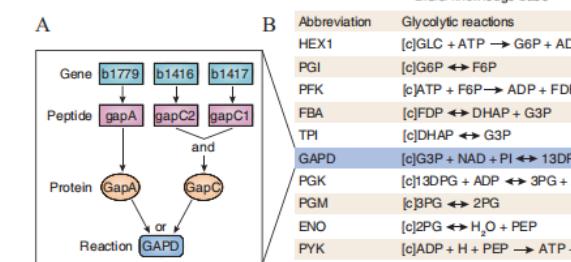


### Network Annotation

Comprehensively represents biochemical activities in all human cells



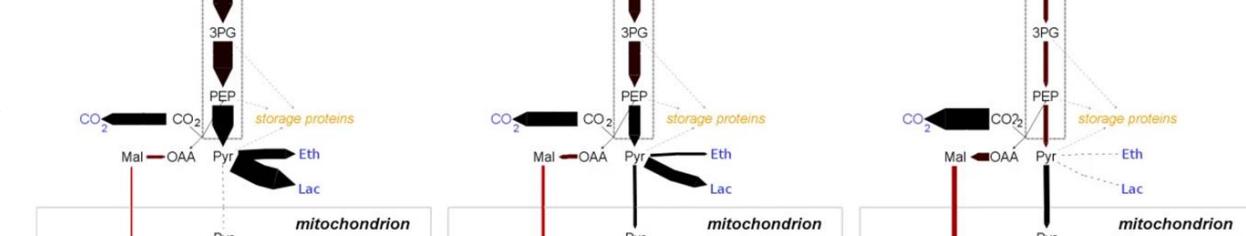
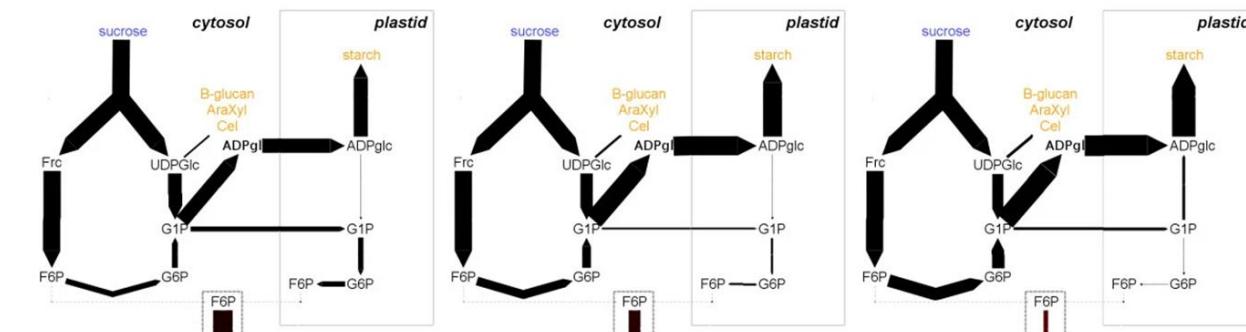
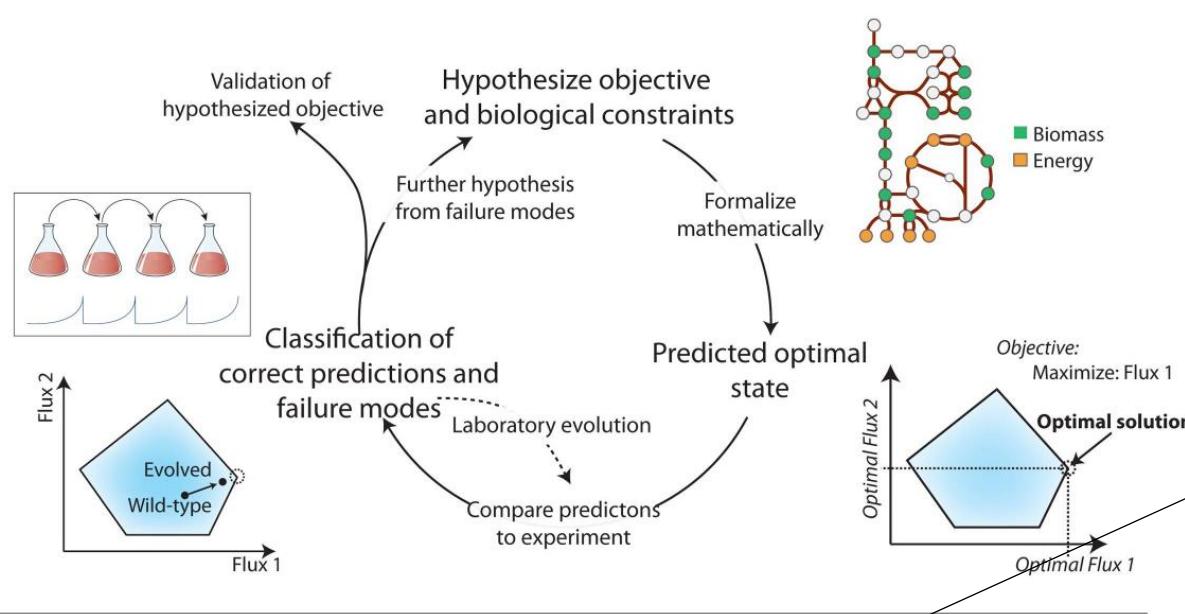
BIGG knowledge base



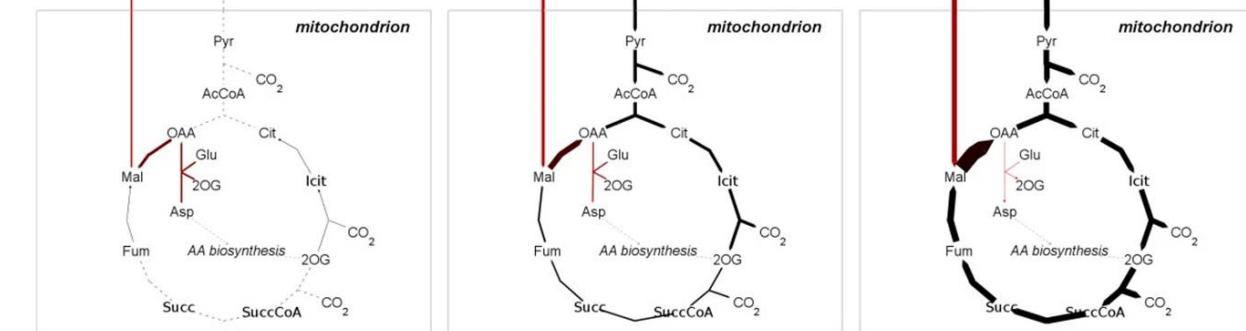
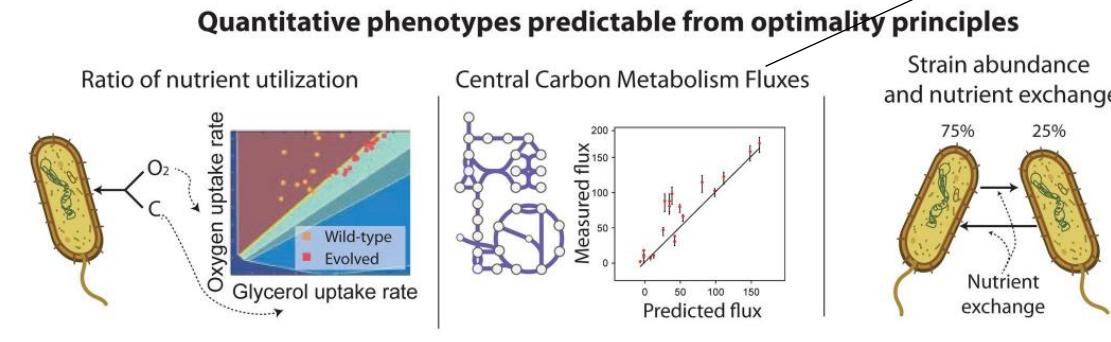
# Background

# Genome-scale metabolic reconstruction

A.



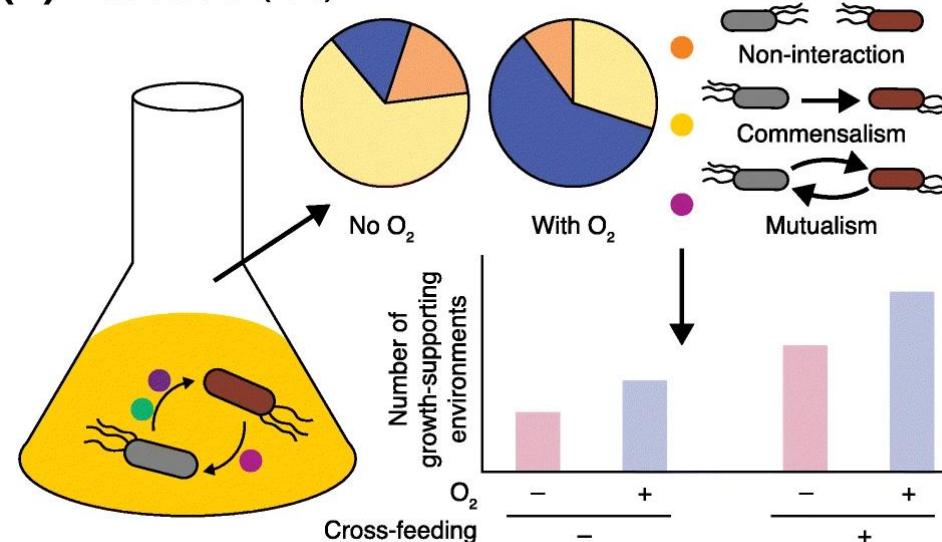
B.



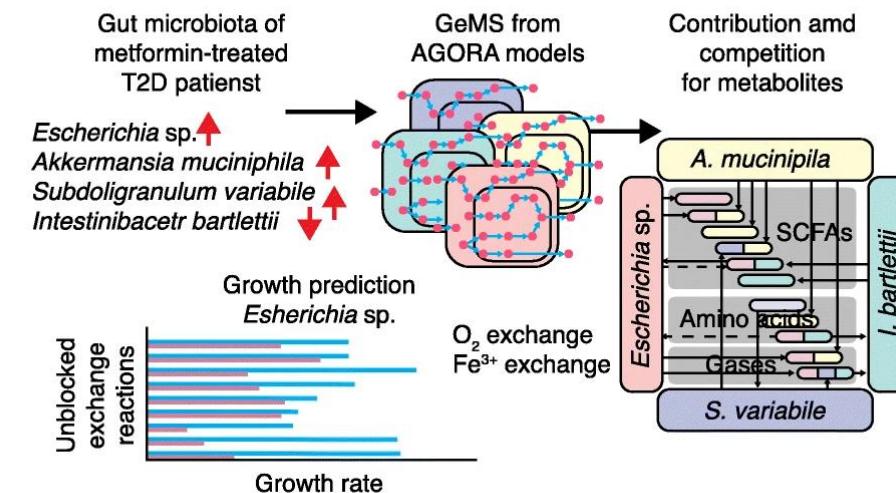
# Background

## Genome-scale metabolic reconstruction

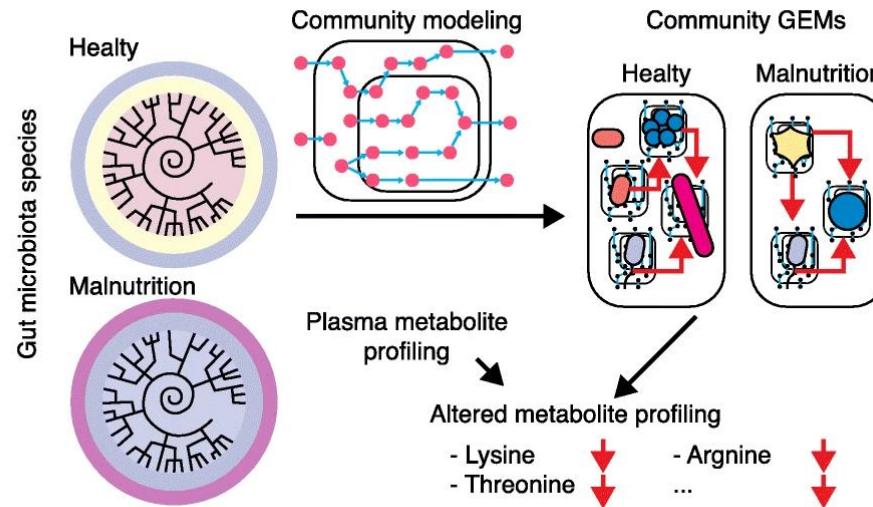
(a) Pacheco et al. (2019)



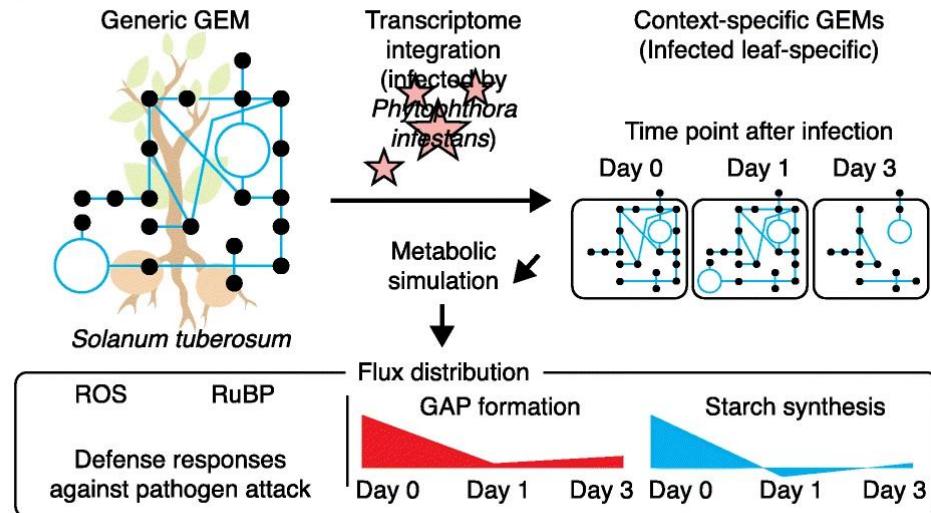
(b) Rosarioet al. (2018)



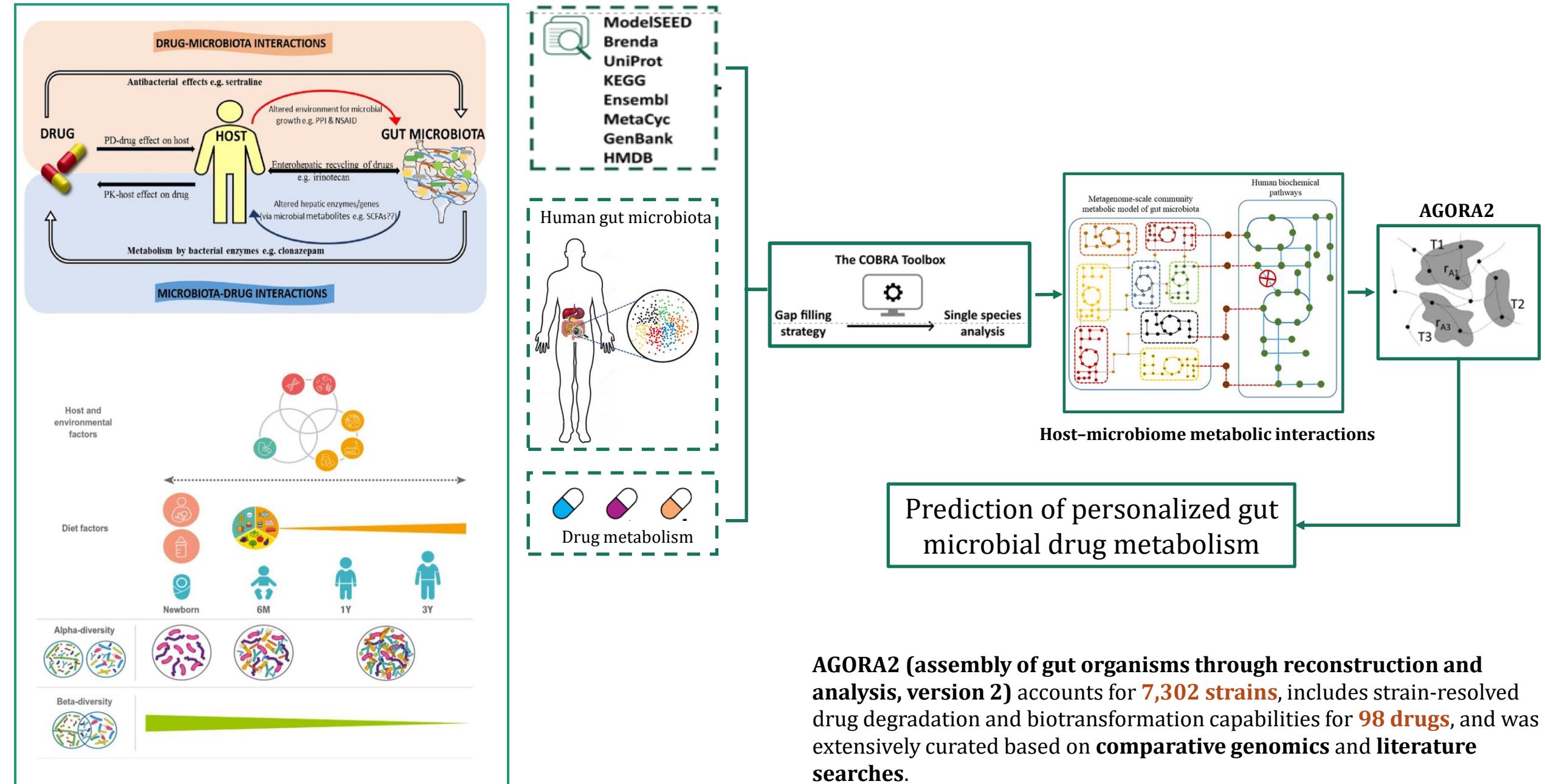
(c) Kumar et al. (2018)



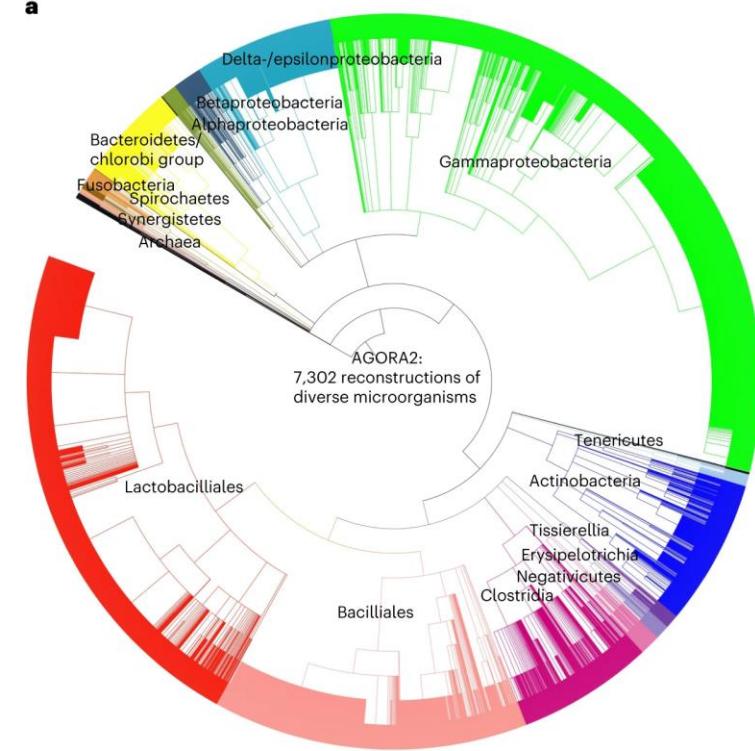
(d) Botero et al. (2018)



# Main



# Results 1 : Data-driven reconstruction of diverse human microorganisms



**b**

Taxonomic coverage	Feature	No. in AGORA2 (AGORA 1.03)
Strain source	Reconstructed strains	7,302 (818)
	Phyla	25 (15)
	Classes	39 (27)
	Order	81 (51)
	Families	163 (106)
	Genera	483 (258)
	Species (characterized)	1,235 (641)
	Species (uncharacterized)	503 (52)
	Human-associated	7,175
	Mouse-associated	127
PubSEED/KBase database	4,411	
Literature research	1,181 (+817)	
Personal communication	132	
Forster et al. (2019) resource	761	

**c**

Feature	Refined	Draft
Reactions	$1,723.13 \pm 817.14$	$1,306.43 \pm 368.19$
Metabolites	$1,538.67 \pm 685.08$	$1,238.41 \pm 327.17$
Genes	$906.71 \pm 336.49$	$944.79 \pm 396.89$
Compartments	c, e, p	c, e
Growth on UM (aerobic)	7,302 (100%)	7,302 (100%)
Growth on UM (anaerobic)	7,302 (100%)	5,629 (77%)
Growth on CM (aerobic)	7,302 (100%)	1,971 (27%)
Growth on CM (anaerobic)	7,294 (100%)	1,736 (24%)
ATP flux on CM (aerobic)	$75.74 \pm 37.55$	$872.2 \pm 291.92$
ATP flux on CM (anaerobic)	$48.82 \pm 27.61$	$863.37 \pm 300.09$

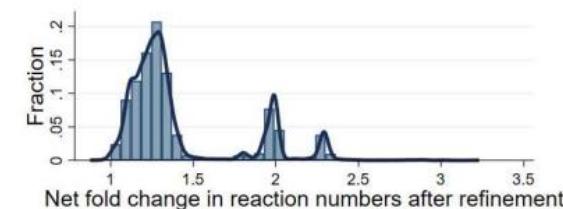
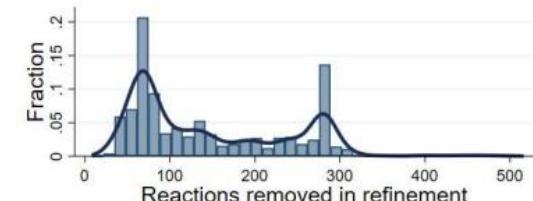
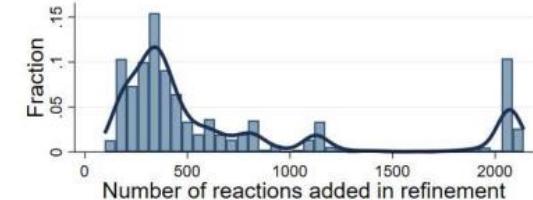
  

**d**

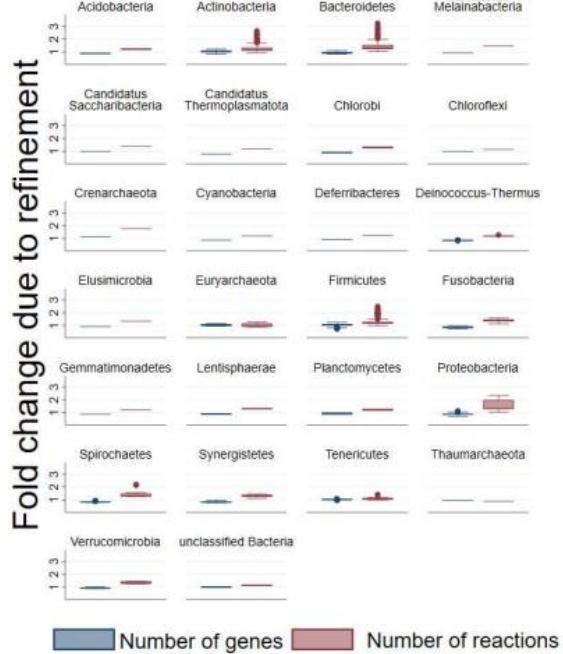
Feature	No. strains with available data	Percentage of strains agreeing with data	
		Refined	Draft
Bile acid metabolism	228	100.0%	0.0%
Carbon sources	6,838	99.99%	4.14%
Drug metabolism	5,373	99.72%	0.00%
Fermentation pathways	6,210	99.94%	0.71%
Growth on defined media	4,667	99.70%	0.84%
Metabolite secretion	4,117	96.55%	9.36%
Metabolite uptake	3,997	99.40%	13.08%
Putrefaction pathways	393	99.75%	1.78%

## Supplemental Figures

Histograms for change in reaction numbers due to refinement across reconstructions



Box plots across phyla for change in model size due to refinement



# Results 1 : Data-driven reconstruction of diverse human microorganisms

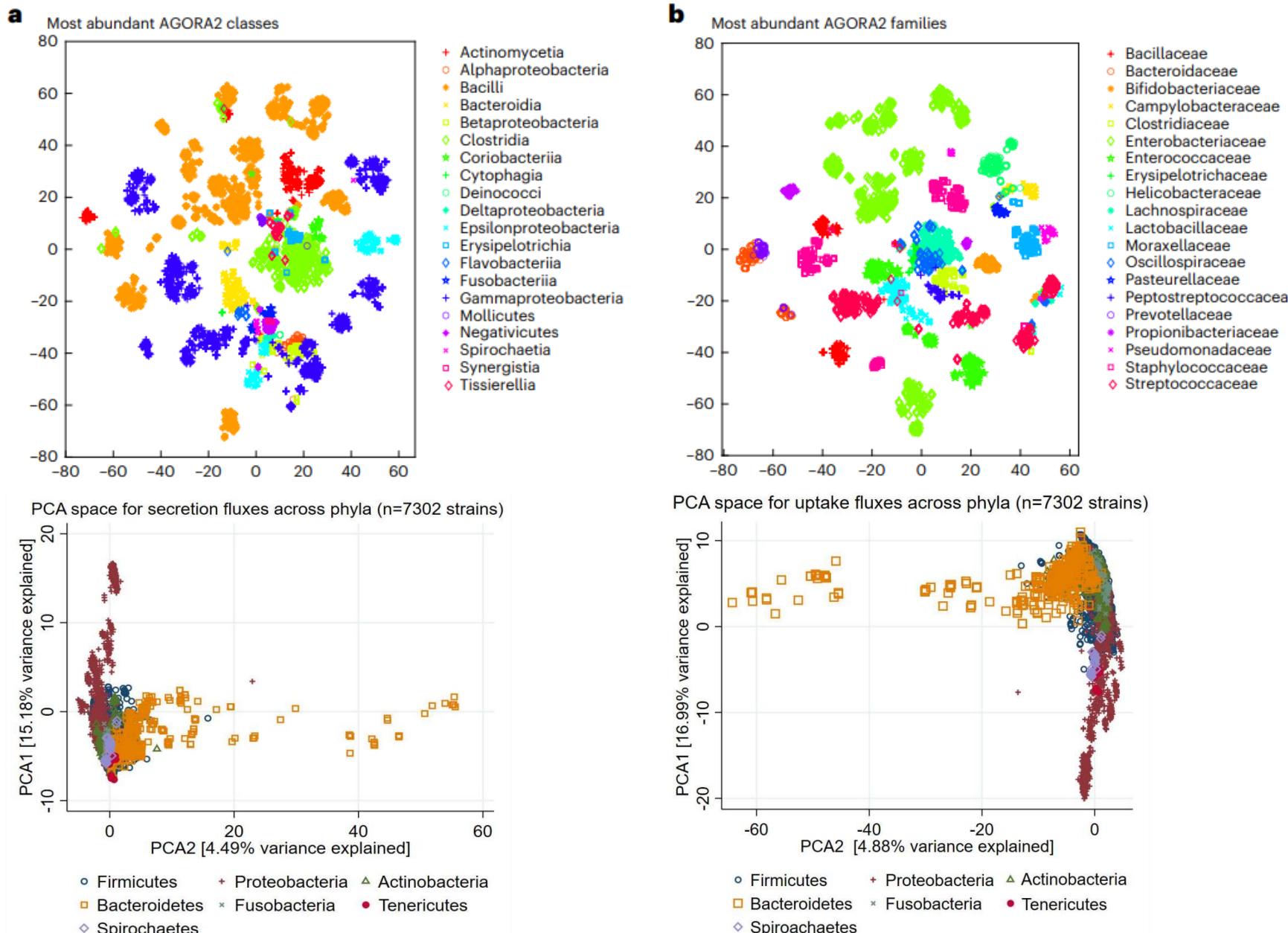


Fig. 2: Taxonomically related strains are similar in their AGORA2 reconstruction content.

# Results 1 : Data-driven reconstruction of diverse human microorganisms

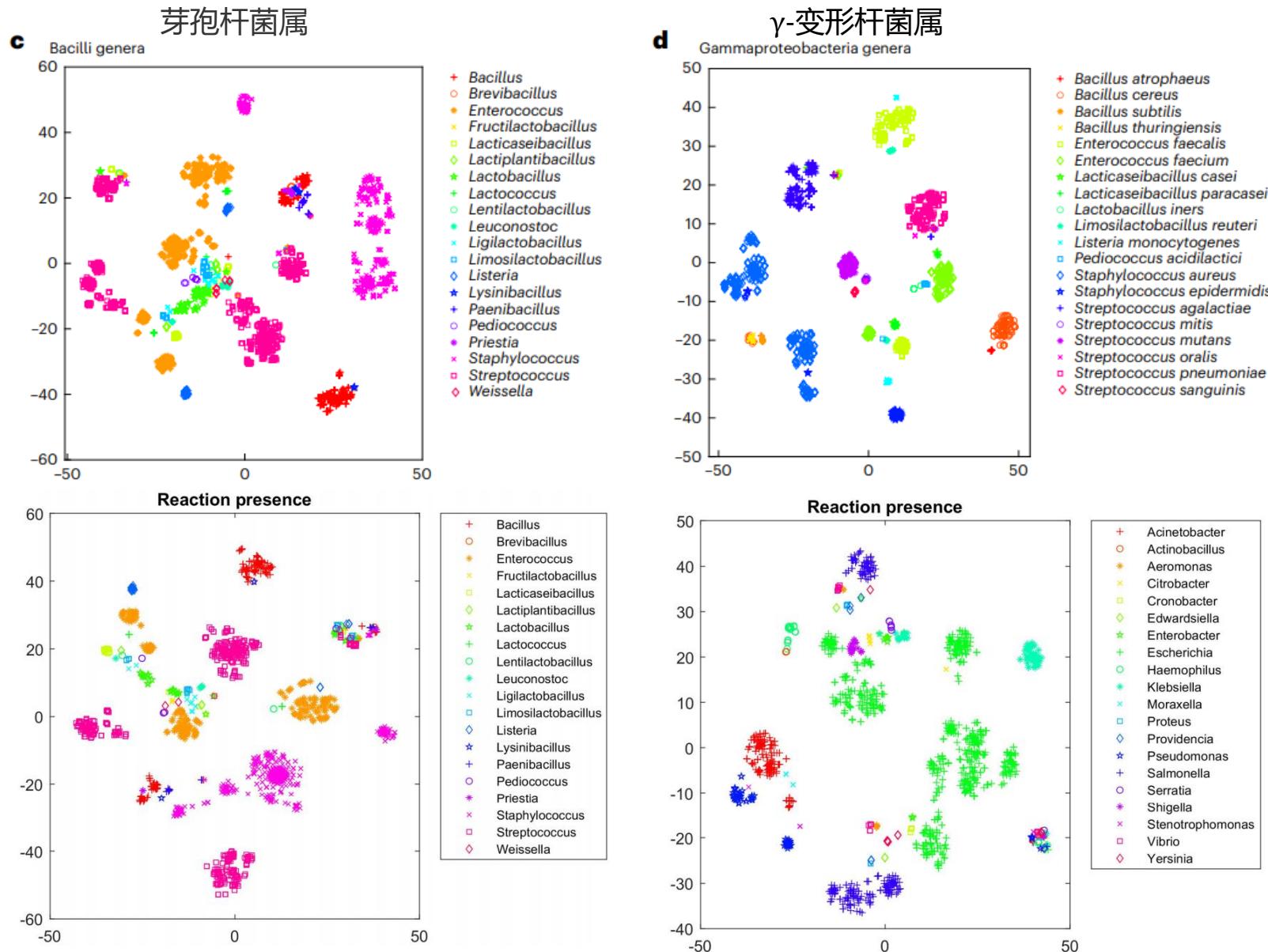
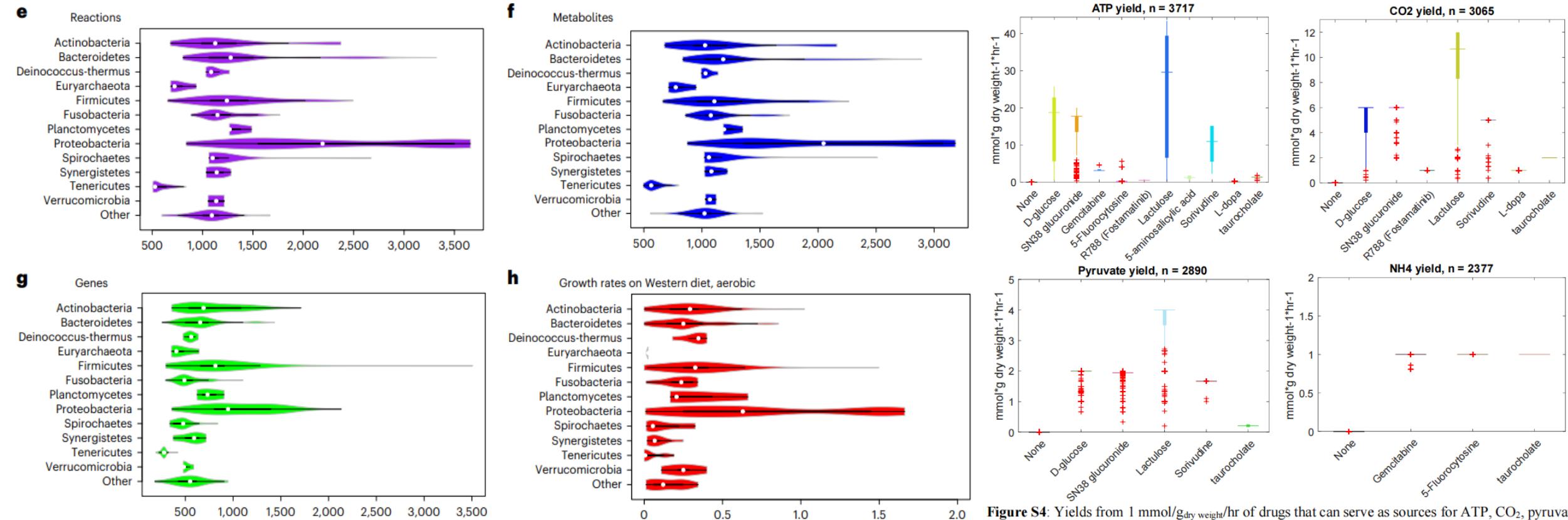


Fig. 2: Taxonomically related strains are similar in their AGORA2 reconstruction content.

# Results 1 : Data-driven reconstruction of diverse human microorganisms



e-h, Features of all AGORA2 reconstructions across phyla(growth rate in  $\text{hr}^{-1}$  on aerobic Western diet.)

**Figure S4:** Yields from 1  $\text{mmol/g}_{\text{dry weight}}/\text{hr}$  of drugs that can serve as sources for ATP,  $\text{CO}_2$ , pyruvate, and  $\text{NH}_4$  production. Shown are all microbes that could use at least one drug to produce the respective source. Flux with glucose and with no compound added are shown as controls. One example drug per enzyme was tested.

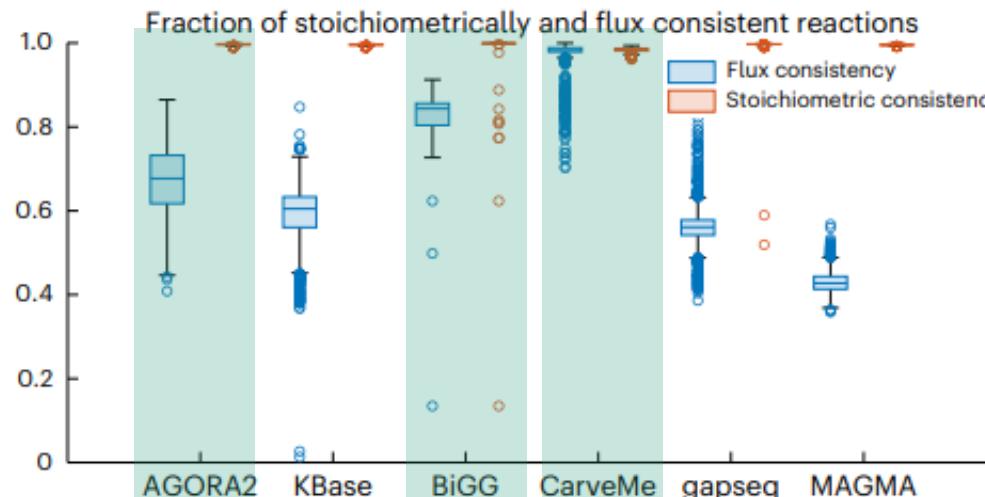
## Results 1 : Data-driven reconstruction of diverse human microorganisms

---

- 作者生成了AGORA2，一个包含7,302个人类微生物基因组水平代谢重建的资源，覆盖了25个门、39个纲、81个目、163个科、483个属和1,738个种。
- AGORA2反映了重建微生物的分类多样性和代谢特征，因为它们根据反应覆盖率按纲和科进行聚类，并在门之间在重建大小、生长速率和代谢物消耗和分泌潜力方面有所不同。

## Results 2 : AGORA2 is predictive against three independent datasets

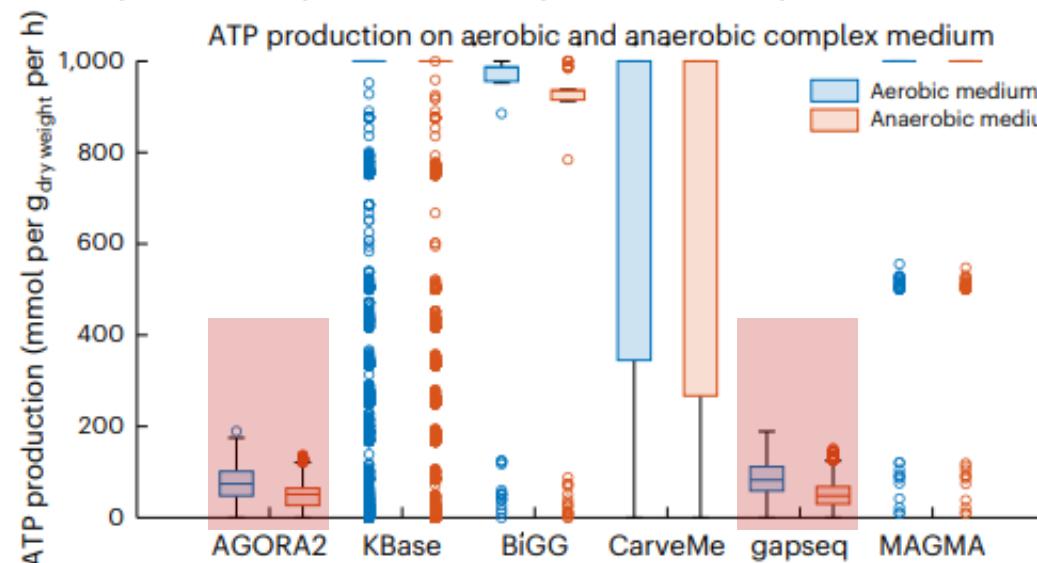
**a** Stoichiometric and flux consistency computed for the compared resources



**c** Overview of reconstruction features for the compared resources

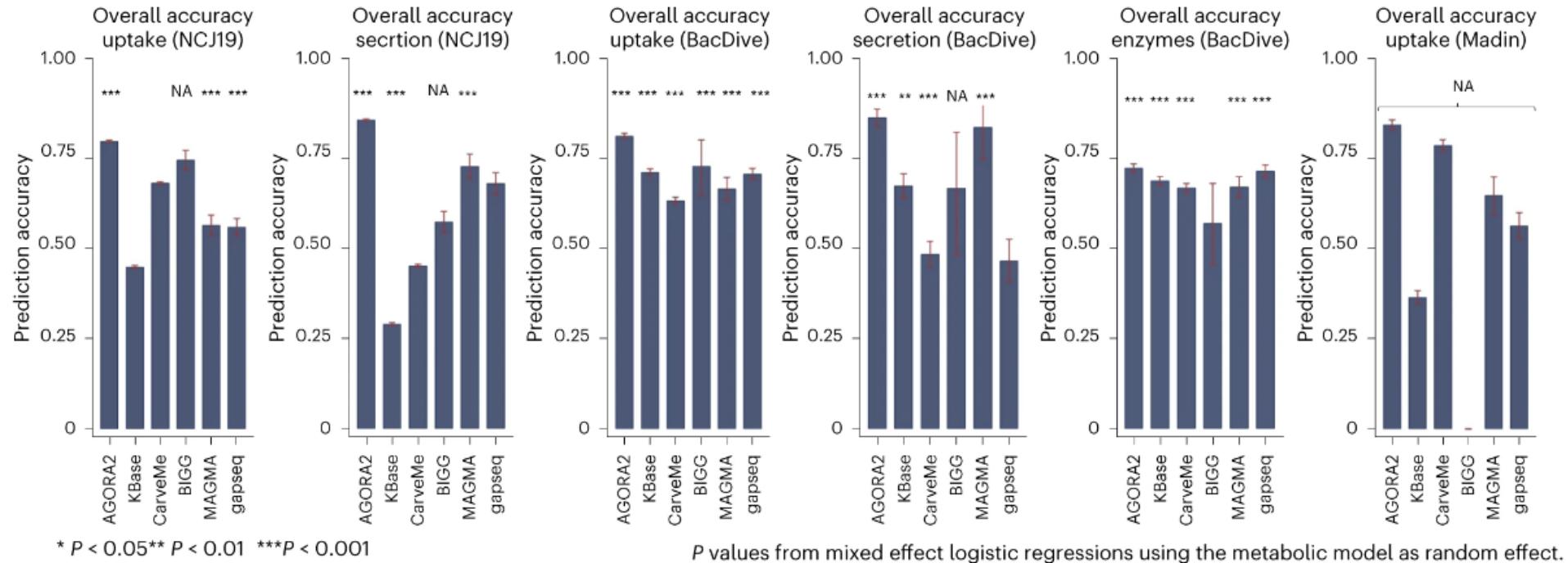
	AGORA2	KBase	BiGG	CarveMe	gapseq	MAGMA
No. of reconstructions	7,302	7,302	72	7,279	8,075	1,333
Average no. of reactions	1,723.12	1,305.10	2,348.47	1,834.30	1,895.03	1,461.73
Average no. of metabolites	1,538.67	1,238.41	1,716.21	1,231.92	1,693.58	1,450.44
Average no. of genes	906.71	944.79	1,188.50	955.83	817.16	0.00
Average no. of stoichiometrically consistent reactions	1.00	1.00	0.96	0.98	1.00	1.00
Average no. of flux consistent reactions	0.68	0.60	0.82	0.98	0.56	0.43
ATP yield on aerobic CM	75.74	871.83	778.74	709.27	85.60	984.67
ATP yield on anaerobic CM	48.81	862.67	740.36	684.66	49.94	984.64

**b** ATP yields on complex medium computed for the compared resources



# Results 2 : AGORA2 is predictive against three independent datasets

**e** Accuracies in qualitative prediction of metabolite uptake and secretion of the various resources on the three experimental datasets (NJC19, BacDive, Madin)



**f** Comparison of qualitative prediction accuracies per model of the various resources on the three experimental datasets (NJC19, BacDive, Madin) with AGORA2

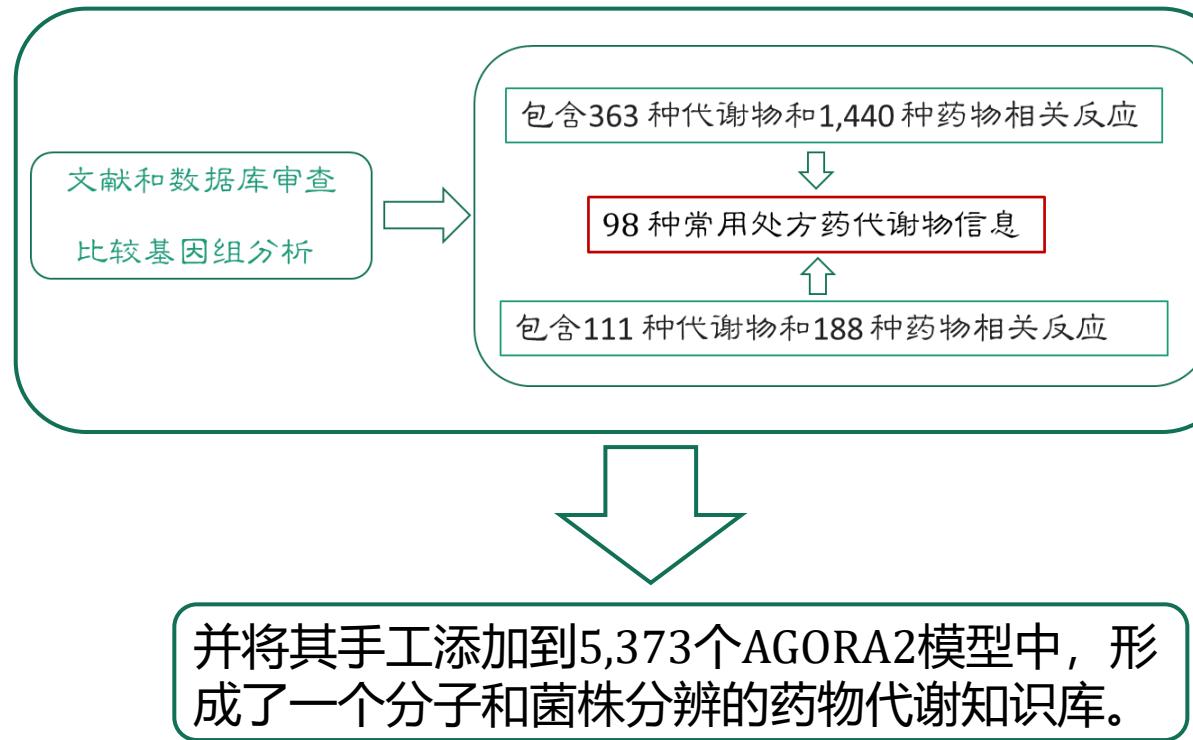
	AGORA2 versus KBase		AGORA2 versus CarveMe		AGORA2 versus BiGG		AGORA2 vs MAGMA		AGORA2 versus gapseq	
	No. of AGORA2 models with higher (+), equal (=) or worse (-) accuracy	P value sign rank test	No. of AGORA2 models with higher (+), equal (=) or worse (-) accuracy	P value sign rank test	No. of AGORA2 models with higher (+), equal (=) or worse (-) accuracy	P value sign rank test	No. of AGORA2 models with higher (+), equal (=) or worse (-) accuracy	P value sign rank test	No. of AGORA2 models with higher (+), equal (=) or worse (-) accuracy	P value sign rank test
NJC19 (Uptake)	+ 4178; = 935; -17	<1.0 × 10 <sup>-80</sup>	+ 2593; = 852; -1675	7.58 × 10 <sup>-35</sup>	+ 54; =0; -5	2.25 × 10 <sup>-07</sup>	+63; =16; -24	1.98 × 10 <sup>-05</sup>	+112; =37; -20	6.10 × 10 <sup>-15</sup>
NJC19 (Secretion)	+ 4763; = 218; -28	<1.0 × 10 <sup>-80</sup>	+4338; = 546; -114	<1.0 × 10 <sup>-80</sup>	+50; =9; 0	1.70 × 10 <sup>-11</sup>	+56; =45; -16	1.62 × 10 <sup>-06</sup>	+88; =78; -17	2.88 × 10 <sup>-12</sup>
BacDive (Uptake)	+344; =247; -84	2.30 × 10 <sup>-35</sup>	+525; =94; -55	6.09 × 10 <sup>-80</sup>	+5;=6; -1	0.063	+52; =27; -12	7.71 × 10 <sup>-07</sup>	+126; =77; -62	6.24 × 10 <sup>-06</sup>
BacDive (Secretion)	+170; =325; -40	6.22 × 10 <sup>-19</sup>	+263; =249; -22	1.37 × 10 <sup>-44</sup>	+4; =8; 0	0.125	+13; =64; -4	0.024	+116; =89; -7	1.82 × 10 <sup>-22</sup>
BacDive (Enzymes)	+263; =135; -143	2.94 × 10 <sup>-10</sup>	+290; =142; -108	1.82 × 10 <sup>-24</sup>	+6;=3; -2	0.313	+38; =17; -17	0.003	+79; =64; -56	0.104
Madin (Uptake)	+280;=46; -0	3.11 × 10 <sup>-53</sup>	+147; =96; -83	3.26 × 10 <sup>-06</sup>	NA	NA	+31;=7; -8	6.18 × 10 <sup>-4</sup>	+61; =22; -12	1.22 × 10 <sup>-09</sup>

## Results 2 : AGORA2 is predictive against three independent datasets

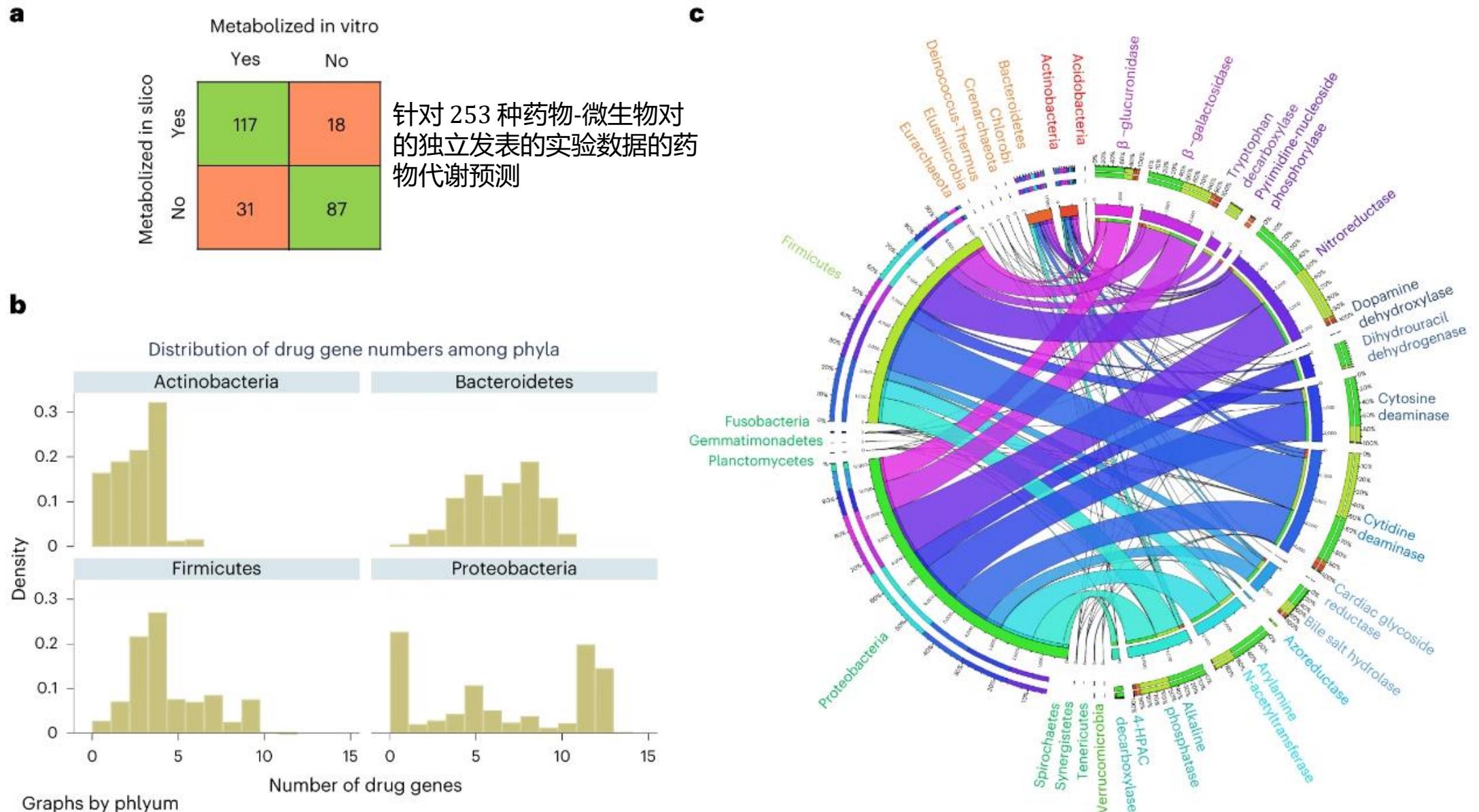
---

- AGORA2 重建很好地捕获了各个生物体的已知特征，超过了其他半自动生成的重建，可与手动策划的重建相媲美。
- 作者证明了AGORA2可以捕捉人类微生物的物种特异性和菌株特异性的代谢特征，以及它们在不同环境条件下代谢不同底物的潜力。

## Results 3 : Microbial drug metabolism guided by genome and bibliome

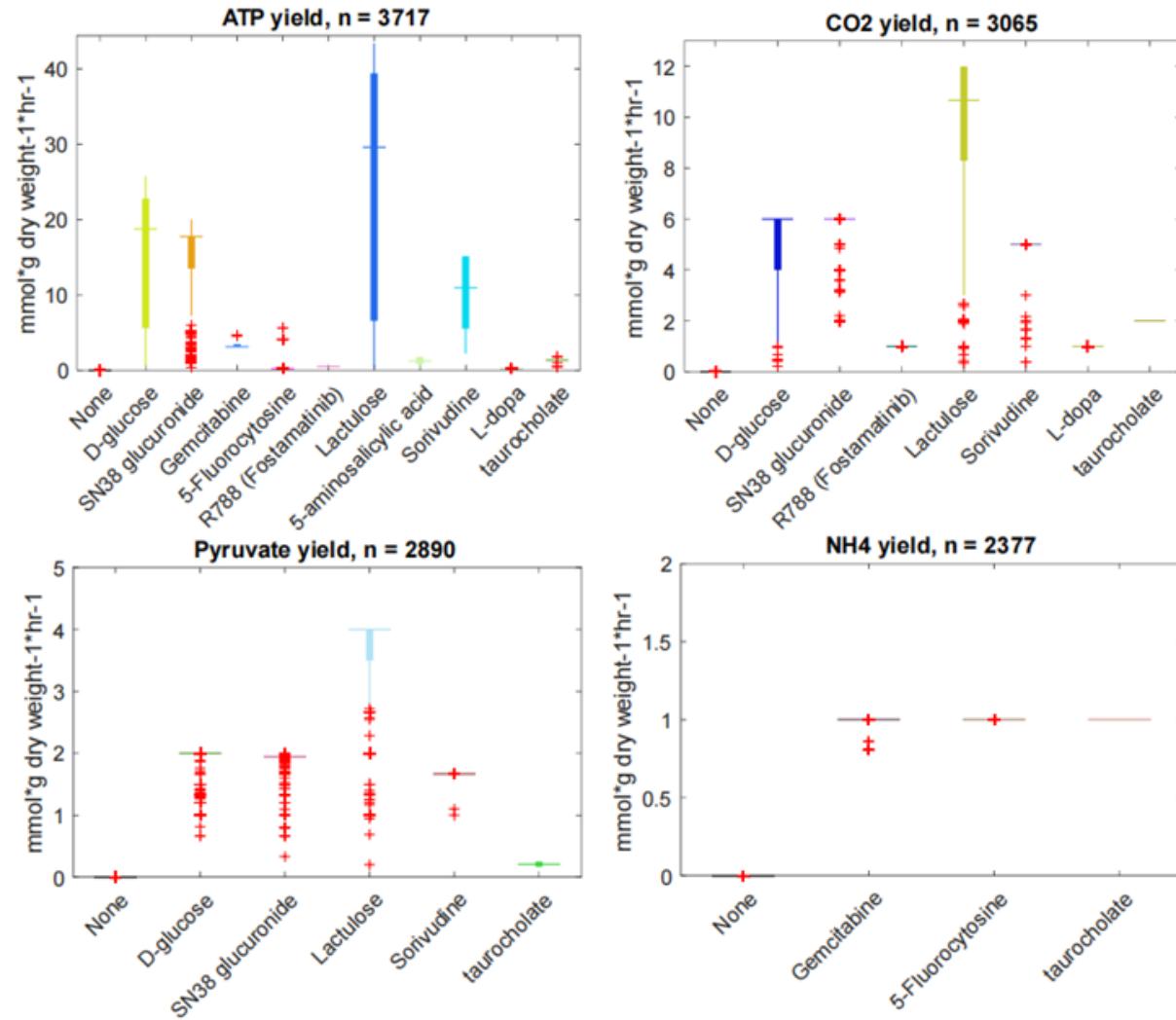
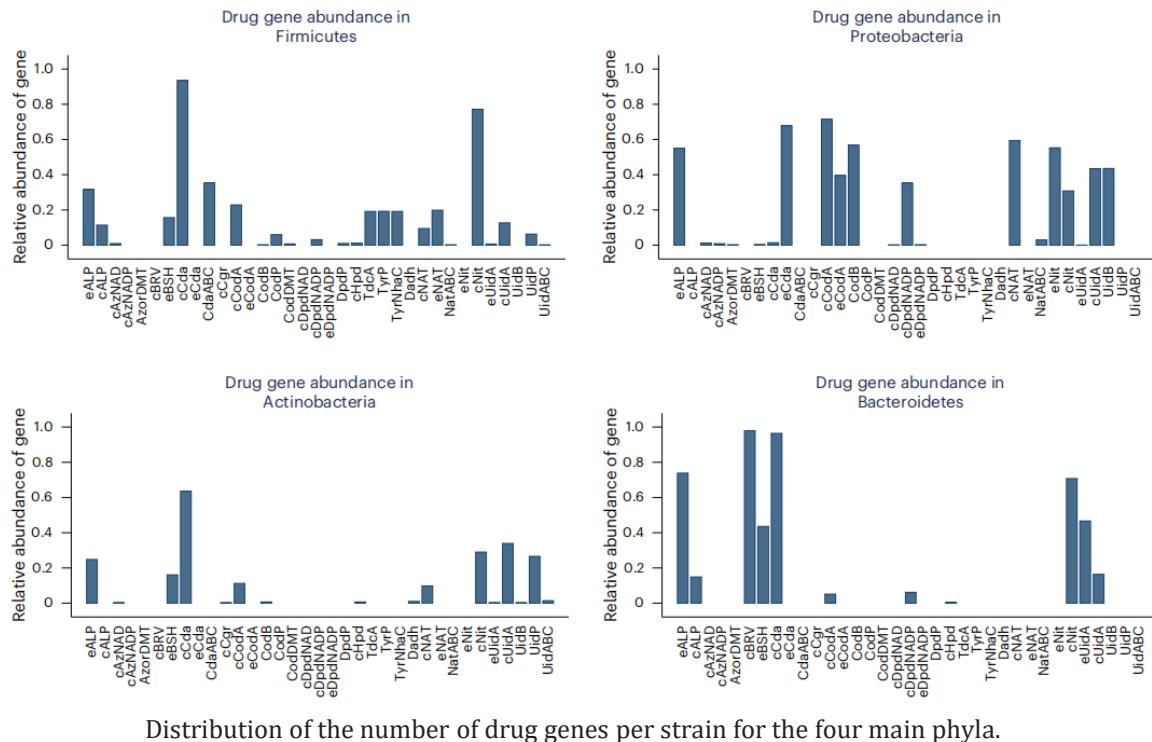


# Results 3 : Microbial drug metabolism guided by genome and bibliome



# Results 4 : Taxonomic distribution of drug-metabolizing capabilities

**d**



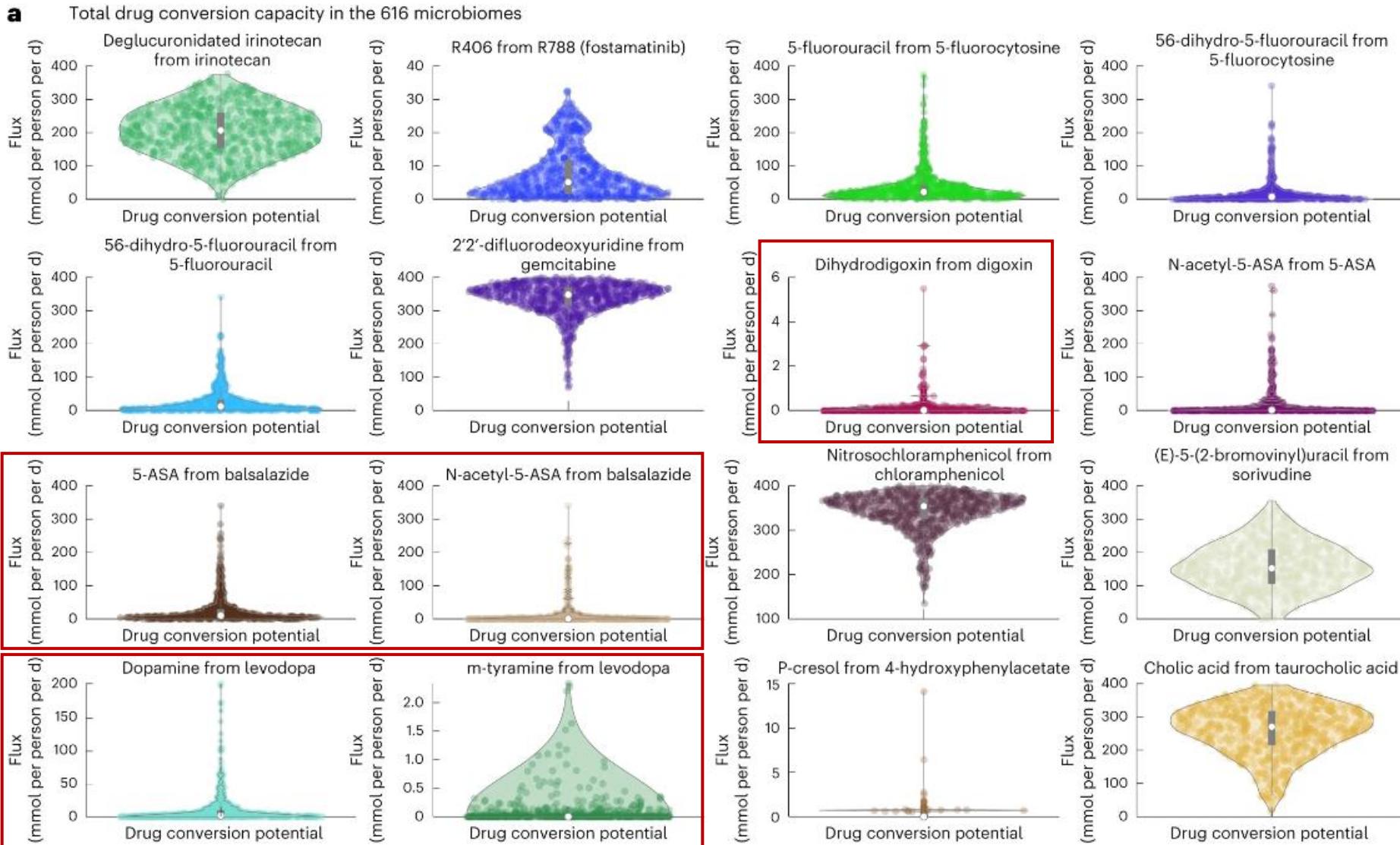
**Figure S4:** Yields from 1 mmol/g<sub>dry weight</sub>/hr of drugs that can serve as sources for ATP, CO<sub>2</sub>, pyruvate, and NH<sub>4</sub> production. Shown are all microbes that could use at least one drug to produce the respective source. Flux with glucose and with no compound added are shown as controls. One example drug per enzyme was tested.

## Results 4 : Taxonomic distribution of drug-metabolizing capabilities

---

- 作者通过对AGORA2中的5,373个菌株进行聚类分析，发现了不同分类单元之间在药物代谢能力上的差异。
- 作者通过对AGORA2中的98种药物进行聚类分析，发现了不同药物之间在微生物代谢潜力上的差异。

## Results 5 : Personalized modeling of drug-metabolizing capacities



a, Drug conversion potential in the microbiomes of 365 Japanese patients with CRC(colorectal cancer ) and 251controls on the Average Japanese Diet. The violin plots show the distribution of drug metabolite flux in mmol per person per d..

## Results 5 : Personalized modeling of drug-metabolizing capacities

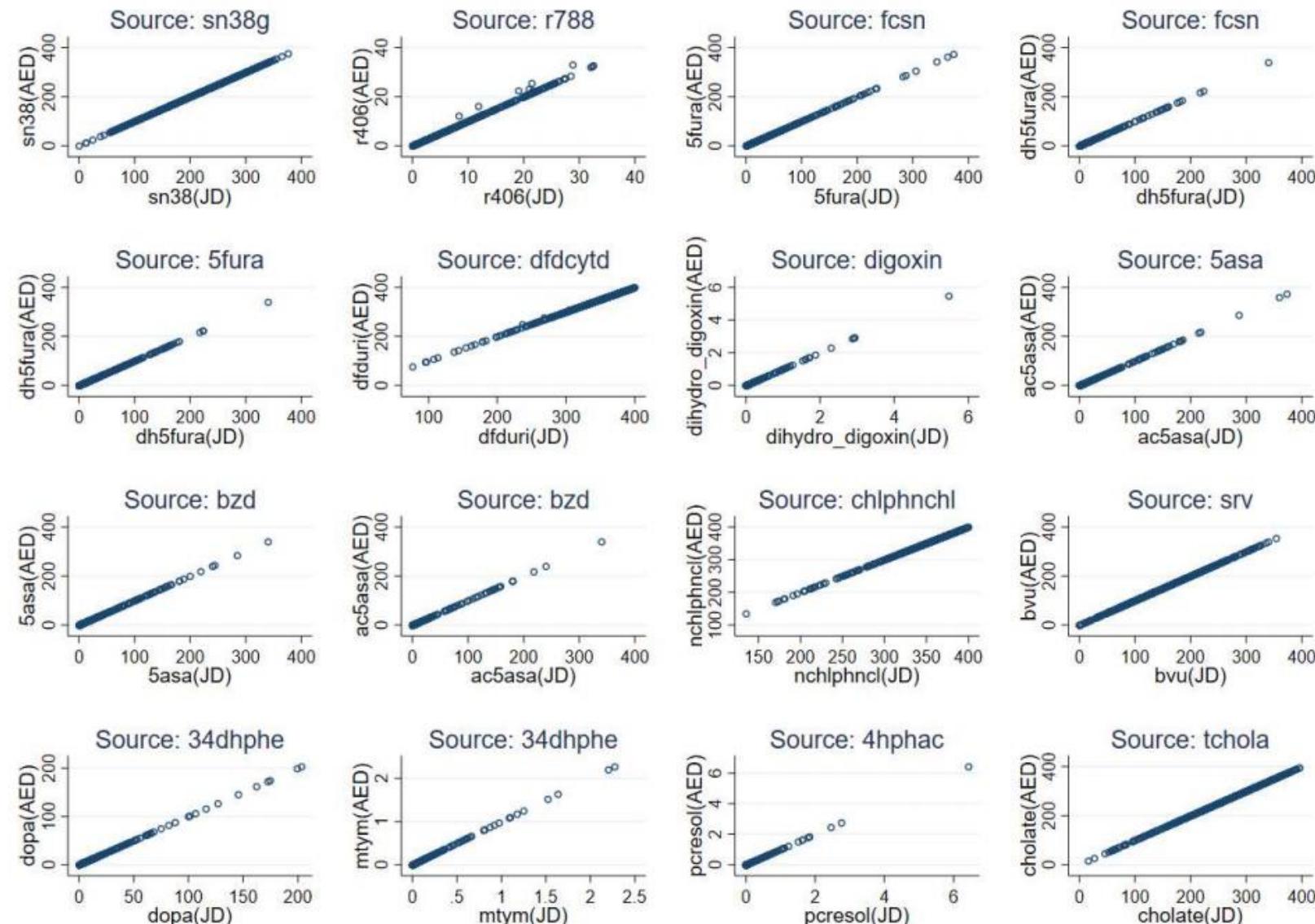
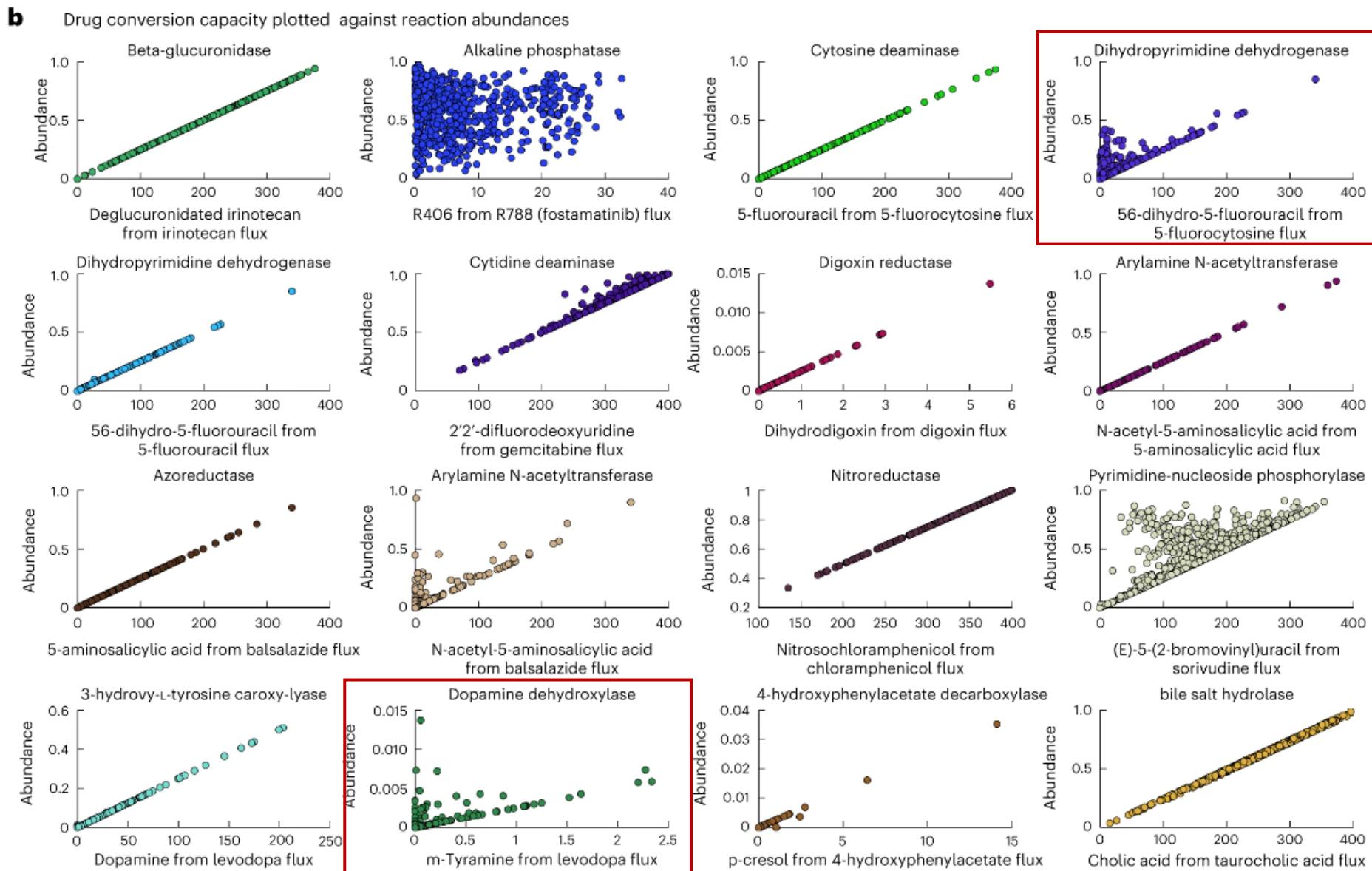


Figure S7: Scatter plots of drug-metabolising potentials under a European average diet (AED) against the metabolising potential under a Japanese diet (JP).

## Results 5 : Personalized modeling of drug-metabolizing capacities



b, Drug conversion potential (mmol per person per d) plotted against the total relative abundance of the reaction producing the shown drug metabolite in the 616 microbiomes.

# Results 5 : Personalized modeling of drug-metabolizing capacities

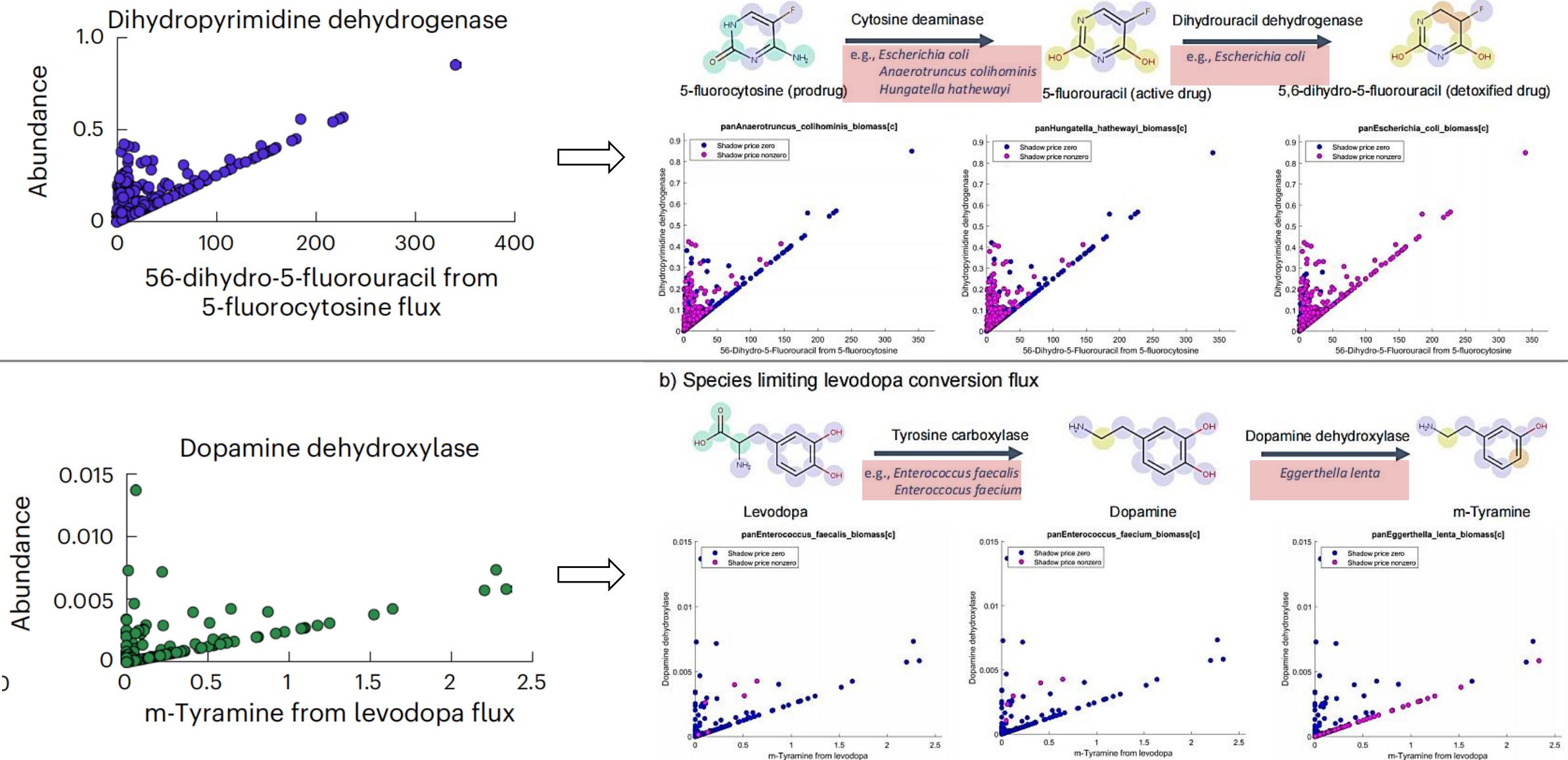


Figure S6: Bottlenecks limiting drug-metabolising capacity in 616 microbiomes.

# Results 6 : Microbiome-level fluxes are sensitive to clinical parameters

## a Descriptive statistics for drug community modeling

Output metabolite (input metabolite)	Cancer		Controls		P value
	Flux mean (s.d.)	Share of producing microbiomes	Flux mean (s.d.)	Share of producing microbiomes	
SN-38 (deglycuronitiated Irinotecan)	202.49 (70.58)	1.00	209.71 (73.74)	1.00	$2.30 \times 10^{-01a}$
R406 (Fostamatinib)	7.28 (7.30)	1.00	8.28 (8.09)	1.00	$3.06 \times 10^{-01a}$
5-Fluorouracil (5-fluorocytosine)	36.48 (45.77)	1.00	39.04 (56.29)	1.00	$9.39 \times 10^{-01a}$
5,6-Dihydro 5-Fluoro Uracil (5-fluorocytosine)	18.46 (34.74)	0.98	15.87 (19.02)	0.98	$1.99 \times 10^{-01a}$
5,6-Dihydro 5-Fluoro Uracil (5-Fluorouracil)	27.90 (40.61)	0.98	24.79 (36.22)	0.98	$1.91 \times 10^{-01a}$
2',2'-Difluorodeoxyuridine (Gemcitabine)	338.43 (46.44)	1.00	335.83 (52.46)	1.00	$6.88 \times 10^{-01a}$
Dihydrotigoxin (Digoxin)	0.10 (0.28)	0.51	0.15 (0.51)	0.57	$2.46 \times 10^{-01a}$
N-Acetyl-5-Aminosalicylic Acid (5-Aminosalicylic Acid)	16.81 (41.14)	0.88	14.34 (40.91)	0.87	$4.33 \times 10^{-01a}$
5-Aminosalicylic Acid (Balsalazide)	26.09 (41.48)	0.99	24.56 (39.96)	0.98	$4.80 \times 10^{-01a}$
N-Acetyl-5-Aminosalicylic Acid (Balsazide)	11.92 (35.01)	0.87	10.84 (30.31)	0.86	$4.73 \times 10^{-01a}$
Nitrosochloramphenicol (chloramphenicol)	340.42 (44.62)	1.00	348.37 (36.88)	1.00	$6.39 \times 10^{-03a}$
(E)-5-(2-Bromovinyl)Uracil (Sorivudine)	159.09 (67.72)	1.00	151.22 (73.96)	0.99	$2.29 \times 10^{-01a}$
Dopamine (Levodopa)	9.09 (22.03)	0.86	11.04 (24.80)	0.86	$5.26 \times 10^{-01a}$
m-tyramine (Levodopa)	0.07 (0.19)	0.43	0.09 (0.25)	0.51	$6.02 \times 10^{-02b}$
p-Cresol (4-hydroxyphenylacetate)	0.10 (0.43)	1.00	0.06 (0.24)	1.00	$1.02 \times 10^{-01a}$
Cholate (Taurocholate)	254.12 (76.20)	1.00	267.67 (75.62)	1.00	$1.62 \times 10^{-01a}$

<sup>a</sup>P value from linear regression adjusted for age (nonlinear), sex and BMI

<sup>b</sup>P value from logistic regression adjusted for age (nonlinear), sex and BMI

## b Dependence of microbial drug metabolites on age

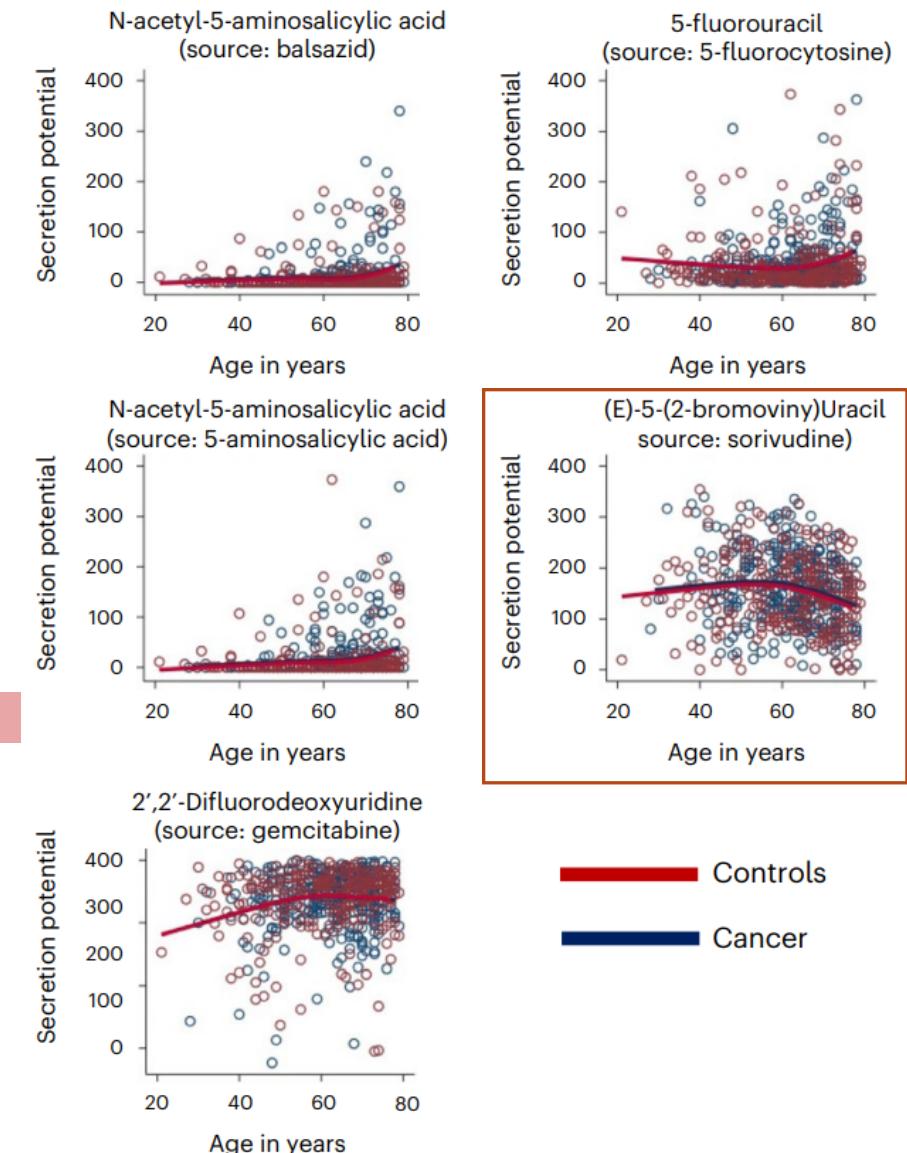


Fig. 6: Descriptive statistics for the modeled drug metabolites and fecal species-metabolite associations.

## Results 6 : Microbiome-level fluxes are sensitive to clinical parameters

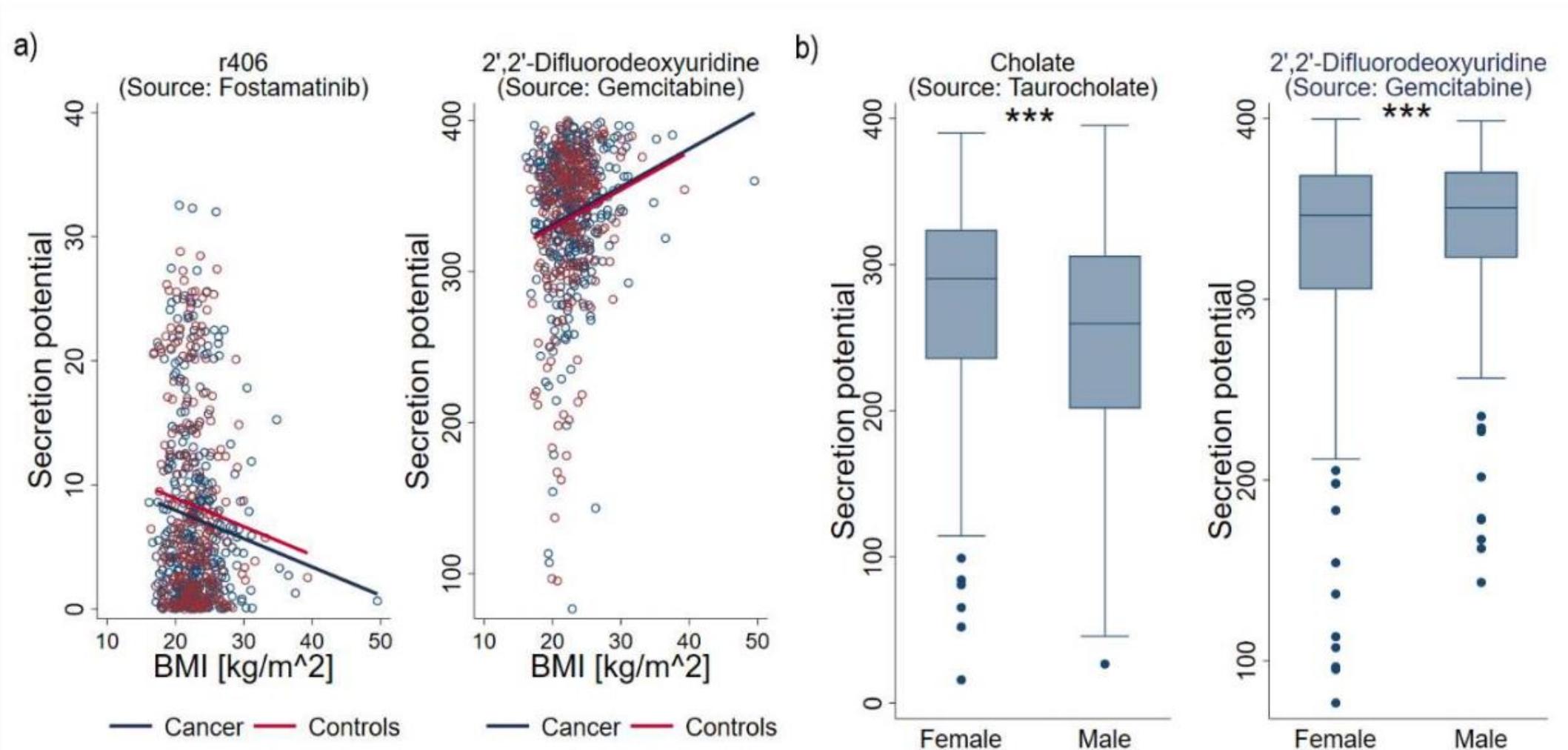


Figure S9: Descriptive statistics for the modelled drug metabolites.

# Results 6 : Microbiome-level fluxes are sensitive to clinical parameters

**c** Examples of fecal species-metabolite sign predictions through AGORA2-based community modeling

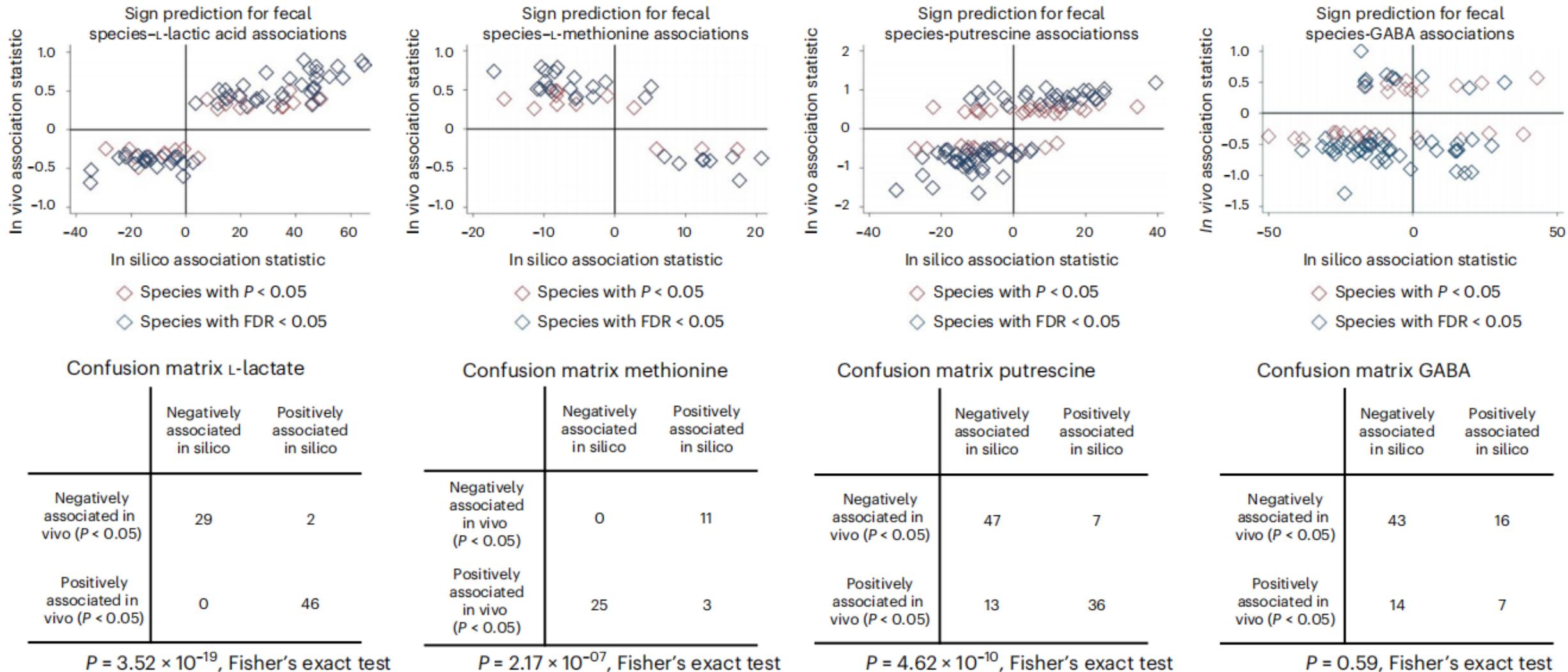


Fig. 6: Descriptive statistics for the modeled drug metabolites and fecal species-metabolite associations.

## Results 5 + Results 6

---

- 作者通过使用AGORA2重建和来自616名结直肠癌患者和对照组的菌群组成数据，构建了个性化的微生物代谢模型，预测了每个个体的肠道微生物对药物的转化潜力。
- 作者发现了不同个体之间在药物代谢能力上的巨大差异，例如，有些个体的微生物组可以代谢多达77种药物，而有些个体的微生物组则只能代谢少于10种药物。
- 作者发现了不同药物之间在微生物代谢潜力上的差异，例如，有些药物可以被大多数个体的微生物代谢，而有些药物则只能被少数个体的微生物代谢。
- 作者发现了影响个体药物代谢能力的一些因素，例如饮食习惯、年龄、性别、体重指数等，以及与药物代谢相关的菌株和基因。

## ◆ AGORA2的主要特点和作用:

它是目前最大、最全面、最准确的人类微生物代谢重建资源，它包含了大量的药物代谢信息和原子映射信息，它可以与人类代谢重建和其他微生物代谢重建兼容，它可以用于预测个性化的微生物代谢和宿主-微生物共代谢。

## ◆ AGORA2的一些局限性:

它仍然存在一些未知或不确定的代谢反应和基因功能，它没有考虑微生物之间的基因转移和表型可塑性，它没有包括一些重要的非代谢因素，如pH、温度、氧气、抗生素等，它需要更多的实验数据和新技术来验证和完善。

## ◆ AGORA2在精准医学和药物开发中的潜在应用:

它可以用于发现新的药物靶点和生物标志物，它可以用于评估药物的效果和安全性，它可以用于设计个性化的营养和治疗方案，它可以用于探索微生物与人类健康和疾病之间的机制。

请批评指正

汇报结束

