



Oct 31, 2025

Protocol for whole shotgun metagenomics pipeline for the study of stool human digestive microbiota

DOI

<https://dx.doi.org/10.17504/protocols.io.5qpvo9ox7v4o/v1>

Victoria Meslier¹, Yani Ren¹, Mallia Geiger¹, Marine Gilles¹, Alexandre Famechon¹, Aymeric David¹, Christian Morabito¹, Benoit Quinquis¹, Mathieu Almeida¹

¹Université Paris Saclay, INRAE MetaGenoPolis, 78350 Jouy-en-Josas, France



Victoria Meslier

Université Paris Saclay, INRAE MetaGenoPolis, 78350 Jouy-en-...

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.5qpvo9ox7v4o/v1>

Protocol Citation: Victoria Meslier, Yani Ren, Mallia Geiger, Marine Gilles, Alexandre Famechon, Aymeric David, Christian Morabito, Benoit Quinquis, Mathieu Almeida 2025. Protocol for whole shotgun metagenomics pipeline for the study of stool human digestive microbiota. **protocols.io** <https://dx.doi.org/10.17504/protocols.io.5qpvo9ox7v4o/v1>

License: This is an open access protocol distributed under the terms of the **[Creative Commons Attribution License](#)**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: December 17, 2024

Last Modified: November 14, 2025

Protocol Integer ID: 115788

Keywords: ASAPCRN, metagenomics, human gut microbiota, dna extraction, shotgun sequencing, bioinformatics, human digestive microbiota this protocol, protocol for whole shotgun metagenomics pipeline, human digestive microbiota, digestive microbiota, whole shotgun metagenomics pipeline, metagenomic read mapping, human gut, microbial composition, functional potentials of the human gut, study of stool, whole dna extraction, procedures for whole dna extraction, stool

Funders Acknowledgements:

Agence Nationale de la Recherche

Grant ID: ANR-11-DPBS-0001

Aligning Science Across Parkinson's

Grant ID: ASAP-000420

Abstract

This protocol describes the procedures for whole DNA extraction, shotgun high throughput sequencing, metagenomic read mapping and bioinformatical pre-processing to determine the microbial composition and functional potentials of the human gut digestive microbiota.

Attachments



[Metagenomic_procedur.](#)



191KB

Guidelines

NA

Materials

NA

Safety warnings

 NA

Ethics statement

The protocols.io team notes that research involving animals and humans must be conducted according to internationally-accepted standards and should always have prior approval from an Institutional Ethics Committee or Board.

Before start

Ensure proper sampling and samples conservation before proceeding to DNA extraction.



Wet lab section : DNA extraction and Shotgun sequencing

- 1 DNA isolation procedure, adapted from [dx.doi.org/10.17504/protocols.io.dm6gpjm11gzp/v1](https://doi.org/10.17504/protocols.io.dm6gpjm11gzp/v1)
- 2 High throughput WGS sequencing as described in [1]

Dry lab section: Bioinformatical procedures

- 3 QC validation and read mapping was performed using the parametres described in the attached document using the 10.4M gut gene catalog and the 8.4M oral gene catalog.
- 4 Supplementary QC check was performed using CroCoDeEL tool <https://github.com/metagenopolis/CroCoDeEL>
- 5 Determination of the microbial composition of high quality samples using MSPminer [10]
- 6 Determination of the functional potentials of high quality samples using in-house pipeline as described in [15]

Protocol references

1. Meslier
V, Quinquis B, Da Silva K, Plaza Oñate F, Pons N, Roume H, et al. Benchmarking second and third-generation sequencing platforms for microbial metagenomics. *Sci Data*. 2022;9(1):694.
2. Criscuolo
A, Brisse S. AlienTrimmer: A tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. *Genomics*. 2013;102(5-6):500-6.
3. Langmead
B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-9.
4. Nurk
S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, et al. The complete sequence of a human genome. *Science*. avr 2022;376(6588):44-53.
5. Wen
C, Zheng Z, Shao T, Liu L, Xie Z, Le Chatelier E, et al. Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis. *Genome biology*. juill 2017;18(1):142.
6. Le Chatelier E, Almeida M, Plaza Oñate F, Pons N, Gauthier F, Ghoulane A, et al. A catalog of genes and species of the human oral microbiota. [Internet]. 2021. Disponible sur: <https://data.inrae.fr/citation?persistentId=doi:10.15454/WQ4UTV>
7. Amine Ghoulane, Florence Thirion, Florian Plaza Oñate et al. Accurate profiling of microbial communities for shotgun metagenomic sequencing with Meteor2, 05 March 2025, PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-6122276/v1>]
8. Le Chatelier Emmanuelle, Prifti Eddi. Mining Metaomics Data In R. Retrived from <https://forgemia.inra.fr/metagenopolis/momr>.
9. Nielsen
HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnology*. 2014;32(8):822-8.
10. Plaza Oñate F, Le Chatelier E, Almeida M, Cervino ACL, Gauthier F, Magoulès F, et al. MSPminer: abundance-based reconstitution of microbial pan-genomes from shotgun metagenomic data. *Bioinformatics*. 1 mai 2019;35(9):1544-52.
11. Plaza
Oñate F, Le Chatelier E. Metagenomic Species Pan-genomes (MSPs) of the human gastrointestinal microbiota. *Portail Data INRAE Recherche Data Gouv*. 2020;

12. Parks
DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research*. 7 janv 2022;50(D1):D785-94.
13. Sayers
EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*. 8 janv 2019;47(D1):D23-8.
14. Altschul
SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. oct 1990;215(3):403-10.
15. Thirion
F, Speyer H, Hansen TH, Nielsen T, Fan Y, Le Chatelier E, et al. Alteration of Gut Microbiome in Patients With Schizophrenia Indicates Links Between Bacterial Tyrosine Biosynthesis and Cognitive Dysfunction. *Biological psychiatry global open science*. avr 2023;3(2):283-91.
16. Kanehisa
M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 1 janv 2000;28(1):27-30.
17. Huerta-Cepas
J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res*. 4 janv 2016;44(D1):D286-93.
18. Haft
DH. The TIGRFAMs database of protein families. *Nucleic Acids Research*. 1 janv 2003;31(1):371-3.
19. Buchfink
B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. janv 2015;12(1):59-60.
20. Vieira-Silva
S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S, et al. Species–function relationships shape ecological properties of the human gut microbiome. *Nat Microbiol*. 13 juin 2016;1(8):16088.
21. Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, et al. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol*. 4 févr 2019;4(4):623-32.