

中文版 methods

仅供客户在文章写作时参考，分析内容和方法请以结题报告为准，请客户自行承担文章查重等相关风险。

1 实验流程

1.1 DNA 样品检测

详见样本检测报告。

1.2 文库构建和上机测序

取样本的 1 μg 基因组 DNA, 用 Covaris 超声波破碎仪随机打断成长度约为 350 bp 的片段后进行文库的构建, 经末端修复、加 A 尾、加测序接头、纯化、PCR 扩增等步骤完成整个文库制备。文库构建完成后, 先使用 AATI 检测文库片段的完整性及插入片段大小, 符合预期后, 使用 Q-PCR 方法对文库的有效浓度进行准确定量(文库有效浓度 $>3 \text{ nM}$), 以保证文库质量。库检合格后, 把不同文库按照有效浓度及目标下机数据量的需求 pooling 后进行测序。

2 生物信息分析

使用 fastp (<https://github.com/OpenGene/fastp>)对测序获得的原始数据(raw data)进行预处理, 获取用于后续分析的有效数据(clean data)。具体处理步骤如下:a)当任一测序 read 中含有接头序列, 去除此 paired reads;b)当任一测序 read 中含有的低质量($Q \leq 5$)碱基数超过该条 read 碱基数的 50%时, 去除此 paired reads;c)当任一测序 read 中 N 含量超过该 read 碱基数的 10%时, 去除此 paired reads。如果样品存在宿主污染, 需与宿主序列进行比对, 过滤掉可能来源于宿主的 reads, 默认采用 Bowtie2 软件(<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>), 参数设置- -end-to-end, --sensitive, -I 200, -X 400 (Karlsson FH et al., 2012; Karlsson FH et al., 2013; Scher JU et al., 2013)。

使用 HUMAnN 3 (Franzosa EA et al.,2018)调用 Metaphlan 4 (Blanco-Míguez A et al.,2023)与 Uniprot 数据库(<https://www.uniprot.org/>)比对得到物种分类结果和 UniRef90 ID, 根据 UniRef90 ID 和其他功能数据库中的对应关系, 得到各个功能库的注释结果。

从物种和功能的各个分类层级上的丰度表出发, 进行 top10 相对丰度概况展示, 丰度聚类热图展示。并进行相对丰度聚类分析, PCA (R ade4 package) (Rao C R et al., 1964), PCoA(R ade4 package)和 NMDS (R vegan package)降维分析(Legendre P, 1998); 使用 Anosim 分析(R vegan package)检验组间的差异情况; 然后使用 MetaGenomeSeq 和 LEfSe 分析寻找组间差异物种, MetaGenomeSeq 分析(Paulson JN et al.,2013)对各个分类层级做组间的假设检验得到 p 值与 Q 值, LEfSe 分析使用 LEfSe 软件(LDA Score 默认为 4) (Segata N et al., 2011)。

3 参考文献

Blanco-Míguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat Biotechnol.* 2023;41(11):1633-1644. doi:10.1038/s41587-023-01688-w

Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods.* 2018;15(11):962-968. doi:10.1038/s41592-018-0176-y

Karlsson FH, Fåk F, Nookaew I, et al. Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nat Commun.* 2012 ; 3:1245. doi:10.1038/ncomms2266

Karlsson FH, Tremaroli V, Nookaew I, et al. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature.* 2013;498(7452):99-103. doi:10.1038/nature12198

Le Chatelier E, Nielsen T, Qin J, et al. Richness of human gut microbiome correlates with metabolic markers. *Nature.* 2013;500(7464):541-546. doi:10.1038/nature12506 Legendre P, Legendre L. Numerical ecology, 2nd edition[J].1998.

Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013;10(12):1200-1202. doi:10.1038/nmeth.2658

Rao C R. The Use and Interpretation of Principal Component Analysis in Applied Research[J]. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 1964, 26(4):329-358.

Scher JU, Szczesnak A, Longman RS, et al. Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife*. 2013; 2: e01202. Published 2013 Nov 5. doi:10.7554/eLife.01202

Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6): R60. Published 2011 Jun 24. doi:10.1186/gb-2011-12-6-r60