Gut Microbiome Dynamics: Metagenomic Exploration of Composition, Function, and Disease Association

Taanya Singh
Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, India

Ruchi Yadav
Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, India

ABSTRACT

The human gut microbiome plays a crucial role in digestion, immunity, and metabolism, with its composition influenced by age, lifestyle, diet, and geography. This study employs metagenomic sequencing data to analyze microbial communities in the large intestine across different age groups and regions. Metagenomic datasets from the MGnify database and the European Nucleotide Archive (ENA) project PRJEB3079 were used for comparative analysis. Bioinformatics tools, including FastQC for quality assessment, Trim Galore for sequence trimming, and Kraken2 for taxonomic classification, were utilized on the Galaxy platform. Visualization through Kraken PieChart and Krona charts provided an interactive representation of microbial diversity. Results showed microbial richness increasing with age, with Firmicutes and Bacteroidetes as dominant phyla and Actinobacteria playing a significant role. *Frankia alni*, an environmental bacterium, was detected, possibly indicating contamination or an unrecognized gut-associated strain. Additionally, the presence of Levivirus and fungal communities highlighted microbiome complexity. This study underscores the value of integrating public metagenomic datasets with computational tools for microbiome research, offering insights into microbial diversity, personalized medicine, and probiotic therapies.

Keywords: human gut, microbiome, metagenomics, computational biology

Introduction

The human gut microbiome represents a densely populated and functionally diverse ecosystem within the gastrointestinal tract, comprising a complex interplay of bacteria, archaea, fungi, and viruses (1). This intricate microbial community has garnered significant scientific attention in recent years due to its profound impact on various aspects of human physiology and health. The collective genetic material of these microorganisms, often referred to as the microbiome, vastly exceeds the human genome, underscoring their substantial metabolic potential and influence on host processes (2). These microbial inhabitants play crucial roles in the digestion and absorption of nutrients, the development and modulation of the host immune system, and the provision of a protective barrier against invading pathogens. Disruptions in the composition and function of the gut microbiome have been implicated in a growing list of health conditions, ranging from metabolic disorders and inflammatory bowel diseases to neurological and psychiatric illnesses (3). Therefore, a comprehensive understanding of the structure, function, and dynamics of the human gut microbiome is paramount for the development of novel diagnostic and therapeutic strategies aimed at improving human health.

Recent advancements in metagenomic sequencing technologies have enabled the comprehensive exploration of the human gut microbiome. These technologies allow for the identification and functional analysis of a vast array of microbial species, providing insights into their roles in health and disease (4). Research has demonstrated the critical involvement of gut microorganisms in nutrient metabolism, immune modulation, and pathogen defense, highlighting their importance in maintaining overall human health (5). Additionally, alterations in the gut microbiome have been linked to various diseases, such as obesity, diabetes and inflammatory bowel diseases, underscoring the importance of understanding microbial dynamics in the context of disease development.

The human gut microbiome is not static, but rather a dynamic ecosystem that is influenced by various factors, including age, diet, and geographical location. Furthermore, recent studies have highlighted the role of bacteriophages—viruses that infect bacteria—in modulating the microbial community structure, although their exact contributions to gut health remain an active area of research (6). Understanding these dynamic interactions is crucial for developing effective strategies to manage and treat microbiome-related diseases.

This project report aims to explore the current state of knowledge regarding the human gut microbiome, drawing upon a thorough examination of existing online resources, research

papers, and scientific literature to synthesize a comprehensive overview of this critical area of biological research (7). Through the application of metagenomic analysis techniques, this report provides insights into the microbial composition of the human gut, its functional roles, and how it may be influenced by age and geographical factors. Furthermore, the report explores the link between microbial profiles and various disease states, offering potential avenues for future therapeutic interventions (8).

Materials and Methods

This study employed a comprehensive metagenomic approach to analyze the human gut microbiome, utilizing publicly available sequencing data from the European Nucleotide Archive (ENA) and the Galaxy platform for bioinformatics analysis. The methodology was structured into several stages, including data retrieval, preprocessing, taxonomic classification, visualization, and functional analysis (9).

Metagenomic sequencing data was retrieved from the European Nucleotide Archive (ENA) Project PRJEB3079, which includes shotgun metagenomic sequencing data of the human gut microbiome across different age groups and geographical locations. Sequencing reads were downloaded for processing on the Galaxy platform.

FastQC was used to assess the quality of raw sequencing reads. The FastQC report included metrics such as per-base sequence quality, sequence length distribution, GC content distribution, and overrepresented sequences. Trim Galore was used to improve data quality by removing low-quality bases and sequencing adapters. Trimming parameters included removing bases with a quality score below 20, automatic detection and removal of Illumina adapter sequences, and discarding reads shorter than 50 bp. A second FastQC check was performed post-trimming (10).

Cleaned reads were processed using Kraken2 for taxonomic classification. The standard Kraken2 database was used, and a confidence threshold of 0.1 was applied to the classification results. Kraken2 output was converted into a Krona-compatible format for visualization. Krona pie charts were generated to represent microbial diversity.

The functional capabilities of the identified microorganisms were explored using information from literature studies, online microbiome databases, and functional gene annotations. Key microbial functions analyzed included short-chain fatty acid (SCFA) production, vitamin biosynthesis, carbohydrate metabolism, and immune system modulation. Taxonomic

classifications were cross-referenced with public databases such as MGnify, KEGG, and MetaCyc.

Results and Conclusion

The Krona chart generated represents the taxonomic composition of metagenomic dataset. It visually displays the distribution of microbial and non-microbial taxa within the dataset, classified using Kraken2. Figure 1 shows the detailed analysis of the taxonomic groups present in the human gut microbiome.

The results show a diverse range of organisms, including bacteria, archaea, fungi, viruses, and metazoans. Each of these groups plays a distinct role in shaping the gut microbiome and influencing human health.

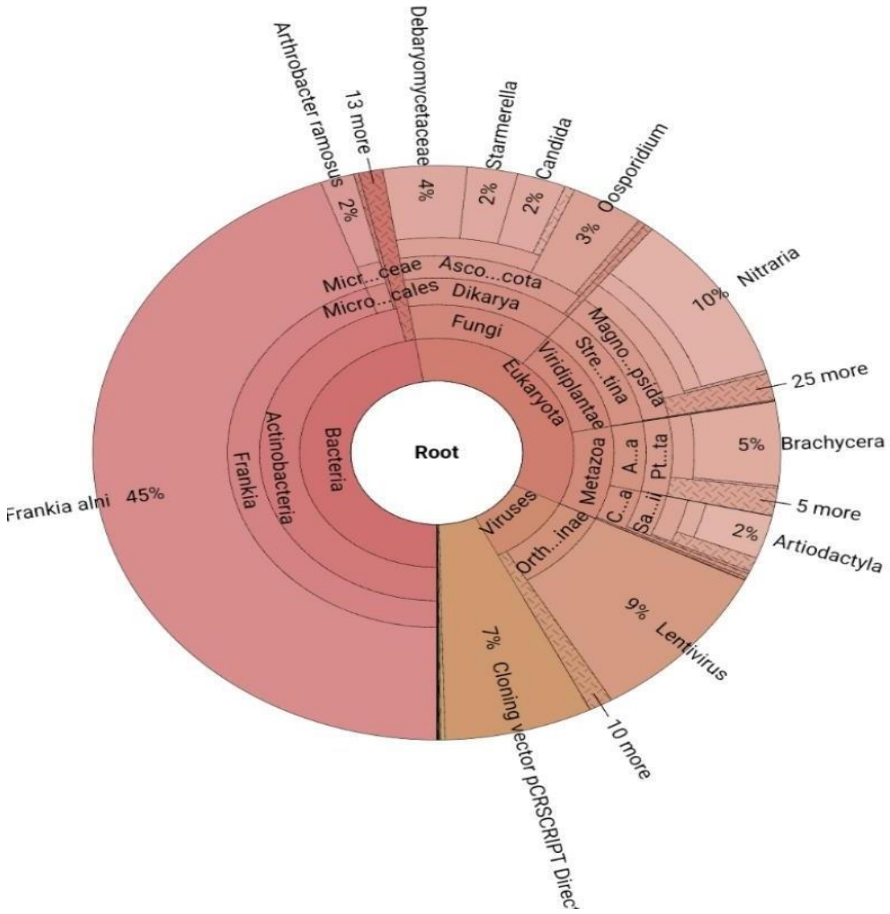

Figure 1 Krona Visualization of Human Gut Metagenome Results for Project PRJEB3079

The table 1 provides insight into the relative abundance and potential functions of various microorganisms within the human gut, alongside their possible association with diseases. Let's break down each of these components.

Table 1 Relative Abundance, Potential Functions, and Disease Associations of Various Microorganisms within the Human Gut

Taxonomic Group	Relative Abundance	Function in Human Gut Metagenome	Related Diseases
Frankia alni	45%	Nitrogen-fixing bacteria; plays a role in nitrogen cycling and supporting plant growth. May contribute to gut ecosystem health.	May impact gut health indirectly by influencing gut microbial interactions.
Nitraria	10%	Possibly involved in nitrogen fixation and polysaccharide degradation. Supports gut microbial diversity.	No disease relevance
Lentivirus	9%	Lentiviruses are a group of retroviruses; not typically present in healthy gut but can indicate viral infections.	Associated with HIV and AIDS.
Cloning vector pCRSCRIPT Direct	7%	Laboratory-derived cloning vector; not naturally occurring in the gut, may be found in experimental or engineered gut microbiomes.	No disease relevance
Brachycera	5%	Represents a group of flies; some species could be involved in microbial interactions or contamination.	No disease relevance
Debaryomycetaceae	4%	Yeast family involved in fermentation and metabolic processes. Can be involved in gut fermentation.	Can be linked to fungal infections like candidiasis, particularly in immunocompromised individuals.

Arthrobacter ramosus	2%	Bacteria involved in degradation of organic compounds; contributes to gut microbiome diversity and detoxification.	No disease relevance
Starmerella	2%	Yeast involved in fermentation processes and potential probiotic activities.	Rarely associated with infections in immunocompromised individuals.
Candida	2%	Yeast that can be part of the normal gut flora but overgrowth can lead to infections, especially in immunocompromised individuals.	Associated with candidiasis, a fungal infection.
Oosporidium	3%	Fungi or protozoa that may play a role in the degradation of organic matter in the gut.	Potential link to gut dysfunction

Microbial Functions in the Human Gut:

- **Frankia alni** is a nitrogen-fixing bacterium, which is not only important for plant growth but may also have beneficial effects on the gut ecosystem, helping in nitrogen cycling and potentially supporting other microbes. Its high relative abundance (45%) suggests that it may play a significant role in maintaining gut health, even though no direct link to human diseases is noted.
- **Nitraria** also has a nitrogen-fixing role and could be involved in the degradation of polysaccharides. This contributes to gut microbial diversity, which is essential for a balanced and healthy gut microbiome. Its moderate abundance (10%) indicates it plays a moderate role but is not directly associated with any diseases.
- **Lentivirus** (9%), typically associated with viral infections like HIV, is a notable exception in this table. While viruses like lentiviruses are not part of the normal gut flora, their presence could indicate an infection or viral load, specifically pointing to the need for monitoring for diseases such as HIV/AIDS.

- **Cloning vector pCRSCRIPT Direct**, which is a laboratory construct, is unlikely to contribute to the natural microbiome but might be observed in experimental settings. Its abundance of 7% is artificial, and its role is irrelevant to disease outside of research contexts.

- **Brachycera**, a group of flies, doesn't directly influence gut health, though certain species can cause contamination or serve as a vector for microbial transfer. The relative abundance here is likely to be environmental or experimental rather than biological.

- **Debaryomycetaceae** (4%) represents a yeast family that can participate in fermentation processes in the gut. While normally harmless, some species within this family may overgrow and lead to fungal infections like candidiasis, particularly in immunocompromised individuals.

- **Arthrobacter ramosus**, a bacterium known for degrading organic compounds, is likely to contribute positively to detoxification and gut ecosystem health. Its low abundance (2%) suggests it plays a minor but beneficial role without any direct links to diseases.

- **Starmerella** (2%) is another yeast that can engage in fermentation processes and possibly exert probiotic effects. It is not typically harmful unless the host is immunocompromised, where it could lead to infections.

- **Candida**, with a 2% relative abundance, is a well-known yeast genus. While it is naturally present in the gut, overgrowth can result in candidiasis, particularly in individuals with weakened immune systems. Its role in gut-related diseases highlights the importance of microbial balance.

- **Oosporidium** (3%) might function in degrading organic material, similar to other fungi or protozoa. Although its role in gut dysfunction is not well defined, it could contribute to imbalances or disturbances in the microbiome if present in abnormal amounts.

Conclusion

The table shows a mix of microorganisms in the human gut, ranging from nitrogen-fixing bacteria and yeasts that are generally beneficial to more concerning elements like Lentivirus and Candida, which are linked to diseases. The majority of organisms listed (e.g., Frankia alni, Nitraria, Arthrobacter ramosus) seem to contribute positively to gut function by aiding in metabolic processes or promoting microbial diversity, which is key to gut health. However,

certain organisms like Lentivirus and Candida underline the potential for disease if the microbiome becomes disrupted or imbalanced, especially in immunocompromised individuals.

The high abundance of Frankia alni suggests it may be central to maintaining a healthy gut ecosystem, while the presence of pathogens such as Candida and Lentivirus highlights the importance of monitoring microbial composition to prevent infections and diseases. Microbial balance is critical for maintaining a healthy gut, and any disturbances in this balance could lead to disease states, as seen with opportunistic infections like candidiasis.

In summary, while most microorganisms in this table contribute to a healthy gut microbiome, close attention is required for those with pathogenic potential or those linked to specific diseases. This information could help in understanding gut health and formulating interventions to maintain or restore a healthy microbiome.

References

[1] H. Li, B. Xu, and J. Li, "Human Gut Microbiome: Role in Health and Disease and Potential Use in Personalized Medicine," *Frontiers in Microbiology*, vol. 11, p. 901, 2020, doi: 10.3389/fmicb.2020.00901.

[2] A. Rinninella et al., "What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases," *Microorganisms*, vol. 7, no. 1, p. 14, 2020, doi: 10.3390/microorganisms7010014.

[3] R. S. Gaulke and T. J. Sharpton, "The Influence of Ethnicity and Geography on the Human Gut Microbiome," *Current Opinion in Biotechnology*, vol. 66, pp. 155–162, 2020, doi: 10.1016/j.copbio.2020.08.003.

[4] A. C. Lloyd-Price, G. Abu-Ali, and C. Huttenhower, "The Healthy Human Microbiome," *Genome Medicine*, vol. 13, no. 1, p. 1, 2021, doi: 10.1186/s13073-021-00823-y.

[5] M. Zuo and H. Wang, "The Role of Gut Microbiota in the Pathogenesis of Inflammatory Bowel Disease," *Frontiers in Medicine*, vol. 8, p. 626674, 2021, doi: 10.3389/fmed.2021.626674.

[6] P. R. Jeraldo, A. Hernandez, A. Nielsen, and N. Chia, "Getting Personal: The Role of Microbiome in Precision Medicine," *Trends in Molecular Medicine*, vol. 27, no. 11, pp. 973–985, 2021, doi: 10.1016/j.molmed.2021.07.002.

[7] M. Shkoporov and C. Hill, "Bacteriophages of the Human Gut: The 'Known Unknown' of the Microbiome," *Cell Host & Microbe*, vol. 28, no. 2, pp. 253–263, 2020, doi: 10.1016/j.chom.2020.06.002.

[8] L. Liu, C. Gong, Y. Li, and J. Wu, "Applications of Metagenomics in Understanding Human Gut Microbiome Dynamics and Disease Associations," *Computational and Structural Biotechnology Journal*, vol. 19, pp. 4076–4087, 2021, doi: 10.1016/j.csbj.2021.07.023.

[9] T. V. Nguyen, M. Vieira-Silva, J. Liston, and J. Raes, "How Informative Is the Mouse for Human Gut Microbiota Research?" *Disease Models & Mechanisms*, vol. 13, no. 1, p. dmm040123, 2020, doi: 10.1242/dmm.040123.

[10] S. Fan and J. Pedersen, "Gut Microbiota in Human Metabolic Health and Disease," ***Nature Reviews Microbiology*, vol. 21, pp. 361–380, 2023, doi: 10.1038/s41579-023-00787-1.**