GENETIC INSIGHTS BREAK INFECTIOUS PATHOGEN BARRIERS

- **Page 2**
  A Strong Research Foundation Supports Infectious Pathogen Discoveries
- **Page 3**
  Addressing Infectious Pathogens with Genetic Technologies
- **Page 4**
  Hijacking Host Genetics to Improve Tuberculosis Outcomes
- **Page 5**
  Uncovering HIV in Hiding: Techniques for Finding Pathogen Reservoirs
- **Page 6**
  Technology Overview

Sponsored by

ThermoFisher SCIENTIFIC





A STRONG RESEARCH FOUNDATION SUPPORTS INFECTIOUS PATHOGEN DISCOVERIES

Medical breakthroughs, accessible healthcare, and improved sanitation have reduced the overall mortality and morbidity of certain infectious pathogens over the past two decades, but the global infectious disease burden remains significant. Basic research on prevalent and neglected infectious diseases and pathogens, such as human immunodeficiency virus (HIV), tuberculosis, malaria, and influenza, is crucial for new treatment, diagnostic, and prevention strategies, translating advancements from the laboratory to the clinic.[1]

Infectious Disease Burden and Research Priorities

The economic and health costs of infectious diseases vary based on pathogen and affected demographics. For instance, in low-income countries, eight out of the 10 leading causes of death in 2021 were due to communicable pathogens, including malaria and HIV/acquired immunodeficiency syndrome (AIDS), as well as lower respiratory infections and tuberculosis, which also remain in the global top 10 leading causes of death.[2] Beyond mortality, infectious diseases are a leading cause of disability, with the three leading pathogens in 2019–(tuberculosis, malaria, and HIV)–each being responsible for more than 50 million disability-adjusted life-years (DALY).[3]

The past twenty years demonstrated increased academic interest in infectious pathogen research, spurred by changes in pathogen emergence, dynamics, and spread associated with shifting climates, technologies, and demographics.[1] Global spending on infectious disease research between 2000 and 2017 was roughly 105 billion dollars US, with nearly 75% of funding allotted to preclinical science.[3] Infectious diseases are among the top 10 most funded disease research areas by the National Institutes of Health (NIH), reflecting the importance of investigating the numerous conditions encompassed within the field.[4] The NIH National Institute of Allergy and Infectious Diseases' featured diseases and conditions include coronaviruses, HIV/AIDS, influenza, malaria, mpox, and tuberculosis.[5]

Addressing Research Challenges

Basic and translational researchers alike face similar methodological challenges identifying infectious microorganisms quickly, accurately, and with adequate sensitivity. Traditional approaches that rely on morphological, physiological, and biochemical characterization are time consuming, labor intensive, and materially costly. In addition, phenotypic detection assays cannot always unambiguously identify microorganism at the species or strain level.[6]

For instance, conventional bacterial and viral identification methods often require lengthy manual culture processes. *Mycobacterium tuberculosis* culture is considered the gold standard method for tuberculosis diagnosis, but the time required to culture certain strains can exceed months, hindering rapid intervention and delaying appropriate treatment windows for drug-resistant bacteria.[7]

Additionally, newly emergent infectious agents require continuous research and adaptation of existing detection methods to improve accuracy and sensitivity when searching for ever-evolving pathogens. For example, foundational studies on HIV revealed how the virus evades the immune system, leading to the development of antiretroviral therapies that have transformed the prognosis for those living with the virus. Still, HIV infection requires lifelong management, and acquired treatment resistance fuels ongoing research efforts to better detect viral variants and elucidate host genetic susceptibilities.[8,9]

Genetic Tools for Studying Infectious Pathogens

To obtain the most efficient and precise pathogen identification, researchers must choose appropriate tools.[6] Genetic workflows are solutions to different challenges, including time- and labor-intensive identification methods, and test accuracy and sensitivity limitations. Molecular detection technologies such as polymerase chain reaction (PCR), Sanger sequencing, microarrays, and next-generation sequencing (NGS) help scientists find and address infectious agents quicker and more reliably than culture-based approaches.[10]

Viruses can be identified by RNA or DNA sequencing, using next-generation or Sanger sequencing methods. Scientists can also determine the presence of specific viral genotypes in samples with targeted PCR-based techniques, which provide the advantage of less human DNA interference and higher sensitivity.[10] These approaches are effective for investigating microorganisms related to common infections affecting the respiratory, gastrointestinal, and urogenital systems, and help researchers study viral reservoirs than enable persistent host infection.[8]

Genomic methods also offer a precise way to identify bacteria, ranging from extensive whole-genome sequencing to straightforward quantitative PCR assays. Both techniques are useful for detecting potential pathogens in food, water, and environmental samples. In contrast to traditional microbiological methods, which require culturing individual organisms, genomic approaches can identify hard-to-culture or unculturable organisms, broadening the detection range of known pathogens and treatment-resistant strains.[6]

See references on page 12

ADDRESSING INFECTIOUS PATHOGENS WITH GENETIC TECHNOLOGIES

Infectious Pathogens Collectively

Global Burden:
In 2019, roughly 704 million DALYs* were associated with 85 different infectious pathogens.[1]

US Prevalence:
2,868,805† new cases in 2019[2]

Publication Trends (2009 to 2023):
138% increase

Estimated NIH 2025 Budget:
8.331 billion USD[3]

HIV/AIDS

Global Burden:
In 2019, 52.1 million DALYs* were associated with HIV.[1] In 2023, roughly 39.9 million people were living with HIV and 1.3 million new infections were acquired.[4]

US Prevalence:
36,398 new cases in 2019[5]

Publication Trends (2009 to 2023):
21% increase

Estimated NIH 2025 Budget:
3.294 billion USD[3]

Influenza

Global Burden:
In 2019, 16.7 million DALYs* were associated with influenza.[1] Around one billion cases of seasonal influenza are diagnosed annually, with 3-5 million severe cases.[6]

US Prevalence:
35 million new cases in 2019-2020[7]

Publication Trends (2009 to 2023):
19% increase

Estimated NIH 2025 Budget:
1.152 billion USD[3]

Tuberculosis

Global Burden:
In 2019, roughly 65.1 million DALYs* were associated with tuberculosis.[1] In 2023, roughly 7.5 million new infections were diagnosed.[8]

US Prevalence:
8,916 new cases in 2019[2]

Publication Trends (2009 to 2023):
70% increase

Estimated NIH 2025 Budget:
734 million USD[3]

Comparison of Genetic Technologies Used for Infectious Disease Research[9,10]

Method	Throughput	Identifies new pathogens/strains
PCR	Low (1 to 100s)	No
Sanger sequencing	Low (1 to 100s)	Yes
Microarray	High (1 to 1000s)	No
NGS	High (1 to 1000s)	Yes

See references on page 7

*Overall burden of disease can be represented by disability-adjusted life-years (DALYs), which combines years of life lost due to premature mortality and time lived in states of less than full health or disability.

†This is the summation of selected nationally notifiable new disease cases and does not represent all infectious diseases.

HIJACKING HOST GENETICS TO IMPROVE TUBERCULOSIS OUTCOMES

An interview with Larry S. Schlesinger, MD
Professor, Host Pathogen Interactions, President and CEO
Texas Biomedical Research Institute



Image taken by Josh Huskin, Courtesy of Texas Biomedical Research Institute

Physician scientist Larry Schlesinger from Texas Biomedical Research Institute has dedicated his career to infectious disease research, with a specific interest in tuberculosis (TB) and lung biology. He and his research team use genetic techniques to investigate how TB's fascinating but devastating pathogen, *Mycobacterium tuberculosis* (Mtb), interacts with human immune cells, translating fundamental host cell processes into host-directed therapeutic targets.

Q: What is unique about TB?

Mtb is human adapted. It is the closest thing to a virus, in my opinion, in its ability to exploit the immune system to its advantage. It is a very successful pathogen because it can get into the body and grow silently. Infection occurs in stages. Stage one is actually inhaling the bacteria into your lungs, followed by the initial battlefield between the bacterium and the immune system. But the major outcome from airborne infection is often a healthy human, this period in which you have no clinical symptoms or signs of infection but the bacterium is hanging out in your body, which is called latency. Active TB can come after latency or right after initial infection in babies, very old individuals, or immunocompromised patients. Mtb is a human-to-human transmissible bacterium, and because people might not know they have TB when it becomes active, it spreads.

TB presents symptoms primarily in the lung over a two-to-three-month period, and it is extremely difficult to treat. It may require multiple drugs, which can cause toxicity, and there is a major problem with drug resistance. It is critical that we develop therapy to reduce TB incidence, reduce treatment duration, and handle drug resistance. My lab focuses on novel therapeutic interventions called host-directed therapies as adjuncts to antibiotics.

Q: How does Mtb adapt to the human immune system?

Mtb lives inside resident macrophages in the lungs called alveolar macrophages, which do not look like macrophages in any other tissue compartment. I like to say that these macrophages are like couch potatoes. They eat beautifully to get rid of inhalants that happen to get deep into the lung, but they dampen signaling pathways that would normally be highly inflammatory. We learned early on that Mtb seems to exploit that process, further enhancing signaling pathways that dampen the inflammatory response. That is part of the magic of stage one, where the bacterium hijacks the macrophage for its intracellular survival, allowing it to grow and then ultimately break out and disseminate.

Q: Which genetic technologies do you use to study TB?

We use sequencing to look at host cell epigenetic and genetic changes occurring in response to TB. This allows us to broadly screen and identify bona fide targets. The sequencing approach allows for straightforward analysis of about 800 immune proteins. We can do it in the laboratory and we do not need bioinformatics. Then we specifically enhance or inhibit a target and study its effects. We use silencing RNA to reduce expression of certain proteins, manipulating macrophages to study the pathways involved in Mtb intracellular survival.

We identified a host effector protein called induced myeloid leukemia cell differentiation protein (MCL-1), which is anti-apoptotic. Apoptosis is important because it is a type of cell death that can enhance antigen presentation and the immune response without causing a lot of inflammation. It turns out that Mtb can drive MCL-1 expression in host cells, downregulating macrophage apoptosis for survival in the latent phase. We found that we could control TB by stimulating infected macrophage apoptosis using anticancer drugs that inhibit MCL-1 and another anti-apoptotic protein called B cell leukemia/lymphoma 2 protein (BCL-2).

Host-directed therapy has the potential advantage of potently inhibiting Mtb and being used in combination with antibiotics. We found that targeting MCL-1 and BCL-2 together can inhibit both drug-susceptible and resistant Mtb by 90%, and it is synergistic with antibiotic treatment.

Q: Which techniques do you use to validate effectors such as MCL-1?

When we identify a specific target based on transcriptomic data, we must validate the results using qPCR and Western blot. To better understand the importance of that target in terms of biological pathways, we take several approaches such as use of specific agonists and antagonists, siRNA to knockdown the gene of interest and, most recently, development of CRISPR in primary human immune cells to knockout the gene of interest.

Q: What is the greatest knowledge gap in the TB research field?

The challenge in understanding human global infectious disease is that despite humans being 99.9% identical in our DNA, that 0.1% makes us completely different. Genetic diversity is a bottleneck. I think the gap in the field is trying to find a needle in the haystack. Because Mtb is so adapted to humans and exploits so many host pathways, it is really difficult to find the critical pathways that will be the targets for therapies and vaccines. Understanding the complexity of the immune response is our challenge.

This interview has been condensed and edited for clarity.

UNCOVERING HIV IN HIDING: TECHNIQUES FOR FINDING PATHOGEN RESERVOIRS

An interview with Charles Wood, PhD
Cancer Crusader Professor, Department of Interdisciplinary Oncology
Associate Director of Basic Science, Stanley S. Scott Cancer Center
Louisiana State University Health Science Center



Image courtesy of LSU LCMC Cancer

As an interdisciplinary virologist, Charles Wood's academic interest in human immunodeficiency virus (HIV) began at the outset of the acquired immunodeficiency syndrome (AIDS) epidemic in 1981. Contributing to some of the first available diagnostic tests early in his career, he has since devoted his work to elucidating how HIV affects the immune system and enables cancer-causing co-infections. His current research at Louisiana State University employs a range of cellular and molecular technologies, collecting immunological and genomic insights into HIV subtypes and host reservoirs such as the brain to improve prevention strategies and curative approaches.

Q: Why is it important to study HIV subtypes?

There are a lot of different subtypes in the world. In Europe and in the US, subtype B is the most common. But in Africa and India, it is predominantly subtype C, which now makes up about half of the infected population. In the early days of HIV vaccine development, researchers often focused on the subtypes in the US because that was where the research was being conducted, but now people are looking into vaccine development and prevention globally, especially for subtype C. It is really important to focus on a vaccine that will apply globally.

The differences between the subtypes are subtle but biologically important. For example, how rapid the disease course is can differ between the different subtypes. Also, we believe that there are some neurotropic differences between strains, which is the ability to infect the brain. Subtype C's neuropathogenesis is much milder compared to subtype B, even before antiretroviral treatment (ART). I think the biggest question we are facing now is, how do we get rid of the virus to cure the disease? Because even for people undergoing treatment, the virus is still there in some of the reservoirs.

Q: How do you identify subtypes and viral reservoirs such as the brain?

The subtype is determined by people taking the virus and sequencing it. Each subtype has signature sequences, so we can line them up based on the RNA sequence and tell what subtype it belongs to.

It is a little tricky to look at reservoirs because, especially for individuals undergoing treatment, the virus is suppressed. Only one in a million white blood cells are infected by HIV. There are some approaches, but none of them are ideal. Sequencing is one way to detect viruses in the blood. It requires a lot of blood to find out if HIV is there, and even if the virus is present, you do not know whether it is defective or not. Because the virus mutates and a single mutation in a whole genome could inactivate the virus, you have to make sure that the whole viral genome is there and there are no mutations or deletions. Another way to detect suppressed active viruses is an outgrowth assay, which is very tricky to do. For example, people can perform outgrowth assays with blood or tissues such as lymph nodes. If you biopsy or collect the right reservoir and remove the ART drug in culture, the active virus will replicate and become detectable.

You cannot perform biopsies on the brain as you would other tissues for outgrowth assays. Our approach to finding reservoirs is conducting autopsies of HIV-positive individuals, taking the bodily tissues from the brain all the way down to lymph nodes. However, it is almost impossible to do outgrowth assays from autopsy tissue because the virus is not stable. We look for viral sequences instead. So, the different techniques have limitations, but I think we have to use all approaches possible to identify potential reservoirs.

Each subtype has signature sequences, so we can line them up based on the RNA sequence and tell what subtype it belongs to.

Q: How do you amplify viral sequences for detection, and why is amplification important?

Viral sequence amplification was carried out by a technique called polymerase chain reaction or better known as PCR. It uses primers, which are short single strand nucleotide sequences of 15 to 30 bases that are complimentary to each end of the gene of interest. It can amplify one copy of a single gene to millions of copies after a number of cycles of amplifications. You cannot detect or sequence a single copy gene if it is not amplified.

Q: What excites you about the future of this field?

This is not limited to HIV research, but I think new research technology is amazing and that keeps me going. New information comes out every day, such as high-throughput analysis data, and people are using omics approaches. Those advances are exciting because they tell us a lot at a very fine resolution and we can make use of those data by translating them into real life-saving approaches.

This interview has been condensed and edited for clarity.

TECHNOLOGY OVERVIEW

Written by Thermo Fisher Scientific Staff

In our globally interconnected world, infectious pathogens are a problem and will continue to be a problem for the foreseeable future. Since infectious pathogens are usually caused by living organisms (in the case of bacteria and fungi) or require living organisms for transmission (viruses), sequence analysis of pathogens and host response is a critical tool for combatting these pathogens. The Human Genome Sequencing project was instrumental for generating tools that both generated sequence information and used that sequence information to understand human health and pathogens. The techniques and data generated by these tools have been critical for contributing to the knowledge of human biology, but have also been critical for understanding infectious pathogens. Thermo Fisher Scientific has been at the forefront of development of these tools and has a deep portfolio of solutions that are used for infectious pathogen research.

The Applied Biosystems portfolio can be thought of as a continuum that gives researchers options based on the types of questions being asked (Figure 1). For example, some investigators might be interested in discovery-based experimentation, using unbiased queries spanning the entire genome to uncover new sequences or relationships between genes. These types of experiments are facilitated by Thermo Fisher Scientific's Ion Torrent™ Ampliseq™ next generation sequencing (NGS) workflows or by Thermo Fisher Scientific's Applied Biosystems™ microarray solutions. Once a set of sequences of interest is decided, investigators might move to a focus-based experimental model, where these selected sequences are analyzed in medium-throughput analyses such as Sanger sequencing to focus on a defined region across many samples, or Thermo Fisher Scientific's Applied Biosystems TaqMan™ panels on array cards to focus on sets containing many genes. Finally, if a very specific set of mutations or genes need to be analyzed in large number of samples with



Figure 1. The genetic analysis continuum. In neuroscience research, the tools and techniques used depend on the question being asked, the scale of the data that need to be analyzed, and the logistical needs of the investigator. Applied Biosystems™ genetic analysis systems fall into a continuum, from very large-scale discovery-based research to medium-scale focused research to very targeted detection of specific genes and mutations.

minimal effort, individual TaqMan assays may be used for detecting these sequences.

Importantly, these approaches and technologies are complementary to each other. For example, a researcher might start with a discovery-based approach to determine the nature of an infectious pathogen, then focus on genetic drift in that organism by Sanger sequencing different isolates in more samples, and finally designing a TaqMan probe that specifically and sensitively detects the presence of a highly pathogenic variant. The Applied Biosystems solutions continuum is meant to ease the research path from discovery to insight.

In this chapter, we give a brief overview of the technologies in the Thermo Fisher Scientific continuum and how they could be used for infectious pathogen research. However, this is not a complete or comprehensive list. The flexibility of these solutions to cancer research depends only on the researcher's experimental creativity.

Microarray Methods and Tools

When a cell is infected with a pathogen, its response is often modulated by the genetic makeup of that cell. The collection of genotypes dictating this response could include single nucleotide polymorphisms, small indels, larger copy number changes, and large chromosomal rearrangements. One way to characterize these genomic features is to use high-density DNA microarrays. For these analyses, hundreds of thousands of probes tiled across the genome are arrayed onto a single chip. Sample hybridization to these arrays can determine which sequences are present and in how many copies (Figure 2). The advantage of microarrays is that hundreds of thousands to millions of sequences can be interrogated in a single experiment. If the appropriate probes that detect specific mutations are present on the chip, they can also detect common SNP variants. Microarrays are also useful for detecting copy number changes. The resolution of the copy number sequences is determined by the number of probes available; when probes are directly designed for predefined common deletions or duplications, these can detect anomalies as small as 1,000 nucleotides. Whole-genome microarrays that cover both polymorphic (e.g., single-nucleotide polymorphisms or SNPs) and nonpolymorphic regions of the genome can be used to assess DNA copy number alterations at a much higher resolution than conventional cytogenetic analyses.

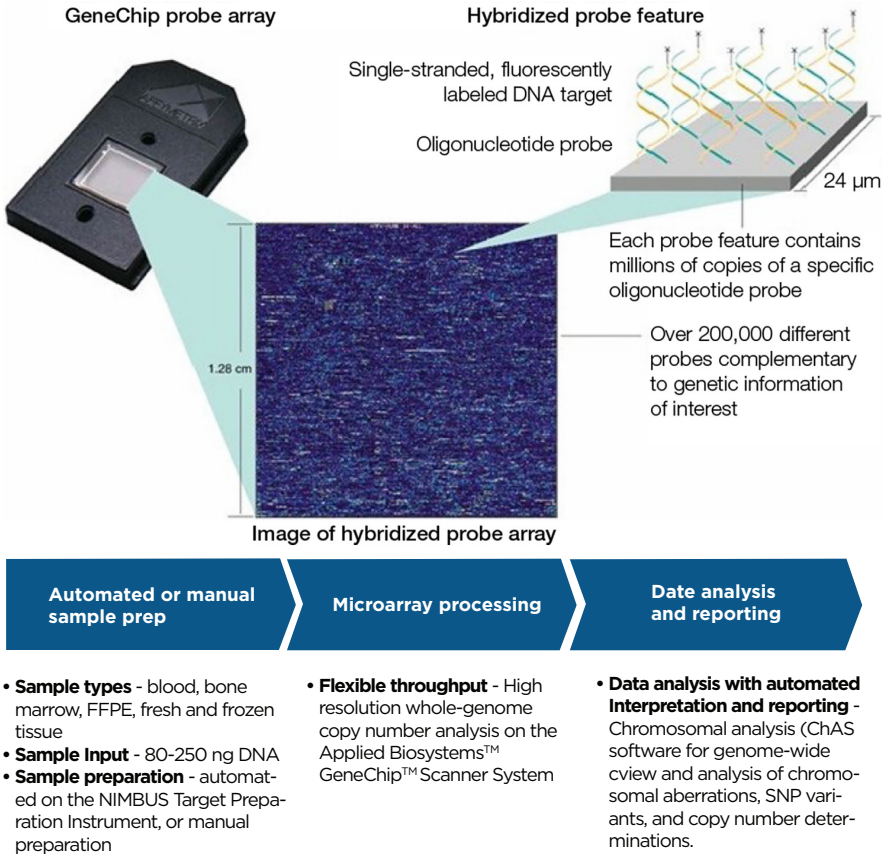


Figure 2. Basics of a microarray assay.

Genotyping microarrays are useful for discovering variants that influence sensitivity and severity of infectious pathogens. For GWAS studies, the Applied Biosystems UK Biobank Axiom™ Array facilitates high-throughput, high-value genotyping of large sample cohorts with a single comprehensive low-cost solution. This assay contains one 96-array plate that allows for genome-wide analysis of large sample collections such as those screened at biobanks, genome centers, and core labs. Genotypes collected using the UK Biobank array provided insights into SNPs that influenced SARS-CoV-2 sensitivity, tuberculosis as well as other infectious pathogens.[1,2] The Applied Biosystems Axiom™ Precision Medicine Diversity Research Array includes content that was selected to emphasize variants that are commonly seen in clinical research. The markers were chosen from the list of published variants associated with phenotypes identified via GWAS, as per NHGRI-EBI GWAS catalog as well as some recently published and unpublished cancer-associated SNPs. The clinically actionable variants include relevant variants from ClinVAR and enables assessment of actionable genetic risk across a wide range of populations.

miRNAs and other non-coding RNAs are important regulators of gene activity. It is estimated that more than 30% of protein translation of coding genes is regulated by miRNA. Importantly, miRNAs can be found in circulating fluids, such as blood or cerebral spinal fluid (CSF), and can be analyzed in liquid biopsies.[3] miRNAs are important biomarkers.[4] Because of their small size, stability, potential for regulating several different pathways, and access blood samples, miRNAs are ideal biomarkers that can provide insights into an infectious pathology. To keep pace with the discovery of

new and novel miRNA, Thermo Fisher Scientific offers the Applied Biosystems GeneChip™ miRNA 4.0 Array. This array is designed to interrogate all mature miRNA sequences in miRbase, including human, mouse and rat miRNAs. The analysis files contain host gene ID, predicted and validated miRNA target genes, and clustered miRNA information, facilitating downstream analysis.

The Applied Biosystems Clariom™ microarrays provide whole transcriptome analysis solutions. Human Clariom S Assay design facilitates extensive coverage of all known well-annotated genes, is compatible with most research sample types, and is a tool of choice for finding gene expression patterns with known function as quickly, easily, and cost-effectively as possible. The more advanced Clariom D Assays are based on coding and noncoding sequences culled from 16 different databases, representing over 540,000 transcripts in over 6.7 million probes.[5] Included in these sequences are probes for different splice isoforms, noncoding RNAs (pre-miRNAs, lincRNAs, Piwi-interacting RNAs (piRNAs)), and circular RNAs, as well as annotated and speculative sequences. Clariom D Assays allow translational research scientists to generate high-fidelity biomarker signatures quickly and easily.[5] Querying the multiple transcript models on the Clariom D chip in a single experiment helps ensure that important pathways and potential biomarkers are not missed.

Analysis of transcriptome analysis is facilitated by the Applied Biosystems Transcriptome Analysis Console (TAC) Software data. A report of all transcripts and the associated relative expression, using statistical and visualization tools, is provided from the data. Additionally, information regarding the role of these transcripts in biological pathways is provided. Together, the combination of microarrays and analysis software facilitates rapid, easy, and economical tools for identifying meaningful gene expression signature biomarkers.

Fragment Analysis Methods and Tools

Fragment analysis is a highly flexible method that can be applied to a wide variety of research fields. The flexibility afforded with the choice of PCR primers (PCR being a necessary step in any fragment analysis experiment) means that a specifically sized fragment corresponding to a PCR target sequence is straightforward to generate (Figure 3). Coupled with the ability to label fragments with four or more different fluorophors, researchers have great flexibility in their fragment analysis experimental design. Below, we describe some examples of how fragment analysis can be used in infectious pathogen research problems.

Mutation Detection

As mentioned above, the response to a pathogen can depend on the allelic makeup of the host. This allelic configuration can be made up of few variants in one or more genes of interest. The Applied Biosystems SNaPshot™ Multiplex System is a versatile and economical method to perform SNP genotyping. Up to ten SNP markers (on different genes) can be investigated simultaneously by using PCR amplification, followed by dideoxy single-base extension (SBE) using an unlabeled primer, and then capillary electrophoresis of the resulting fragments. After electrophoresis and fluorescence detection, the alleles of a single marker appear as different colored peaks at roughly the same size in the electropherogram plot.[6]

Pathogen Detection

For the infectious pathogen researcher, fragment analysis can be used to detect the presence of a single or multiple pathogens in a single PCR reaction. This is often manifest as a multiplexed PCR assay, with different organisms or strains detected by different sized and different fluor-labelled fragments. For example, a panel that can detect 15 targets by multiplexed PCR followed by fragment analysis was used to demonstrate how this works for respiratory viruses.[7] A similar approach was used to detect respiratory bacterial pathogens.[8] Similarly, panels can be designed to query and classify the microbiome in health and disease (for example see references 9 and 10). The ability to design and detect amplicons based on length and fluorophore label

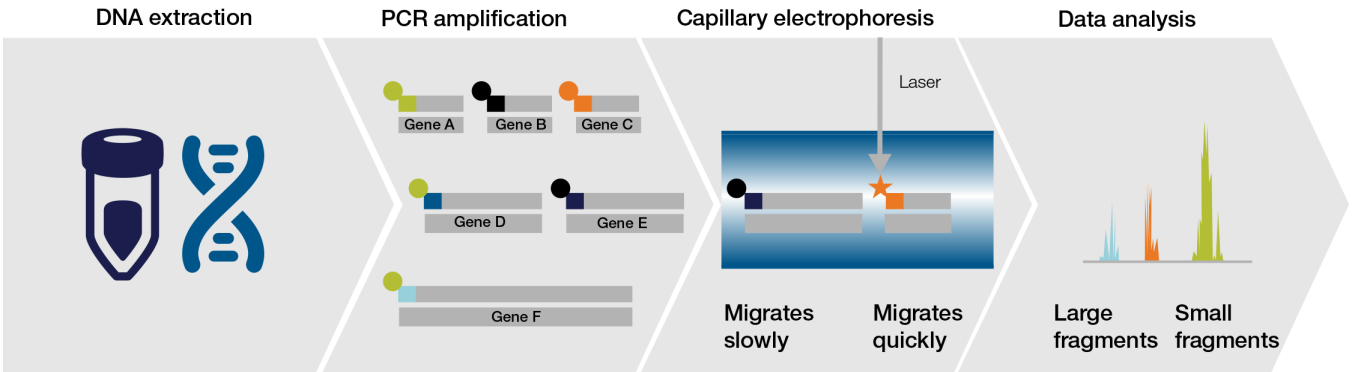


Figure 3. Basics of a fragment analysis assay. PCR primers are designed such that different target sequences are amplified with primers with different fluorophore labels that will generate different sized amplicons. Following PCR, the amplicons are electrophoretically separated by size in a capillary. A laser excites the fluorophores as the fragments migrates through the capillary. The peak intensity of the resulting fragments reflects the abundance of the target sequence in the sample.

* Public database sources for Clariom D Assays: RefSeq, ccdsGene, mgcGenes, UCSC Known Genes, lincRNATranscripts, VEGA, GENCODE, NONCODE, MiTranscriptome, RNAcentral, Ensembl, circBase, Guo CircRNAs, AceView, lncRNAWiki, lncRNADB.

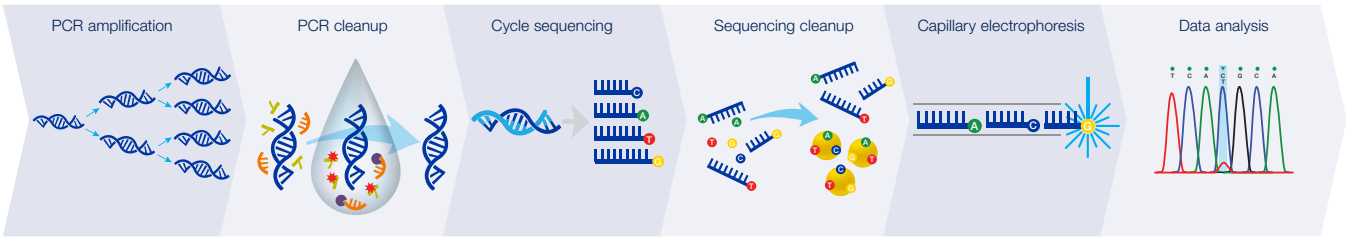


Figure 4. Basics of a Sanger sequencing workflow. The target region that will be sequenced is amplified by PCR. The primers are PCR reagents are removed before a second linear amplification is performed. This step fenerates the fragments that are chain-terminated with a fluorescent dideoxynucleotide. The cycle sequencing reaction is purified, and the resulting fragments are separated by capillary electrophoresis and detected with a laser. The sequence can be read from the lengths of the fragments and the colors of the dideoxynucleotide terminators.

provides an opportunity for very high degrees of target multiplexing. Furthermore, by designing amplicons with same sequence but labelled with different fluorophores, samples may be multiplexed by combining them after PCR but before CE, greatly increasing the number of samples that can be analyzed at one time.

Copy Number Variability

Multiplex Ligation Dependent Probe Amplification (MLPA), a technology developed and commercialized by MRC Holland B.V., is a flexible technique (www.mlpa.com) that is commonly used to detect aberrations in gene copy number such as duplications or deletions. It is based on the ligation and PCR amplification of up to 50 multiplexed pairs of probe oligonucleotides, which hybridize to the loci of interest. Each oligonucleotide pair is designed to give an amplification product of a specific length; and by using sequence-tagged ends, all ligated probes can be amplified with a single primer pair in a PCR reaction. The forward PCR primer carries a fluorescent label, allowing for the detection and quantification of size-separated probes on an automated capillary electrophoresis system.[11] MLPA can facilitate the analysis of genes where copy number variation may modulate pathogenic responses.

Epigenetic Modification

When infected with a bacterial or viral pathogen, host cells undergo extensive changes in gene expression. Part of this transcriptional response may involve DNA methylation changes in genes regulating the immune response.[12,13] Genome-wide hypomethylation or hypermethylation, particularly of cytosines 5' to guanosines (CpG) within the promoter regions of genes, has been shown to be associated changes in gene expression that could lead to a gene expression signature biomarker (for example, reference 14). The SNaPShot assay can also be used to investigate DNA methylation patterns following bisulfite treatment of genomic DNA. Sodium bisulfite converts unmethylated cytosines into uracil, which are subsequently replaced as thymine in PCR; in contrast, methylcytosines remain unchanged. This feature can be used to distinguish base differences by PCR amplification, T/C genotyping using SNaPshot SBE followed by fragment analysis.

MLPA can also be used to detect epigenetic modifications involving DNA methylation.[15] The methylation specific- MLPA (MS-MLPA) assay has probes that bind and ligate over a GCGC sequence, which is also a cleavage site for a methylation-sensitive restriction enzyme HhaI. The enzyme cleaves the probes ligate and hybridizes to unmethylated DNA while the probes bound to methylated DNA remain intact and are subsequently amplified by PCR.

Cell Line Authentication and Human Sample Matching

Human cell lines are heavily relied upon for biomedical research studying infectious pathogens. To assist with reproducibility of scientific research, cell line authentication (CLA) is of paramount importance to confirm the cell's origin as well as check for contamination and genomic instability. CLA is performed by generating a profile of highly variable short tandem repeat (STR) markers from microsatellite loci with varying number of repeats for a particular cell line/type and then comparing it with the allelic profiles present at these loci against known standards. A study aimed at authenticating 278 human tumor cell lines used in China found that nearly 46% of the samples were either cross contaminated or misidentified.[16] Needless to say, such findings can have massive implications on the conclusion of studies that utilize such cell lines. The Applied Biosystems Global Filer™ CLA kit generates a molecular fingerprint for 24 different STR loci, while the Identifiler Plus™ CLA kit analyzes 16 STR loci.

Sanger Sequencing Methods and Tools

Sanger sequencing is the trusted standard for obtaining DNA sequence information. It powered the Human Genome Project and investigators continue to rely on this method to generate highly accurate reliable sequencing results. Sanger sequencing is a specialized form of fragment analysis: it relies on chain-terminating fluorescent nucleotides to generate a series of fragments that differ by one nucleotide (Figure 4). Applied Biosystems' fast and straightforward Sanger sequencing workflows enable a high degree of accuracy, long-read capabilities, and simple data analysis. Applied Biosystems BigDye™ Terminator v1.1 and v3.1 cycle sequencing chemistries are widely trusted for Sanger sequencing by CE. After cycle sequencing, various options exist for clean-up before electrophoresis, including Applied Biosystems ExoSAP-IT™ enzyme mix , BigDye Xterminator™ kit, and the newly introduced Dynabeads™ Sequencing Clean-Up Kit. An entire sequencing workflow can be completed in less than one workday, from sample to answer, offering the flexibility to support a diverse range of applications in many research areas.

The simplicity of Sanger sequencing facilitates the analysis of methylated DNA sequences (for example, reference 17).

Primers that can be used for sequencing can be designed using the freely-available Methyl Primer Express software.[18] After bisulfite conversion, the DNA can be sequenced using the standard Sanger sequencing workflow.[19] Data analysis is also straightforward, since the converted sequence can be analyzed directly without needing to compare to a specialized reference genome.

Discovery-based genomic research, such as next generation sequencing (NGS), often uncovers novel or unexpected variants or other sequence anomalies. Investigators look for ways to verify these new discoveries using orthogonal methods. Sanger sequencing is often the method of choice for confirming NGS results studies because of its workflow simplicity and unambiguous results. For these confirmatory studies, short amplicons, usually covering only the region to be confirmed, need to be sequenced in these studies. Moreover, minor allelic variants, present in a heterogeneous sample, can be identified and confirmed by Sanger sequencing. Applied Biosystems' Minor Variant Finder Software is simple, easy-to-use desktop software designed for the accurate detection and reporting of minor variants in Sanger sequencing traces with a detection level of minor alleles as low as 5%. On a test set of 632,452 base positions, it exhibited a 5% limit of detection with 95.3% sensitivity and 99.83% specificity.[20] Minor Variant Finder Software can also readily align sequences with the human reference genome and VCF files from NGS experiments, supporting a smooth workflow for NGS confirmation with annotations in the dbSNP database.

PCR Methods and Tools

The TaqMan Assay family is the most comprehensive set of real-time PCR (qPCR) products available for analyzing gene expression, miRNA levels, protein abundance, copy number variation, SNP genotyping, and rare allele mutation detection. The broad library of predesigned Applied Biosystems assays helps laboratories accurately detect many pathogenic organism targets for various applications including clinical testing and infectious pathogen research.

The typical TaqMan assay consists of forward and reverse primers to amplify a target by PCR (Figure 5). TaqMan assays also contain a sequence-specific fluorescent probe, designed with a minor groove binder (MGB) moiety at the 3' end that increases the melting temperature (Tm) of the probe and stabilizes probe–target hybrids. This means that TaqMan MGB probes can be significantly shorter than traditional probes, allowing better sequence discrimination and flexibility to accommodate more targets. In addition, TaqMan probes have a non-fluorescent quencher (NFQ-MGB) that binds to the fluor until cleaved by the polymerase and minimizes background. TaqMan assays come with a performance guarantee* such that if they don't perform as expected, they can be replaced or refunded.[21]

One strategy that facilitates high-quality PCR data is that investigators follow standardized practices for performing and reporting RT-PCR results. The MIQE recommendations (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) are set of guidelines for real-time PCR (RT-PCR) experimental designs and data reporting practices, as well as standards for sharing experimental information with colleagues.[22] The standards are designed to ensure that published real-time PCR research is meaningful, accurate, and provides researchers with the information necessary to faithfully reproduce results. Applied Biosystems supports these guidelines and provides the information necessary to support MIQE compliance when publishing the results of experiments with TaqMan assays. Information requested under the qPCR target, qPCR oligonucleotide, and qPCR protocol sections of the guidelines are readily available in the Applied Biosystems reagent protocol or from the Applied Biosystems website.

The TaqMan portfolio is made up of over 27 million predesigned assays. These assays can be ordered as individual assays in tubes. Alternatively, TaqMan Arrays contain TaqMan assays dried down in three array formats: 384-well TaqMan Array microfluidic cards (TACs), 96- and 384-well TaqMan Array plates, and OpenArray formats. The TaqMan collection of assays includes genotyping and gene expression assays. In addition, the highly-specific TaqMan chemistry can also be used to analyze miRNAs for infectious pathogen response biomarker research.

To help with choosing the right TaqMan assays for an experiment, Applied Biosystems developed an assay search wizard.[23] Inputting a gene name, keyword, pathway or pathogen can return a list of all the assays that meet the search criterium. The results page highlights the "best coverage" assay for each

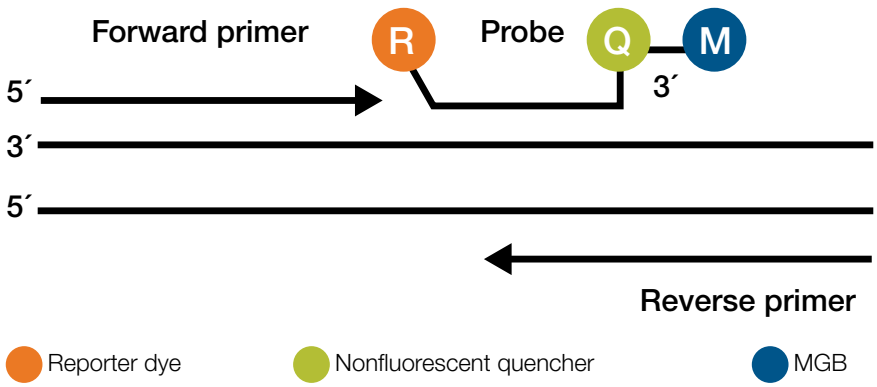


Figure 5. Basics of a TaqMan Assay. Unlabeled forward and reverse oligonucleotide primers define the region that will be queried. A third oligonucleotide binds to the region between the primers. The probe has an end-label fluorophore reporter, a nonfluorescent quencher, and a minor groove binder (MGB) moiety that stabilizes binding of the short probe sequence. When the probe is hybridized to its target sequence the quencher greatly reduces the fluorescence of the reporter. When the sequence is copied during the PCR reaction, the nuclease activity of the polymerase cleaves the reporter from the rest of the probe, releasing the reporter be detected. The amount of fluorescence is therefore directly related to the amount of target amplified.

* Terms and conditions apply. For complete details, go to TaqMan Assays qPCR Guarantee.

gene, publications that used the assay highlighted, and provides an opportunity to select and design a combination of assays for an array format. Finally, if an assay is not present in our portfolio, our TaqMan Custom Design Assay Tool can be used to design an assay based on any target sequence of interest.[24] Each custom assay is a mix of forward primer, reverse primer, and FAM and/or VIC dye-labeled TaqMan MGB probe.

Applied Biosystems has also developed a portfolio of pre-designed panels targeting common pathogens. These have been prespotted with gene expression assays targeting common pathways, pathogens and gene families, and include panels for respiratory track pathogens, vaginal and STI pathogens, urinary tract pathogens, and wound infections.[25]

dPCR Methods and Tools

Digital PCR (dPCR) is a method that quantifies sequences present in a sample by counting them. The basis of dPCR is that a nucleic acid sample is segregated into thousands of parallel PCR reactions, such that each reaction well contains one target molecule on average (Figure 6). In this scenario, some reactions will not contain the target molecule at all, while others will contain one or more copies. The collection of segregated reactions is subjected to endpoint PCR, and the number of wells containing a positive signal and no signal are tallied. The number of target copies can be calculated from the fraction of negative reactions, based on the assumption that the segregation follows a Poisson distribution (thus accounting for the possibility that multiple target molecules occupy the same reaction). The number of individual reactions influences the sensitivity of the assay—the more reactions there are, the lower the limit of detection and the higher the accuracy.

dPCR provides advantages to detecting rare target sequences in a sample. Because the entire sample is segregated into individual wells, detection of targets is not affected by an overabundance of alternate, unrelated targets. Secondly, dPCR reactions are endpoint PCR reactions. The ability to detect postive signals is therefore less effected by common PCR inhibitors. Finally, TaqMan assays are compatible with dPCR workflows; therefore an assay that is used in qPCR can also be used for dPCR.

The Applied Biosystems QuantStudio™ Absolute Q™ Digital PCR System consists of just one instrument, consolidating all steps required into a single plate and transforming a multi-step multi-instrument workflow into a one-step qPCR-like workflow. The utility of dPCR to monitor pathogens in wastewater was demonstrated during the SARS-CoV-2 pandemic, and may be more generally applicable to other pathogens affecting public health.[26,27] In addition, the Absolute Q system was recently used to quantify the amount of HIV2 virus in circulating plasma.[28]

For more information, contact your sales representative.

See references on page 12

For Research Use Only. Not for use in diagnostic procedures. © 2025 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. TaqMan is a trademark of Roche Molecular Systems, Inc., used under permission and license. MLPA is a trademark of MRC Holland. **COL35660 0423**

References

Article 1: A Strong Research Foundation Supports Infectious Pathogen Discoveries

1) Baker RE, et al. Infectious disease in an era of global change. *Nat Rev Microbiol.* 2021;20(4):193-205.
2) The top 10 causes of death. World Health Organization. 2024. Accessed October 9, 2024.
3) IHME Pathogen Core Group. Global burden associated with 85 pathogens in 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet Infect Dis.* 2024;24(8):868-895.
4) Estimates of funding for various Research, Condition, and Disease Categories (RCDC). National Institutes of Health. 2024. Accessed October 9, 2024.
5) Disease and conditions. National Institute of Allergy and Infectious Diseases. Accessed October 9, 2024.
6) Franco-Duarte R, et al. Advances in chemical and biological methods to identify microorganisms—from past to present. *Microorganisms.* 2019;7(5):130.
7) Wang HY, et al. Development of a high sensitivity TaqMan-based PCR assay for the specific detection of *Mycobacterium tuberculosis* complex in both pulmonary and extrapulmonary specimens. *Sci Rep.* 2019;9(1).
8) Kwok AJ, et al. Host genetics and infectious disease: new tools, insights and translational opportunities. *Nat Rev Genet.* 2020;22(3):137-153.
9) Lambrechts L, et al. Emerging PCR-based techniques to study HIV-1 reservoir persistence. *Viruses.* 2020;12(2):149.
10) Hilt EE, Ferrieri P. Next generation and other sequencing technologies in diagnostic microbiology and infectious diseases. *Genes.* 2022;13(9):1566.

Infographic: Addressing Infectious Pathogens with Genetic Technologies

1) IHME Pathogen Core Group. Global burden associated with 85 pathogens in 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet Infect Dis.* 2024;24(8):868-895.
2) National Center for Health Statistics. Health, United States, 2020-2021: Table IDNotif. Hyattsville, MD. Available from: https://www.cdc.gov/nchs/hus/data-finder.htm
3) Estimates of funding for various Research, Condition, and Disease Categories (RCDC). National Institutes of Health. 2024. Accessed October 9, 2024.
4) HIV data and statistics. World Health Organization. 2024. Accessed October 9, 2024.
5) National Center for Health Statistics. Health, United States, 2020-2021: Table HIV. Hyattsville, MD. Available from: https://www.cdc.gov/nchs/hus/data-finder.htm
6) Influenza (seasonal). World Health Organization. 2023. Accessed October 9, 2024.
7) Estimated flu disease burden 2019–2020 flu season. US Centers for Disease Control and Prevention. 2024. Accessed October 9, 2024.
8) Global tuberculosis report 2023. World Health Organization. 2023. Accessed October 9, 2024.
9) Das S, et al. The new frontier of diagnostics: Molecular assays and their role in infection prevention and control. *Am J Infect Control.* 2017;45(2):158-169.
10) Nafea AM, et al. Application of next-generation sequencing to identify different pathogens. *Front Microbiol.* 2024;14.

Article 4: Technology overview

1) Lin S, et al. Genome-wide epistasis study highlights genetic interactions influencing severity of COVID-19. *Eur J Epidemiol.* 2023;38(8):883-889.
2) Kerner G, et al. Homozygosity for TYK2 P1104A underlies tuberculosis in about 1% of patients in a cohort of European ancestry. *Proc Natl Acad Sci U S A.* 2019;116(21):10430-10434.
3) Ozturk EA, Caner A. Liquid biopsy for promising non-invasive diagnostic biomarkers in parasitic infections. *Acta Parasitol.* 2022;67(1):1-17.
4) Wang L, et al. MicroRNAs as immune regulators and biomarkers in tuberculosis. *Front Immunol.* 2022;13:1027472.
5) Clariom D microarrays provide a deep view of our transcriptome. Thermo Fisher Scientific. 2019.
6) Using the SNaPshot® Multiplex System with the POP-7™ Polymer on Applied Biosystems 3730/3730xl DNA Analyzers and 3130/3130xl Genetic Analyzers. Thermo Fisher Scientific. 2005.
7) Application Note: Multiplex Detection of Respiratory Pathogens using Fragment Analysis by Capillary Electrophoresis. Thermo Fisher Scientific. 2022.
8) Xi JL, et al. Simultaneous detection of 13 key bacterial respiratory pathogens by combination of multiplex PCR and capillary electrophoresis. *Biomed Environ Sci.* 2017; 30(8): 549-561.
9) Lai L, et al. A capillary electrophoresis method for analyzing the diversity of vaginal microbiome. *Clin Lab.* 2022;68(3).
10) inbiome Homepage. Accessed December 9, 2024.
11) MLPA assays on the SeqStudio Genetic Analyzer. Thermo Fisher Scientific. 2019.
12) Qin W, et al. The role of host cell DNA methylation in the immune response to bacterial infection. *Front Immunol.* 2021;12:696280.
13) Anderson D, et al. Zika virus changes methylation of genes involved in immune response and neural development in Brazilian babies born with congenital microcephaly. *J Infect Dis.* 2021;223(3):435-440.
14) Arama C, et al. Epigenetics and malaria susceptibility/protection: A missing piece of the puzzle. *Front Immunol.* 2018;9:1733.
15) Nygren AO, et al. Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucleic Acids Res.* 2005;33(14):e128.
16) Huang Y, et al. Investigation of cross-contamination and misidentification of 278 widely used tumor cell lines. *PLoS One.* 2017;12(1):e0170384.
17) Salcedo-Tacuma D, et al. Differential methylation levels in CpGs of the BIN1 gene in individuals with Alzheimer Disease. *Alzheimer Dis Assoc Disord.* 2019;33(4):321-326.
18) Methyl Primer Express™ Software v1.0. Thermo Fisher Scientific. Accessed December 9, 2024.
19) Analyzing methylated DNA in colon tumor samples for translational research. Thermo Fisher Scientific. 2024.
20) Low-level somatic variant detection in tumor FFPE samples by Sanger sequencing. Thermo Fisher Scientific. 2016.
21) TaqMan Assays qPCR Guarantee. Thermo Fisher Scientific. Accessed December 9, 2024.
22) Bustin SA, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem.* 2009;55(4):611-22.
23) TaqMan Assay Search Wizard. Thermo Fisher Scientific. Accessed December 9, 2024.
24) Custom TaqMan® Assay Design Tool. Thermo Fisher Scientific. Accessed December 9, 2024.
25) Real-time PCR for the molecular detection of infectious disease pathogens and resistance genes. Thermo Fisher Scientific. Accessed December 9, 2024.
26) Tiwari A, et al. Application of digital PCR for public health-related water quality monitoring. *Sci Total Environ.* 2022;837:155663.
27) Environmental Surveillance using Digital PCR. Thermo Fisher Scientific. Accessed December 9, 2024.
28) Falces-Romero I, et al. Quantification of human immunodeficiency virus type 2 (HIV-2) viral load in plasma: Comparison of three commercial assays. *J Clin Virol.* 2024; 175:105745.