



HAL
open science

Highlighting discriminating functional annotations in patients' gut microbiota through Machine Learning classification applied to functional descriptions of the microbiome

Baptiste Ruiz, Arnaud Belcour, Samuel Blanquart, Isabelle Luron Le Huërrou-Luron, Sylvie Buffet-Bataillon, Yann Le Cunff, Anne Siegel

► To cite this version:

Baptiste Ruiz, Arnaud Belcour, Samuel Blanquart, Isabelle Luron Le Huërrou-Luron, Sylvie Buffet-Bataillon, et al.. Highlighting discriminating functional annotations in patients' gut microbiota through Machine Learning classification applied to functional descriptions of the microbiome. International Conference on Clinical Metagenomics, Oct 2022, Geneve, Switzerland. hal-04155604

HAL Id: hal-04155604

<https://hal.inrae.fr/hal-04155604v1>

Submitted on 7 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Highlighting discriminating functional annotations in patients' gut microbiota through Machine Learning classification applied to functional descriptions of the microbiome

Baptiste RUIZ¹, Arnaud BELCOUR¹, Samuel BLANQUART¹, Isabelle LE HUËROU-LURON², Sylvie BUFFET-BATAILLON^{2,3}, Yann LE CUNFF¹ and Anne SIEGEL¹

1 Univ Rennes, Inria, CNRS, IRISA, Campus de Beaulieu, 263 Av. Général Leclerc, 35042, Rennes, France

2 Institut NuMeCan, INRAE, INSERM, Univ Rennes, Rennes-Saint-Gilles, France

3 Department of Clinical Microbiology, CHU Rennes, Rennes, France

Presenter: baptiste.ruiz@inria.fr

1 Background

Understanding the gut microbiota and its mechanisms has become a major point of interest in the medical field, with more and more studies correlating it to a variety of pathologies. Beyond information relative to its microbial composition, the focus is currently shifting towards the exploration of the microbiota on a mechanistic scale. Machine Learning methods have been applied to this issue, approaching the microbiome as a predictor of the subjects' health. From these works, the micro-organisms' functional annotations have been suggested as a promising supplement to augment the data and enhance the comprehension of the microbiota as a metabolic network.

2 Methods

In line with this approach, we propose a new method to shift the representation of the gut microbiota from relative OTU abundances to a numeric mapping of the associated functional annotations, creating a mechanistic description of the microbial community. We have then explored the application of Random Forest classifiers, a classic Machine Learning approach for microbiota classification, to publicly available data converted to this new paradigm.

3 Results

This approach allowed us to automatically highlight important metabolic mechanisms that characterised the differences between the microbiotas of healthy and unhealthy individuals in the context of a variety of diseases (IBD, Obesity, Type 2 Diabetes, Cirrhosis, CRC). Furthermore, we have shown that it is possible to improve classification performances after selecting these important variables, allowing for OTU and abundance profiles to both output equivalent metrics.

4 Discussion

Exploiting this method yields more thorough and complete results than what can be gathered through the standard approach based on finding OTUs that make a difference between classes of subjects. Though these outputs' overall relevance should still be overviewed in light of the relevant bibliography, they could provide insight on physiopathologies that could be exploited to develop or complement treatments.