

# Evaluation

Huang LiChuang of Wie-Biotech

## Contents

<b>1</b>	<b>摘要</b>	<b>2</b>
<b>2</b>	<b>研究设计流程图</b>	<b>3</b>
<b>3</b>	<b>材料和方法</b>	<b>3</b>
<b>4</b>	<b>分析结果</b>	<b>3</b>
4.1	网络药理学和 PPI . . . . .	3
4.2	肠道菌分析 . . . . .	4
4.3	代谢物分析 . . . . .	7
4.4	肠道菌和代谢物整合 . . . . .	8
4.5	药物和肠道菌的相互作用 . . . . .	9
4.5.1	通过机器学习预测 . . . . .	9
4.5.2	药物被肠道菌代谢 . . . . .	10
4.5.3	分子相似性 . . . . .	12
4.6	分子对接 . . . . .	13
<b>5</b>	<b>结论</b>	<b>14</b>
	<b>Reference</b>	<b>14</b>

## List of Figures

1	PPI network for targets . . . . .	4
2	Enrichment analysis . . . . .	4
3	Flow chart of qiime2 processing . . . . .	5
4	Qiime2 quality control . . . . .	6
5	Gut microbiome abundance . . . . .	6
6	Use MetaboAnalyst for analysis of metabolites . . . . .	7
7	Identify compounds with SIRIUS 4 OPTIONAL . . . . .	8
8	Correlation of metabolites with microbiota . . . . .	9
9	Machine learning prediction of drug towards microbiota . . . . .	10
10	Drug metabolized by microbiota . . . . .	11

11	Chemical modification by microbiota . . . . .	12
12	Molecule similarity . . . . .	12
13	Autodock vina binding affinity . . . . .	13

## List of Tables

## 1 摘要

以下为需要评估的内容以及相应的答复：

- 用网络药理学分析中药方的主要活性成分，对比文献研究现状挑选 4 个左右的活性成分作为候选药物成分
  - 可分析主要成分，结合目标疾病筛选主要活性成分。
  - 可从 HERB (<http://herb.ac.cn/Download/>) 获取成分信息（该数据库整合了较多的其他数据库）。
  - 目标疾病是否为糖尿病肾病？且有代谢组数据？如果为（中药方成分的）非靶向代谢组，则能根据非靶向代谢组数据鉴定更多的化合物用于网络药理学分析，而不是只通过数据库筛选。
  - 可进一步通过 PPI 网络和通路富集分析筛选活性成分<sup>1</sup>。
  - 可查阅文献（较为主观，可作为辅助手段）
- 分析糖尿病肾病肠道差异菌群
  - 需要 16s RNA 数据，或从 GEO 公共数据库获取 (<https://www.ncbi.nlm.nih.gov/gds/?term=16s>)。
  - 可用 qiime2 (<https://qiime2.org/>) 筛选肠道差异菌群<sup>2,3</sup>。
- 分析糖尿病肾病代谢组学差异
  - 需明确，是人的肾脏的代谢组，还是肠道菌的代谢组（因为上述有肠道菌分析，容易混淆）。
  - 可分为挖掘公共数据库（GNPS: <https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp>）和客户代提供代谢数据的情况。
  - 如果是客户提供，请考虑：
    - 包含对照组和模型组的数据，需要生物学重复，最好为非靶向代谢组数据（这种情况下，能鉴定和找到最多的差异代谢物）
    - 如果是非靶向代谢组，需要数据鉴定。可分为谱图匹配性鉴定（常规方法），和预测性鉴定（例如，SIRIUS<sup>4</sup>）。
    - 如果是靶向代谢组，已知目标代谢物，则不需要额外的鉴定，根据分子量比对即可。
    - 需确认代谢组数据的采集是否包含 MS<sup>2</sup>。如果仅包含 MS<sup>1</sup>，则鉴定准确度会相对偏低。如果是靶向代谢组，仅有 MS<sup>1</sup> 亦可。
  - 以标准的方法：PCA 聚类，OPLS-DA 聚类、VIP、P 值筛选差异代谢物。
  - 还可以结合不同 Feature selection 算法进一步筛选，例如 LASSO，EFS<sup>5</sup> 等。
  - 差异代谢物可通路富集分析，结合疾病，进一步筛选。可用方法为 MetaboAnalyst (<https://www.metaboanalyst.ca/MetaboAnalyst/ModuleView.xhtml>)<sup>6</sup>。
  - 根据上述情况不同，工作量会大不相同。可做大量分析，也可仅做少量分析。
- 分析差异菌群与差异代谢物的相关性（桑基图）
  - 16s RNA 和代谢组的多组学分析<sup>7</sup>。
  - 对差异菌群和差异代谢物关联性分析，桑基图、热图均可（热图可能更直观，见 Fig. 8）。
- 结合药物成分筛选目标菌群和相关代谢物

- 首先需要获取药物成分的靶点。可通过 Binding DB 获取 <https://www.bindingdb.org/rwd/bind/index.jsp>。
- 通过（人的）代谢物的通路富集结果（步骤 3）得到相应蛋白，可对药物靶点取交集。
- 通过机器预测药物和肠道菌互作<sup>8,9</sup>。目前似乎存在一系列方法，需要探索找到合适的方法。
- 若上述方法简便容易，以下或许可以不用考虑：
  - 关于药物-肠道菌互作<sup>10</sup>，药物作用于菌，得到代谢物（细菌的代谢物，而非人）。
  - 想要从药物中筛选，需要细菌的代谢物信息（这一步骤可能存在一定苦难，需要结合实际筛选的细菌考虑）。
  - 上述，肠道菌代谢物的获取是关键步骤之一，可以从 gutMGene (<http://bio-annotation.cn/gutmgene/home.dhtml>) 获取。
  - 通过计算药物和肠道菌代谢物的分子相似性，推测是否存在药物-菌作用关系<sup>11</sup>（具有不确定性）
- 6. 代谢小分子靶点蛋白分析（这个看能不能用分子对接的方式获取）
  - 只要上述（步骤 5）的药物和靶点都具备，即可以分子对接方式分析。
- 7. 广泛的靶点蛋白与糖尿病肾病差异基因取交集，筛选候选基因
  - 结合疾病的公共数据库筛选进一步筛选，例如 genecards <https://www.genecards.org/>。
  - 需要明确，这里还能进一步通过分析其他 GEO 数据筛选差异基因，再结合筛选。是否需要？

总体上，工作量较大，视情况可能需要 1-3 周。

## 2 研究设计流程图

## 3 材料和方法

## 4 分析结果

以下内容仅为示例。

### 4.1 网络药理学和 PPI

Figure 1为图 PPI network for targets 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 11-56-42.png)

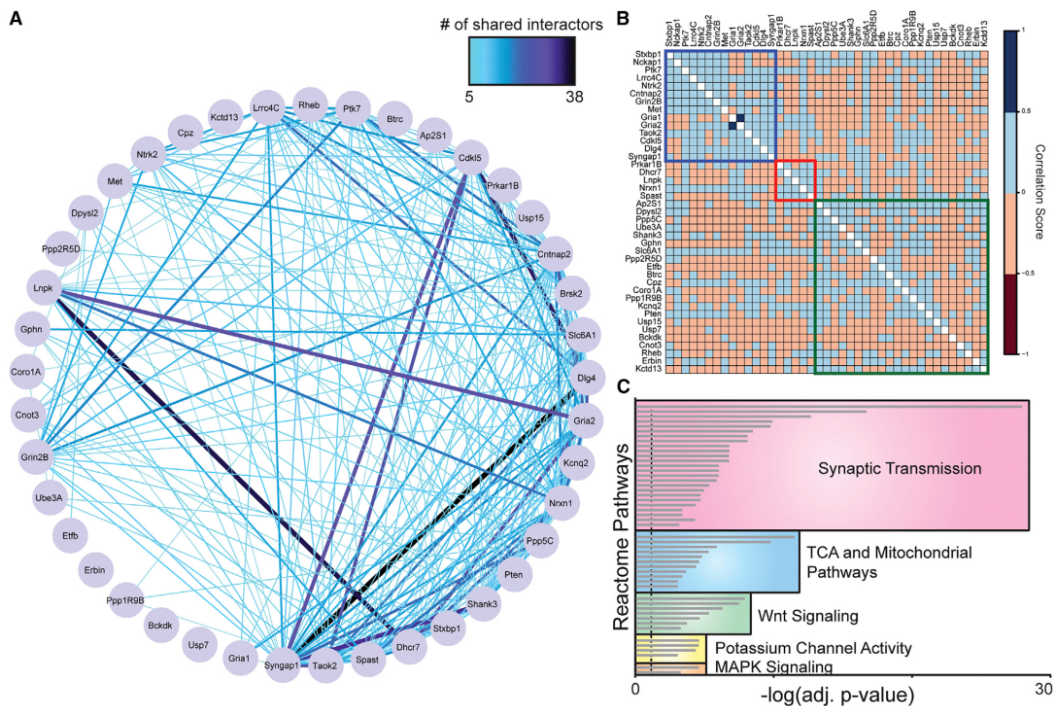


Figure 1: PPI network for targets

Figure 2为图 enrichment analysis 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 13-03-48.png)

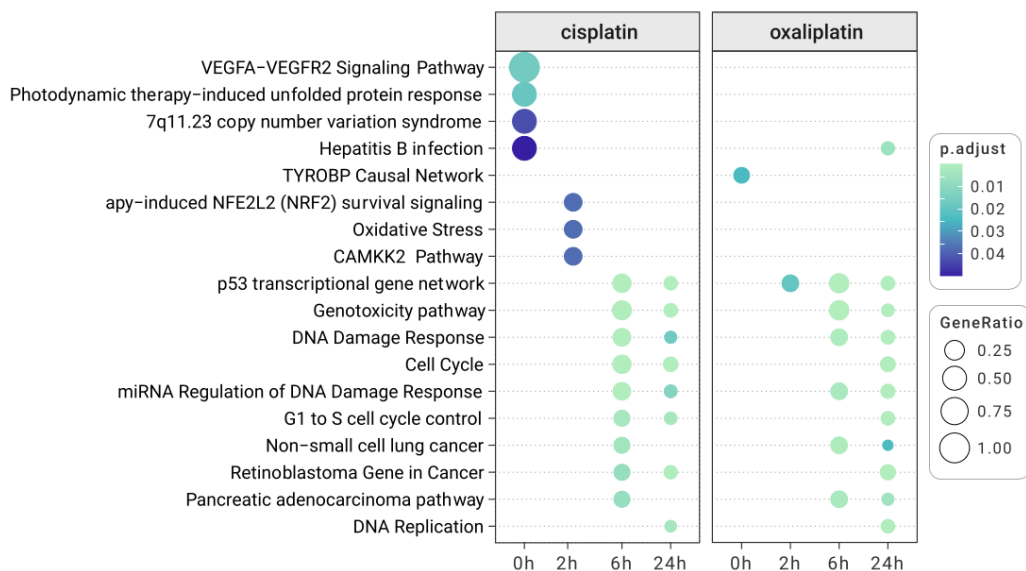


Figure 2: Enrichment analysis

## 4.2 肠道菌分析

Figure 3为图 flow chart of qiime2 processing 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 11-26-30.png)

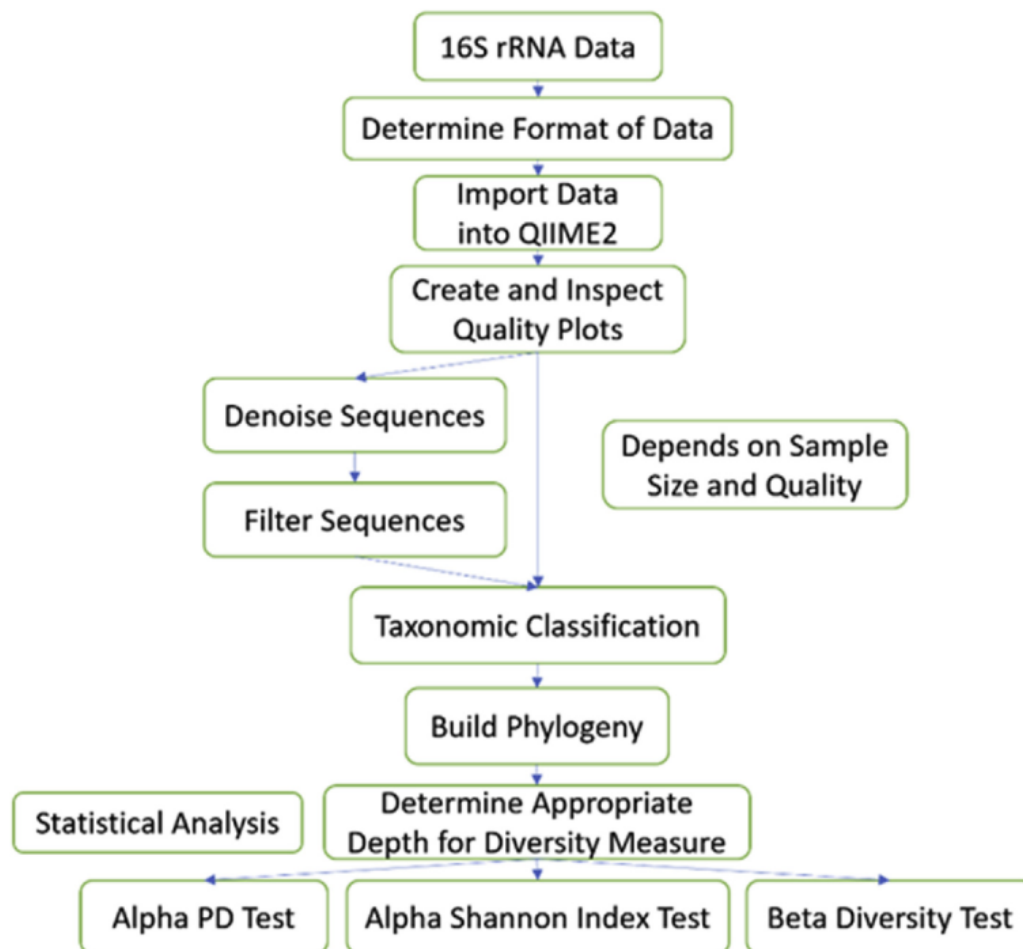


Figure 3: Flow chart of qiime2 processing

Figure 4为图 qiime2 quality control 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 11-36-35.png)

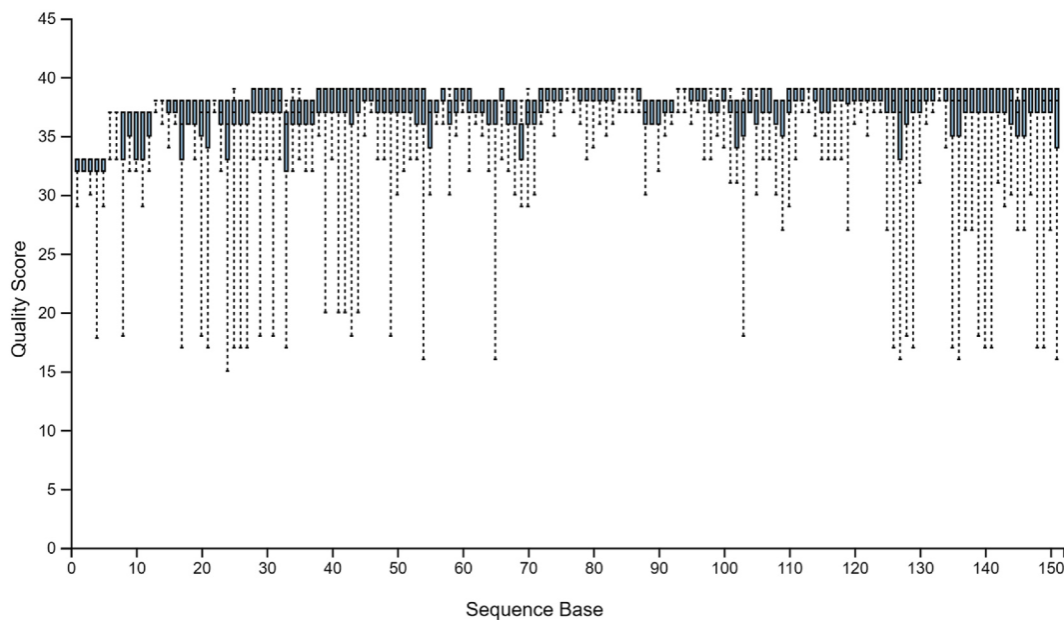


Figure 4: Qiime2 quality control

Figure 5为图 gut microbiome abundance 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 11-40-38.png)

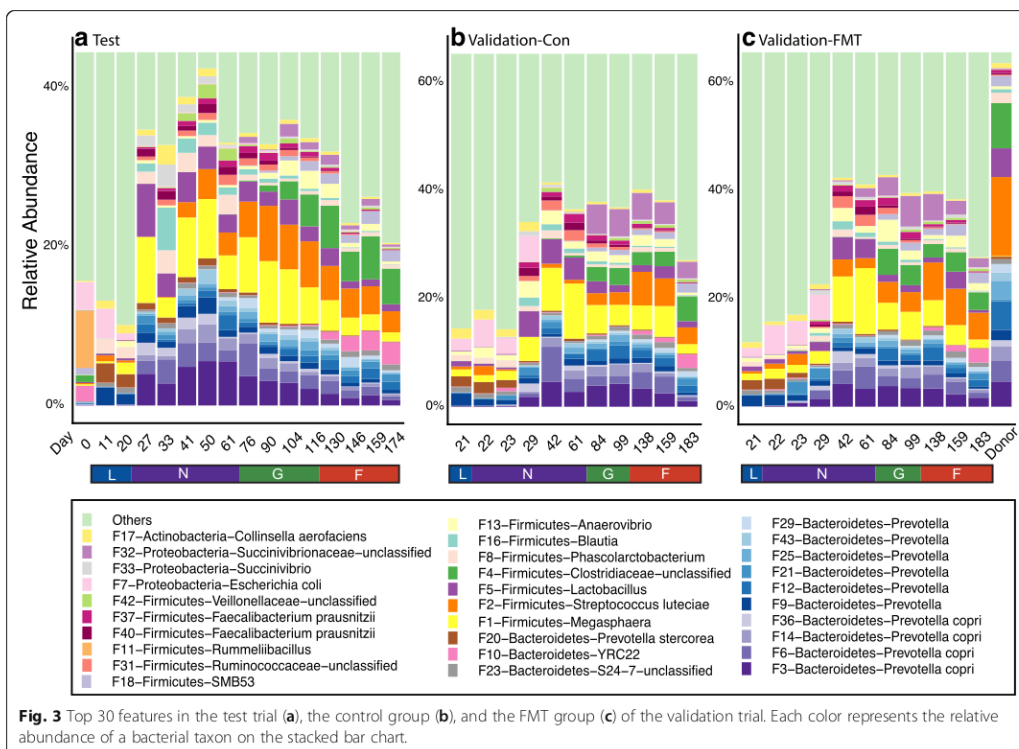


Figure 5: Gut microbiome abundance

### 4.3 代谢物分析

Figure 6为图 use MetaboAnalyst for analysis of metabolites 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 11-32-20.png)

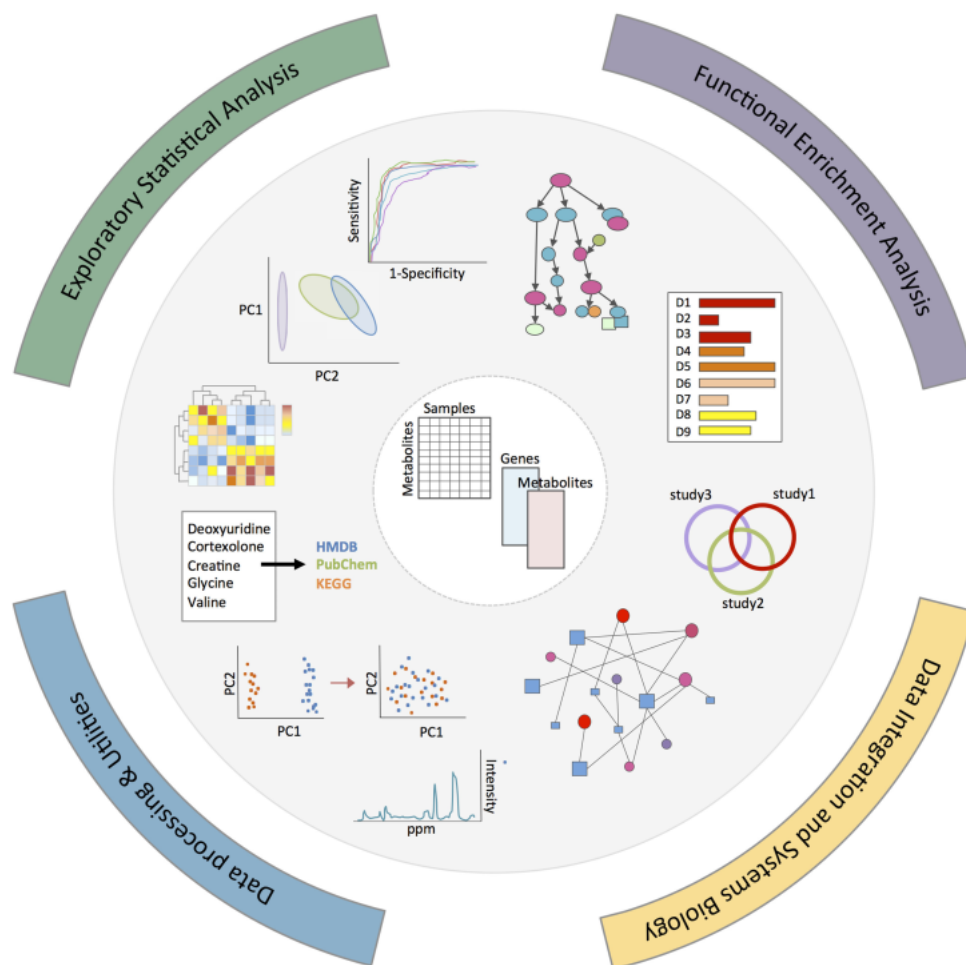


Figure 6: Use MetaboAnalyst for analysis of metabolites

Figure 7为图 identify compounds with SIRIUS 4 OPTIONAL 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 11-28-15.png)





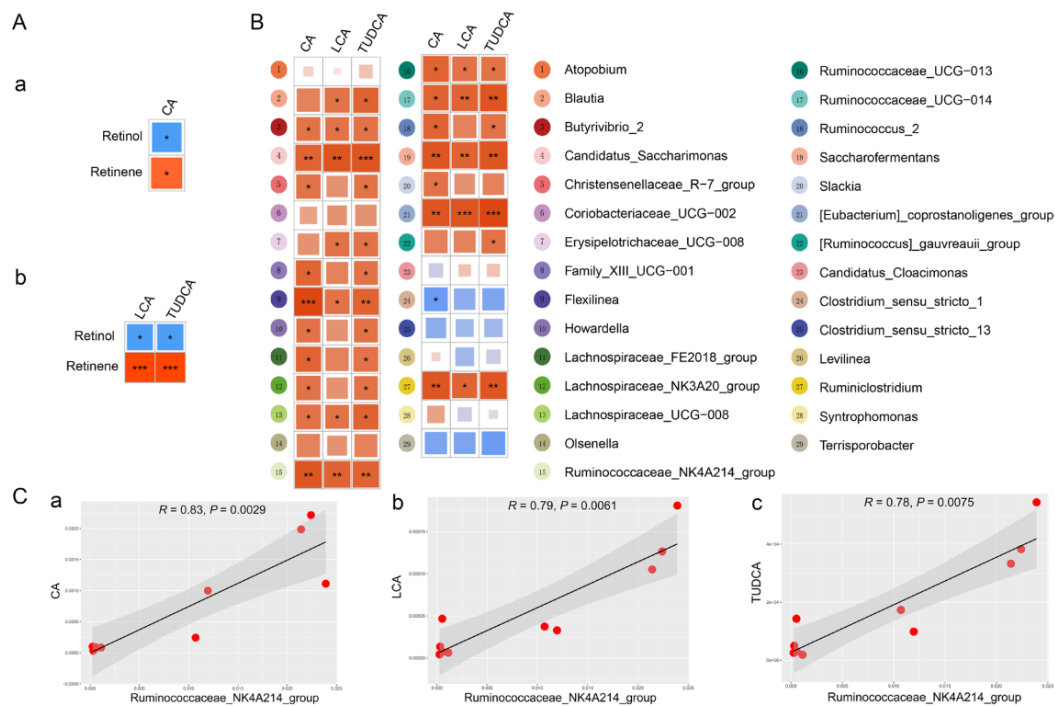


Figure 8: Correlation of metabolites with microbiota

## 4.5 药物和肠道菌的相互作用

### 4.5.1 通过机器学习预测

来自于综述文章，需要找到合适的方法。

Figure 9为图 machine learning prediction of drug towards microbiota 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 13-38-39.png)

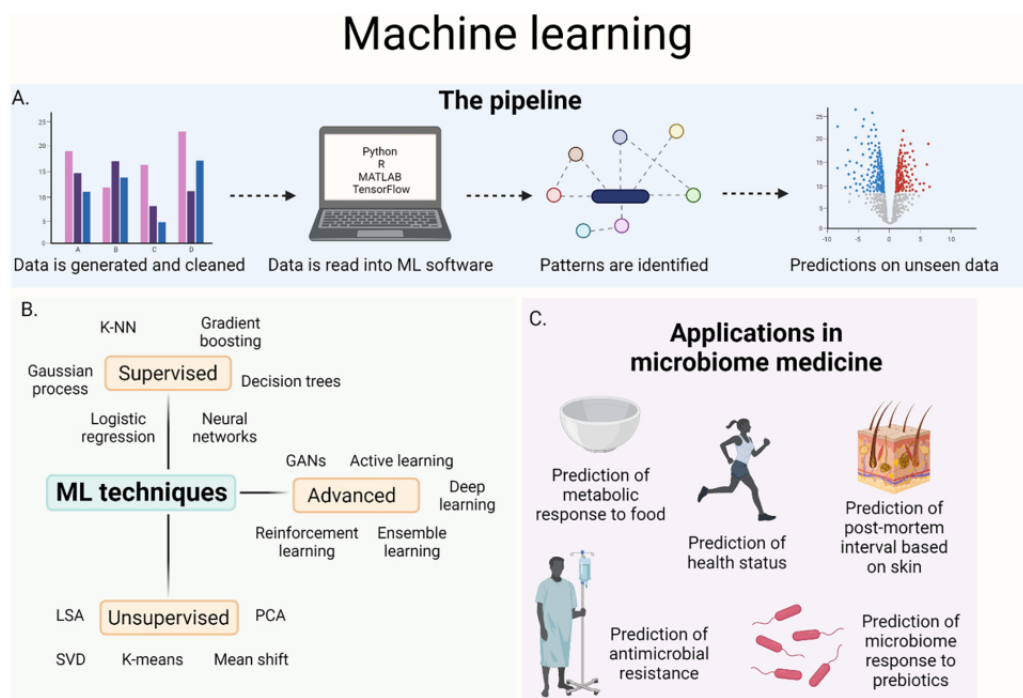


Figure 9: Machine learning prediction of drug towards microbiota

## 4.5.2 药物被肠道菌代谢

Figure 10为图 drug metabolized by microbiota 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 13-13-48.png)

## Graphical Abstract

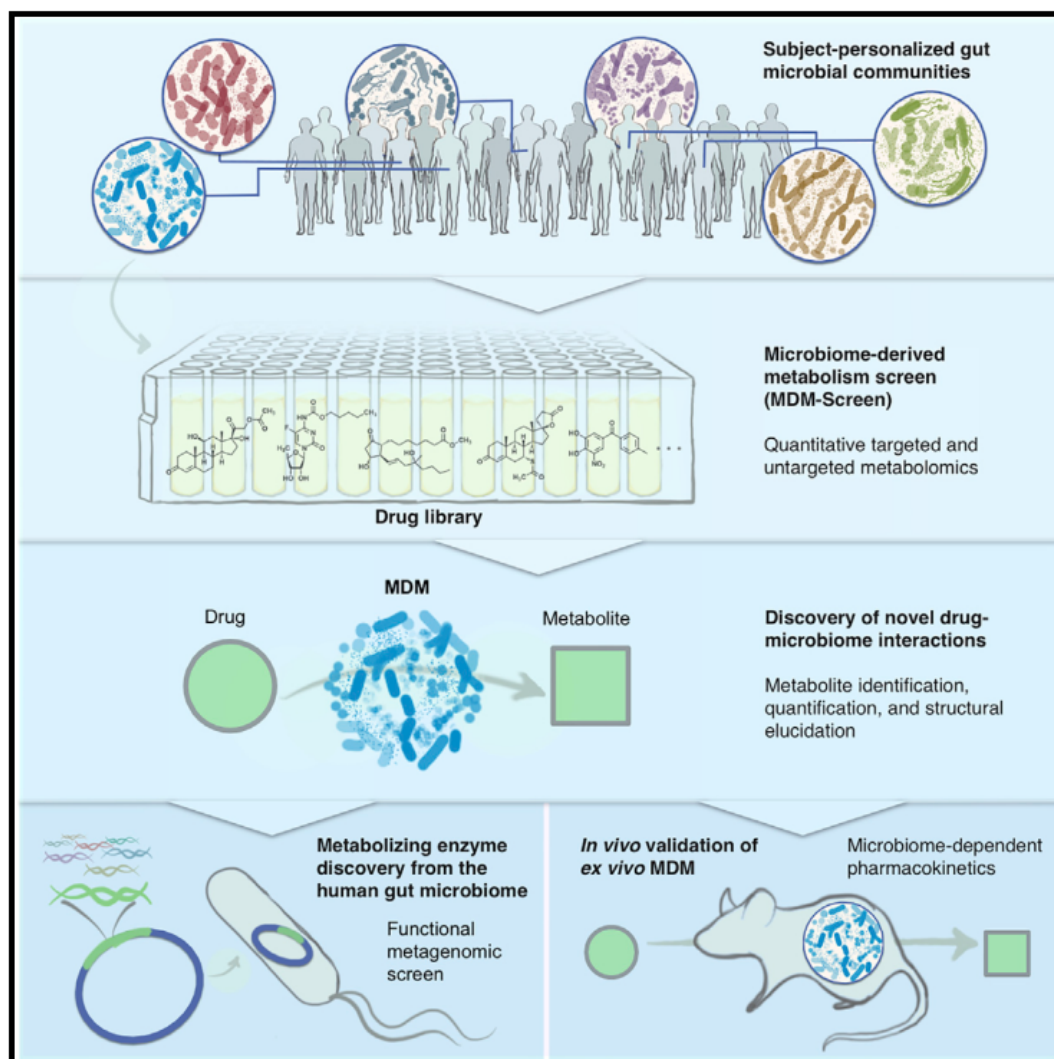


Figure 10: Drug metabolized by microbiota

Figure 11为图 chemical modification by microbiota 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 13-16-48.png)

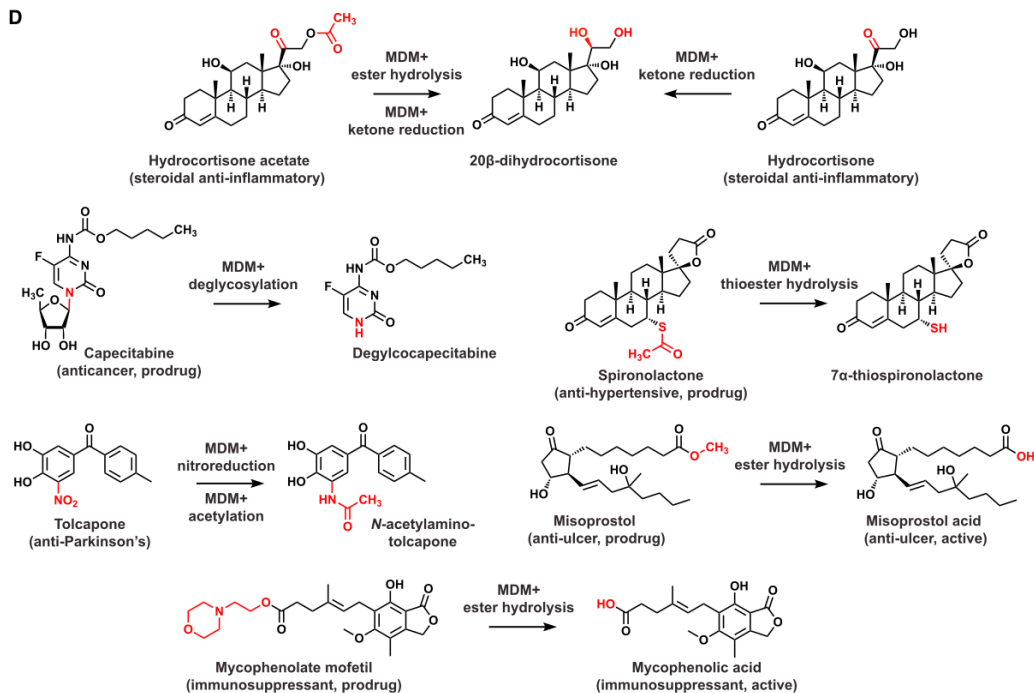


Figure 11: Chemical modification by microbiota

#### 4.5.3 分子相似性

Figure 12为图 molecule similarity 概览。

(对应文件为 ~/Pictures/Screenshots/Screenshot from 2023-08-24 12-52-35.png)

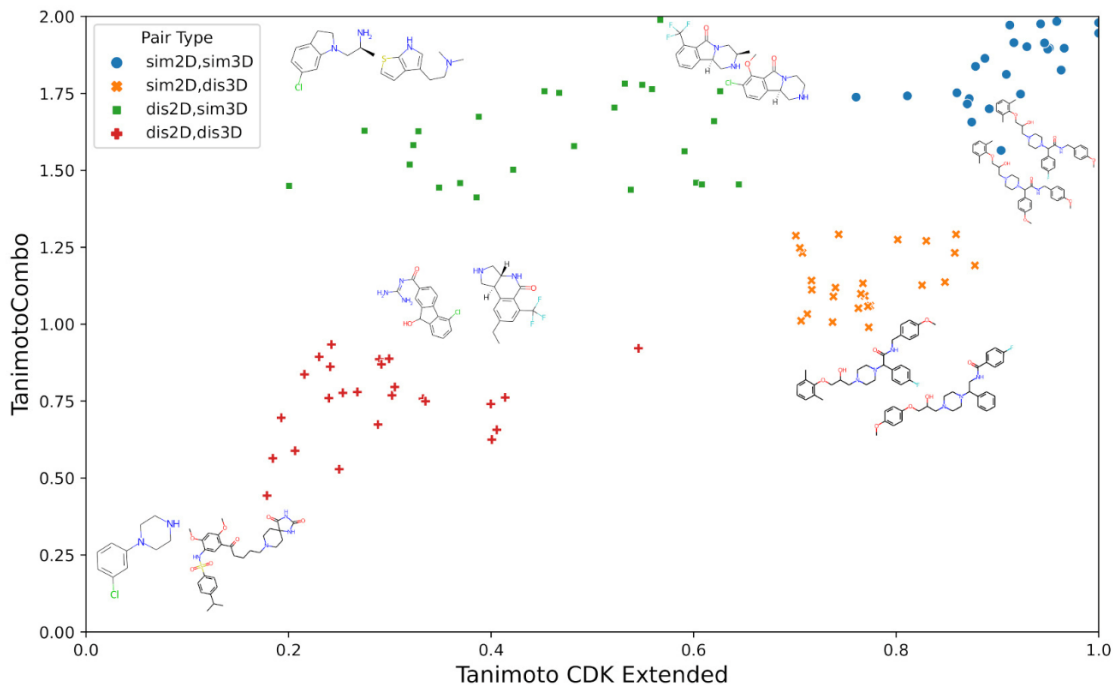


Figure 12: Molecule similarity

## 4.6 分子对接

Figure 13为图 autodock vina binding affinity 概览。

(对应文件为 ../2023\_06\_30\_eval/figs/Docking\_Affinity.pdf)

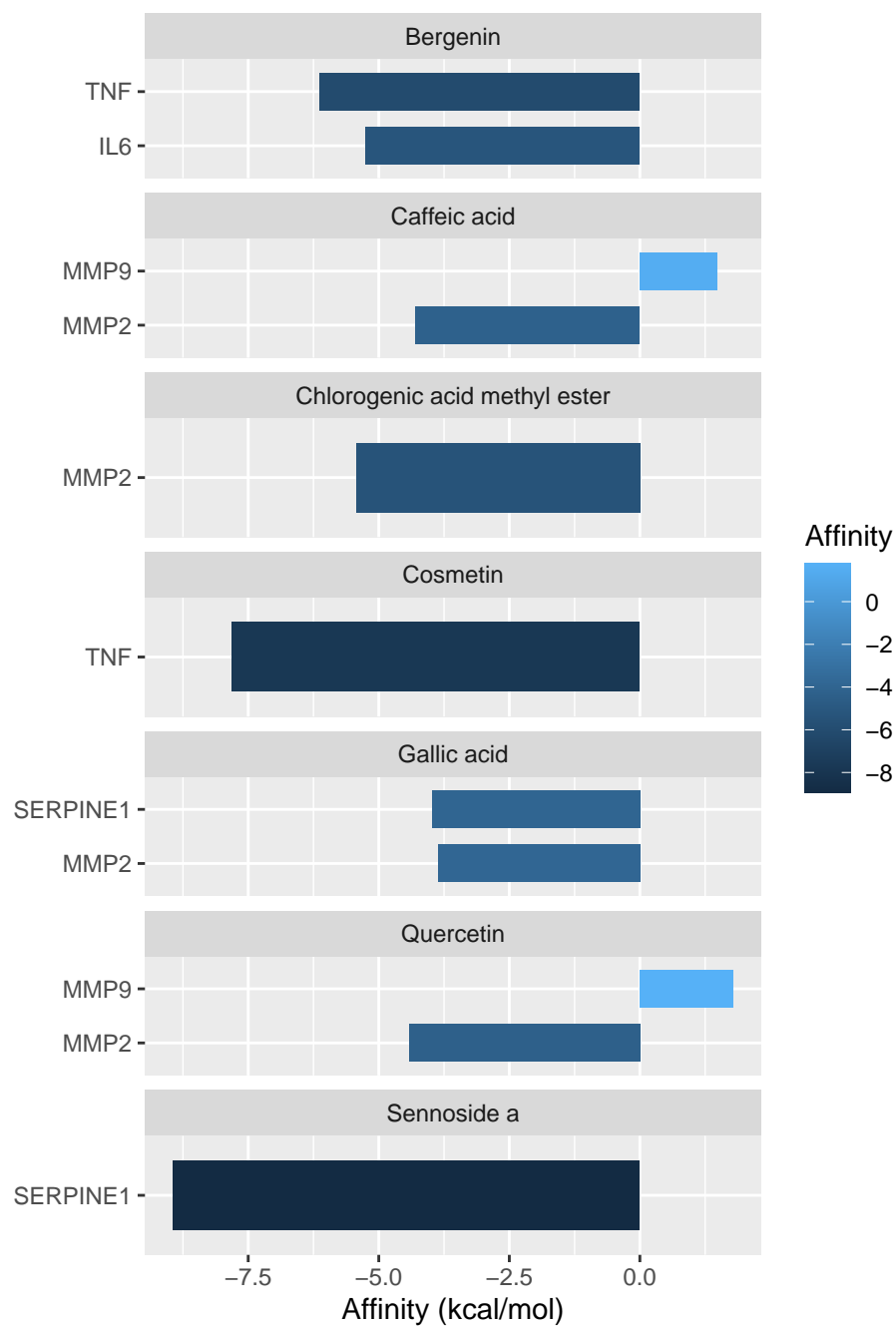


Figure 13: Autodock vina binding affinity

## 5 结论

### Reference

1. Murtaza, N. *et al.* Neuron-specific protein network mapping of autism risk genes identifies shared biological mechanisms and disease-relevant pathologies. *Cell Reports* **41**, (2022).
2. Rai, S. N. *et al.* Microbiome data analysis with applications to pre-clinical studies using qiime2: Statistical considerations. *Genes & Diseases* **8**, (2021).
3. Wang, X. *et al.* Longitudinal investigation of the swine gut microbiome from birth to market reveals stage and growth performance associated bacteria. *Microbiome* **7**, (2019).
4. Dührkop, K. *et al.* SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods* **16**, 299–302 (2019).
5. Neumann, U., Genze, N. & Heider, D. EFS: An ensemble feature selection tool implemented as r-package and web-application. *BioData Mining* **10**, 21 (2017).
6. Pang, Z., Chong, J., Li, S. & Xia, J. MetaboAnalystR 3.0: Toward an optimized workflow for global metabolomics. *Metabolites* (2020) doi:10.3390/metabo10050186.
7. Zhang, T. *et al.* Disrupted spermatogenesis in a metabolic syndrome model: The role of vitamin a metabolism in the guttestis axis. *Gut* **71**, (2021).
8. McCoubrey, L. E., Gaisford, S., Orlu, M. & Basit, A. W. Predicting drug-microbiome interactions with machine learning. *Biotechnology Advances* **54**, (2022).
9. McCoubrey, L. E., Elbadawi, M., Orlu, M., Gaisford, S. & Basit, A. W. Machine learning uncovers adverse drug effects on intestinal bacteria. *Pharmaceutics* **13**, (2021).
10. Javdan, B. *et al.* Personalized mapping of drug metabolism by the human gut microbiome. *Cell* **181**, (2020).
11. Gandini, E. *et al.* Molecular similarity perception based on machine-learning models. *International Journal of Molecular Sciences* **23**, (2022).