A multi-omic Knowledge Graph targeting inflammatory bowel disease (IBD) biomarker discovery

samuel.chaffron@univ-nantes.fr, audrey.bihouee@univ-nantes.fr, alban.gaignard@univ-nantes.fr

Nantes, décembre 2021

Keywords: IBD, gut microbiome, multi-omics data, knowledge graph.
Skills: good programming skills, knowledgeable in semantic web technology, curiosity.
Gratification: 15 % du plafond horaire de la sécurité sociale, environ 540€/mois.

Context

Inflammatory Bowel Diseases (IBDs) are multifactorial chronic conditions of the gastrointestinal tract, of which the most common are Crohn's disease (CD) and Ulcerative Colitis (UC). These disorders are notably associated with the environment, the patient's genetics and its gut microbiota [1]. The microbiota is known to play an important role in host's health by protecting them from pathogens [2], helping digest and absorb nutrients [3], and shaping the immune system [4]. Although a core group of essential bacteria is shared across individuals [5], the gut microbiota contains individual-specific strains stable over time, which makes it harder to identify bacteria linked to the host, the environment, or a given pathology.

Today, various 'omics' technologies are available to measure molecules at all levels of cellular organisation in complex microbial communities. DNA sequences can be determined (metagenomics), transcripts levels can be measured (metatranscriptomics), metabolites can be detected (metametabolomics), and proteins can be catalogued and quantified (metaproteomics). Computational biology advances enabling the description of environmental genomes and their expression in situ have accompanied these new technologies [6]. The field of environmental omics (or meta-omics) has drastically expanded our knowledge about microbial communities [7], prompting a change of paradigm in which not a single species is considered but rather a complete microbial community. The importance of ecological interactions among microorganisms is also now recognized, and they need to be included in a global framework to further develop models of community eco-systems functioning [8].

Challenges

Multi-omics technologies in gut microbiome research provide a global view of changes in genetic, metabolic and biochemical processes, and have recently been applied to the gut microbiota in the context of IBD [9]. These data have provided a first comprehensive view of functional dysbiosis in the gut microbiome during IBD activity and first steps towards a functional understanding of host–microbe interactions during disease pathogenesis. However, major challenges remain to systematically integrate microbial, biochemical, and host factors in order to identify systems-level biomarkers, from genes to communities, of IBD onset and progression.

Objectives

The aim of this project is i) to **develop a workflow for the graph-based transformation and integration of meta-omics data** (metagenomic and metatranscriptomic), and ii) **feed an integrated**

knowledge graph connecting longitudinal genomic observations as well as phenotypes and environmental context. Through a catalogue of semantic queries, this knowledge graph will empower biologists to more easily assemble genome co-activity networks towards identifying novel systems-level (from genes to communities) IBD biomarkers and patient stratification strategies.

Roadmap

- Task 1 will consist in reviewing literature and ontology-based standards to represent and share longitudinal observations on metagenomics and metatranscriptomics data. In line with biologist expectation, motivating research questions will be formalised into graph patterns.
- Through the development of computational workflows, Task 2 will address the transformation of meta-omics observations into RDF knowledge graphs following the recommendations of Task 1.
- Task 3 will consist in serving a query-able knowledge graph leveraging graph database systems such as Neo4J or GraphDB.
- Finally Task 4 will provide IBD biologists and clinicians with a catalogue of queries targeting both (i) the assembly of co-activity matrices feeding predictive models and (ii) the assembly of networks linking host phenotypes, physiological parameters and their microbiomes.

Candidate profile and working environment

We are looking for an enthusiastic bioinformatician or data scientist, with excellent programming skills, and real interest in Life Science applications. The recruited data scientist will work in the context of an interdisciplinary collaboration, involving health, computational research teams and the Bioinformatics core facility, under the supervision of Audrey Bihouée (Engineer at BiRD facility), Samuel Chaffron (CNRS Researcher at LS2N) and Alban Gaignard (CNRS Research Engineer at Institut du Thorax). The developments resulting from this internship will constitute a framework for a generalisation of the model to other pathologies.

References

[1] Wlodarska, M., Kostic, A.D., & Xavier, R.J. (2015). An integrative view of microbiome-host interactions in inflammatory bowel diseases. Cell host & microbe, 17 5, 577-91 . https://pubmed.ncbi.nlm.nih.gov/25974300/

[2] Stecher, B., & Hardt, W. D. (2011). Mechanisms controlling pathogen colonization of the gut. Current opinion in microbiology, 14(1), 82–91. https://doi.org/10.1016/j.mib.2010.10.003

[3] Walter, J., & Ley, R. (2011). The human gut microbiome: ecology and recent evolutionary changes. Annual review of microbiology, 65, 411–429. https://doi.org/10.1146/annurev-micro-090110-102830

[4] Jarchum, I., & Pamer, E. G. (2011). Regulation of innate and adaptive immunity by the commensal microbiota. Current opinion in immunology, 23(3), 353–360. https://doi.org/10.1016/j.coi.2011.03.001

[5] Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M. J., Valles-Colomer, M., Vandeputte, D., Tito, R. Y., Chaffron, S., Rymenans, L., Verspecht, C., De Sutter, L., Lima-Mendez, G., D'hoe, K., Jonckheere, K., Homola, D., Garcia, R., … Raes, J. (2016). Population-level analysis of gut microbiome variation. Science (New York, N.Y.), 352(6285), 560–564. https://doi.org/10.1126/science.aad3503

[6] Segata, N., Boernigen, D., Tickle, T. L., Morgan, X. C., Garrett, W. S., & Huttenhower, C. (2013). Computational meta'omics for microbial community studies. Molecular systems biology, 9, 666. https://doi.org/10.1038/msb.2013.22

[7] Waldor, M. K., Tyson, G., Borenstein, E., Ochman, H., Moeller, A., Finlay, B. B., Kong, H. H., Gordon, J. I., Nelson, K. E., Dabbagh, K., & Smith, H. (2015). Where next for microbiome research?. PLoS biology, 13(1), e1002050. https://doi.org/10.1371/journal.pbio.1002050

[8] Raes, J., & Bork, P. (2008). Molecular eco-systems biology: towards an understanding of community function. Nature reviews. Microbiology, 6(9), 693–699. https://doi.org/10.1038/nrmicro1935

[9] Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., Andrews, E., Ajami, N. J., Bonham, K. S., Brislawn, C. J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T. G., Hall, A. B., Lake, K., Landers, C. J., Mallick, H., Plichta, D. R., Prasad, M., … Huttenhower, C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature, 569(7758), 655–662. https://doi.org/10.1038/s41586-019-1237-9