Current Trends in Biotechnology and Pharmacy
Vol. 18(1) 1616-1628, January 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
DOI: 10.5530/ctbp.2024.1.10

Human Metagenome Analysis with COVID-19 Infectious Disease

Kashish Jain and Ruchi Yadav*

Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, UP, India

Corresponding author: ryadav@lko.amity.edu

Abstract

Coronavirus disease (Covid-19) is an infectious disease which is caused by a virus named as SARS-COV. This virus belongs to the Coronavirus family that causes a variety of diseases including head and chest colds, severe acute respiratory syndrome (SARS), and Middle East respiratory syndrome (MERS). Metagenomics is the study of genetic material that is recuperated from the environmental samples directly. By examining and elucidating microbial genomes in healthy and infected samples, the metagenomics branch has allowed us to discover the significance of microbial genomes. The metagenomic approach used in the covid-19 infectious disease will serve as a prominent tool for explicating the relationship between host-associated microbial communities and host phenotype. European Nucleotide Archive (ENA) was used to retrieve the metagenome data of the Gut Microbiome of Covid-19 patients with the project id PRJDB12349. For metagenomics analyses, Galaxy server was used identification and classification of microbial communities furhter taxonomic analysis and functional analysis was also performed. Different tools from the galaxy like FastQC, MultiQC, Trim Galore, Megahit, Kraken 2, Convert Kraken, Krona Pie Chart and HUMANn2 were used for complete metagenomic analysis. Metagenomic Analysis shows that there is difference in the percentage and abundance of bacteria in all the covid-19 patients sample. The fungi that were commonly present in all the samples were identified as Eurotiales 78-93% and Mycospaerellales 7-20% and the other fungi that were not common but were observed are Dipodasceae 10% and Sordariales 4-6%. Further, the gene abundance and the families of the samples were also observed. This study highlights the metagenome of the Covid-19 infectious disease. The metagenomic analysis of the gut microbiome of the Covid-19 patients shows difference in the microbial community among all the samples.

Keywords: Human Metagenome, Galaxy server, Gut Microbiome, Krona Pie Chart, COVID-19.

Introduction

As a result of the Coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) a global pandemic has been unleashed, with subsequent restrictions on public and private life. On March 11, 2020, the World Health Organization proclaimed Coronavirus disease 2019 (COVID-19) a pandemic (1). This disease has essentially been elucidated as a respiratory illness, but the growing affirmations suggests

that the gastrointestinal tract may also be involved (2). Due to this phenomenon, clinical trials have been challenged worldwide, but there are also distinctive suggestions for research into the gut microbiome, which is promptly progressing (3). Microorganisms and genes colonizing the gastrointestinal tract are known collectively as the gut microbiome, and it is estimated that there are 100 billion or more microorganisms (4). In addition to playing an essential role in the immune system and metabolism, they are also considered preeminent in the pathogenesis of diseases due to their role in the immune and metabolic pathways (5). A COVID-19 infection has an expansive spectrum of clinical manifestations, from asymptomatic illness accompanied by cough and fever to severe acute pneumonia with multi-organ failure and acute respiratory distress syndrome (ARDS). As part of the cytokine cascade associated with ARDS, interleukin (IL), granulocyte colony-stimulating factor (G-CSF), and tumor necrosis factor are all implicated. A previous look at the patients revealed that intensive care unit (ICU) sufferers with COVID-19, including ARDS, had an abundance of proinflammatory cytokines, which includes IL-2, IL-7, IL-10, GCSF, IP10, MCP1, MIp1A, and TNFα, in comparison to non-ICU patients. These pro-inflammatory cytokines have been proclaimed to correspond with distinct patterns of the gut microbiota (6).

The human body's greatest immunological organ is the gastrointestinal tract (GIT) (7). Recent indications from other respiratory diseases suggests that the gut microbiota influences lung immunity and inflammation (8). The host and the gut bacteria coexist in symbiotic harmony. It promotes vitamin production, carbohydrate and other undigested food fermentation, and assists in the supplying colonic epithelial cells with vital nutrients like short-chain fatty acids (SCFAs) (9). Moreover, it controls mucosal permeability and acts as a deterrent for harmful bacteria(10). More crucially, by modifying the host's local and systemic immune responses, the microbiota plays a crucial role in maintaining intestinal homeostasis (11). Through whole-genome sequencing (WGS) approaches, several SARS-CoV-2 variants have been identified, including variants of concern (VOC) Alpha, Beta, Gamma, Epsilon, and Delta (12).

Currently, booming evidences stipulates that the gut microbiome ecology in patients with COVID-19 is altered and that gut microflora configurations are related to immune responses. As a result of SARS-CoV-2 infection, there is a significant impact on the environment and dynamics of the human gut microbiome over the short and long term, which directly affects the health of the patient (13). The gut environment that is infected by the active SARS-CoV-2 virus and affected microbiome may also lead to an antagonistic environment, which facilitates the bloom of pathogenic bacteria and fungi, further limiting the formation and function of the gut microbiome, thereby crippling the immunity of the host (14).

The study of genetic material directly extracted from environmental or clinical samples using a technique called sequencing is known as metagenomics (15). In this study, the genomic composition of an entire organism is scrutinized, as well as the genomic configuration of each microbe that inhabit within it (16). An important approach is the notion that microbes and hosts are mutually dependent and should be perceived as a community rather than contemplated to be different entities. All life on earth depends on microorganisms because they play an essential role in transforming key elements of life (e.g. carbon, nitrogen, oxygen, sulphur) into forms that can be used by living organisms. Because of the innumerable microbes and their own gene complements that live inside of them, mammals are metagenomic (17). When compared to other microbiota members, bacteria are more numerous and consequently have a greater impact on mammal gut health, well-being and disease. Despite being essential to human and animal health, the overwhelming majority of the microbe population that resides within humans

has not been distinguished or studied crucially (18). In the field of metagenomics, this opens up advanced possibilities. Utilizing metagenomics, bioinformatics analysis is an important step in the process of discovering and defining novel microorganisms in the samples of the sufferer suffering from coronavirus disease (19).

The more reliable approach is laboratory-based surveillance, where a variety of techniques can be used to find and confirm the presence of pathogens. When a causal agent cannot be found, samples in the field of conventional pathogen detection are frequently left undiagnosed (20). Metagenomics has the ability to completely redesign this process by detecting the presence of all microorganisms with a single sequencing technique and without the requirement for culture, as opposed to utilising specialised tests to look for a specific pathogen (21). Metagenomics is relatively new because microbes have consistently been studied in laboratories rather than within hosts. During the COVID-19 pandemic, the use of metagenomic methods has been crucial for effective outbreak control and traceback. As a result, our latest understanding of microbes in their natural habitat is limited. A key goal of metagenomics is the improved understanding of environmental and clinical microbes, despite significant barriers such as the difficulty of making microbes and their diversity genomically (22). It is anticipated that improved understanding about the essence of microbes in the surrounding could have a great impact on other research-based and technological areas biotechnology, ecology, medicine etc.

In the early stages of a pandemic, an unbiased metagenomics approach to virus identification may be crucial. With the help of metagenomic analysis all the microorganisms can be discovered and recognised that are present in the gut microbiome samples of the covid-19 patients through analysis of the genomic data obtained from these samples (23). To find SARS-CoV-2 mutations and respiratory co-pathogens possibly causing COVID-19 symptoms, metagenomics can be used. It will also provide knowledge and understanding about the species that are present in the sample of the covid-19 patients as well as it allows to isolate the information concerning the functionality and the performance of the microbial communities in their natural habitat (24). An additional and effective approach for developing new theories about how bacteria function is metagenomics. Phylogenetic surveys, which frequently only consider the diversity of one gene, are remarkably more confined in scope than metagenomics, which enables access to the functional gene composition of microbial communities (25).

The metagenomic approach used in the covid-19 infectious disease will serve as a prominent tool for explicating the relationship between host-associated microbial communities and host phenotype (26).

Materials and Methods

Data collection

I have retrieved the data from some of the nucleotide sequence databases such as European Molecular Biology Laboratory(EMBL) (https://www.ebi.ac.uk/) and European Nucleotide Archive Database (ENA) (https://www.ebi.ac.uk/ena/browser).

Data analysis

For analysis of data some Galaxy Tools (https://usegalaxy.org/ , https://usegalaxy.eu/ and https://huttenhower.sph.harvard.edu/galaxy/) were used such as FastQC, MultiQC, Trim Galore!, Megahit. For Taxonomic Analysis of the data tools such as Kraken2, Convert Kraken, Krona pie Chart were used and for Functional analysis of the data the tool HUMANn2 was used.

Retrieval of the metagenomic data

The raw paired-end data was retrieved from a database known as ENA (European Nucleotide Archive).

Table 1 Metagenome data of the gut Microbiome of Covid-19 patients retrieved from ENA Database. FastQ files of all the five samples were used for metagenomic analysis.

S.No.	Sample Accession	Experiment Accession	Run Accession	Scientific Name	FASTQ Files
1	SAMD00408333	DRX317592	DRR328588	Human gut metagenome	DRR328588_1.Fastq.gz
2	SAMD00408413	DRX317593	DRR328589	Human gut metagenome	DRR328588_2.Fastq.gz
3	SAMD00408415	DRX317594	DRR328590	Human gut metagenome	DRR328589_1.Fastq.gz
4	SAMD00408417	DRX317595	DRR328591	Human gut metagenome	DRR328589_2.Fastq.gz
5	SAMD00408419	DRX317596	DRR328592	Human gut metagenome	DRR328590_1.Fastq.gz

The instrument model for sequencing was Illumina Miseq Paired- end Sequencing, organism: Human gut metagenome (Taxonomy ID- 408170). The project ID of the metagenomic samples is PRJDB12349.

Methodology for NGS data analysis

Figure 1:Workflow of Metagenomic Data Analysis

Computational softwares and tools are used for the Metagenomic analysis of the gut microbiome of the Covid-19 patients. Figure1 shows the structural outline used for the metagenomic data analysis for taxonomic and functional analysis.

Pre-processing of metagenomic data

FastQC

Quality control analysis (FastQC) is performed on Fastq files. Fastq files are taken from a web-based platform ENA. The main purpose of FastQC is to provide a QC report which can detect the problems which emerge either in the starting library materials or in the sequencers (27). This tool checks the quality of raw sequence data coming from high throughput sequencing pipelines in a simple and straightforward way. This is achieved by running a set of standard analyses on one or more raw sequence files either in the BAM or fastq format (28). This is further followed by a report compiling the results and underlining any anomalies in the library's operations. It also provides summary graphs and tables for rapid computation of data and exports the results in HTML form.

MultiQC

After performing FastQC, MultiQC was

performed on the samples. MultiQC is a modular tool that is used to combine the results of bioinformatics analyses to be viewed in a single report based on a variety of samples (29). It is basically an outreach tool that interpret the summary statistics from results and log files generated by other computational tools. An HTML report is generated by multiQC which can be shared and observed in any of the modernized web browsers (30). By elongating and exporting data, MultiQC acts as a central collection point for analysis pipeline.

TrimGalore

Now Trim Galore is used after MultiQC for data cleaning. Trim Galore! is a shell script to imbrute quality and adapter trimming along with quality control, with some added efficacity to remove bigoted methylation positions for RRBS Sequence files for single-end and paired-end sequencing (31). If a sequence becomes too short during trimming, Trim Galore! Can remove it. If one or both of the reads are shorter than a set length cut-off for paired-end files, Trim Galore! removes both sequence pairs. Fastq files can be produced in standard or gzip compressed format By Trim Galore! (32).

Megahit

Now after data cleaning MEGAHIT is used as data assembler. Megahit is a brisk and sustainable anew assembly tool for next generation sequencing data that assembles large and complicated metagenomic datasets (33). It is a high-speed and memory efficient metagenome assembler. In addition to metagenome assembly MEGAHIT also works on generic single genome assembly (small or mammalian size) and single cell assembly (34). To achieve low memory usage, MEGAHIT uses Succinct de Bruijn graph.

For taxonomic analysis

Kraken2

After using Megahit, taxonomic analyses are to be done by Kraken 2. Kraken 2 was executed on all the five samples that have been collected. Kraken 2 is the advanced version of Kraken. By using this tool, it is possible to achieve high fidelity and fast categorization speeds in taxonomic classification by using accurate k-mer matches (35). By using this classifier, all k-mers within a query sequence are coordinated to the lowest common ancestor of all genomes that contain the specified k-mer. The classification accuracy of Kraken 2 is further improved by the use of spaced seeds for stowing and querying minimizers. Kraken 2 also gives assistant and support to databases that are not entrenched on NCBI's taxonomy (36).

Convert κraken

This tool is performed followed by Kraken 2. Convert Kraken is used separately on all the outputs of Kraken 2. To convert Kraken metagenomic classification results in the form of NCBI's taxonomic representations, convert Kraken is used (37). This is done by utilizing taxonomic ID field provided by Kraken 2. The output of this tool can be instantly envisaged by Krona tool.

Krona pie chart

Now the next step after performing Convert Kraken is to perform metagenomic profiling. For this, a tool is required which is known as Krona Pie Chart. This tool is both an authoritative metagenomic envision tool and a display of the capable HTML5 for eminently attainable bioinformatics visualization (38). Krona is used to demonstrate zoomable pie charts that are derived from metagenomic profiling. It authorizes stratified data to be scrutinized by means of multi-layer pie charts. The associated charts are self-governing and can be visualized using any modernized web-browser.

For functional analysis

HUMANn2

Now for functional analysis of the samples another tool is used that is HUMANn2. This tool is more perceptive, accurate and effectual than other existing methods. HUMANn2 is a pipe-

line for proficiently and correctly describing the presence/absence and affluence of microbial passages in an association from metagenomic and meta transcriptomic sequencing data (39). HUMANn2 tool generates three files the abundance of gene families, abundance of pathways and coverage of pathways. This tool also provides a simple user interface.

Results and Discussion

Quality control analysis of metagenomic data

The quality control analysis was done on 4 paired-end Fastq files i.e., Human Gut Metagenome from the database known as ENA. The tool FastQC of an open source platform known as Galaxy was used for quality assessment of data (40). An HTML output file and basic text is generated by FastQC tool which conceals the quality control plots which incorporates per base sequence quality, basic statistics, per base sequence content, per sequence quality scores, per base GC content, per sequence GC content, per base N content, sequence, sequence duplication levels, overrepresented sequences, kmer content. After this, MultiQC was executed on raw FastQC data files to accumulate the outcomes from the four FastQC done on the four FastQC raw data.

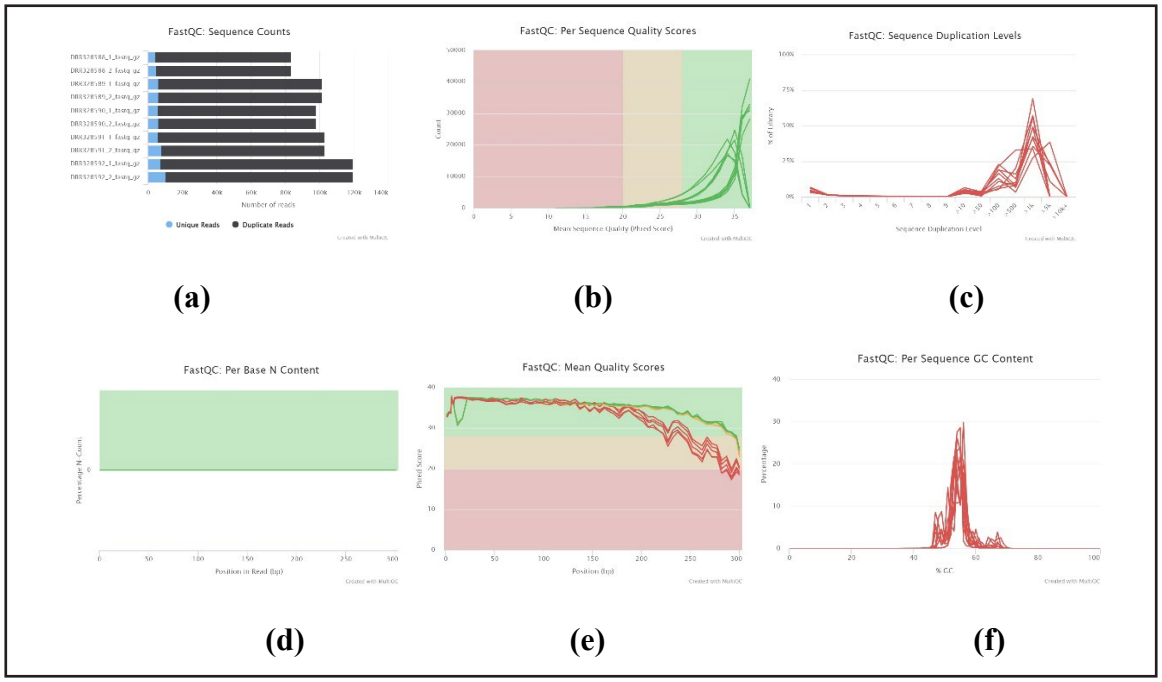

Figure 2: Quality control plots by MultiQC (a) FastQC: Sequence counts (b) FastQC: Per Sequence Quality Scores (c) FastQC: Sequence Duplication Levels (d) FastQC: Per Base N Count (e) FastQC: Mean Quality Score (f) Per Sequence GC Count

After QC data cleaning and trimming of the sequences were done and then all the pair end sequences of the five samples were assembled in 5 single files with the help of MEGAHIT. After this Kraken2 was used for Taxonomic analyses followed by Convert Kraken was used for converting the metagenomic classification data in to NCBI's taxonomic representations. Then the results were amalgamated in the form of Krona Pie Chart which shows the profuse microorganisms that were in accordance with all the samples.

Community structure visualization

A visualization tool was used for performing visual analysis of community profile. This tool collectively gives an HTML output. Krona Pie Chart is basically an interactive tool for visualizing metagenomic composition within a web browser (41). Krona Pie charts of all the five samples are given below respectively.

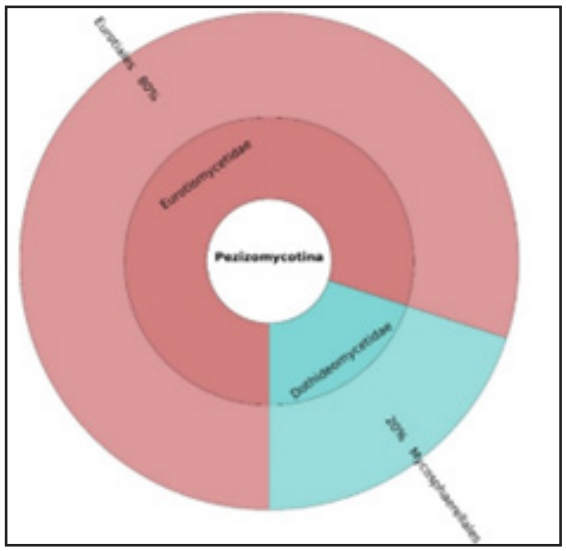

Figure 3: Krona Visualization of community profile of metagenomic dataset 1

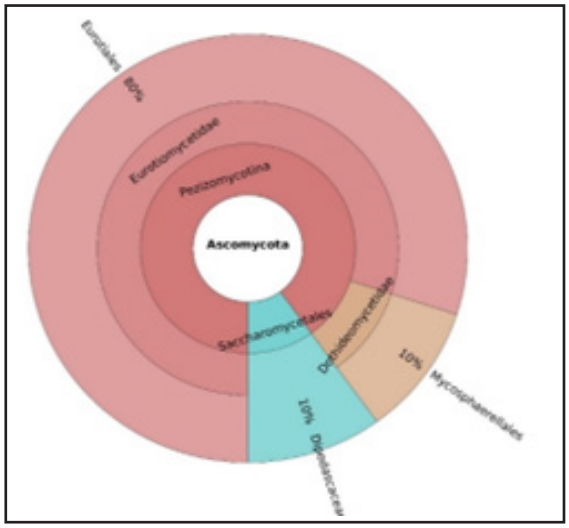

Figure 4: Krona Visualization of community profile of metagenomic dataset 2.

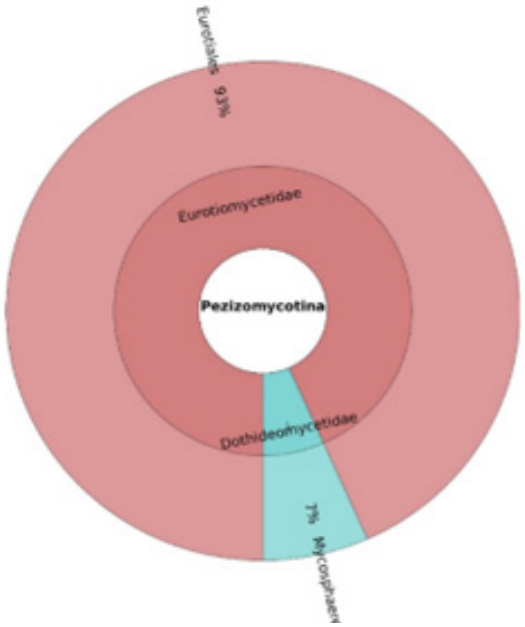

Figure 5: Krona Visualization of community profile of metagenomic dataset 3

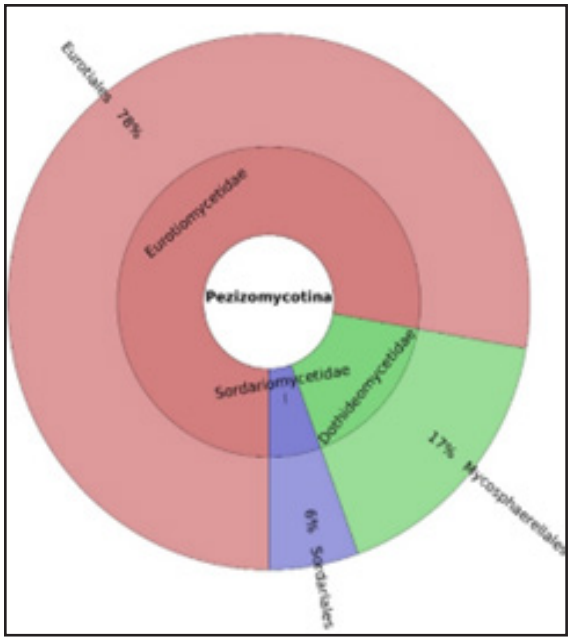

Figure 6: Krona Visualization of community profile of metagenomic dataset 4

Figure 7: Krona Visualization of community profile of metagenomic dataset 5

Figure 3, 4,5,6 and 7 illustrates the result of taxonomic community profile outcomes of each of the five samples i.e., Human Gut Metagenome 1, Human Gut Metagenome 2, Human Gut Metagenome 3, Human Gut Metagenome 4, Human Gut Metagenome 5 which are envisaged using Krona Pie Chart tool. The taxonomy nodes are represented by the nested sectors organised from the top level of hierarchy to the outermost level. At the top left, we will find navigation controls, and at the top right you will find details about the selected nodes. Krona's website offers an interactive version of this chart.

Figure 3 portrays the Krona Pie Chart for the sample 1 of Human Gut metagenome. Here, it can be observed that maximum abundance of Eurotiales is seen while the lowest abundance was of Mycospaerellales.

Figure 4 displays the Krona Pie Chart for the sample 2 of Human Gut Metagenome. Here, it can be observed more than half of the portion was taken by Eurotiomycetideae which consists of Pezizomycotina and the other half portion was taken by Dothideomycetidae which consists of Saccharomycetales

Figure 5 depicts the Krona Pie Chart for the sample 3 of Human Gut Metagenome. Here, it can be observed that the maximum abundance was found of Eurotiales while the lowest abundance was of Mycospaerellales.

Figure 6 shows the Krona Pie Chart for the sample 4 of Human Gut Metagenome. Here, it can be seen that more than half of the portion belong to Eurotiomycetidae and the other half portion belong to Sordariomycetidae which is further divided to Dothideomycetidae.

Figure 7 illustrates the Krona Pie Chart for the sample 5 of Human Gut Metagenome. As it can be observed that the maximum abundance was of Eurotiales which is 83% followed by Mycospaerellales which is 13% and the lowest abundance was of Sordariales which is 4%.

Table 2: The percentage composition of fungi present in all the five samples of Gut microbiome of Covid-19 patients.

SAMPLES	EUROTIALES	MYCOSPAERELLALES	DIPODASCACEAE	SORDARIALES
1	80%	20%	-	-
2	80%	10%	10%	-
3	93%	7%	-	-
4	78%	17%	-	6%
5	83%	13%	-	4%

Table 2 shows the common fungi that are present in all the samples. There are two fungi that were found in all the five samples they are, Eurotiales and Mycospaerellales. And the other fungi that were not common in all the samples were Dipodascaceae and Sordariales. Among all the four bacteria the highest percentage was of Eurotiales 78-93%.

Eurotiales which has the maximum percentage are a Ubiquitous and plentiful fungi (42). They are considerably large order of Ascomycetes with members repeatedly having adverse impact on human activities. Eurotiales is widely used in various research fields food, biotechnology, medical mycology etc. Followed by Mycospaerellales, one of the largest genera of plant pathogenic fungi. It can cause a wide variety of diseases in both humans and plants. The spores of the molds of these fungi are transmitted by inhalation through various transcutaneous routes or by ingestion of spores (43). Some diseases caused by these fungi are Diabetes mellitus, Breathing problems, Neutropenia etc. The family Dipodascaceae includes yeasts in the order Saccharomycetales (44). The fungi belonging to this class are economically important in the food industry. Then Sordariales, is one of the most diverse taxonomic groups within the Sordariomycetes (45). These have a broad range of sustainable diversification. Ascomycetes belonging to the Sordariales order contain a wide range of potential beneficial secondary metabolites. These are all the fungi that were present in the gut of Covid-19 Patients.

Functional data analysis

Using metagenomics data, we can learn which genes are expressed by the community. It can be used to recognise the gene families and their abundance and to look into their contribution to the community using HUMANn2. A new approach to functional profiling of metaomically sequenced microbial communities is HUMAnN2. In addition to identifying species in communities, HUMANN2 aligns reads to local pangenomes, translates unclassified reads, and quantifies gene families and their abundance. For this an output from the Megahit file of all the samples separately were taken as a input to HUMANn2. It generated three files- Gene families and their abundance, Pathways and their Coverage, Pathways and their abundance.

Each gene family in the community has an abundance value contained in the gene and family abundance file. Often, protein-coding sequences within gene families perform identical functions because they are evolutionarily related. It is here that we get the UniRef50 gene families: sequences of a gene family must have at least 50% sequence identity.

The pathway abundance file shows the abundance of each pathway. Using community- and species-level gene abundances along with pathway structure, pathway abundance is evaluated once for each community and once for each species. A pathway coverage file gives an alternative representation of the existence and absence of pathways in a community, independent of their quantitative abundance.

Gene Abundance and their Families

To combine Taxonomic Data with Microbial community functional profiles, and to control the hindrance constrained by transcribed search, HUMANn2 is used as a next generation metagenomic functional profiling approach (46).

Table 3: Genes Based on their abundance value

SAMPLE	GENE	ABUNDANCE
Sample 1	UNMAPPED	167.0000000000
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)	20.9425257607
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)|unclassified	20.9425257607
	UniRef50_K7EWA4	2.2865411250
	UniRef50_K7EWA4|unclassified	2.2865411250
	UniRef50_A0A023FUY2: Putative secreted protein (Fragment)	0.6178137497
	UniRef50_A0A023FUY2: Putative secreted protein (Fragment)|unclassified	0.6178137497

Sample 2	UNMAPPED	304.0000000000
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)	36.5604914986
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)|unclassified	36.5604914986
Sample 3	UNMAPPED	317.0000000000
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)	28.3144400556
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)|unclassified	28.3144400556
Sample 4	UNMAPPED	340.0000000000
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)	47.4056133915
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)|unclassified	47.4056133915
	UniRef50_U6L5R8	1.3192612137
	UniRef50_U6L5R8|unclassified	1.3192612137
	UniRef50_Q3KSS4: Epstein-Barr nuclear antigen 1	0.6286503941
	UniRef50_Q3KSS4: Epstein-Barr nuclear antigen 1|unclassified	0.6286503941
	UniRef50_Q1HVF7: Epstein-Barr nuclear antigen 1	0.5591847891
	UniRef50_Q1HVF7: Epstein-Barr nuclear antigen 1|unclassified	0.5591847891
Sample 5	UNMAPPED	577.0000000000
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)	70.0240751151
	UniRef50_A0A023GNY3: Putative secreted protein (Fragment)|unclassified	70.0240751151

Table 3 shows the genes of all the five samples with their abundance value. This table only shows those genes whose HUMANn2 abundance value is 0.5 or above 0.5.

Conclusion

The infection caused by SARS-COV-2 leads to complex pathophysiologic and immunologic responses in the host. In addition to the phenotypic changes in the host, the gut microbiome is widely transformed in Covid-19. Furthermore, consequent growth of expedient microorganisms corresponding to the Infection caused by SARS-COV-2 and quiescent gut inflammation in Covid-19 create further hazards to host health and well-being and gut flora rejuvenation. Such enlargements in certain microbial species and decline in the microbial biome diverseness in amalgamation with the weakened host immunity may obstruct reconstruction of the gut microflora post Covid-19. Subsequently, the altered gut microbiome bionomics continues even after the infection is resolved. Overall, the complicated microbiome biological network in a stable state is considerably weakened in Covid-19 shifting to one predominated by Covid-19 enriched microbes.

It is familiar that important factors such as diet and treatment can substantially affect gut microflora composition. Due to the excruciating nature of Covid-19 controlling of these important factors or including treatment-naïve Covid-19 patients seems impractical. Here it was observed that different types of microorganisms mainly fungi was found in the gut microbiome of the Covid-19 patients. Different percentage of the fungi present in all the samples can be observed with the help of taxonomic analysis. The maximum abundance of the fungi that was seen in the samples was Eurotiales followed by Mycospearellales, Dipodasceae and Sordariales. Further functional analysis was done to identify the genes and their abundance.

References

1. Yamamoto, S., Saito, M., Tamura, A., Prawisuda, D., Mizutani, T., & Yotsuyanagi, H. (2021). The human microbiome and COVID-19: A systematic review. *PloS*

one, *16*(6), e0253293.

2. Nag, A., Arora, S., & Sethi, S. (2021). Gastroenteritis: Symptoms and Epidemiology of SARS CoV 2. *Coronaviruses: Volume 2*, *2*, 97.

3. Rauch, M., & Lynch, S. V. (2012). The potential for probiotic manipulation of the gastrointestinal microbiome. *Current Opinion in Biotechnology*, *23*(2), 192-201.

4. Pickard, J. M., Zeng, M. Y., Caruso, R., & Núñez, G. (2017). Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunological reviews*, *279*(1), 70-89.

5. Jahani, M., Dokaneheifard, S., & Mansouri, K. (2020). Hypoxia: A key feature of COVID-19 launching activation of HIF-1 and cytokine storm. *Journal of inflammation*, *17*(1), 1-10.

6. Marietta, M., Ageno, W., Artoni, A., De Candia, E., Gresele, P., Marchetti, M., ... & Tripodi, A. (2020). COVID-19 and haemostasis: a position paper from Italian Society on Thrombosis and Haemostasis (SISET). *Blood Transfusion*, 167-169.

7. Burkey, T. E., Skjolaas, K. A., & Minton, J. E. (2009). Board-invited review: porcine mucosal immunity of the gastrointestinal tract. *Journal of animal science*, *87*(4), 1493-1501.

8. Konturek, P. C., Harsch, I. A., Neurath, M. F., & Zopf, Y. (2020). COVID-19-more than respiratory disease: a gastroenterologist's perspective. *J Physiol Pharmacol*, *71*(2), 179-89.

9. Murphy, E., Murphy, C., & O'Mahony, L. (2009). Influence of the Gut Microbiota with Ageing. *Microbiology and Aging: Clinical Manifestations*, 153-17

10. Hussain, I., Cher, G. L. Y., Abid, M. A., & Abid, M. B. (2021). Role of gut microbiome in COVID-19: an insight into pathogenesis and therapeutic potential. *Frontiers in immunology*, *12*, 4164.

11. Pickard, J. M., Zeng, M. Y., Caruso, R., & Núñez, G. (2017). Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunological reviews*, *279*(1), 70-89.

12. Lin, L., & Zhang, J. (2017). Role of intestinal microbiota and metabolites on gut homeostasis and human diseases. *BMC immunology*, *18*, 1-25.

13. Nasereddin, A., Golan Berman, H., Wolf, D. G., Oiknine-Djian, E., & Adar, S. (2022). Identification of SARS-CoV-2 Variants of Concern Using Amplicon Next-Generation Sequencing. *Microbiology Spectrum*, *10*(4), e00736-22.

14. Dellino, M., Cascardi, E., Vinciguerra, M., Lamanna, B., Malvasi, A., Scacco, S., ... & La Forgia, D. (2022). Nutrition as personalized medicine against SARS-CoV-2 infections: clinical and oncological options with a specific female groups overview. *International Journal of Molecular Sciences*, *23*(16), 9136.

15. Harper, A., Vijayakumar, V., Ouwehand, A. C., Ter Haar, J., Obis, D., Espadaler, J., ... & Day, R. (2021). Viral infections, the microbiome, and probiotics. *Frontiers in Cellular and Infection Microbiology*, *10*, 596166.

16. Tringe, S. G., & Rubin, E. M. (2005). Metagenomics: DNA sequencing of environmental samples. *Nature reviews genetics*, *6*(11), 805-814.

17. Behl, A., Nair, A., Mohagaonkar, S., Yadav, P., Gambhir, K., Tyagi, N., ... & Sharma, N.

(2022). Threat, challenges, and preparedness for future pandemics: A descriptive review of phylogenetic analysis based predictions. *Infection, Genetics and Evolution*, 105217.

18. Peixoto, R. S., Harkins, D. M., & Nelson, K. E. (2021). Advances in microbiome research for animal health. *Annual Review of Animal Biosciences*, *9*, 289-311.

19. Gupta, A., Gupta, R., & Singh, R. L. (2017). Microbes and environment. *Principles and applications of environmental biotechnology for a sustainable future*, 43-84.

20. Mokili, J. L., Rohwer, F., & Dutilh, B. E. (2012). Metagenomics and future perspectives in virus discovery. *Current opinion in virology*, *2*(1), 63-77.

21. Lake, M. A. (2020). What we know so far: COVID-19 current clinical knowledge and research. *Clinical medicine*, *20*(2), 124.

22. Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in plant science*, *5*, 209.

23. National Research Council. (2007). The new science of metagenomics: revealing the secrets of our microbial planet.

24. Li, S., Yang, S., Zhou, Y., Disoma, C., Dong, Z., Du, A., ... & Xia, Z. (2021). Microbiome profiling using shotgun metagenomic sequencing identified unique microorganisms in COVID-19 patients with altered gut microbiota. *Frontiers in Microbiology*, *12*, 712081.

25. Sun, Z., Song, Z. G., Liu, C., Tan, S., Lin, S., Zhu, J., ... & Chen, Y. M. (2022). Gut microbiome alterations and gut barrier dysfunction are associated with host immune homeostasis in COVID-19 patients. *BMC medicine*, *20*(1), 24.

26. Shu, W. S., & Huang, L. N. (2022). Microbial diversity in extreme environments. *Nature Reviews Microbiology*, *20*(4), 219-235.

27. Zhang, Y., Thompson, K. N., Branck, T., Yan, Y., Nguyen, L. H., Franzosa, E. A., & Huttenhower, C. (2021). Metatranscriptomics for the human microbiome and microbial community functional profiling. *Annual Review of Biomedical Data Science*, *4*, 279-311.

28. Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*, *33*(19), 3137-3139.

29. Tischler, G., & Leonard, S. (2014). biobambam: tools for read pair collation based algorithms on BAM files. *Source Code for Biology and Medicine*, *9*(1), 1-18.

30. Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048.

31. Wee, S. K., & Yap, E. P. H. (2021). GALAXY Workflow for Bacterial Next-Generation Sequencing De Novo Assembly and Annotation. *Current Protocols*, *1*(9), e242.

32. Krueger, F. (2015). Trim Galore!: A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. *Babraham Institute*.

33. Chuan, J., Zhou, A., Hale, L. R., He, M., & Li, X. (2021). Atria: an ultra-fast and accurate trimmer for adapter and quality trimming. *bioRxiv*, 2021-09.

34. Korobeynikov, A., Lapidus, A., Almeida, A., Finn, R., Richardson, L., & Sakharova, E.

Bioinformatics: from Algorithms to Applications 2020.

35. Islam, R., Raju, R. S., Tasnim, N., Shihab, I. H., Bhuiyan, M. A., Araf, Y., & Islam, T. (2021). Choice of assemblers has a critical impact on de novo assembly of SARS-CoV-2 genome and characterizing variants. *Briefings in bioinformatics*, *22*(5), bbab102.

36. Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome biology*, *15*(3), 1-12.

37. Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome biology*, *20*, 1-13.

38. Davis, M. P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N., & Enright, A. J. (2013). Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods*, *63*(1), 41-49.

39. Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC bioinformatics*, *12*(1), 1-10.

40. Franzosa, E. A., McIver, L. J., Rahnavard, G., Thompson, L. R., Schirmer, M., Weingart, G., ... & Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nature methods*, *15*(11), 962-968.

41. Patel, R. K., & Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PloS one*, *7*(2), e30619.

42. Sudarikov, K., Tyakht, A., & Alexeev, D. (2017). Methods for the metagenomic data visualization and analysis. *Current issues in molecular biology*, *24*(1), 37-58.

43. Ribes, J. A., Vanover-Sams, C. L., & Baker, D. J. (2000). Zygomycetes in human disease. *Clinical microbiology reviews*, *13*(2), 236-301.

44. Deak, T. (1999). Molecular taxonomy of yeasts. *Acta microbiologica et immunologica hungarica*, *46*(2-3), 181-186.

45. Zhang, N., Castlebury, L. A., Miller, A. N., Huhndorf, S. M., Schoch, C. L., Seifert, K. A., ... & Sung, G. H. (2006). An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. *Mycologia*, *98*(6), 1076-1087.

46. Shakya, M., Lo, C. C., & Chain, P. S. (2019). Advances and challenges in metatranscriptomic analysis. *Frontiers in genetics*, *10*, 904.