Annals of Medicine and Medical Sciences (AMMS)
Volume 04, 2025, *Page No.: 1606 to 1622*
Available at: *http://ammspub.com*

Review Article

OPEN ACCESS

Microbial Informatics in Association with Disease Biology: A Multi-Omics Approach on Computational Microbiology Applications

Debaleena Samanta [1], MSc, Malavika Bhattacharya [2], PhD

Department of Biotechnology, Techno India University, West Bengal, India.

*Corresponding Author: Malavika Bhattacharya; *malavikab@gmail.com*

Abstract

Background: Bioinformatics is a scientific sub-discipline that covers the area of biology at multi-disciplinary levels of microbiology, immunology, molecular biology, biochemistry, genetics, proteomics etc. This is a technical analysis of computational biology that collect and convert raw data into storage to analyse and disseminate this biological information such as, De-oxy-ribonucleic Acid, protein, amino acid sequences, alpha-fold 3-Dimensional structure, taxonomic hierarchy, polypeptides, messenger Ribonucleic Acid, chemical origin into annotated databases. ***Aim:*** Bioinformatics tools and programs are used to interpret and develop biological raw data and efficiently manage through accession database to create statistical approaches for evaluating relationships between large datasets. ***Scope:*** The advancement of genome sequencing technologies and metagenomic analysis have well developed microbial informatics applications to study microorganisms and their functionality based on microbial interactions in natural and artificial environments. This large amount of microbial information has been structured, indexed and correlated with existing experimental evidences. The bioinformatics solutions of overwhelming research parameters are cross-connected with information science and microbiology in relevance. In the recent scenario, artificial intelligence techniques, aggregated statistical analysis methods and large-scale data input system have provided microbiology to evolve with data science. ***Key Insights:*** "Clinical Microbiology Informatics" deals with detection, identification and confirmation of antimicrobial susceptibility testing and communicates with clinicians about the importance of clinically relevant microbes. Clinical microorganisms in pathogenic diseases provide insights into genetic and molecular mechanisms of pathogen through bioinformatics analysis. Microbial bioinformatics applications are used for pathogen identification, characterization, mutation, detection of virulence factors, microbial engineering, vaccine targets, drug and vaccine combinations, disease regulation, host-immune evasion and diagnostic approaches. ***Future Perspective:*** This study highlights the critical significance of computational microbiology in comprehending microbial pathogens and their interactions with hosts by examining the nexus between microbial informatics and disease biology. Researchers can examine microbial populations, analyze genomic data, and pinpoint virulence factors linked to a range of illnesses by utilizing cutting-edge bioinformatics methods. We go on recent applications that improve our comprehension of microbial dynamics in illness contexts, including metagenomics, machine learning, and network analysis. The analysis highlights the potential of microbial informatics to guide treatment plans, enhance diagnostic techniques, and support vaccine research, all of which will ultimately lead to improvements in public health.

Keywords: ***Clinical Microbiology, Microbial Informatics, Microbial Engineering, Pathogenic Diseases, Computational Biology***

1. Introduction

The diversity matrices that are computed to describe microbial diversity, as opposed to macroorganisms like plants and animals, usually make use of genomics-based technologies, particularly amplicon-based evaluations that demand for some sort of "bioinformatics tools" in addition to standard statistical methods. When evaluating functional diversity in addition to phylogenic diversity, it becomes imperative to incorporate additional omics-based data [1-3]. The word "ome" means "many," and the field of "omics" has up to now focused on figuring out how complex biological things are. Given the massive volume of biological data being produced by the "omics" era. As a field of biology, bioinformatics has become increasingly significant throughout time. Much work has been done to advance the tools used in biota research, and this has resulted in the development of several "omics," including transcriptomics, metabolomics, proteomics, genomes, and so on. Multiple "omics" approaches were employed at various levels because it was initially evident that a single "omics" approach was insufficient to adequately describe the complexity. Beginning with the creation of a large number of databases, online resources, and software, high-throughput genome sequencing has given microbiology new insights and enabled the tools to be used to a broad range of tasks [4-6].

The use of bioinformatics tools and methods to analyse Next Generation Sequencing (NGS) data is growing in popularity for infectious disease monitoring and diagnosis. Molecular identification, genotyping, microbiome research, antimicrobial resistance analysis, and the identification of unknown disease-associated pathogens in clinical specimens are among the topics of

Received: November 16, 2025; **Revised:** December 15, 2025; **Accepted:** December 18, 2025

interest when reviewing the use of bioinformatics tools, frequently used databases, and NGS data in clinical microbiology [7-9]. This came when genomic techniques were widely used to diagnose and treat bacterial, fungal, and viral illnesses. Bioinformatics has been applied to strain typing, resistance analysis, virulence factor detection, and pathogen identification. NGS, or next-generation sequencing, ascertains the DNA sequence of an entire bacterial genome in a single sequence run, and from this data, virulence and resistance information, together with typing information, are gathered, which is helpful for investigating outbreaks. An overview of NGS is provided in this review, along with information on library preparation and the key features of the most used NGS platforms [10-12]. Even though NGS has a lot of potential for clinical infectious disease testing, there are still a lot of obstacles to overcome before NGS can be widely used in clinical microbiology labs. These obstacles include automation, standardising technical protocols and bioinformatics pipelines, enhancing reference databases, putting in place proficiency testing and quality control measures, and cutting costs and turnaround times. Additionally, NGS applications in the clinical setting is explained, including how to manage outbreaks, find molecular cases, characterise and monitor pathogens, quickly identify bacteria using taxonomy, use metagenomics techniques on clinical samples, and determine whether zoonotic microorganisms are transmitted from animals to humans. Lastly, we outline our expectations for the near-term application of NGS in personalised microbiology while highlighting some needs [13-15].

2. Review of Literature

2.1 Computational Approaches

2.1.1 Next-Generation Sequencing: Large genomic segments are now described in a whole new way thanks to quick and inexpensive sequencing techniques. Sequence analysis can be used to forecast routine diagnostics and public health care in addition to answering the traditional questions of a microbe's evolutionary status or functional classification. 16S rRNA sequence analysis is one method of functional classification for microbial populations [**Fig. 1**]. Comparative genomics, intraspecies diversity, the creation of specialised databases, pathway analysis for the dissemination of complex diseases, microbial community modelling, microbial evolution, genome restructuring, and whole genome comparisons are just a few of the many applications that have been discussed recently. A dramatic shift towards DNA-based classification was observed as a result of the explosion of sequencing data. As a result, numerous tools for analysing sequencing information to taxonomical data were developed. One of the main tools for determining evolutionary relationships based on DNA is a phylogenetic tree [16-19].

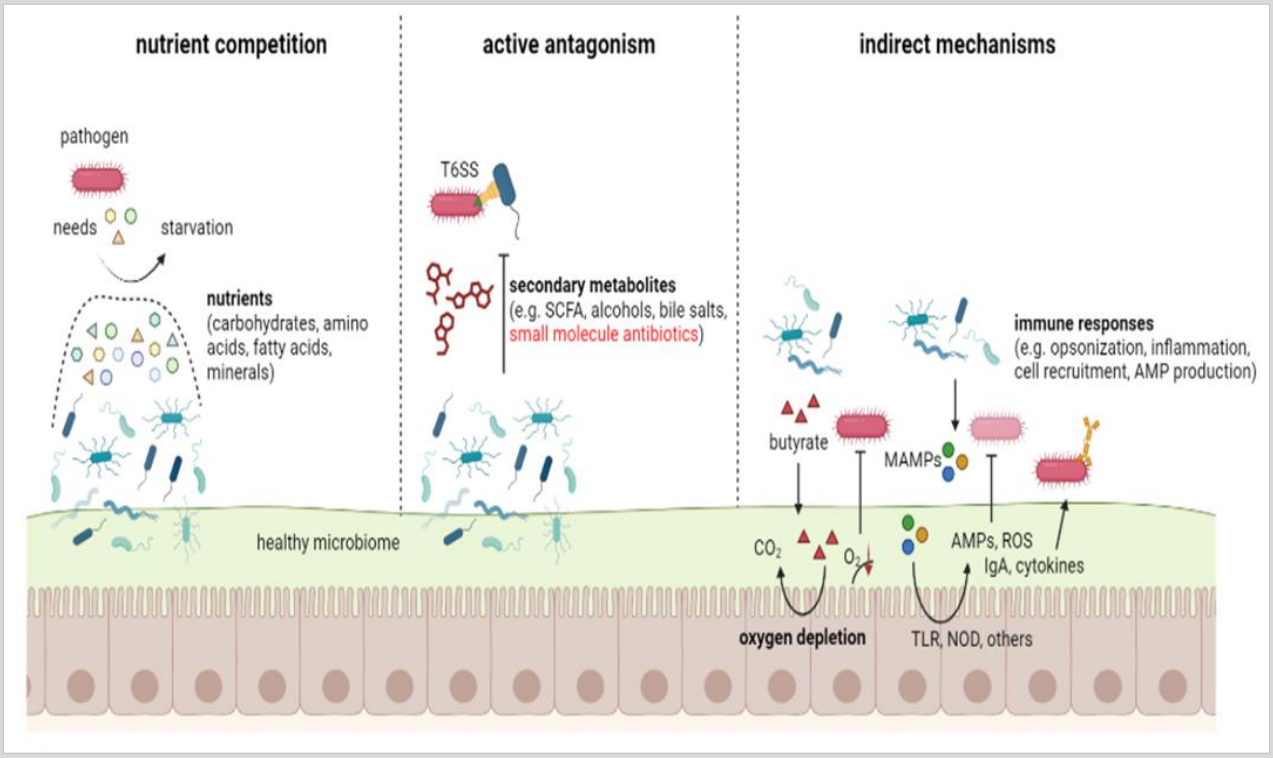

Figure 1: Microbial Systematic Hierarchy at Different level of Resolution. [https://app.biorender.com]

2.1.2 Microbial Whole Genome Assembly: Numerous platforms, including Oxford Nanopore, Pacific Biosciences, Ion Torrent, and Illumina, have flourished since the introduction of the 454's NGS technology. Illumina is one of the most popular of them all, and it has arguably been used for short-read high-throughput sequencing that is deep enough to reliably identify the sequence of a bacterial genome. However, because genomes generate several sequences, gathering contingency and reads, putting them together, filling in the gaps, and other tasks are laborious and manual. De novo sequencing is made possible by the lack of precise coverage and alignment mapped to whole genome reference assemblies [20-22].

2.1.3 Annotation: The exponential growth of prokaryotic genomes in recent years has led to a significant decrease in the time and financial investment required for each research, which is comparable to Moore's law. [23,24]. It would be easier to predict strain-specific changes and homologs linked (genetically) to its neighbouring relatives if it had coding sites (CDS), ribosomal-binding sites (RBSs), and start/termination sites. Thankfully, gene banks have well-annotated, if not well curated, genomes that might potentially prevent the introduction of additional mistakes. The annotation site for GenBank is one such pipeline that has been utilised extensively. It would still be difficult to annotate several genomes simultaneously

and determine whether the annotation is innovative in terms of quality [25-27].

2.1.4 Microbial Pan-Genomics: A core genome, which contains genes shared by all strains, and a dispensable genome, which contains genes specific to a single strain (sometimes called a singleton) and accessory genes conserved in two or more strains, make up a microbial pan-genome [28]. While the accessory genes and single tonnes typically relate to supplemental biochemical pathways and functions that may confer selective advantages like ecological adaptation, virulence mechanisms, antibiotic resistance, or colonisation

Table 1: Bioinformatics tools for Microbial Pan-genomics Analysis [29]

Database	URL	Salient Features
BPGA	http://sourceforge.net/projects/bpgatool/	Gene clustering and related variants, KEGG pathway
CAMBer	http://bioputer.mimuw.edu.pl/camber/index.html	Study of Related Variants
GET_HOMOLOGUES	http://www.eead.csic.es/compbio/soft/gethoms.php	study of related variants, and conniving homologs
Harvest	https://github.com/marbl/harvest	Study of Related Variants
ITEP	https://price.systemsbiology.net/itep	Connecting pan-genomic biographies, function-based searches, phylogeny, and related variations
PanCake	https://bitbucket.org/CorinnaErnst/pancake/wiki/Home	Study of Related Variants
PanCGHweb	http://bamics2.cmbi.ru.nl/websoftware/pancgh/pancgh_start.php	A structure with a visualizer Pan-genome
PanGP	http://PanGP.big.ac.cn	Study of Related Variants
PANNOTATOR	http://bnet.egr.vcu.edu/pannotator/index.html	Function-grounded searching, connecting variants
Panseq	https://lfz.corefacility.ca/panseq/	Relating variants, conniving pan-genomic biographies
PGAP	http://pgap.sourceforge.net/	Relating Variants, conniving pan-genomic biographies, function grounded searching, clusters of orthologous gene (COG) analysis
PGAT	http://nwrce.org/pgat	Relating Variants, Conniving pan-genomic Biographies, COG Analysis
ROARY	https://sanger-pathogens.github.io/Roary/MicrobialMetagenomics.docx	Conniving pan-genome biographies and Gene Clustering. A Channel Program
Spine and AGEnt	http://vfsmspineagent.fsm.northwestern.edu/index_age.html	Study of Related Variants

of a new host, the core genes are in charge of the fundamental aspects of the species' biology and its main phenotypic traits [30-32]. Analysis of pan-genomes is done at species level and can be informative at any taxonomic level spanning the entire bacterial domain. Numerous software programs and tools have been created that can be used to profile core, shared, and isolate-specific genes, find single-nucleotide polymorphisms (SNPs), build phylogenies, and cluster orthologous genes. [**Table. 1**] A few of the well-known pan-genome analysis programs and online servers that provide the ability to calculate pan-genome analyses for genomes. A web server named PanGeneHome was recently created that provides a thorough and consistent framework for large-scale precomputed pan-genome analysis of genomes that have previously been sequenced [33-35].

2.1.5 Microbial Metagenomics: Metagenomics, which translates to "beyond the genome," is the study of genomic DNA from an entire population. This is not the same as genomics, which focusses on analysing each individual organism's genetic DNA. Although this word has also been used in literature to refer to studies in which only a portion of 16S rRNA has been chosen for study, metagenomics can be interpreted as the random shotgun sequencing of microbial DNA without selecting any particular gene [36,37]. The phrase was initially used in a study on soil bacteria, but it has since been used in a number of settings, such as environmental 16S rRNA gene diversity [**Fig. 2**]. One of the limitations of marker gene sequencing, like 16S rRNA, is that it does not capture viruses [**Table. 2**]. Additionally, recent research showed that >50% of species were overlooked by 16S or 18S rRNA gene-based amplicon sequencing [38-40].

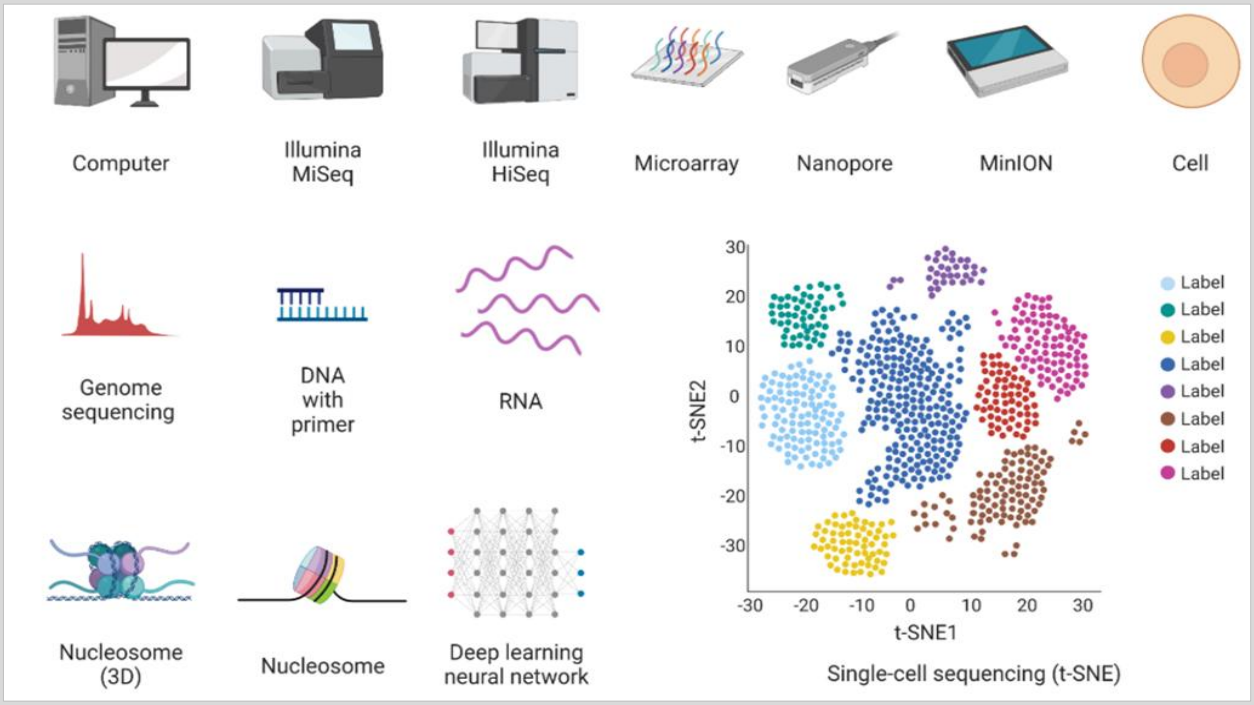

Figure 2: Computational Biology Applications in Microbial Disease Prediction. [https://app.biorender.com]

Finding the best phylogenetic anchor is essential for determining the phylogenetic affiliation of a genome's gene sources [41-43]. One can continue by i) locating anchors on a DNA fragment and ii) locating genome fragments connected to it via anchors. ii) The metagenome would be of the order of 1 Gbp DNA, assuming 200 species per millilitre. This therefore necessitates extensive DNA isolation and cloning. iii) Sequencing library insert sizes are crucial for capturing the complexity of communities and microbial diversity comprising the metagenome. Because harsh methods can be used to lyse microorganisms, smaller insert libraries are more convenient. iv) Ten thousand of additional genomes, many of which are extremely fragmented and incomplete, have been created by the rapid speed of "draft" genome sequencing [44-45].

2.1.6 Gut Metagenomics: The majority of human intestinal microorganisms are essential to nutritional metabolism, which supports physiological function. Only microorganisms that reside in the human gut are able to collect energy from food; any disruption of this microbial community causes illnesses like obesity or bowel disorders [46-47]. In order to maintain human health and performance, it is crucial to identify and understand the composition, variety, and mode of operation of these gut microorganisms. Bacteroidetes and Firmicutes are the two types of bacteria that make up the gut microbiota. Based on phylogenetic data, 16S ribosomal sequencing identified the former. Regarding gut metagenomics, two significant initiatives were completed: the American Human Microbiome Project and the European project MetaHIT [48,49].

2.1.7 Oral Metagenomics: In terms of microbial community richness, the human mouth is second only to the gastrointestinal system. It is commonly recognised that a large number of

Table 2: Bioinformatics Software tools for Microbial Metagenomics Analysis [50]

Tool	URL and Reference	Salient Features
MIRA	https://sourceforge.net/p/mira-assembler/wiki/Home/	Sanger, 454, Illumina, ion alluvion, PacBio; reference grounded
AMOS and MetaAMOS	https://github.com/treangen/MetAMOS	Genomic pulpits, open reading frames, and taxonomic or functional reflections
SOAP	http://www.soap.genomics.org.cn	Single- genome *De novo* assembly, reference grounded assembly
ESOM	http://databionic-esom.sourceforge.net/	Composition- grounded binning; binning
MEGAN	http://ab.inf.uni-tuebingen.de/software/megan6/	Similarity- grounded binning; binning
WIMP	https://doi.org/10.1101/030742	Complete metagenome channel allowing data collection from undressed samples to species strain-position bracket
Centrifuge	https://ccb.jhu.edu/software/centrifuge/	Classification of metagenomes
Kaiju	https://bioinformatics-centre.github.io/kaiju/	Taxonomic categorization

bacterial species inhabit the human oral cavity; some of these species are beneficial, while the majority cause oral illnesses. Oral-friendly bacteria coexist and are necessary to maintain the balance of the oral environment, but they also cause periodontal disease and tooth cavities. The enlarged Human Oral Microbiome Database (eHOMD), a database devoted to human oral metagenomics, includes all of the mouth cavity's microbial data. Based on 16S rDNA profiling, the healthy oral bacteria are divided into six major phyla: Firmicutes, Actinobacteria, Proteobacteria, Fusobacteria, Bacteroidetes, and Spirochaetes [51-52]. These phyla account for 96% of all oral bacteria. Numerous taxa are included in the CORE database, which is based on the phylogenetically curated 16S rDNA of the human oral microbiota. Research on these communities depends on precise sequencing data identification. A thorough and less redundant list of bacteria linked to the human mouth cavity can be found in this database [53].

2.1.8 Insect-Microbe Metagenomics: This claim is being more and more supported by research, which shows that the microbial community stimulates many facets of insects' lives, just like plants do. In the present situation, metagenomics and metatranscriptomics have both improved research on the interaction between insects and microbes [54,55]. It used a short nucleotide fragment known as a barcode (e.g. 16S, 18S, ITS, COI) as a proxy for identification to reconstruct biological communities based on taxonomy. In addition to providing information on the relative taxonomical abundance and the existence of certain genes in a sample, metagenomic and metabarcoding techniques are helpful for qualitatively assessing the diversity of organisms within that sample. But these methods only tell us what is happening (gene identification) and (taxonomic reconstruction) [56-58].

Table 3: A simplified Software list for Microbial Metagenome Analysis [59]

Tool	URL	Salient Features
PHAST/ PHASTER	http://phaster.ca/	Useful in chancing prophages in bacterial genomes
QIIME	http://qiime.org/	A channel for analysing raw metagenome sequencing data. Support for generating publication quality figures
MetaPhlAn2	http://segatalab.cibio.unitn.it/tools/metaphlan2/	A channel for sketching microbial composition. Depends on the curated MetaPhlAn2 database
StrainPhlAn	http://segatalab.cibio.unitn.it/tools/strainphlan/	Strain- position resolution in sketching known microbial specie s. Depends on the curated MetaPhlAn2 database
MetaTrans	https://manichanh.vhir.org/metatrans.org/index.html	Tool for mapping taxonomic and gene expression analysis which also integrates quality control and rRNA junking
Tbooster	http://tbooster.erc.monash.edu	Ensemble styles for prognosticating effector proteins in bacterial systems, specifically for T3SS, T4SS, and T6SS

2.1.9 Soil-Microbial Metagenomics: The soil has the most diversity of microorganisms and may continue to be used for agricultural, industrial, and environmental purposes. Soil has a much more varied microbial community than other microbial sources, particularly when it comes to prokaryotes. [**Table. 3**] According to the study, one gram of soil contains ten billion microorganisms and thousands of distinct species. NGS technologies have also helped to quantify and depict the diversity of soil microbes. Several significant studies were conducted to investigate the soil taxonomic [60-62].

2.1.10 Microbial Genetic Engineering: Through genetic engineering, microbes like Bacillus subtilis and Escherichia coli can create ethanol, vitamins, enzymes, carbohydrates, and chemical precursors. Riboflavin (vitamin B2) is produced by the fungus Ashbya gossypii, which, when compared to chemical synthesis, enhanced production efficiency and decreased environmental protection expenditures by 43% and 30%, respectively [63,64].

Table 4: Bioinformatics Software tools for Microbial Resources [65]

Microbial resources	URL	Salient Features
IMG (Integrated Microbial Genomes)	https://img.jgi.doe.gov/	Reflection and analysis of microbial genomes and metagenomes
MicrobesOnline	http://www.microbesonline.org	Gate for relative and functional microbial genomics
ModelSEED	http://modelseed.org/	Gate for curated genomic data and automated reflection of microbial genomes
GOLD (Genomes Online Database)	https://gold.jgi.doe.gov/	Resource for comprehensive information about genome and metagenome sequencing projects
CDD	https://www.ncbi.nlm.nih.gov/cdd	Conserved sphere database
Pfam	https://pfam.xfam.org/	Database of protein families
STRING	https://string-db.org/	Database of protein association networks
Ribosomal Database Project (RDP)	https://rdp.cme.msu.edu/	16S rRNA gene database
SILVA	https://www.arb-silva.de/	rRNA gene database
GREENGENES	http://greengenes.lbl.gov/	16S rRNA gene database
Bacterial Isolate Genome Sequence Database (BigSdb)	https://pubmlst.org/software/database/bigsdb/	Bacterial insulate genome sequence database
EBI metagenomics	https://www.ebi.ac.uk/metagenomics/	Portal for submission and analysis of metagenomics data
EcoCyc	https://ecocyc.org/	*E. coli* genome and metabolism knowledge base
RegulonDB	http://regulondb.ccg.unam.mx/	*E. coli* transcriptional regulation resource Pseudomonas genome database
Pseudomonas Genome Database	http://www.pseudomonas.com/	Pseudomonas genome database
PATRIC	https://www.patricbrc.org/	Gate for numerous prokaryotic pathogen
EuPathDBs	https://eupathdb.org/	Gate for numerous eukaryotic pathogen
TBDB	http://genome.tbdb.org/tbdb_sysbio/MultiHome.html	Integrated platform for tuberculosis exploration
TCDB	http://www.tcdb.org/	Transporter bracket database

TransportDB	http://www.membranetransport.org/transportDB2/index.html	Transporter protein analysis database
MetaCyc	https://metacyc.org/	Metabolic pathway database
Kyoto Encyclopedia of Genes and Genome	https://www.genome.jp/kegg/	Genome database with emphasis on metabolism
MiST	http://fgrtools.hms.harvard.edu/MIST/help.jsp	Microbial signal transduction database
SwissRegulon	http://www.swissregulon.unibas.ch/	Genome-wide reflections of nonsupervisory spots in model organisms
RegPrecise	http://regprecise.lbl.gov	Database of regulons in prokaryotic genomes
Orione	http://orione.crs4.it/	Web- grounded frame for NGS analysis

In order to combat significant threats to ecosystems and to meet our future demand for refined industrial products, genetically modified organisms may be crucial [**Table. 4**]. With the development of molecular tools, it is now feasible to study entire biological samples at the level of the genome, transcriptome, proteome, metabolome, and interactome following initial phenotypic and characterisation. It is anticipated that genetic engineering methods will play a significant role in creating environmentally friendly methods to lessen the use of chemicals and the manufacturing of agriculturally based goods [66].

2.2 Comparison between Tools

2.2.1 QIIME2 and Roche 454 Genome Sequencing (Pyrosequencing)

Feature	QIIME2	Roche 454 Genome Sequencing (Pyrosequencing)
Type of Technology	*Bioinformatics analysis platform/pipeline.*	*Next-Generation Sequencing (NGS) platform.*
Primary Purpose	Analysis of microbial community data (amplicon, metagenomics, metabolomics).	Experimental generation of DNA sequence reads.
Role in Workflow	**Post-sequencing analysis**—data processing, quality control, taxonomic classification, statistics, visualization.	**Sequencing stage**—converts biological DNA samples into digital reads.
Input	Demultiplexed FASTQ files from any sequencer (Illumina, Ion Torrent, Nanopore, previously 454).	Biological DNA sample (amplicon or whole genome) for sequencing.
Output	OTU/ASV tables, taxonomic profiles, diversity indices, phylogenetic trees, visualizations.	Sequence reads with long read lengths (350–700 bp), typical of amplicon sequencing.
Accuracy	Depends on upstream sequencing quality; uses DADA2/Deblur for error correction to generate high-resolution ASVs.	Moderate accuracy; known issues with **homopolymer errors** (overestimation/underestimation of repeated nucleotides).
Read Length	Not applicable (analysis only).	**Longer reads** compared to early Illumina platforms (up to ~700 bp).
Error Profile	Uses error-modeling tools to remove low-quality sequences.	High error rate in homopolymers; lower throughput than Illumina.
Throughput	Not applicable; scalable for large datasets depending on computational resources.	Low throughput compared to modern NGS (Illumina/Nanopore).
Cost	Free, open-source software.	Expensive per base; platform discontinued after 2016.
Popularity/Status	Actively developed, widely used for microbiome research.	Technology discontinued; historically important but now obsolete.
Supported Analysis	16S/18S/ITS, shotgun metagenomics, metadata integration, machine learning, phylogeny.	Amplicon sequencing (mostly 16S rRNA) and some whole-genome applications.

2.2.3 Practical Relevance of QIIME2:

- Illumina MiSeq/HiSeq/NovaSeq data
- IonTorrent data
- Nanopore data
- Legacy 454 datasets

2.2.2 Practical Relevance of Roche 454 Genome Sequencing:

- Offered long reads ideal for 16S rRNA surveys.
- Limited throughput, high cost, and error-prone homopolymer regions made it less competitive.
- Illumina platforms replaced it with cheaper, higher-quality sequencing.

2.2.4 Strengths of QIIME2:

- Comprehensive Microbiome Analysis Platform
- Plugin-Based, Modular System
- High Reproducibility
- Wide Compatibility
- Open-Source and Community-Driven

2.2.5 Strengths of Roche 454 Genome Sequencing:

- Longer Read Lengths
- Fast Sequencing Chemistry
- Strong Amplicon Sequencing Performance
- Low GC Bias
- Simpler Assembly of Small Genomes

2.2.6 Limitations of QIIME2:

- Dependent on Sequencing Quality
- Computationally Intensive
- Learning Curve
- Limited to Post-Sequencing Analysis
- Plugin Variability

2.2.7 Limitations of Roche 454 Genome Sequencing:

- Homopolymer Errors
- Low Throughput
- High Cost per Base
- Platform Discontinued
- Shorter Reads than Modern Long-Read Platforms

2.2.8 Illumina and Qiagen Gene Reader

Feature	Illumina Sequencing	QIAGEN GeneReader
Type	Widely used high-throughput Next-Generation Sequencing (NGS) platform	Targeted NGS platform optimized for clinical diagnostics
Sequencing Chemistry	Sequencing-by-Synthesis (SBS) using reversible terminators	Sequencing-by-Synthesis with "real-time" fluorescent detection
Primary Use	Research, clinical genomics, whole-genome sequencing, RNA-seq, metagenomics	Targeted gene panel sequencing, particularly oncology
Technology Scale	Broad range: MiSeq, NextSeq, NovaSeq, iSeq	Single integrated platform for targeted sequencing
Throughput	Very high (depending on system); suitable for whole genomes to large studies	Low–moderate; designed for small-scale clinical workflows
Read Length	Typically 2 × 150 bp, up to 2 × 300 bp (MiSeq)	Up to ~2 × 150 bp
Error Rate	Low (~0.1%)	Low–moderate (~1%, platform-dependent)

2.2.9 Practical Relevance of Illumina

- Whole-genome sequencing (WGS)
- Whole-exome sequencing (WES)
- RNA sequencing
- Metagenomics
- 16S/ITS amplicon studies
- Single-cell sequencing

2.2.10 Practical Relevance of QIAGEN GeneReader

- Primarily oncology-focused targeted panels (e.g., lung, breast, colon cancer)
- Variant detection for actionable mutations
- Clinical diagnostics and routine testing
- Companion diagnostics assays

2.2.11 Strengths of Illumina

- Industry standard for NGS
- Extremely high throughput
- Flexible applications (research + clinical)
- Low error rate
- Large ecosystem of kits and tools

2.2.12 Strengths of QIAGEN Gene Reader

- Designed for clinical labs with minimal NGS experience
- Closed, integrated workflow from sample to report
- Strong interpretation layer (QCI Knowledge Base)
- Standardized kits reduce variability

2.2.13 Limitations of Illumina

- Requires more hands-on time and workflow optimization
- Higher cost instruments
- Needs separate bioinformatics systems

2.2.14 Limitations of QIAGEN Gene Reader

- Low throughput
- Limited to targeted gene panels
- Less flexible for research use
- More expensive per sample
- Lower adoption and ecosystem compared to Illumina

2.2.15 SMRT Sequencing and Oxford Nanopore Sequencing

Feature	PacBio SMRT Sequencing	Oxford Nanopore Sequencing (ONT)
Developer	Pacific Biosciences	Oxford Nanopore Technologies
Technology Type	Long-read sequencing	Long-read sequencing
Core Principle	Real-time detection of fluorescently labeled nucleotides	Real-time measurement of ionic current changes as DNA passes through nanopores
Read Length	Avg 10–30 kb; HiFi reads ~15–25 kb	Wide range: 5 kb – >1 Mb; ultra-long reads possible (>2 Mb)
Primary Strength	Very high accuracy (HiFi reads)	Flexible, portable, ultra-long reads
Typical Accuracy	HiFi reads: 99.8–99.9%	Standard reads: 92–97%; duplex >99% but lower yield
Throughput	High (Sequel II/IIe/Revio)	Variable: pocket-size MinION to large-scale PromethION
Instrument Range	Benchtop, high-cost	Portable (MinION), mid-scale (GridION), high-throughput (PromethION)

2.2.16 Practical Relevance of SMRT Sequencing:

- High-quality genome assemblies
- Variant detection (SNVs, indels, structural variants)
- Microbial genomics
- Isoform sequencing (Iso-Seq)
- Clinical/diagnostic-grade sequencing

2.2.17 Practical Relevance of Oxford Nanopore Sequencing

- Long/ultra-long read assembly
- Real-time pathogen detection
- Field/deployment sequencing (e.g., Ebola, COVID-19)
- Metagenomics
- Direct RNA sequencing
- Epigenetic profiling

2.2.18 Strengths of SMRT

- Highest accuracy (HiFi)
- Excellent for Genome assembly, Variant calling, Clinical/precision genomics
- Lower systematic errors

2.2.19 Strengths of Oxford Nanopore

- Portable (MinION), scalable (PromethION)
- Ultra-long reads possible
- Detects epigenetics directly
- Fast, real-time sequencing
- Lower instrument cost

2.2.20 Limitations of SMRT

- Expensive instrument and consumables
- Larger footprint
- No ultra-long reads like ONT
- Less suitable for fieldwork

2.2.21 Limitations of Oxford Nanopore

- Lower raw accuracy than PacBio
- Error profile biased (homopolymers, indels)
- Higher variability between runs
- Duplex mode reduces throughput

2.3 Case Studies

2.3.1 QIIME2: Numerous microbiome researches have made substantial use of QIIME 2. It was utilized to describe gut microbial changes in seditious bowel complaints in human health research, linking decreased diversity and dysbiosis patterns. QIIME 2 was used in environmental research to analyze soil microbiomes under various agricultural techniques, demonstrating shifts in taxa that cycle nutrients. In the field of marine ecology, it facilitated the monitoring of microbial communities in coral reefs to identify infections linked to stress. QIIME 2 was utilized in clinical tests to compare the oral microbiota of healthy and ill individuals. Additionally, QIIME 2 was used in food assiduity research to address turbulent microbiomes and improve product quality [67].

2.3.2 Mothur: In environmental microbiology and microbial ecology, mothur has been widely utilized. It was used to investigate gut microbiome variations in rotundity in mortal health research, exposing changes in Firmicutes-to-Bacteroidetes rates. Mothur was employed in soil microbiology research to monitor changes in the microbial population under crop gyration and fungicide treatment. Mothur assisted in the evaluation of brackish and marine microbial diversity in undersea ecology, which helped identify changes associated with pollution. Mothur was used in clinical research to describe oral and nasal microbiota for infection threat profiling. Additionally, food safety research employed mothur to analyze sources of contamination in the environments of dairy and meat production, refining microbiological monitoring techniques [68].

2.3.3 Qiita: Large-scale and multi-study microbiome analysis in many contexts have been made possible by Qiita [**Fig. 3**]. Qiita made it possible to do meta-analyses of gut microbiota in relation to rotundity, diabetes, and seditious bowel complaints in mortal health research. These analyses revealed harmonious patterns of decreased diversity and changed microbial signatures. Qiita was utilized in environmental studies to encyclopedically integrate soil microbiome datasets and relate changes in nutrient-cycling taxa that are depending on climate. Qiita was utilized in marine exploration to evaluate microbial populations between ocean locations, supporting evaluations of pollution and climate impact. Additionally, Qiita has provided unified workflows and standardized data integration to allow dietary microbiome comparisons, longitudinal host-microbe commerce systems, and forensic microbiome studies [69].

2.3.4 MIMIX: Studies using standardized reporting of microbial sequencing data have made considerable use of MIMIx. MIMIx guidelines improved cross-study comparison of nutrient-cycling microorganisms in soil microbiome investigation by unifying metadata from worldwide soil checks. In coral reef and deep-ocean microbial investigations, MIMIx was employed by marine microbiology systems to standardize slice, sequencing, and ambient factors. In order to promote repeatable detection of antibiotic-resistant bacteria, clinical examinations used MIMIx to unify pathogen sequencing metadata. MIMIx facilitated the harmonized attestation of rhizosphere sequencing trials across crop kinds in agricultural investigation. These case studies show how MIMIx enhances community, reproducibility, and data translucency in microbial sequencing research [70].

2.3.5 Ion Torrent Sequencing: Microbial, environmental, and clinical genomics have all made substantial use of Ion Torrent sequencing. Using targeted gene panels, it was employed in cancer diagnostics to identify physical mutations in lung, bone, and colorectal excrescences, providing quick variant identification for a proven treatment. Ion Torrent was used in clinical microbiology research to sequence viral and bacterial infections, allowing for the profiling of antibiotic resistance and epidemic shadowing. It was employed in food safety to locate sources of contamination in dairy and water. Utilizing Ion Torrent for 16S rRNA-grounded profiling of soil and wastewater microbiomes, environmental research revealed contaminant-driven microbial changes and facilitated ecosystem monitoring [71].

2.3.6 QIAGEN: QIAGEN technologies have been widely applied in molecular, environmental, and clinical research. In oncology, the GeneReader platform and QIAseq panels assisted in identifying feasible mutations in colorectal and lung tumors, supporting ideas about targeted treatments. Contagious complaint research employed QIAamp processes and QIAGEN DNA/RNA birth equipment to detect diseases with high perceptivity, including as SARS-CoV-2, TB, and hepatitis. Microbiome researchers used QIAGEN PowerSoil equipment to describe gut and soil microbial populations from a variety of environmental and complaint perspectives. QIAGEN's Investigator equipment improved crime-scene DNA profiling and trustworthiness in forensics by enabling mortal identification from contaminated samples [72].

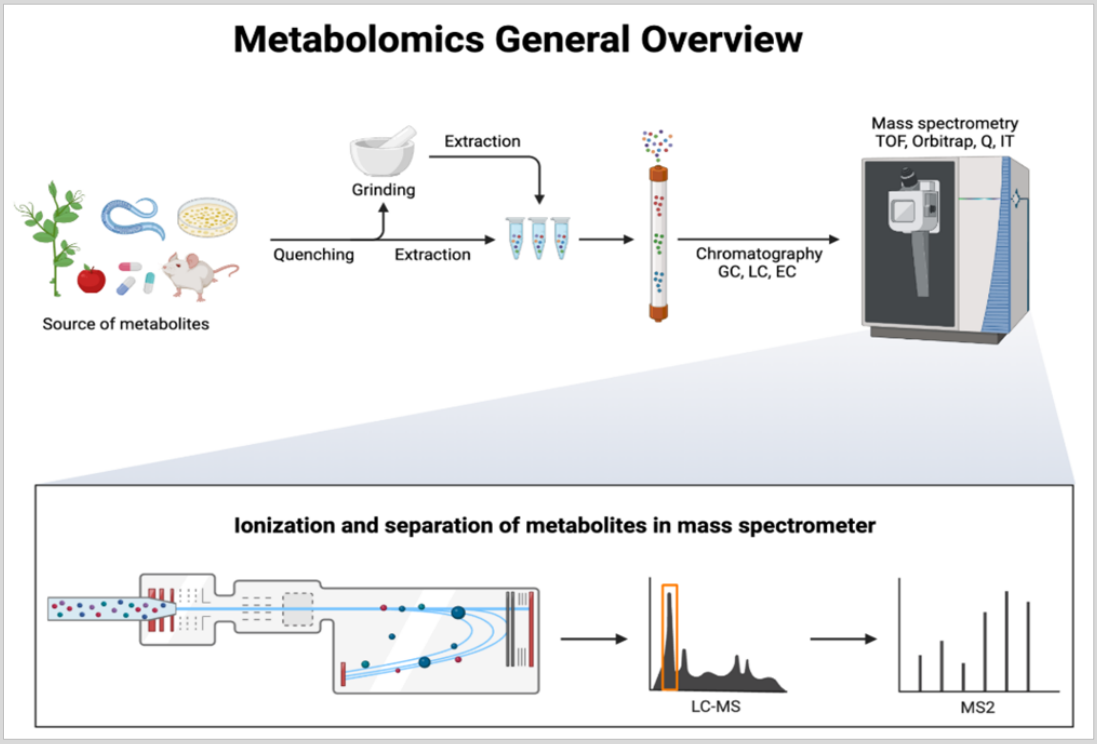

Figure 3: Multi-omics Integration Chart of Metabolomics Study Preview. [https://app.biorender.com]

2.4 Applications in Disease Biology

2.4.1 Gene Ontology and Community Annotations in Cancer: Understanding the diversity of cancer cells, as previously mentioned, is valuable since research has begun to show that there are relationships between these differences and overall patient survival when it comes to liver cancer. The new study unequivocally shows that transcriptome diversity, which is linked to genetic variety, is correlated with patient prognosis. A special bioinformatics workflow that used the Seurat package (version 2.3.0) in R (version 3.4.3) was used to accomplish this. Using single cell RNA sequencing data, the computational technique enables the identification of common populations across different datasets [73]. The creation of algorithms that make it easier to assess and analyse digital pathology images is the focus of this. Deep neural network training models are used in the construction of such an approach, which is a component of a semi-supervised method for effective differentiation between various community cells. In the end, it means building a library of digital images of diseased cells that are fed into a model to detect the cancerous cells in a new environment or when researchers are faced with categorising a new dataset.

2.4.2 Microbial Genomics in Cancer: The elucidation of microbial genomes has been employed mostly in cancer diagnosis, prognosis, and unravelling mechanisms of disease development using technologies like high throughput array panels and 16S rRNA gene sequencing [**Table. 5**]. Increased levels of the salivary microbiota (Capnocytophaga gingivalis, Streptococcus mitis, and Prevotella melaninogenica) have been identified as possible markers of oral cancer. Another study that used 16S rRNA gene sequencing revealed that Veillonella, Dialister, and Streptococcus species were more prevalent in oral tumours. When researchers examined the oral bacteria of 242 controls and 121 patients with oral cancer, they discovered notable variations in the two groups' microbial diversity and abundance [75].

Table 5: Some Web-Based Resources for Microbial Genomics [74]

Website	URL	Application
Integrated Microbial Genomes (IMG)	https://img.jgi.doe.gov/	Relative analysis and reflection of all intimately available genomes
CAMERA (Community Cyber infrastructure for Advanced Marine Microbial Ecology Research and Analysis)	https://bioinformaticshome.com/db/tool/CAMERA	A rich, distinctive data depository and a bioinformatics tools of metagenomic analysis.
DOE JGI Microbial Genomics Database	https://jgi.doe.gov/data-and-tools/software-tools/	Direct access to download sequence files, BLAST, and view reflections.
JCVI Comprehensive Microbial Resource	https://www.jcvi.org/research/software-tools	Display information on all of the intimately available, complete prokaryotic genomes.
Sanger Centre Bacterial Genomes	https://www.sanger.ac.uk/data/bacterial-data/	Concentrated on pathogens and model organisms. The point provides a list of systems funded, underway or completed.
RISSC – Ribosomal Internal Spacer Sequence Collection	http://ulises.umh.es/RISSC	Database of ribosomal 16S- 23S spacer sequences intended substantially for molecular biology studies in codifying, phylogeny and population genetics.

probeBase	http://www.microbial-ecology.net/probebase	Database of annotated rRNA- targeted oligonucleotide examinations and supporting. Rapid access to inquiry, microarray and reference data.
RRNDB	http://rrndb.cme.msu.edu/	Contains annotated information on rRNA operon dupe number among prokaryotes.
Bakta	https://bakta.computational.bio/	Rapid & standardized reflection of bacterial genomes, Metagenome-assembled Genomes & plasmids

2.4.3 Machine Learning and Data Science related to Inflammatory Bowel Disease and Bacterial Vaginosis: The size of datasets in microbiome investigations and the depth of sequencing per sample have grown as sequencing costs have decreased. Higher statistical power studies resulted from this, and as a result, Operational Taxonomic Unit (OUT) tables and functional profiles were transformed from end-goal deliverables into raw data for further analysis, including machine learning (ML) applications. Many people have successfully applied techniques like random forests (RF) in the context of diseases, such as correctly predicting BV and IBD based on taxonomic characteristics [76].

2.5 Microbiome Big Data for Pathogenic Diseases

2.5.1 QUIIME 2: The pipeline Quantitative Insights into Microbial Ecology (QIIME) 2 is specifically designed to analyse data obtained from amplicon sequencing of the 16S rRNA marker gene. The pipeline annotates marker gene sequences based on similar sequences in reference databases after grouping them into Operational Taxonomic Units (OTU) at a specific phylogenetic level [77].

2.5.2 Mothur: A software program called Mothur makes it possible to analyse sequence data. Because it incorporates a number of tools for sequence alignment, pairwise distance calculation, sequence quality screening, taxonomic analysis, diversity analysis, and functional analysis, among other things, its flexibility and user-friendliness are what give it its power. As a result, it operates as a powerful, stand-alone executable program that can execute on any platform.

2.5.3 Qiita: For conducting meta-analyses, Qiita offers a web-based platform that is very user-friendly for non-bioinformaticians. It gives access to widely used pipelines like GNPS and QIIME 2. The program's user-friendly interface, which enables users to move from data to statistical analysis and figure output, has several advantages over conventional pipelines [78].

2.5.4 MEGAN: For analytical purposes, MEGAN is especially interested in the use of shotgun sequencing as opposed to amplicon sequencing. As was previously said, this approach produces in-depth functional and taxonomic analysis, but only for very big data sets. Furthermore, the algorithm uses protein reference sequences to capture a higher percentage of evolutionary diversity.

2.5.5 Microbiome Analyst: The goal of Microbiome Analyst is to enable the study of the most prevalent, but frequently misunderstood, outputs generated from microbiome research by fusing an intuitive user interface with statistical and visualisation methods. A metabolic network visualisation tool, comparative analysis, and diversity profiling are all supported by the program [79].

2.5.6 iMAP: Another tool designed to give the microbiome research community an easy-to-use program that makes bioinformatics analysis and data visualisation easier is the Integrated Microbiome Analysis Pipeline (iMAP). Through wrapping functionalities, it enables diversity analysis, sequence classification, metadata profiling, and read quality control.

2.5.7 MIMIX: A new framework called the MIMIX aims to solve the difficulties of dealing with sparse, distributed microbiome data that is compositional rather than absolute, which could lead to improper correlations being produced by conventional approaches.

Table 6: Software Packages for Microbial NGS Data Analysis [80]

Application	Software	URL
Annotation	Prokka	https://sapac.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/prokka-genome-annotation.html
	RAST	https://www.mg-rast.org/
Assembly	BioNumerics	https://www.bionumerics.com/applications
	CLC Workbench	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/
	SeqSphere	https://www.ridom.de/seqsphere/
	SPAdes	https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/algorithmic-biology-lab-spades-genome-assembler.html
	Velvet	https://bioinformaticshome.com/tools/wga/descriptions/Velvet.html#gsc.tab=0
Data quality check	BaseSpace	https://sapac.illumina.com/products/by-type/informatics-products/basespace-sequence-hub.html
	BioNumerics	https://www.bionumerics.com/applications
	CLC Workbench	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/
	FastQC	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Identification	K-mer Finder	https://bio.tools/kmerfinder
	NCBI BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Metagenomics	MEGAN	https://bio.tools/megan
Phylogeny	FASTTree	https://bio.tools/fasttree
	RAxML	https://bio.tools/raxml

	SeqSphere	https://www.ridom.de/seqsphere/
	SNPTree	https://bioinformaticshome.com/tools/descriptions/snpTree.html#gsc.tab=0
Resistance	ARDB	ftp://ftp.cbcb.umd.edu/pub/data/ARDB/ARDBflatFiles.tar.gz
	CARD	http://arpcard.mcmaster.ca/
	ResFinder	https://bio.tools/resfinder
SNP calling	BioNumerics	https://www.bionumerics.com/applications
	CLC Workbench	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/
	Samtools	https://bio.tools/samtools
	SeqSphere	https://www.ridom.de/seqsphere/
Typing (wgMLST)	BIGSdb	https://pubmlst.org/software/bigsdb
	BioNumerics	https://www.bionumerics.com/applications
	CLC Workbench	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/
	EnteroBase	https://enterobase.warwick.ac.uk/
	SeqSphere	https://www.ridom.de/seqsphere/
Virulence	VFDB	https://www.mgc.ac.cn/VFs/
	Virulence Finder	https://cge.food.dtu.dk/services/VirulenceFinder/
Visualisation & Comparative Study	ACT	https://sanger-pathogens.github.io/Artemis/ACT/
	Artemis	https://www.sanger.ac.uk/tool/artemis/
	BRIG	https://sourceforge.net/projects/brig/
	ClustalW	https://www.genome.jp/tools-bin/clustalw
	DNA plotter	http://genocat.tools/tools/dnaplotter/
	WebACT	https://db.cngb.org/search/literature/16076890/?u_if=NWUwY2M0NzM1NDNlMDE1M2FhNjM1YmUyOWI1ZmU2YWR8Y25nYnwoQWJib3R0IEpDKVtBdXRob3Jd&pos=14

MIMIX software tool aims to comprehend how the environment affects alpha and beta diversity metrics.

2.6 Sequencing Techniques for Microbial Diseases

2.6.1 Roche 454 Genome Sequencing: Although availability to the technique is restricted and can be costly, 454 sequencing enables the creation of vast volumes of small sequences (less than 300 bp) per run due to early treatment with restriction enzymes [**Table. 6**]. Using adaptor fragments with complementary sequences, the procedure anneals each of the target DNA strands to a single resin bead [81].

2.6.2 Illumina: As a corporation, Illumina has a large selection of the most widely used sequencing technologies, a lengthy read-length, and affordable prices. It works by using fluorescently labelled nucleotides in sequencing by synthesis. A developing chain of nucleic acids is supplemented with deoxynucleoside triphosphates (dNTPs), which reversibly break the chain.

2.6.3 Qiagen Gene Reader: With fluorescently labelled nucleotides, the Qiagen gene reader operates similarly to Illumina. The gene reader combines many Qiagen technologies to enable sample preparation, sequencing, and analysis to be included in a single workflow [82].

2.6.4 Ion Torrent Sequencing Technology: Instead of using optical signals, ion torrent sequencing technology uses a tiny semiconductor to convert chemical data into detectable electrical data. The enzymatic cascade occurs in microwells, each of which has a nucleotide template that is used to catalyse additions in parallel.

2.6.5 SMRT Sequencing: Compared to conventional sequencing systems, single-molecule real-time (SMRT) sequencing has some special benefits. The method uses fluorescent dyes and sequencing by synthesis. There is a difference, though, in that the emissions from integration are picked up instantly. For this, zero mode wave guides are used [80].

2.6.6 Oxford Nanopore Sequencing: In order for nanopore sequencing to work, DNA molecules must be directed through a nanopore while an ionic current flows across it. As a result, a potential gradient is created throughout the solution that can be used to determine the length and shape of DNA. With comparatively small, inexpensive, and fast hardware, this method can avoid the requirement for PCR amplification or sample labelling [83].

3. Current Challenges

Understanding how bacteria affect mortal health and complaints is now mostly dependent on microbial informatics. The development of transcriptomics, proteomics, metagenomics, metabolomics, and high-outturn sequencing has made it possible for researchers to record microbial dynamics at unexplored resolution. However, there are still a number of scientific, computational, and translational issues with employing multi-omics for complaint-associated microbiology, which continue to restrict precision, repeatability, and clinical relevance. Data variety, which results from variations in experimental procedures, sequencing platforms, sample runs, and logical channels, is a significant difficulty. Cross-study comparisons are challenging because microbial populations differ among individuals, body locations, diets, and environmental exposures. Moreover, the scale, complexity, and noise position of multi-omics datasets vary. Strong normalizing and cross-modality mapping techniques, which are presently lacking in microbial systems, are needed to integrate these layers, which include genetic variants, transcriptional signatures, protein networks, and metabolic biographies. Reproducibility is made even more difficult by the absence of consistent metadata reporting. Inadequate microbiological reference databases are another significant drawback. There are many unidentified microbial species in clinical or environmental samples, which results in significant portions of "unknown" readings in metagenomic data. Accurate taxonomic profiling and functional reflection are hampered. Because standard reference-grounded alignment is unable to capture microdiversity

relevant to pathogenicity or antimicrobial resistance, vertical gene transfer, strain-position variation, and genomic malleability introduce additional complication [84,85]. Pangenome-apprehensive methods and high-resolution strain reconstruction are still developing, and their computing costs are still significant. Additionally, the field has a lot of computational difficulties. Advanced machine literacy, network modeling, and systems biology techniques that can manage high-dimensional, sparse, and often longitudinal data are necessary for multi-omics integration. Microbial datasets, however, are inherently compositional and defy the hypotheses of many traditional statistical approaches. Although deep literacy has demonstrated promise, it is constrained by the shortcomings of labeled microbiological complaint datasets and the dangers of overfitting models [86,87]. Additionally, improved pall-grounded structures, memory-efficient assembly techniques, and high-performance computing are needed for spanning multi-omics operations. Natural interpretation presents another challenge. Although there are several links between microorganisms and complaint characteristics, the cause is yet unknown. The relationships between microbes and host impunity, metabolism, signaling pathways, and epigenetic changes are generally intricate. Large point sets are produced by multi-omics, but connecting molecular signatures to mechanistic pathways sometimes necessitates reciprocal trials, gnotobiotic models, or CRISPR-based manipulation—all of which are impractical for many organisms [88,89]. Additionally, the microbiota is highly dynamic, making it difficult to discern between changes brought on by complaints and those that occur naturally. Clinical microbiology continues to face translational difficulties. Only a small portion of the results of extensive multi-omics research are applied to individual or corrective operations. Relinquishment is hampered by regulatory concerns, the absence of clinical confirmation cohorts, and the challenge of interpreting multi-omics biomarkers. Complexity is increased by concerns about data sequestration, concurrence, and the moral use of metagenomic data [90,91]. Additionally, pathogen discovery, outbreak shadowing, and antibiotic resistance surveillance continue to be hampered by fragmented data architectures and insufficient real-time analytics. Interdisciplinary moxie will eventually become urgently needed. Microbiologists, physicians, bioinformaticians, data scientists, statisticians, and systems biologists must work together for computational microbiology to be effective. The skills gap is increased by the lack of experimenters with multi-omics analytics training as well as the quick speed of technology development. Microbial informatics has enormous potential in complaint biology, but fully using multi-omics for improving mortal health requires overcoming obstacles in data integration, computation, mechanistic sapience, and clinical restatement [92].

4. Future Directions

The role of Microbial Communities in Disease is the human body harbour a diverse community of microorganisms, collectively referred to as the microbiome. These communities play significant roles in maintaining health but are also implicated in various diseases. For instance, dysbiosis a microbial imbalance has been associated with conditions such as inflammatory bowel disease (IBD), diabetes, and even obesity. Microbial informatics enables researchers to analyze the complex datasets generated from microbial sequencing and metagenomic studies, facilitating the identification of microbial signatures associated with these diseases. High-Throughput Sequencing and Data Analysis depicts high-throughput sequencing technologies, such as 16S rRNA sequencing and whole-genome shotgun sequencing, have transformed microbial research by providing extensive genetic information about microbial communities. These technologies generate vast amounts of data, necessitating advanced computational tools for analysis. Taxonomic Profiling and Diversity Assessment through Microbial informatics employs computational methods to characterize microbial communities, uncoupling the intricate relationships among microbes and their functions. Tools like QIIME (Quantitative Insights Into Microbial Ecology) and Mothur allow researchers to perform taxonomic profiling, revealing alpha and beta diversity within microbiomes [93,94]. By comparing microbial community structures between healthy and diseased states, researchers can identify specific taxa that may contribute to disease. Metagenomic approaches empower researchers to analyze the collective genomic content of all microorganisms within a defined environment. Computational tools such as METAWATT and HUMAnN (The Hierarchical- and Unified-Metagenomic-Analysis-for-Natural-and-Nurtured-Organisms) enable the functional profiling of microbial communities, linking specific gene functions to disease pathogenesis. Whole-genome sequencing (WGS) of pathogens can reveal genetic variations and virulence factors involved in infection. Tools like PathogenWatch and Genome Detective utilize genomic sequences to track pathogen outbreaks and predict antimicrobial resistance, contributing to public health responses. Immune Response Analysis and Microbial informatics also enhances our understanding of the host's immune response to microbial invasions. The integration of transcriptomic and proteomic data with microbial genomics provides insights into how host immune pathways are modulated by microbial presence. Computational models can simulate these interactions, helping to identify novel therapeutic targets and strategies to manipulate the host immune response. Disease Dynamics and Epidemiology underlie the dynamics of disease transmission is another vital application of microbial informatics. Computational modeling and epidemiological tools are used to predict outbreaks and assess the impact of interventions. Phylogenetic Analysis and Evolutionary Tracking allow researchers to trace the evolutionary relationships between pathogens, identifying sources and routes of transmission [95]. Software like MEGA (Molecular Evolutionary Genetics Analysis) and RAxML (Randomized Axelerated Maximum Likelihood) facilitate the construction of phylogenetic trees based on genomic data, aiding in disease outbreak investigations. Predictive modeling frameworks enable researchers to simulate the spread of infectious diseases and evaluate control strategies [96].

5. Conclusion

Microbial informatics stands at the forefront of disease biology, offering unprecedented insights into the complex interactions between microorganisms and human health. Through the integration of advanced computational tools, high-throughput sequencing, and interdisciplinary approaches, researchers are better equipped to understand the microbial underpinnings of diseases. As the field continues to evolve, addressing the challenges of data integration, ethical considerations, and the need for collaborative approaches will be crucial in harnessing the full potential of microbial informatics for public health benefits. The journey ahead promises not only to deepen our understanding of disease mechanisms but also to inform the development of innovative therapeutic strategies and interventions. Since full-length marker gene sequencing improves taxonomic resolution and more contiguous metagenome assemblies improve functional analyses, we anticipate that long-read sequencing methods will soon be widely employed in microbiome

studies as costs decline. Developments in this area are probably going to concentrate on converting bioinformatics protocols that have already been developed for short reads to long read equivalents, such as read classification and denoising techniques. Additional difficulties arise from shotgun metagenome analysis in large research, where capacity problems may arise from costly computations required in de novo assembly and annotation. The cloud offers a flexible alternative for businesses unable to finance massive on-premise compute infrastructures, where proficiency in cloud computing becomes crucial.

6. List of Abbreviations

DNA: de-oxy-ribonucleic acid, NGS: next generation sequencing, 16S rRNA: 16 subunit ribosomal ribonucleic acid, PCR: polymerase chain reaction, PCR-RAPD: polymerase chain reaction random amplification of polymorphic DNA, AFLP: amplified fragment length polymorphism, CDS: coding sites, RBS: ribosomal binding sites, BPGA: Bacterial Pan Genome Analysis, CAMBER: Comparative Analysis of Multiple Bacterial Strains, ITEP: Integrated Toolkit for Exploration of Pan-genomes, PanCGHweb: Pangenome Comparative Genome Hybridization web tool: PANGP: Pan-Genome Profile Tool, PanSeq: Pan-genome sequence analysis, PGAP: Prokaryotic Genome Annotation Pipeline, PGAT: Prokaryotic-genome Analysis Tool, ROARY: Rapid large-scale prokaryote pan genome analysis, SNP: single-nucleotide polymorphisms, GBP: Giga-bases, METAHit: METAgenomics of the Human Intestinal Tract, AMOS: A modular open source whole genome assembler, METAMOS: A modular framework for metagenomic assembly, analysis and annotation, SOAP: Short Oligonucleotide Analysis Package, ESOM: Emergent Self-Organizing Map, MEGAN: MEtaGenome Analyzer, WIMP: Windows, Icons, Menus, Pointers, eHOMD: enlarged Human Oral Microbiome Database, ITS: Internally Transcribed Spacer, COI: Cytochrome c oxidase subunit I, PHAST and PHASTER: PHAge Search Tool and PHAge Search Tool Enhanced Release, QIIME: Quantitative Insights Into Microbial Ecology, Metaphlan: Metagenomic Phylogenetic Analysis, Strainphlan: Strain PHyLogenetic Analysis, Metatrans: Metatranscriptomics, GOLD: Genomes Online Database, CDD: Conserved Domain Database, Pfam: Protein Families, RDP: Ribosomal Database Project, STRING: Search Tool for Retrieval of the Interacting Genes or Proteins, SILVA: SILva ribosomal RNA gene database project, GREENGENES: 16S rRNA gene database, BigSdb: Bacterial Isolate Genome Sequence Database, EcoCyc: Encyclopedia of Escherichia coli Genes and Metabolism, EBI Metagenomics: European Bioinformatics Institute Metagenomics, PATRIC: Pathosystems Resource Integration Center, EUPathDBs: Eukaryotic Pathogen Database, TBDB: T-box riboswitch annotation database, TCDB: Transporter Classification Database, Metacyc: Curated Database of Metabolic Pathways and Enzymes, KEGG: Kyoto Encyclopedia of Genes and Genome, MiST: Molecular Interaction Search Tool, IMG: Integrated Microbial Genomes, CAMERA: Community Cyber infrastructure for Advanced Marine Microbial Ecology Research and Analysis, JGI: Joint Genome Institute, RISSC: Ribosomal Internal Spacer Sequence Collection, RRNDB: Ribosomal RNA Operon Copy Number Database, JCVI CMR: J. Craig Venter Institute Comprehensive Microbial Resource, OUT: Operational Taxonomic Unit, ML: Machine Learning, RF: Random Forests, IBD: Inflammatory Bowel Disease, BV: Bacterial Vaginosis, GNPS: Global Natural Products Social Molecular Networking, iMAP: Integrated Microbiome Analysis Pipeline, MIMIX: Microbiome Mixed Model, PROKKA: Prokaryotic Genome Annotation Toolkit, RAST: Rapid Annotation using Subsystem Technology, SPAdes: St. Petersburg genome assembler, FASTQC: Fast Quality control tool, NCBI: National Center for Biotechnology Information, BLAST: Basic Local Alignment Search Tool, RAxML: Randomized Axelerated Maximum Likelihood, SNPTree: Single Nucleotide Polymorphism Detection and Construction of a Phylogenetic Tree, ARDB: Antibiotic Resistance Genes Database, CARD: Comprehensive Antibiotic Resistance Database, SAM Tool: Sequence Alignment/Map Tool, wgMLST: Whole Genome Multilocus Sequence Typing, VFDB: Virulence Factor Database, ACT: Artemis Comparison Tool, BRIG: BLAST Ring Image Generator, dNTPs: deoxynucleoside triphosphates, SMRT: single-molecule real-time, HUMAnN: The Hierarchical- and Unified-Metagenomic-Analysis-for-Natural-and-Nurtured-Organisms, WGS: Whole-genome sequencing, MEGA: Molecular Evolutionary Genetics Analysis.

7. Declarations

Ethical Approval and Consent to participate

The study received ethical approval from the appropriate institutional committee. All authors provided informed consent to participate and agreed to comply with ethical research standards.

Consent for publication

I, the author, hereby give full consent for the publication of my work in all formats and agree to the terms and conditions.

Availability of supporting data

Yes

Competing interests

The authors declare that there is no conflict of interests.

Funding

Not Applicable.

Authors' contributions

D.S: Original Draft Writing, Editing, Formal Analysis and Data Curation.
M.B: Conceptualization and Supervision.

Acknowledgement

The authors are thankful to the Chancellor of Techno India University, Kolkata for providing the necessary infrastructure facilities.

References

[1] Rana S, Singh V, Preeti, Bhardwaj N. Bioinformatics in microbial biotechnology: a genomics and proteomics perspective. Innovations in Information and Communication Technology Series. 2021:53-67.

[2] Liu H, Xin B, Zheng J, et al. Build a bioinformatic analysis platform and apply it to routine analysis of microbial genomics and comparative genomics. Protocol Exchange. Published April 25, 2022. doi:10.21203/rs.2.21224/v6

[3] Antipov D, Korobeynikov A, McLean J, Pevzner P. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics. 2015;32(7):1009-1015. doi:10.1093/bioinformatics/btv688

[4] Arboleya S, Bottacini F, O'Connell-Motherway M, et al. Gene-trait matching across the Bifidobacterium longum pan-genome reveals considerable diversity in carbohydrate catabolism among human infant strains. BMC Genomics. 2018;19(1):33. doi:10.1186/s12864-017-4388-9

[5] Ayling M, Clark M, Leggett R. New approaches for metagenome assembly with short reads. Brief Bioinform. 2019;20(6):1-11. doi:10.1093/bib/bbz020

[6] Beck D, Foster J. Machine learning techniques accurately classify microbial communities by bacterial vaginosis characteristics. PLoS One. 2014;9(2):e87830. doi:10.1371/journal.pone.0087830

[7] Bishara A, Moss E, Kolmogorov M, et al. High-quality genome sequences of uncultured microbes by assembly of read clouds. Nat Biotechnol. 2018;36(11):1067-1075. doi:10.1038/nbt.4266

[8] Yadav BS, Chauhan P, Kushwaha S. Bioinformatics resources for microbial research in biological systems. In: Tripathi V, Kumar P, Tripathi P, Kishore A, Kamle M, eds. Microbial Genomics in Sustainable Agroecosystems. Singapore: Springer; 2019. doi:10.1007/978-981-32-9860-6_3

[9] Zhou A, Xie S, Zhang Y, et al. Environmental Science and Pollution Research. 2021;28(29):38839. doi:10.1007/s11356-021-13231-y

[10] Azuaje FJ, Heymann M, Ternes AM, et al. Bioinformatics as a driver, not a passenger, of translational biomedical research: perspectives from the 6th Benelux bioinformatics conference. J Clin Bioinform. 2012;2(1):7.

[11] Zhang C, Xu S, Xu D. Bioinformatics applications in clinical microbiology. In: Wei DQ, Ma Y, Cho W, Xu Q, Zhou F, eds. Translational Bioinformatics and Its Application. Translational Medicine Research. Dordrecht: Springer; 2017. doi:10.1007/978-94-024-1045-7_15

[12] Du RF, Li JY, Liu JL, Zhao JZ. Application of bioinformatics in microbial ecology. Adv Mater Res. 2014;955–959:276-280. doi:10.4028/www.scientific.net/AMR.955-959.276

[13] Rhoads DD, Sintchenko V, Rauch CA, Pantanowitz L. Clinical microbiology informatics. Clin Microbiol Rev. 2014;27(4):1025-1047. doi:10.1128/CMR.00049-14

[14] El-Kareh R, Roy C, Williams DH, Poon EG. Impact of automated alerts on follow-up of post-discharge microbiology results: a cluster randomized controlled trial. J Gen Intern Med. 2012;27(10):1243-1250. doi:10.1007/s11606-012-2054-1

[15] Campos JM. Clinical microbiology informatics—its time has come. Rev Med Microbiol. 2003;14(3):105-108.

[16] Kim H. Bioinformatics tools in clinical microbiology and infectious disease prevention algorithms. Adv Biotech Micro. 2018;11(4):555820. doi:10.19080/AIBM.2018.11.55582

[17] Agüero F, Al-Lazikani B, Aslett M, et al. Genomic-scale prioritization of drug targets: The TDR Targets database. Nat Rev Drug Discov. 2008;7(11):900-907.

[18] Akarsu H, Bordes P, Mansour M, et al. TASmania: a bacterial toxin-antitoxin systems database. PLoS Comput Biol. 2019;15(4):e1006946. doi:10.1371/journal.pcbi.1006946

[19] Anand S, Kuntal BK, Mohapatra A, et al. FunGeCo: a web-based tool for estimation of functional potential of bacterial genomes and microbiomes using gene context information. Bioinformatics. 2020;36(8):2575-2577. doi:10.1093/bioinformatics/btz961

[20] Arango-Argoty GA, Dai D, Pruden A, et al. NanoARG: a web service for detecting and contextualizing antimicrobial resistance genes from nanopore-derived metagenomes. Microbiome. 2019;7(1):88. doi:10.1186/s40168-019-0703-9

[21] Aziz RK, Bartels D, Best A, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:1-15. doi:10.1186/1471-2164-9-75

[22] Besser J, Carleton HA, Gerner-Smidt P, et al. Next-generation sequencing technologies and their application to the study and control of bacterial infections. Clin Microbiol Infect. 2018;24(4):335-341. doi:10.1016/j.cmi.2017.10.013

[23] Boetzer M, Pirovano W. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. BMC Bioinformatics. 2014;15(1):211. doi:10.1186/1471-2105-15-211

[24] Boisvert S, Raymond F, Godzaridis É, Laviolette F, Corbeil J. Ray Meta: scalable de novo metagenome assembly and profiling. Genome Biol. 2012;13(12):R122. doi:10.1186/gb-2012-13-12-r122

[25] Tabish M, Azim S, Hussain MA, Rehman SU, Sarwar T, Ishqi HM. Bioinformatics approaches in studying microbial diversity. In: Malik A, Grohmann E, Alves M, eds. Management of Microbial Resources in the Environment. Dordrecht: Springer; 2013. doi:10.1007/978-94-007-5931-2_6

[26] Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012;6(8):1621-1624. doi:10.1038/ismej.2012.8

[27] Novais RC, Thorstenson Y. The evolution of Pyrosequencing® for microbiology: from genes to genomes. J Microbiol Methods. 2011;86(1):1-7. doi:10.1016/j.mimet.2011.03.002

[28] Caporaso JG, Lauber CL, Walters WA, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A. 2011;108(Suppl 1):4516-4522. doi:10.1073/pnas.1000080107

[29] Pawar S, Tan YC, Lahiri C. Bioinformatics algorithms and software for predicting microbiomes. In: Bramhachari PV, ed. Microbiome in Human Health and Disease. Singapore: Springer; 2021. doi:10.1007/978-981-16-3156-6_14

[30] Bansal AK. Bioinformatics in microbial biotechnology: a mini review. Microb Cell Fact. 2005;4:19. doi:10.1186/1475-2859-4-19

[31] Aloy P, Russell RB. Ten thousand interactions for molecular biologists. Nat Biotechnol. 2004;22(10):1317-1321. doi:10.1038/nbt1018

[32] Apweiler R, Attwood TK, Bairoch A, et al. InterPro: an integrated documentation resource for protein families, domains, and functional sites. Bioinformatics. 2000;16(12):1145-1150. doi:10.1093/bioinformatics/16.12.1145

[33] Prabha R, Rai A, Singh DP. Bioinformatics-driven big data analytics in microbial research. In: Wang B, Li R, Perrizo W, eds. Big Data Analytics in Bioinformatics and Healthcare. Hershey, PA: IGI Global; 2015:265-283. doi:10.4018/978-1-4666-6611-5.ch012

[34] Upadhyayula RS, Solanki PS, Suravajhala P, Medicherla KM. Bioinformatics tools for microbial diversity analysis. In: Satyanarayana T, Johri B, Das S, eds. Microbial Diversity in Ecosystem Sustainability and Biotechnological Applications. Singapore: Springer; 2019. doi:10.1007/978-981-13-8315-1_2

[35] Narayan A, Singh A, Kumar S. Understanding microbiome science through big data analysis. In: Singh S, ed. Metagenomic Systems Biology. Singapore: Springer; 2020. doi:10.1007/978-981-15-8562-3_3

[36] Ghosh A, Firdous S, Saha S. Bioinformatics for human microbiome. In: Singh V, Kumar A, eds. Advances in Bioinformatics. Singapore: Springer; 2021. doi:10.1007/978-981-33-6191-1_17

[37] Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L, Ellis JT. Machine learning and applications in microbiology. FEMS Microbiol Rev. 2021;45(5):fuab015. doi:10.1093/femsre/fuab015

[38] Turnbaugh PJ, Ley RE, Hamady M, et al. The Human Microbiome Project. Nature. 2007;449(7164):804-810. doi:10.1038/nature06244

[39] Wise MJ, Wanner BL. Welcome to Microbial Informatics and Experimentation. Microb Inform Exp. 2011;1(1):1. doi:10.1186/2042-5783-1-1

[40] Galloway-Peña J, Hanson B. Tools for analysis of the microbiome. Dig Dis Sci. 2020;65(3):674-685. doi:10.1007/s10620-020-06091-y

[41] Schirmer M, Smeekens SP, Vlamakis H, et al. Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. Cell. 2016;167(4):1125-1136.e8. doi:10.1016/j.cell.2016.10.020

[42] Xu S, Li X, Zhang S, et al. Oxidative stress gene expression, DNA methylation, and gut microbiota interaction trigger Crohn's disease: a multi-omics Mendelian randomization study. BMC Med. 2023;21(1):179. Published 2023 May 11. doi:10.1186/s12916-023-02878-8

[43] Ning L, Zhou YL, Sun H, et al. Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. Nat Commun. 2023;14(1):7135. Published 2023 Nov 6. doi:10.1038/s41467-023-42788-0

[44] Zhernakova A, Kurilshikov A, Bonder MJ, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science. 2016;352(6285):565-569. doi:10.1126/science.aad3369

[45] Huang J, Liu S, Li P, et al. Multi-omics analysis of gut-brain axis reveals novel microbial and neurotransmitter signatures in patients with arteriosclerotic cerebral small vessel disease. Pharmacol Res. 2024;208:107385. doi:10.1016/j.phrs.2024.107385

[46] Vich Vila A, Collij V, Sanna S, et al. Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. Nat Commun. 2020;11(1):362. Published 2020 Jan 17. doi:10.1038/s41467-019-14177-z

[47] Li X, Hu S, Shen X, et al. Multiomics reveals microbial metabolites as key actors in intestinal fibrosis in Crohn's disease. EMBO Mol Med. 2024;16(10):2427-2449. doi:10.1038/s44321-024-00129-8

[48] Huang Y, Lu W, Zeng M, et al. Mapping the early life gut microbiome in neonates with critical congenital heart disease: multiomics insights and implications for host metabolic and immunological health. Microbiome. 2022;10(1):245. Published 2022 Dec 30. doi:10.1186/s40168-022-01437-2

[49] Rozera T, Pasolli E, Segata N, Ianiro G. Machine Learning and Artificial Intelligence in the Multi-Omics Approach to Gut Microbiota. Gastroenterology. 2025;169(3):487-501. doi:10.1053/j.gastro.2025.02.035

[50] Stenton SL, Pejaver V, Bergquist T, et al. Assessment of the evidence yield for the calibrated PP3/BP4 computational recommendations. Genet Med. 2024;26(11):101213. doi:10.1016/j.gim.2024.101213

[51] Thomas AM, Manghi P, Asnicar F, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat Med. 2019;25(4):667-678. doi:10.1038/s41591-019-0405-7

[52] Wang Y, Zhou J, Ye J, et al. Multi-omics reveal microbial determinants impacting the treatment outcome of antidepressants in major depressive disorder. Microbiome. 2023;11(1):195. Published 2023 Aug 28. doi:10.1186/s40168-023-01635-6

[53] Dai Z, Coker OO, Nakatsu G, et al. Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. Microbiome. 2018;6(1):70. Published 2018 Apr 11. doi:10.1186/s40168-018-0451-2

[54] Li T, Zhao Z, Peng M, et al. Multi-omics analysis reveals the interplay between intratumoral bacteria and glioma. mSystems. 2025;10(1):e0045724. doi:10.1128/msystems.00457-24

[55] Jacobs JP, Lagishetty V, Hauer MC, et al. Multi-omics profiles of the intestinal microbiome in irritable bowel syndrome and its bowel habit subtypes. Microbiome. 2023;11(1):5. Published 2023 Jan 10. doi:10.1186/s40168-022-01450-5

[56] Climacosa FMM, Anlacan VMM, Gordovez FJA, et al. Monitoring drug Efficacy through Multi-Omics Research initiative in Alzheimer's Disease (MEMORI-AD): A protocol for a multisite exploratory prospective cohort study on the drug response-related clinical, genetic, microbial and metabolomic signatures in Filipino patients with Alzheimer's disease. BMJ Open. 2024;14(11):e078660. Published 2024 Nov 27. doi:10.1136/bmjopen-2023-078660

[57] Murai H, Kodama T, Maesaka K, et al. Multiomics identifies the link between intratumor steatosis and the exhausted tumor immune microenvironment in hepatocellular carcinoma. Hepatology. 2023;77(1):77-91. doi:10.1002/hep.32573

[58] Kim H, Nelson P, Nzabarushimana E, et al. Multi-omic analysis reveals transkingdom gut dysbiosis in metabolic dysfunction-associated steatotic liver disease. Nat Metab. 2025;7(7):1476-1492. doi:10.1038/s42255-025-01318-6

[59] Gao X, Sun R, Jiao N, et al. Integrative multi-omics deciphers the spatial characteristics of host-gut microbiota interactions in Crohn's disease. Cell Rep Med. 2023;4(6):101050. doi:10.1016/j.xcrm.2023.101050

[60] Farias Amorim C, Lovins VM, Singh TP, et al. Multiomic profiling of cutaneous leishmaniasis infections reveals microbiota-driven mechanisms underlying disease severity. Sci Transl Med. 2023;15(718):eadh1469. doi:10.1126/scitranslmed.adh1469

[61] Yang Y, Olah P, Radai Z, et al. Exploratory multi-omics analysis reveals host-microbe interactions associated with disease severity in psoriatic skin. EBioMedicine. 2024;105:105222. doi:10.1016/j.ebiom.2024.105222

[62] Huang L, Rao Q, Wang C, et al. Multi-omics joint analysis reveals that the Miao medicine Yindanxinnaotong formula attenuates non-alcoholic fatty liver disease. Phytomedicine. 2024;135:156026. doi:10.1016/j.phymed.2024.156022

[63] Xiong R, Aiken E, Caldwell R, et al. AI-driven multi-omics modeling of myalgic encephalomyelitis/chronic fatigue syndrome. Nat Med. 2025;31(9):2991-3001. doi:10.1038/s41591-025-03788-3

[64] Ali RO, Quinn GM, Umarova R, et al. Longitudinal multi-omics analyses of the gut-liver axis reveals metabolic dysregulation in hepatitis C infection and cirrhosis. Nat Microbiol. 2023;8(1):12-27. doi:10.1038/s41564-022-01273-y

[65] Wang Q, Sun LN, Shi H, et al. Multi-omics analysis reveals gut microbiota-metabolite interactions and their association with liver function in autoimmune overlap syndrome. World J Gastroenterol. 2025;31(25):106371. doi:10.3748/wjg.v31.i25.106371

[66] Gwak HJ, Lee SJ, Rho M. Application of computational approaches to analyze metagenomic data. J Microbiol. 2021;59(3):233-241. doi:10.1007/s12275-021-0632-8

[67] Wang D, Cui Y, Cao Y, He Y, Chen H. Human Microbe-Disease Association Prediction by a Novel Double-Ended Random Walk with Restart. Biomed Res Int. 2020;2020:3978702. Published 2020 Aug 10. doi:10.1155/2020/3978702

[68] Shi H, Wu M, Wu X, et al. Multi-omics integration reveals functional signatures of gut microbiome in atherosclerosis. Gut Microbes. 2025;17(1):2542384. doi:10.1080/19490976.2025.2542384

[69] Song Y, Zhou X, Zhao H, et al. Characterizing the role of the microbiota-gut-brain axis in cerebral small vessel disease: An integrative multi omics study. Neuroimage. 2024;303:120918. doi:10.1016/j.neuroimage.2024.120918

[70] Bai J, Eldridge R, Houser M, et al. Multi-omics analysis of the gut microbiome and metabolites associated with the psychoneurological symptom cluster in children with cancer receiving chemotherapy. J Transl Med. 2024;22(1):256. Published 2024 Mar 9. doi:10.1186/s12967-024-05066-1-98o

[71] Liu Z, Ma A, Mathé E, Merling M, Ma Q, Liu B. Network analyses in microbiome based on high-throughput multi-omics data. Brief Bioinform. 2021;22(2):1639-1655. doi:10.1093/bib/bbaa005

[72] Shan M, Xiao M, Xu J, et al. Multi-omics analyses reveal bacteria and catalase associated with keloid disease. EBioMedicine. 2024;99:104904. doi:10.1016/j.ebiom.2023.104904

[73] Wang Y, Zeng T, Tang D, Cui H, Wan Y, Tang H. Integrated Multi-Omics Analyses Reveal Lipid Metabolic Signature in Osteoarthritis. J Mol Biol. 2025;437(6):168888. doi:10.1016/j.jmb.2024.168888

[74] Ungaro F, Massimino L, D'Alessio S, Danese S. The gut virome in inflammatory bowel disease pathogenesis: From metagenomics to novel therapeutic approaches. United European Gastroenterol J. 2019;7(8):999-1007. doi:10.1177/2050640619876787

[75] Tiew PY, Mac Aogain M, Ali NABM, et al. The Mycobiome in Health and Disease: Emerging Concepts, Methodologies and Challenges. Mycopathologia. 2020;185(2):207-231. doi:10.1007/s11046-019-00413-z

[76] Du L, Zhang K, Liang L, et al. Multi-omics analyses of the gut microbiota and metabolites in children with metabolic dysfunction-associated steatotic liver disease. mSystems. 2025;10(4):e0114824. doi:10.1128/msystems.01148-24

[77] Jiang T, Li C, Li Y, et al. Multi-omics and bioinformatics for the investigation of therapeutic mechanism of roucongrong pill against postmenopausal osteoporosis. J Ethnopharmacol. 2025;337(Pt 2):118873. doi:10.1016/j.jep.2024.118873

[78] Tiew PY, Meldrum OW, Chotirmall SH. Applying Next-Generation Sequencing and Multi-Omics in Chronic Obstructive Pulmonary Disease. Int J Mol Sci. 2023;24(3):2955. Published 2023 Feb 3. doi:10.3390/ijms24032955

[79] Pita-Galeana MA, Ruhle M, López-Vázquez L, de Anda-Jáuregui G, Hernández-Lemus E. Computational Metagenomics: State of the Art. Int J Mol Sci. 2025;26(18):9206. Published 2025 Sep 20. doi:10.3390/ijms26189206

[80] Ben Khedher M, Ghedira K, Rolain JM, Ruimy R, Croce O. Application and Challenge of 3rd Generation Sequencing for Clinical Bacterial Studies. Int J Mol Sci. 2022;23(3):1395. Published 2022 Jan 26. doi:10.3390/ijms23031395

[81] Rhoads DD, Sintchenko V, Rauch CA, Pantanowitz L. Clinical microbiology informatics. Clin Microbiol Rev. 2014;27(4):1025-1047. doi:10.1128/CMR.00049-14

[82] Sabir M, Sohail A, Saeed A, Abbas S, Khan KA, Khalil M. Bioinformatics applied to microbiology. In: Microbiology in the Era of Artificial Intelligence. Boca Raton, FL: CRC Press; 2025:209-243.

[83] Kim H. Bioinformatics tools in clinical microbiology and infectious disease prevention algorithms. J Med Biomed Appl Sci. 2018;6(9):159-165.

[84] Ghosh R, Bera RD, Chandra S. An overview of bioinformatics and its application: a ray of hope towards microbial biotechnology. Int J Multidiscip Res. 2024;6(1):1-23.

[85] Jiang Y, Luo J, Huang D, Liu Y, Li DD. Machine learning advances in microbiology: a review of methods and applications. Front Microbiol. 2022; 13:925454. doi:10.3389/fmicb.2022.925454

[86] Kim H. Bioinformatics technology in clinical and public health microbiology applying computational methods. J Med Biomed Appl Sci. 2018;6(9):136-140.

[87] Sharma P, Dahiya S, Kaur P, Kapil A. Computational biology: role and scope in taming antimicrobial resistance. Indian J Med Microbiol. 2023;41:33-38. doi:10.1016/j.ijmmb.2023.02.008

[88] Raslan MA, Raslan SA, Shehata EM, Mahmoud AS, Sabri NA. Advances in the applications of bioinformatics and chemoinformatics. Pharmaceuticals (Basel). 2023;16(7):1050. doi:10.3390/ph16071050

[89] Durmuş S, Çakır T, Özgür A, Guthke R. A review on computational systems biology of pathogen–host interactions. Front Microbiol. 2015;6:235. doi:10.3389/fmicb.2015.0023

[90] Wen Z, Yan C, Duan G, Li S, Wu FX, Wang J. A survey on predicting microbe–disease associations: biological data and computational methods. Brief Bioinform. 2021;22(3):bbab001. doi:10.1093/bib/bbab001

[91] Sintchenko V. Informatics for infectious disease research and control. In: Infectious Disease Informatics. New York, NY: Springer; 2009:1-26. doi:10.1007/978-1-4419-0879-1_1

[92] Qu K, Guo F, Liu X, Lin Y, Zou Q. Application of machine learning in microbiology. Front Microbiol. 2019;10:827. doi:10.3389/fmicb.2019.00827

[93] Kuznetsov V, Lee HK, Maurer-Stroh S, et al. How bioinformatics influences health informatics: usage of biomolecular sequences, expression profiles and automated microscopic image analyses for clinical needs and public health. Health Inf Sci Syst. 2013;1(1):2. doi:10.1186/2047-2501-1-2

[94] Ali MA, Zahoor A, Niaz Z, et al. Bioinformatics and computational biology. In: Trends in Plant Biotechnology. 2024:281-334.

[95] Rukan M, Ejaz M, Hamza M, Munir HMW. Computational biology and its role in life science research. In: Microbiology in the Era of Artificial Intelligence. Boca Raton, FL: CRC Press; 2025:157-186.

[96] Evans RS, Pestotnik SL. Applications of medical informatics in antibiotic therapy. Adv Exp Med Biol. 1994;349:87-96. doi:10.1007/978-1-4757-9206-5_8.