**Chapter 22**

**Genomics and Proteomics in Pathogen Typing**

**Abubakar Yaro1, 2, 3**

1. Dept of Infectious Diseases and Public Health, AHRO Center for Academic Research, London, UK
2. Africa Health Research Organization, Accra, Ghana
3. Dr Yaro Laboratory ltd, Accra Ghana & Kano, Nigeria

**Abstract**

Pathogen typing is a technique that aids in identifying and characterization of different strains of pathogens including viruses, bacteria, and fungi in which the genetic material of other characteristics is used. Although in the past traditional methods have been used for pathogen typing, advances in omics technology has provided us with accurate and rapid approaches of pathogens typing. This chapter review the role of genomics and proteomics in addressing key issues with regards to identification and characterising pathogens. With the threat of emerging and re-emerging pathogens on the increase as experienced with COVID-19, one the most devastating pandemic to ever face humanity, rapid identification and typing of novel pathogens are critical. By building our knowledge on this approaches, we can understand how pathogens are transmitted, track outbreaks and able to develop ultimate control strategies. The advantages and limitations of these approaches would be discussed.

**Keywords** genomics, proteomics, pathogens, typing

**Introduction**

With the continuous threat posed by emerging and re-emerging pathogens which is further being compounded by the emergence of resistant pathogens which have impacted the global effort to control and manage infectious diseases, there is the need for the rapid and accurate identification and classification of infectious agents with pandemic and epidemic potential to enable optimised control of infectious diseases and reduce the costs associated with these pathogens. Pathogen typing would therefore provide the tool for epidemiological surveillance and disease-management (1).

Genomic and proteomics are part of the omic technology that is now becoming useful tools in infectious diseases research which is now changing the face of medical diagnosis, treatment and drug development. “OMICS” (Figure 1) is the process of investigating and evaluating large amount of data consisting of the structure and function of a given biological system at multiple levels. These include genomics, proteomics, metabolomics, transcriptomics, lipidomics, etc. Omics technology has therefore being used to evaluate genomic changes, temporal transcriptomics changes and different splicing and spatio-temporal dynamics and post-translational modifications (PTMs) (2). In oncology, utilization of omics technology has being used to types different types of cancers which has resulted in the identification of molecular characteristics that underlines the pathogenesis of different range of cancer types and also provided details of patient subgroups and the molecular characterization of these cancer subtype (3). Therefore in oncology, multi-omics tools have been utilized in several clinical studies for better evaluation of clinical subtypes, prediction of effective therapeutic interventions, profiling cancer drug resistance, and identification of novel biomarkers that are essential in diagnosis, monitoring of drug responses and prognosis. Furthermore, multiomics have important role to play in precision medicine which is also referred to as personalized medicine where an individual’s genome, and other factors such as environment, lifestyle is taking into consideration in order to develop a strategic healthcare plan that can be used for prevention, diagnosis and therapeutic interventions. Therefore by combining genomics, metabolomics, microbiomics, an understanding of individual’s risk at the molecular level can be ascertain for more effective interventions (4). From these, it can be suggested that infectious diseases research would benefit from the utilization of omic technology.

This chapter will review the concept of omic technology with focus on genomic and proteomics in pathogen typing. We will focus on some basic and clinical studies that have gained from the use of genomics and proteomics to type pathogens of human interest. In addition, we will also some of the challenges associated with genomics and proteomics approaches and then suggest how genomics and proteomics can positively impact infectious diseases research.

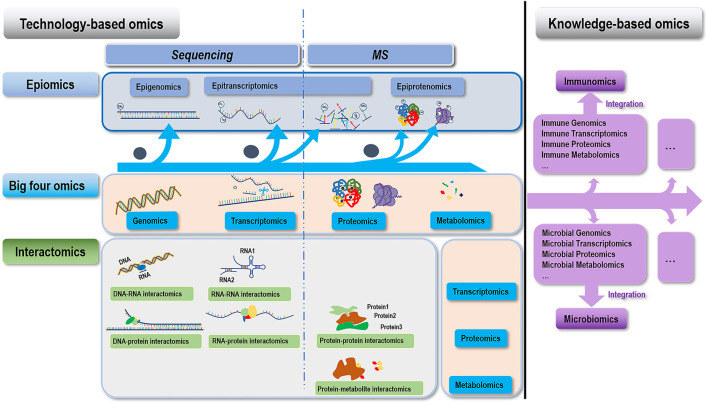


Figure 1: Illustration of omic technology (Source: Dai & Shen, 2022)

**Application of Genomics in Pathogen Typing**

Genomics tool compares the sizes and numbers of different DNA fragments which can be separated by gel electrophoresis such as pulse field gel electrophoresis (PFGE) and nucleic acid amplification-dependent typing techniques like resitriction fragment length polymorphism (RFLP). Genomics have been used in the identification of novel drug targets for infectious agents. The strategies used in this are based on identifying genes that are associated with a pathogen’s virulence or viability. In a study, De Becker and Van Dijick used functional genomics to identify antifungal agents (5). Kiu et al also used whole genomic analysis to profile the virulence factors, strain tracking as well as plasmid evaluation of Clostridium perfringens isolates and the preterm infant disease called necrotizing enterocolitis. They identified gene encoding toxin perfringolysin O, pfoA+ which highlights the importance of C pergfringens as gut pathogen in preterm infants (6). Allard et al utilized genomics tool to describe foodborne pathogens (7). Hammarlof et al used genomics to identify gene function in bacterial pathogens using Salmonella functional genomics (8). By combining two different functional genomics tools, they found that certain genes of unknown function (FUN) during the pathogenic process of mammalian infection and also changes in FUN genes were essential for disease pathogenesis. By adopting this model in their study, similar tool can be utilized in other bacterial pathogens by identifying novel changes that are associated with disease progression and whether mutagenesis have any effect on the infective process. In the field of virology, metagenomics have been used to provide profile viral pathogens (9). Elucidating host-virus interaction is important in understanding the infectious pathogenicity of viruses. Wang et al identified high variability in influenza A virus (IAV) transcription among cells which included increased presence of defective viral genomes that had an impact on IAV infection. Genomic analysis suggested that it is important to understand host-virus interaction at single cell level (10). Although very promising, these methods are less reliable than direct sequence –based tools because these methods lack precision and reproducibility (1).   
  
Over the past years, different genomic approaches have been exploited in pathogen typing. Below are brief descriptions of some of the tools.

***Whole Genome Sequencing (WGS)***

WGS is a tool that is used to sequence the whole (entire) genome of several organisms within a species utilizing a known genomic sequence (11). This mean WGS is used in obtaining genetic information so that genomic variation across species can be studied. In the past, this was performed using Reduced-Representation Genomic Sequencing (RRGS) which is now replaced by WGS because of limitation in genomic coverage. Depending on the depth of sequencing, the size of the sample, and the coverage, WGS is classified into three: Individual sequencing where high depth coverage is dependent on the haplotype resolution, high-depth sequence where population genomics is evaluated by combining equal number of moles of unclassified individual DNA, and low-depth sequencing where several individual in a population are evaluated (11, 12, 13, 14, 15) High-depth sequencing is the “gold standard when high-quality data is required for accurate identification of coding and non-coding DNA (11). The WGS tool start (Figure 2) starts from preparation of genomic sample to the evaluation of the data. It also consist of different phases including DNA extraction, sequencing, referencing, genomic alignment, and the evaluation of variant.

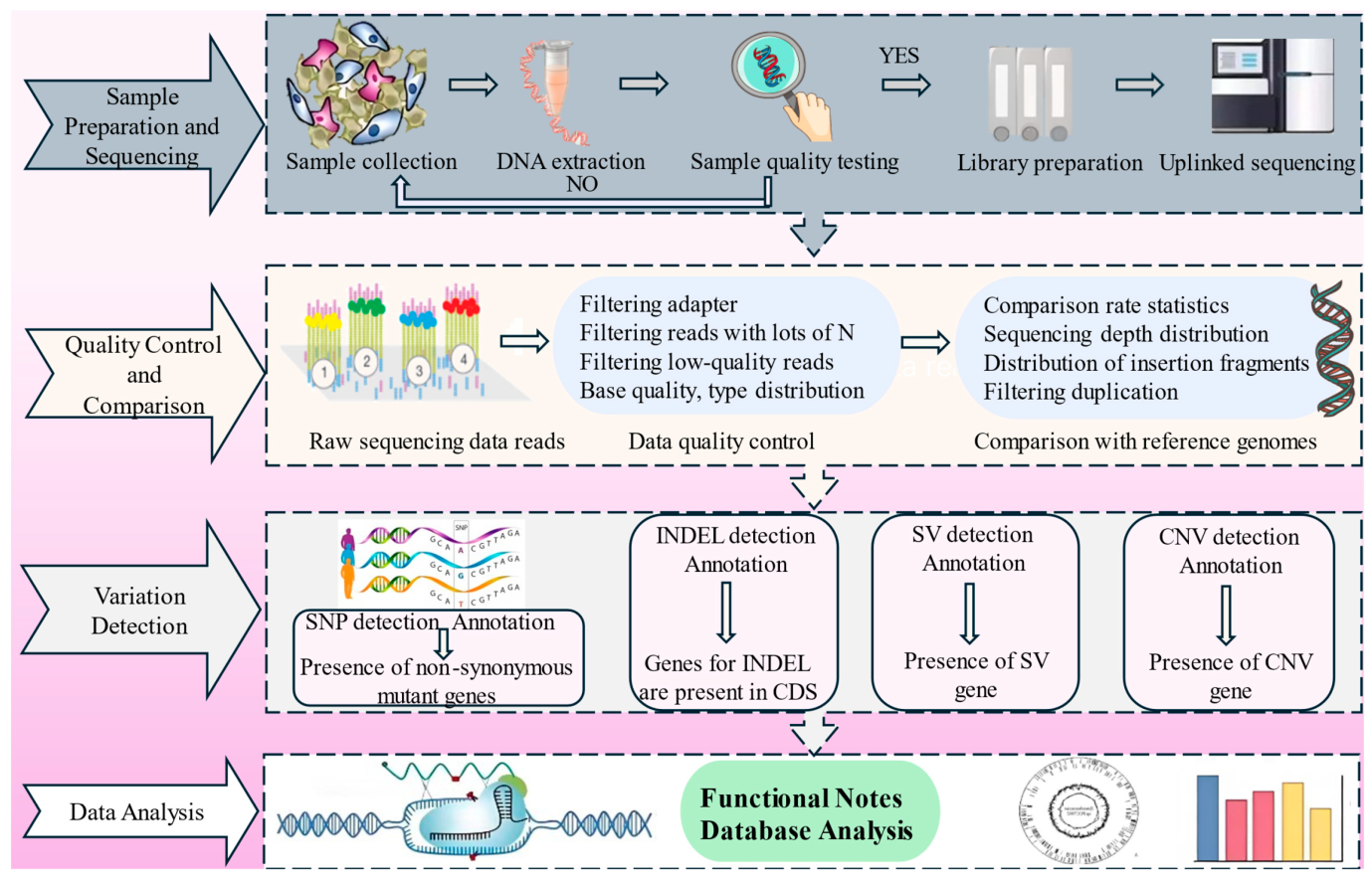


Figure 2: The work flow of WGS (Source: Lu et al, 2025)

Genomic sequencing tools have been optimized over the years after the introduction of the first-generation sequencing technique. This generation of sequencing offered high accuracy and medium lengths reads. However, it has limitations of having low throughput and high cost. The Sanger Sequencing tool is the well known type first-generation sequencing (16). The current improved versions include next-generation sequencing (NGS) and third-generation sequencing (TGS) techniques (11). Figure 3 outlines the evolution of sequencing technology.The advantage of NGS is it is fast and cost-effective which is able to generate large sequencing data (17). NGS tools are now utilized in studying molecular characteristics of unknown pathogens, understanding of pattern of infectious diseases through molecular epidemiology studies, and evolution of quasispecies of pathogens. Furthermore, NGS is used in understanding host-pathogen interaction at the genomics, transcriptomics, and proteomics level.

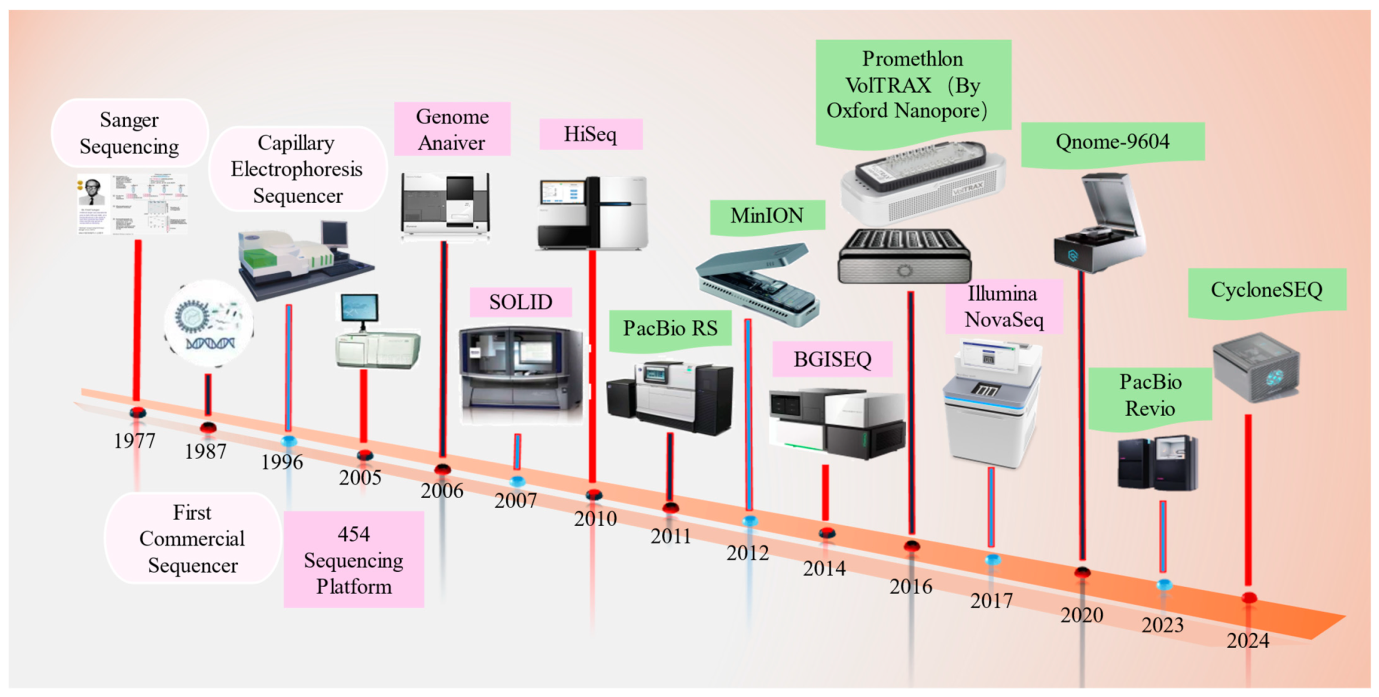


Figure 3: The journey of sequencing technology (Source: Lu et al, 2025)

WGS has been used in comparative genomics which is beneficial in understanding phylogenetic correlation between different strains of pathogens. WGS is also useful tool for investigating outbreaks, characteristic features, genomic evaluation and characterizing genomic information that can be used to predict phenotypes of pathogens such as serotyping, understanding antimicrobial resistance, virulence and pathogenicity and the process of biofilm formation (18). For example when estimating the pathogenicity of a novel strain, this is performed based on identifying the virulence factors (VFs). Adherence and colonization, type of secretion system, how the pathogen evade the immune system, type of toxins, the sidophores used for iron absorption, and the process of genes invasion are some of the factors that should be evaluated. WGS is useful in detection of known VFS and are able to also identify genes or gene variants that are unknown but might be involved in virulence of bacteria. It must be added that the presence or absence of VFs should not be sufficient in drawing conclusion to the virulence of a pathogen because of the prospect of mutation in regulation pathways which can change virulence as reported by Tagini & Greub and Cassat et al for *Streptococcus* species and *Staphylococcus aureus*, respectively (19,20). During the COVID-19 outbreak, Want et al used WGS and analysis of host genetic contribution to severity and susceptibility to infection. In their study consisting of 332 COVID-19 patients who were categorized based on different levels of severity from Shenzhen Third People’s Hospital, Using a total of 22.2 million genetic variants, they performed both single-variant and gene-based correlation studies among 5 severity groups. They reported that the most significant locus that was linked with disease severity was found in TMEM189-UBE2V1 that was associated with the IL-1 signalling pathway. HLA-A\*11:01, B\*51:01, and C\*14:02 alleles were found to be associated with patient’s worst outcome. This study therefore provided a genetic insight into the phenotypic difference when categorizing COVID-19 patients based on disease severity (21). Furthermore, Pathak et al used integrative genome-wide association study to find correlation of COVID-19 hospitalization. They reported that 27 genes were associated with inflammation and coagulation pathways whose expression was genetically associated with COVID-19 hospitalization (22). This type of typing would be of benefit when trying to contain outbreaks. In addition, WGS is used to differentiate single nucleotide polymorphisms (SNP level in real-time or during retrospective analysis of outbreaks as reported by Mellmann et al and Quick et al for *E. coli* and *Salmonella enteric*, respectively (23,24).WGS is useful in studying the virulence of pathogens. Staphylococcus aureus is a known adaptable pathogen with quite a number of its genes associated with virulence. WGS was used to screen the genome to identify specific genes associated with virulence such Pantone-Valentine leucocidin (19).

While WGS is beneficial in clinical microbiology laboratories, certain challenges such as how to standardise and validate the technique are some of the recognized limitations which should be addressed to enable it reach the clinical standards for routine use. The technique also does not permit easy comparison between related studies. In addition, need for reference genome or whole genome alignments are some of the associated limitations (19). Based on these limitations, new sequencing tools have been developed to address some of these limitations.

***Multilocus Sequence Typing (MLST)***

The MLST tool is a typing tool that is dependent on sequencing gene fragments from several housekeeping loci. This is fully portable and can store data in a single increasing central multilocus sequence database which can be cross-examined electronically through the internet to construct powerful resource for typing of pathogens; such as for long-term global epidemiology studies and pathogen characterizations. The concept of MLST was based on multilocus enzyme electrophoresis (MLEE) which was used to identify variations that accumulated rather slowly within a population and was neutrally selective. MLEE can therefore identify small number of alleles which are obtained directly from the nucleotide sequences from internal fragments from housekeeping genes instead of comparing eletrophoretic movement of enzymes they encode. However, high levels of differentiation are obtained using this method. MLST is different from MLEE in that

It can be identify more variations which leads to more alleles per locus been acquired than in MLEE. Secondly, the data can easily be compared with other laboratories (25). Maiden et al used developed and validated MLST which was utilized in the identification of virulent lineages of *Neisseria meningitides*. MLST has advantages over other molecular typing techniques in that the sequence data are transferable between laboratories therefore allowing the expansion of global database per species to be uploaded on the World-Wide Web site. This allows exchange of molecular typing data to be used globally. Different optimized versions of MLST have been introduced including core genome MLST (cgMLST), which is a tool used to compare genomes utilizing large number of gene loci. The central application of cgMLST is based on cgMLST scheme which consist of defined set of loci for which each collection of alleles for each locus are typically numbered. This means a scheme is gathered by combining large number of genomes of species and the set of loci are then identified in most of the genomes. There are several cgMLST schemes (Table 1) for various species.

Table 1

Available cgMSLT schemes

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Website | Publically accessible | Example of species |
| PubMLST | [www.pubmist.org](http://www.pubmist.org) | Yes | Neisseria gonorrheoeae, Campylobacter, Streptococcus agalactiae, Streptococcus pneumonia  Brucella species |
| cgMLST | [www.cgMLST.org](http://www.cgMLST.org) | Yes | Acinetobacter baumannii  Brucella species  *Clostridium perfringens*  *Corynebacterium diphtheriariae*  *Enteroccous faecalis*  *Mycobaterium tuberculosis*  *Staphylococcus aureus*  *Klebsiella pnueomoniae* |
| wgMLST | [www.pubmlst.org](http://www.pubmlst.org) | Yes | *Enterobacter cloacae*  *E coli*  *Citrobacter species*  *E cloacae*  *Klebsiella oxytoca*  *Klebsiella pneumonia* |
| Cg/wgMLST | [www.chewbbaca.online](http://www.chewbbaca.online) | Yes | Acinetobacter calcoaceticus/baumannii complex, Legionella pneumophila, Streptococcus pyogenes, Escherichia coli, Yersinia enterocolitica, Campylobacter jejuni, Salmonella |

MLST is therefore can be useful for recognizing pathogens and elucidate their evolutionary process; which can have impact on both human and animal health as well as in plant research (19, 27). MLST tool have been utilized in typing different pathogens. Wagner et al utilized cgMLST in combination with nanopore Q20 with cgMLST for surveillance of bacterial pathogen. Also Kluytmans-van den Bergh et al used to wgMLST to type extended-spectrum beta-lactamase-producing *Enterobacteriaceae* (ESBL) from isolates that were obtained from Dutch hospitals. They concluded that wgMLST scheme led to high-resolution typing of most common EBSL-producing species and this may result in further evaluation of the complex epidemiological profile of AMR-Enterobacteriaceae (26). This suggests that wgMLST can be used in typing pathogens during outbreaks surveillance. While MLST is useful in pathogen typing especially bacterial species, it has certain limitation which is a drawback. For example housekeeping selection loci need a reference genome while absence of diversity in the entire genome or housekeeping genes in some infectious agents, and presence of a currently emerging and re-emerging pathogen may lead to the MLST scheme having restricted discriminatory ability (27).

***Metagenomic Next-Generation Sequencing (mNGS)***

The word “metagenome” was initially used in 1998 to mean “collection of all microbial genomes found in a sample of soil, including sequences from organisms that could not be culture) (29***).*** Over the decade, metagenomic tools have been used in characterizing DNA or RNA that are found in samples. This enables the evaluation of complete microbiome and the human host genome or transcriptome that are found in patient samples. Metagenomic tool have been used in characterizing different environments including marine environment, human microbiome, and disease vectors. Within the world of infectious agents, it has been used in identification of infections, discovering novel pathogens, and characterizing human viruses in healthy and diseased condition as well as utilized in forensic activities (28, 29, 30, 31). Figure 4 outlines a workflow of mNGS. Rapid approaches are needed for the identification of novel and / or emerging and re-emerging pathogens. This was conspicuous during the COVID-19 pandemic when the rapid identification of the virus was key. Therefore, mNGS has become an attractive tool for the identification of infectious agents especially in cases of pathogens of unknown origins because all potential pathogens can be characterized using single assay without requiring specific primers and probes. Tan et al developed and validated automated mNGS assay that was used for diagnosis of novel respiratory viral pathogens. They found that the diagnostic assay attained a mean detection limit of 543 copies/mL and was 93.6%, 93.8% and 93.7% sensitive, specific and accurate, respectively (32). Furthermore, it was superior to RT-PCR. This means mNGS is an ideal tool to type novel pathogens during outbreaks when rapid diagnosis is needed as reported by Palacios et al who identified new arenavirus that was transmitted during solid-organ transplantation while Towner et al novel species Ebola virus from Uganda (33,34)

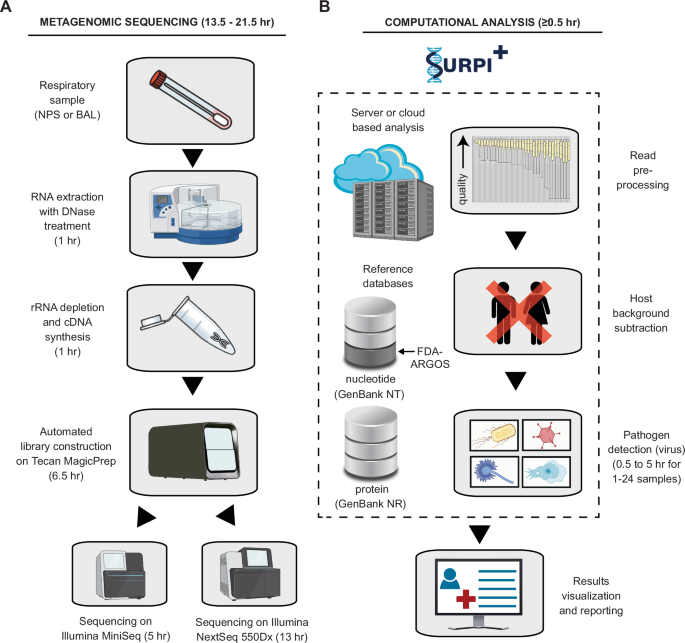


Figure 4: A typical mNGS work flow (Source: Tan et al, 2024)

Furthermore, mNGS is now utilized in pathogens discovery as reported by Smits and Osterhaus who revealed several “orphan viruses” that were not linked with any disease thereby proving the existence of normal human virus microbiome (35). Finally, Chen et al used mNGS to identify pathogens associated with lower respiratory tract infections (LRTI). These pathogens are usually difficult to identify due to the limitations of the traditional microbiological techniques. By enrolling 93 and 69 patients with and without LRTI, respectively. Consistent with the findings of Tan et al, they also reported of 66.7%, 75.4% and 78.5% % sensitiveness, specificity, and positive-predictive values, respectively (36). The authors therefore concluded that metagenomic DNA- and RNA-seq are useful tools for identifying wide range of LRTI pathogens. On this basis, it is recommended that data from host transcriptome can be useful when diagnosing LRTI pathogens. Aside viruses, mNGS can be used in characterizing non-viral pathogens. It can also be used to type pathogenic strains resistant to antimicrobial agents, Table 2 shows example of some utilization of mNGS. Although mNGS is an efficient tool for typing of pathogens, it is limited by high costs, lack of automation, complicated protocols, long turnaround times, absence of regulatory guidelines needed for clinical validation and the overall low sensitivity when compared ro targeted techniques such as in PCR assay (34).

Table 2. Some uses of diagnostic mNGS

|  |  |  |
| --- | --- | --- |
| Pathogen  Lymphocytic choriomeniingitis virus  Ebola virus  Norovirus  Adenovirus, respiratory syncytial virus, rhino virus, and other respiratory tract pathogens  Pandemic influenza  *Chlamydia trachomatis*  *Mycobacterium tuberculosis*  *Plasmodium falciparum* | Circumstance  Fatal infections among transplant recipients  Hospital outbreak of hemorrhagic fever  Fecal samples  Nasopharyngeal aspirates from respiratory tract infections  Nasopharyngeal swabs  Vaginal swab  Historic and prehistoric material  Historic and Prehistoric material | Reference  Palacios et al, 2008 (33)  Towner et al, 2008 (34)  Batty et al, 2013 (37)  Yang et al, 2011 (38)  Greninger et al , 2010 (39)  Andersson et al, 2013 (40)  Chan et al, 2013 (41)  Khairat et al, 2013 (42) |

**Application of Proteomics in Pathogen Typing**

Proteomics is an important tool that can be utilized in typing pathogens for identification and characterization of protein markers that can differentiate species, strains, and even evaluate genetic variation in a single pathogen species. These markers can therefore be used as disease markers or evaluating therapeutic responses. This is because any proteome changes are associated with pathological or biological processes (43). A number of techniques can be used to achieve these objectives. Some of these techniques are reviewed below.

**Bottom-up and top-down Proteomics**

There are two analytical approaches frequently utilized in proteomics: bottom-up and top-down proteomics. In bottom-up proteomics is a process where proteins are initially digested into smaller peptides then these digested peptides are evaluated utilizing mass spectrometry. It is referred to as “bottom-up” because it starts with the evaluation of small peptide fragments which leads to protein identification (44). Top-down proteomic is the different version to bottom-up proteomics in which intact proteins are purified before they are digested into fragments either within a mass spectrometry or 2D electrophoresis and the consequential peptide are evaluated and identified (45). However, it should be noted that the peptide mixture used in bottom-up proteomics consists of thousands of peptides thereby maintaining the multidimensional separation required for in-detailed proteomics evaluation. There are some differences between the two approaches; while top-down proteomics consist of protein extraction from biological samples and sample purification; bottom-up proteomics consist of additional steps including reduction of protein, alkylation, digestion of enzymes, fractionation and desalting (44). In the early phase of proteomics, in-gel sample preparation was used with different versions used including in-solution digestion and filter-aided sample preparation. Single-sell proteomics techniques have been developed to address the issue of sample consisting of low protein levels.

A typical bottom-up proteomic workflow (Figure 5) is made of sample preparation, LC-MS/MS evaluation, and data evaluation. Whereas LS/MS/MS evaluation and data evaluations have been developed comprehensively over the years, sample preparation remains a difficult procedure. This is one of the limitations of this approach in different applications. Sample preparation is the most critical phase of the approach and it can affect the efficiency of any proteomic analysis. However, it is susceptible to errors and has low reliability and throughput. On the other hand, top-down tool has some significant limitations including the solubility of proteins, separation and how to detect large intact proteins. In addition tools for data-analysis have not been developed efficiently (49).

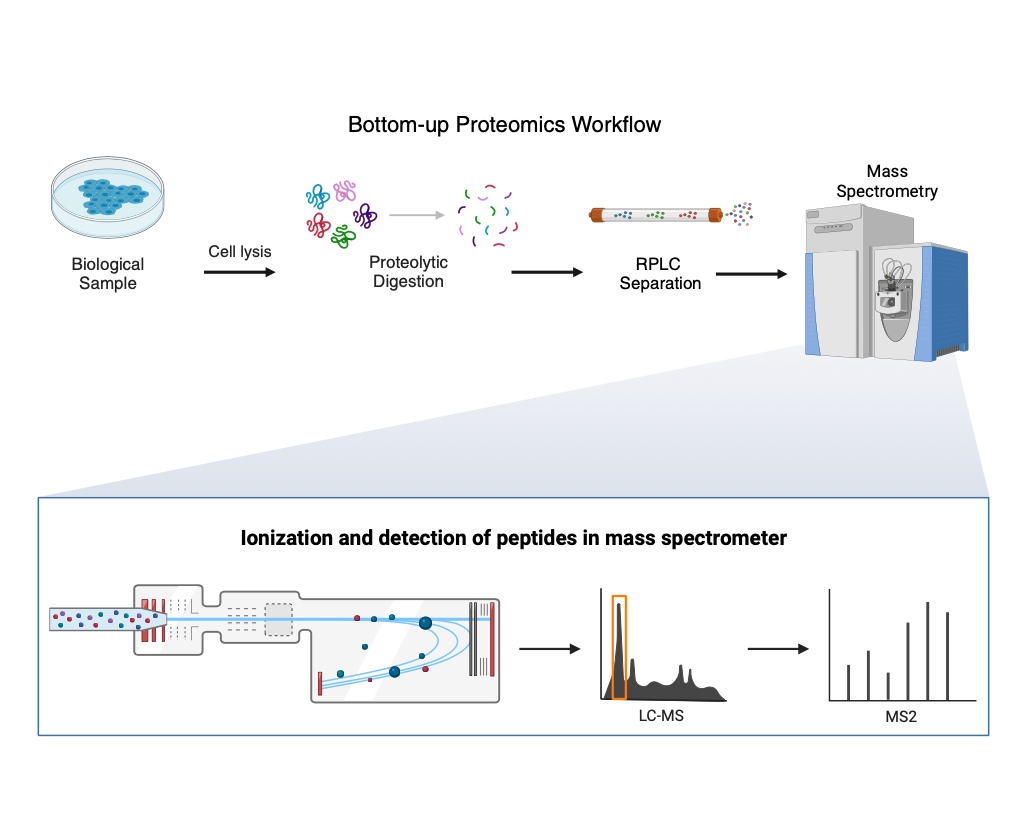


Figure 5: A typical bottom-up proteomic workflow (Source: Biorender)

Several groups have used bottom-up and top-down strategies for identification and typing of pathogens. Boulund et al used bottom-up tandem mass spectrometry proteomics to type and characterize several clinically important bacterial species such as *E coli*, *Pseudomonas aeruginosa* and *Heamophilus influenza* (46). Their result showed that their computational method correctly classified the correct peptides at a rate of between 90.3% and 98.5% when analysed at species level. Furthermore, the technique was utilized to evaluate the relative abundances of individual species found in mixture of cultures. They also reported that β-lacatamse was identified in ESBL *E coli* strain when even antibiotic was not present. Similarly Fulton et al utilized LC-MS for identification characterizing bacterial glycoprotein which is clinically significant due to their special structures and exposure on the cells surface (47). Mass-spectrometry-based proteomics should be used in combination wih bioinformatics approaches as attempted by Dworzanski et al where they utilized this combination for identification and typing of bacteria proteome (48). With regards to top-down proteomics, others have reported of its importance in health research with number of studies associating proteoform changes to phenotype of diseases which shows that top-town proteomics can be useful in the evaluation of proteoform-associated disease (49). Fagerquist and Dodd used top-down proteomic for the identification of plasmid and host proteins that were produced by pathogenic E coli utilizing MALDI-TOF-TOF tandem mass spectrometry (50). They identified bacterial proteins possessing genes that were induced by exposure to antibiotic. It shows that this technique has the potential of identifying biomarkers for antimicrobial exposure. In viral research, the viral capsid is made up to various proteins which may be identified and typed by bottom-up or top-down proteomic techniques. This can be achieved by using MALDI-TOF or ESI for the identification of intact viral proteins. The bottom-up process can be used to proteolytically cleave proteins into peptides which are then grouped base of their sizes. During the COVID-19 pandemic, tandem MS/MS is then used to generate mass spectra. LC-MS/MS or gel electrophoresis is then performed followed by electrospray ionization (ESI)-MS/MS to identify the different type of proteins including spike, membrane, nucleocapsid, and envelope proteins of COVID-19 virus (51)

**Strain-resolved community Proteomics**

Strain-resolved community proteomics (Figure 6) involves evaluation of protein expression within specific strains of pathogens found in complex community of microorganisms (e.g. gut microbiome). With the threat of emerging and re-emerging pathogens on the increase, there is the need to understand the environmental factors that drives such emergence. Our knowledge of how these pathogens can be built through utilization of strain-resolved community proteomic. The aim of such approach is to use genomic data for the identification of strain based on the genome influenced by recombination that involves chromosomal regions that are acquired from two immediate related pathogenic population e.g. inter-related strains. A study found that by utilizing strain-resolved community proteomics, the authors suggested that due to exchange of large block of gene variants, acidophilic bacteria were able to adapt to specific ecological environment within acidic, metal-rich environment (52). This is important because it facilitates our understanding on how pathogen survives in the ecological niche and this can be useful in the identification of virulence factors associated with specific pathogens. Zlitni et al used same approach to describe mobile element that drives competition in bacteria on a clinical timescale (54). Their study also found specific structural genomic changes that are linked with increased antibiotic resistance over the period of therapeutic intervention. Finally, Jochheim et al by using strain-resolved de-novo metagenomic software, PenguiN to describe how viral genome and 16S RNA assembly overlap which can be used as template to understand the key role of viruses within microbial communities (55).

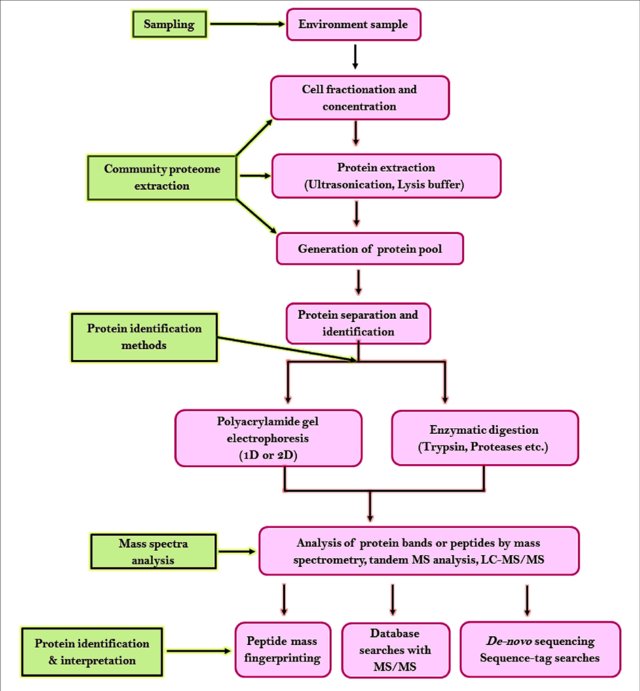


Figure 6: A typical Strain-resolved community Proteomics (Source: Chandran et al, 2020 [53])

**Conclusion and Recommendation**

The rapid identification and typing of pathogens is essential especially during outbreaks when it is essential to not only identify and characterize the novel pathogens but also the clinical source of the infection, its virulence, pathogenic process, drug responses, and transmission dynamic are critical. Also in this era of global burden of antimicrobial resistance, typing of pathogens to understand the mutational mechanism are also essential. The introduction of omic technology has improved the clinical diagnostic of pathogens from clinical samples. This had led to better understanding of disease. Because data from multiple molecular levels are combined to give holistic view which can lead to a more accurate typing of pathogens especially emerging and emerging of infectious agents. Genomics and proteomics which are part of the omic technology are very essential tools for identification and typing of pathogens as they can be utilized in identification of new therapeutic targets, biomarker discovery and predicting the prognosis of an infection process. However, genomics and proteomics should be used with combination with transcriptomics. Furthermore, omics technology should not be soley used for typing pathogens. Although the traditional methods of identification and typing of pathogens can be cumbersome, it should still be used in combination with omic technology for optimal identification and typing of pathogens. This has been effective in cancer and cardio-metabolic disease research. It therefore offer better option for us to deal with emerging anf re-emerging pathogens.

**References**

1. Sintchenko V, Iredell JR, Gilbert GL (2007): Pathogen profiling for disease management and surveillance, Nat Rev Microbiol; 5: 464-470. DOI: [10.1038/nrmicro1656](https://doi.org/10.1038/nrmicro1656)
2. Dai X, Shen L (2022): Advances and trends in Omics technology Development, Front med (Lausanne); 9: 911861. DOI: [10.3389/fmed.2022.911861](https://doi.org/10.3389/fmed.2022.911861)
3. Heo YJ, Hwa C, et al (2021): Integrative multi-omics approaches in cancer research: from biological networks to clinical subtypes, Mol Cells; 44: 433-443. DOI: [10.14348/molcells.2021.0042](https://doi.org/10.14348/molcells.2021.0042)
4. Mani S, Lalani SR, Pammi M (2025): Genomics and multiomics in the age of precision medicine, Pediatric Research. DOI: 10.1038/s41390-025-04021-0
5. De Becker MD, Van Dijick P (2003): Progress in functional genomics approaches to antifungal drug target discovery, Trends Microbiol; 11: 470-478. DOI: 10.1016/j.tim.2003.08.008.
6. Kiu R, Shaw AG, et al (2023): Particular genomic and virulence traits associated with preterm infant-derived toxigenic Clostridium perfringens strains, Nat Microbiol; 8: 1160-1175. DOI: [10.1038/s41564-023-01385-z](https://doi.org/10.1038/s41564-023-01385-z)
7. Allard MW, Bell R, et al (2018): Genomics of foodborne pathogens for microbial safety, Curr Opin Biotechnol; 49: 224-229. DOI: 10.1016/j.copbio.2017.11.002
8. Hammarlof DL, Canals R, Hinton JCD (2013): The FUN of identifyinh gene function in bacterial pathogens; insights from Salmonella functional genomics, Curr Opin in Microbiology; 16: 643-652. DOI: [1016/j.mib.2013.07.009](https://doi.org/10.1016/j.mib.2013.07.009)
9. Edwards AR, Rohwer F (2005): Viral metagenomics, Nat Rev Microbiol; 3: 504-510. DOI: [10.1038/nrmicro1163](https://doi.org/10.1038/nrmicro1163)
10. Wang C, Forst CV, et al (2020): Cell-to-cell variation in defective virus expression and effects on host responses during influenza virus infection, mBio; 11: e02880-19. DOI: [10.1128/mBio.02880-19](https://doi.org/10.1128/mbio.02880-19)
11. Lu Y, Li M, et al (2025): Advances in whole genome sequencing: Methods, tools, and application in population genomics, Int J Mol Sci; 26: 372. DOI: [10.3390/ijms26010372](https://doi.org/10.3390/ijms26010372)
12. Fuentes-Pardo AP, Ruzzante DE (2017): Whole-Genome Sequencing Approaches for conservation biology: Advantages, Limitation and Practical Recommendations, Mol Ecol; 26: 536905406. DOI: [10.1111/mec.14264](https://doi.org/10.1111/mec.14264)
13. Schlotterer C, Tobler R, et al (2014): Sequencing pools of individuals- mining Genome-Wide Polymorphism Data with Big Funding, Nat Rev Genet; 15: 749-763. DOI: 10.1038/nrg3803
14. Nielsen R, Paul JS, et al (2011): Genotype and SNP calling from Next-Generation Sequencing Data, Nat Rev Genet; 12: 443-451. DOI: 10.1038/nrg2986
15. Lou RN, Jacobs A, et al (2021): A beginner’s guide to low coverage Whole Genome Sequencing for population genomics. Mol Ecol; 30: 5966-5993. DOI: [10.1111/mec.16077](https://doi.org/10.1111/mec.16077)
16. Shendure J, Ji H (2008): Next-generation DNA sequencing, Nat Biotechnol; 26: 1135-1145. DOI: 10.1038/nbt1486
17. Van Borm S, Belak S, et al (2015): Next-Generation Sequencing in Veterinary Medicine: How can the massive amount of information arising from high-throughput technologies improve diagnosis, control and management of infectious diseases, Methods Mol Biol; 1247: 415-436. DOI: [10.1007/978-1-4939-2004-4\_30](https://doi.org/10.1007/978-1-4939-2004-4_30)
18. Uelze L, Grutzke J, et al (2020): Typing methods based on whole genome sequencing data, One Health Outlook; 2:3. DOI: [10.1186/s42522-020-0010-1](https://doi.org/10.1186/s42522-020-0010-1)
19. Tagini F, Greub G (2017): Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review, Eur J Clin Microbiol Infect Dis; 36: 2007-2020. DOI: [10.1007/s10096-017-3024-6](https://doi.org/10.1007/s10096-017-3024-6)
20. Cassatt J, Dunman PM, et al (2006): Transcriptional profiling of a Staphylococcus aureus clinical isolate and its isogenic agr and sarA mutant reveals global differences in comparison to the laboratory strains RN6390, Mircrobiology (Reading); 152 (pt 10): 3075-3090. DOI: [10.1099/mic.0.29033-0](https://doi.org/10.1099/mic.0.29033-0)
21. Wang F, Huang S, et al (2020): Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility, Cell Discovery 6, 83. DOI: 10.1038/s41421-020-00231-4
22. Pathak GA, Singh K, et al (2021): Integrative genomics analyses identify susceptibility genes underlying COVID-19 hospitalization, Nat Communications, 12: Article number 4569. DOI: 10.1038/s41467-021-24824-z
23. Mellmann A, Harmsen D, et al (2011): Prospective genomic characterization of the German enterohemorrhagic Escherichia coli 0104:H4 outbreak by rapid next generation sequencing technology, PLoS Pne; 6: e22751. DOI: [10.1371/journal.pone.0022751](https://doi.org/10.1371/journal.pone.0022751)
24. Quick J, Ashton P, et al (2015): Rapid drafts sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella, Genome Biol; 16: 114. DOI: 10.1186/s13059-015-0677-2
25. Maiden MCJ, Bygraves JA, et al (1998): Multilocus sequencing typing: A portable approach to the identification of clones within population of pathogenic microorganisms, PNAS; 95: 3140-3145. DOI: [10.1073/pnas.95.6.3140](https://doi.org/10.1073/pnas.95.6.3140)
26. Kluytmans-van den Bergh MFQ, ROssen JWA, et al (2016): Whole-Genome Multilocus sequence typing of extended-spectrum-beta-lactamase-producing Enterobacteriaceae, J Clin Microbiol; 54: 2919-2927. DOI: [10.1128/JCM.01648-16](https://doi.org/10.1128/JCM.01648-16)
27. Losada-Perez M, Cabezas P, et al (2013): Pathogen typing in the genomics era: MLST and the future of molecular epidemiology, Infect Genet Evol; 16: 38-53. DOI: [10.1016/j.meegid.2013.01.009](https://doi.org/10.1016/j.meegid.2013.01.009)
28. Chiu CY, Miller SA (2019): Clinical metagenomics, Nat Rev Genet; 20: 341-355. DOI: [10.1038/s41576-019-0113-7](https://doi.org/10.1038/s41576-019-0113-7)
29. Pallen MJ (2014): Diagnostic metagenomics: potential application to bacterial, viral and parasitic infections, Parasitology; 141: 1856-1862. DOI: [10.1017/S0031182014000134](https://doi.org/10.1017/S0031182014000134)
30. Moustafa A, Xie C, et al (2017): The blood DNA virome in 8,000 humans, PLoS Pathog; 13: e1006292. DOI: 10.1371/journal.ppat.1006292
31. Hamptom-Marcell JT, Lopez JV, Gilbert JA (2017): The human microbiome: an emerging tool in forensics, Microbial Biotechnology; 10: 228-230. DOI: [10.1111/1751-7915.12699](https://doi.org/10.1111/1751-7915.12699)
32. Tan JK, Servellita V, et al (2024): Laboratory validation of a clinical metagenomic next-generation sequencing assay for respiratory virus detection and discovery, Nat Communication; 15: 9016. DOI: 10.1038/s41467-024-51470-y
33. Palacios G, Druce J, et al (2008): A new Arenavirus in a cluster of fatal transplant-associated diseases, NEJM; 358: 991-998. DOI: 10.1056/NEJMoa073785
34. Towner JS, Sealy TK, et al (2008): Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda, PLoS Pathogens; 4: e1000212. DOI: 10.1056/NEJMoa073785
35. Smits SL, Osterhuas ADME (2013): Virus discovery: one step beyond, Current Opinion in Virology; 3: e1-e6. DOI: [10.1016/j.coviro.2013.03.007](https://doi.org/10.1016/j.coviro.2013.03.007)
36. Chen H, Yin Y, et al (2020): Clinical utility of in-house metagenomic next-generation sequencing for the diagnosis of lower respiratory tract infections and analysis of the host immune response, Clin Infect Diseases; 71, Issue Supplement\_4, :S416-S426. DOI: [10.1093/cid/ciaa1516](https://doi.org/10.1093/cid/ciaa1516)
37. Batty EM, Wong THN, et al (2013): A modified RNA-Seq approach for whole genome sequencing of RNA viruses from faecal and blood sample, PLoS One; 8: e66129. DOI: 10.1371/journal.pone.0066129
38. Yang J, Yang F, et al (0211): Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach, J Clinical Microbiology; 49. DOI: [10.1128/jcm.00273-11](https://doi.org/10.1128/jcm.00273-11)
39. Greninger AL, Chen EC, et al (2010): A metagenomic analysis of pandemic influenza A (2009 H1N!) infection in patients from North America, PLoS One; 5: e13381. DOI: [10.1371/journal.pone.0013381](https://doi.org/10.1371/journal.pone.0013381)
40. Andersson P, Klein M, et al (2013): Sequences of multiple bacterial genomes and a Chlamydia trachomatis genotype from direct sequencing of DNA derived from a vaginal swab diagnostic specimen, Clin Microbiol Infection; 19: E405-E408. DOI: [10.1111/1469-0691.12237](https://doi.org/10.1111/1469-0691.12237)
41. Chan JZM, Sergeant MJ, et al (2013): Metagenomic analysis of tuberculosis in a mummy, NEJM; 369: 289-290. DOI: 10.1056/NEJMc1302295
42. Khairat R, Ball M, e al (2013): First insights into the metagenome of Egyptian mummies using next-generation sequencing, J Appl Genet; 54: 309-325. DOI: 10.1007/s13353-013-0145-1
43. Doung V-A, Lee H (2023): Bottom-up proteomics: Advancements in sample preparation, Int J Mol Sci; 24: 4350. DOI: [10.3390/ijms24065350](https://doi.org/10.3390/ijms24065350)
44. Chait BT (2006): Mass spectrometry: bottom-up or top-down? Science; 314: 65-66. DOI:: [10.1126/science.1133987](https://doi.org/10.1126/science.1133987)
45. Wright EP, Partrigde MA, et al (2014): Top-down proteomics: Enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection, Proteomics; 14: 872-889. DOI: [10.1002/pmic.201300424](https://doi.org/10.1002%2Fpmic.201300424)
46. Boulund F, Karlsson R, et al(2017): Typing and characterization of bacteria using bottom-up tandem mass spectrometry proteomics, Mol Cell Proteomics, 16: 1052-1063.DOI: [10.1074/mcp.M116.061721](https://doi.org/10.1074/mcp.m116.061721)
47. Fulton KM. Li J, et al (2017): Characterizing bacterial glycoprotein with LC-MS, Expert Review of Proteomic, 15: 203-216. DOI: [10.1074/mcp.M116.061721](https://doi.org/10.1074/mcp.m116.061721)
48. Dworzanski J, Deshpanda SV, et al (2005): Mass spectrmmetry-based proteomic combined with bioinformatic tools for bacterial classification, J of Proteome Research; 5: 76-87. DOI: [10.1021/pr050294t](https://doi.org/10.1021/pr050294t)
49. Gregorich ZR, Ge Y (2015): Top-down proteomics in health and disease: challenges and opportunitied, Proteomics; 14: 1195-1210. DOI: [10.1002/pmic.201300432](https://doi.org/10.1002/pmic.201300432)
50. Fagerquist CK, Dodd CE (2021): Top-down proteomic identification of plasmis and host proteins produced by pathogenic Escherichia coli using MALDI-TOF-TOF tandem mass spectrometry, PLoS One; 16: e0260650. DOI: 10.1371/journal.pone.0260650
51. Gupte R: Pathogen proteotyping: A game changer for clinical microbiology on <https://www.clinicallab.com/pathogen-proteotyping-a-game-changer-for-clinical-microbiology-23191> (Accessed on 10/05/2025)
52. VLo I, Denef VJ, et al (2007): Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria, Nature; 446: 537-541. DOI: 10.1038/nature05624
53. Chandran H, Meena M, Sharma K (2020): Microbial biodiversity and bioremediation assessment through omics approaches, Front Environ Chem; 1: 2020. DOI: [10.3389/fenvc.2020.570326](https://doi.org/10.3389/fenvc.2020.570326)
54. Zlitni S, Bishara A, et al (2020): Strain-resolved microbiome sequencing reveals mobile elements that drive bacterial competition on a clinical timescale, Genome Medicine, 50: 2020. DOI: 10.1186/s13073-020-00747-0
55. Jochheim A, Jochheim FA, et al (2024): Strain-resolved de-novo metagenomic assembly of viral genomes and microbial 16 rRNAs, Microbiome; 12: 187. DOI: 10.1186/s40168-024-01904-y