

Jumpstarting Infectious Disease Research With Genomics

ABOUT INFECTIOUS DISEASE

Infectious diseases are the second leading cause of death worldwide and can be attributed to two-thirds of all childhood deaths under the age of five. Tuberculosis alone affects approximately one-third of the world's population in either its latent or active form, causing 1.5 million deaths annually. Treatments for infectious diseases have saved countless lives but researchers and healthcare workers continue to face challenges: the pathogens continue to evolve and adapt to their hosts, leading to antibiotic resistance and creating what is essentially a moving target for drug discovery. Antibiotics designed to kill bacteria can have the secondary effect of allowing those bacteria that are resistant to the antibiotic to flourish, causing resistant strains to develop. Pathogen adaptations to the host environment also add to the list of moving parts shifting the landscape of infectious disease.

Developing innovative therapeutics to treat infectious diseases will require innovative approaches. The Infectious Disease Program at the Broad Institute of MIT and Harvard uses genomic and computational analysis, along with chemical screens, to raise the understanding of pathogen biology to a resolution that will help reveal the next generation of infectious disease therapies. Approaches to searching for drugs and vaccines that rely upon testing potential

compounds or therapies a few at a time have been limited in their success. This is especially true in cases where a good animal model of infection doesn't exist, the pathogen is hard to grow in the laboratory, or infection is mediated by a complex immune response. Genomic and computational approaches can transcend these obstacles, enabling researchers to rapidly identify pathogens and potential therapeutic targets, as well as conduct epidemiological risk analysis and model evolutionary change in order to select drug targets in genomic regions less likely to mutate.

Researchers can also create an infrastructure for the infectious disease community to collectively annotate pathogen genomic data to help everyone "speak the same language." As Deborah Hung, co-director of the Infectious Disease Program, noted, "Given the magnitude of the problem and the incredible diversity of pathogens, even within the same species, creating large high-quality datasets that are obtained and analyzed in a uniform way for all researchers to access will be critical to understanding how pathogens cause disease and thus ways to intervene on disease."

WHY ARE THESE DISEASES SO HARD TO TREAT?

Developing effective therapeutics against infectious diseases requires an understanding of how the pathogen evolves and changes within the context of its host ecosystem over time. Host infection provides a supportive environment for the pathogen to replicate. Replication creates the opportunity for mutation and, significantly, the opportunity for mutations in parts of the pathogen genome that could alter important features, such as its transmission mode or lethality. Pathogen evolution also needs to be taken into account during vaccine development because



Office of Communications

www.broadinstitute.org

617-714-7151 • communications@broadinstitute.org

415 Main Street • Cambridge, MA 02142

targeting areas of the pathogen's genome that have a high mutation rate increases the likelihood of wasted efforts if, for example, the region targeted has already mutated before the vaccine has even been released.

In addition to the pathogen itself, infection relies on successful pathogen-host interaction. The pathogen must manage to evade destruction by the host immune system. The genes encoding these effector proteins are usually extrachromosomal, located on plasmids or bacteriophages, which have an even higher mutation rate than the DNA on chromosomes, and the sequences can undergo rapid selection based on their ability to evade the host immune system.

TUBERCULOSIS: A NIMBLE FOE

Tuberculosis is caused by *Mycobacterium tuberculosis* infection. This bacterium has been among the most intractable to eradicate in part due to its nimble adaptation strategies. It infects the host's macrophage cells in the lungs, it is able to survive for long time periods in cellular compartments for foreign particles called phagosomes, and it can prevent the compartment's acidification, prolonging its survival. The four types of tuberculosis recognized in the world today are: regular, multidrug resistant, extensively drug resistant, and totally drug resistant. New therapeutics are clearly needed to stay ahead of this rapidly adapting bacteria.



Mycobacterium tuberculosis Credit: NIAID

"The fact that about one-third of the world's population is infected with tuberculosis, but only a small fraction actually gets active disease suggests that, in fact, in the majority of people, host immunity is really effective. If we could understand the basis for this host control, we might be able to harness it towards more effective therapy."

— Deborah Hung, M.D., Ph.D.

Tuberculosis researchers at the Broad are using comparative genomic analysis, population genetics of pathogen populations, and chemical screening approaches to try to usher in the next generation of therapeutics. Deborah Hung's laboratory uses chemical and genomic approaches to better understand virulence and host-pathogen interactions. These approaches have led to some encouraging results. Hung and her collaborators designed a high-throughput imaging assay and chemical screen to look for small molecules that act in the infected macrophages and limit tuberculosis bacterial growth by interfering with host pathways. The screen identified several different compounds including the selective serotonin reuptake inhibitor, fluoxetine (more commonly known as Prozac), and the EGF pathway inhibitor, gefitinib. These results point to new compounds for further study, as well as opportunities to repurpose drugs

already in use. They also demonstrate the value of focusing on the host response. "The fact that about one-third of the world's population is infected with tuberculosis, but only a small fraction actually gets active disease suggests that, in fact, in the majority of people, host immunity is really effective. If we could understand the basis for this host control, we might be able to harness it towards more effective therapy," said Hung.

ANTIBIOTIC RESISTANCE

Last spring the World Health Organization released its first-ever report on antibiotic resistance, warning that antibiotic resistance is now a significant threat to public health and proclaiming that without "urgent, coordinated action" the world is headed for a "post-antibiotic era." The report notes that resistance to carbapenem antibiotics, used as a treatment of last resort for a majority of hospital-acquired infections, has now spread worldwide. While there are initiatives being developed to reduce the need for antibiotics, such as improved hygiene and access to clean drinking water to curb the spread of infection, researchers in the Broad's Infectious Disease Program have been looking prospectively, using genomics to understand how bacteria become antibiotic resistant and learn how to make better antibiotics. Additionally, they are working on faster and more sensitive diagnostics to both detect pathogens and determine their antibiotic sensitivities so patients can be treated quickly and with the right antibiotic.

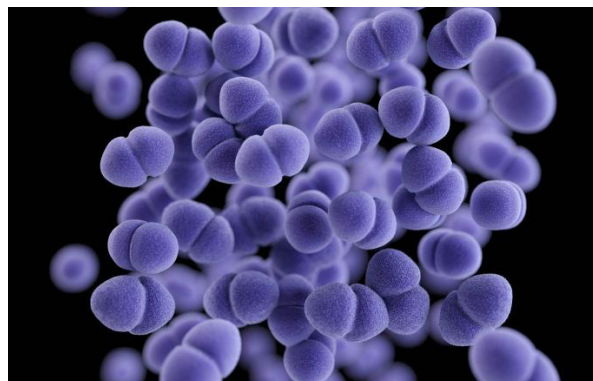
GENOMICS OF ANTIBIOTIC RESISTANCE

Enterococcus bacteria have long been part of our natural microbiome, the delicate ecosystem of microorganisms that live in and on our body. Among the more than 17 species of *Enterococcus*, only a few cause pathogenic infections in humans. With the advent of antibiotics, *Enterococcus* has become a leading cause of multidrug resistant hospital

Last spring the World Health Organization released its first-ever report on antibiotic resistance, warning that antibiotic resistance is now a significant threat to public health and proclaiming that without "urgent, coordinated action" the world is headed for a "post-antibiotic era."

infections. Broad associate member Michael Gilmore and his laboratory have teamed up with the Broad's Genome Sequencing and Analysis Program to understand what changed, together launching the EnteroGenome project (<https://olive.broadinstitute.org/projects/enterogenome>). This collaboration has found that the genomes of *Enterococcus* have changed substantially over the last 75 years.

Antibiotic use has selected for strains of *Enterococcus* that are very good at acquiring antibiotic resistance and other traits that allow them to persist in hospitals



3D illustration of drug-resistant *Enterococcus*.
Credit: CDC/JamesArcher.

and infect patients. For example, the CRISPR system has received a lot of attention as a tool for genome editing. It originated, however, as a defense system in bacteria against invaders such as viruses and plasmid DNA. Since approximately 25% of the DNA in multi-drug resistant hospital strains of the species *Enterococcus faecalis* is made up of mobile DNA elements, the researchers investigated the relationship between antibiotic resistance and the CRISPR system. The group built a genome sequence database that now contains information from over 200 *Enterococcus faecalis* genomes. They found that loss of CRISPR elements correlated with increased antibiotic resistance, suggesting that antibiotic use resulted in evolutionary selection for bacteria that traded this defense of the cell against invading DNA in favor of the ability to acquire DNA. Whereas historically it was advantageous for *Enterococcus* to protect its highly adapted genome for life in the gut, in the antibiotic era it became advantageous to discard that protection so that it could acquire resistance genes on mobile DNA elements and adapt to the antibiotic-rich hospital environment. "Antibiotics and hygiene have arguably done more to extend human life than any other medical advance," said Gilmore. "But our use of antibiotics is permanently changing the bacteria we have to live with, some for the worse. We have to understand these changes so that we can be careful stewards of these critical drugs and preserve their use."

BETTER DIAGNOSTICS KEY FOR CONTAINING AND TREATING OUTBREAKS

Critical aspects of infectious disease management include rapid diagnosis and appropriate antibiotic selection in the case of bacterial infection. Delays in either of these can lead to increased disease transmission as well as an increased mortality rate among patients. Broad Infectious Disease Program researchers, along with the Broad Technology Labs and Genome Sequencing and Analysis Program, have been working to develop RNA-based pathogen and

drug resistance detection methods in order to begin treating patients quickly and with the right drugs.

Currently, many patients are diagnosed using a variety of immunoassays and culturing approaches that can be slow and reveal limited information about the infecting pathogen. DNA-based methods can improve upon the time to diagnosis if the pathogen sequence is known, but require a great deal of knowledge about the genes involved to assess antibiotic resistance. RNA-based detection is ideal to rapidly detect pathogens and determine antibiotic resistance, because even brief antibiotic exposure can quickly cause detectable changes in transcription if the patient is not resistant. Researchers at the Broad have published the proof-of-principle for this approach and they continue to refine and expand its utility. These types of studies capitalize on the breadth of expertise and collaborative structure at the Broad.

THE MICROBIOME: OUR VITAL BIOTA

The microorganisms that inhabit our body exist in a delicate equilibrium. Dysregulation of this equilibrium has a growing list of implications on our health and well-being and is now connected with a range of issues from autoimmune diseases, such as Crohn's disease, allergies, and obesity, underscoring the importance of a thorough understanding of the microbiome and its constituents.

In order to better understand the critical constituents of our microbiome, scientists at the Broad who embarked on the Human Microbiome Project have now built an expanded toolbox for studying the internal microbial commu-

nities and the host-microbe interactions. Since many of the microbes are difficult to culture in the laboratory, they use culture-independent sequence-based



approaches to identify and determine the lineages of the microbes, approaches that are complementary with other 'omics assays (see also ibdmdb.org).

Infectious diseases have traditionally been studied by focusing in on individual organisms. However, infections should be considered from the perspective of complex interactions between a host, a pathogen, and the commensals of the host. For example, overgrowth with the bacteria *C. difficile*, a toxin producer that attacks the walls of the intestine and causes infectious diarrhea, typically follows an antibiotic therapy that results in an imbalanced intestinal microbiome. Instead of treating this by only eradicating the pathogen with more antibiotics, a fundamentally different approach is to replenish the gut with a complex microbiome through a fecal transplant from a healthy donor. Despite a proven record of high efficacy linked to this method, many questions still remain before this crude method can be more refined.

GENOMIC EVOLUTION OF GLOBAL SCOURGES

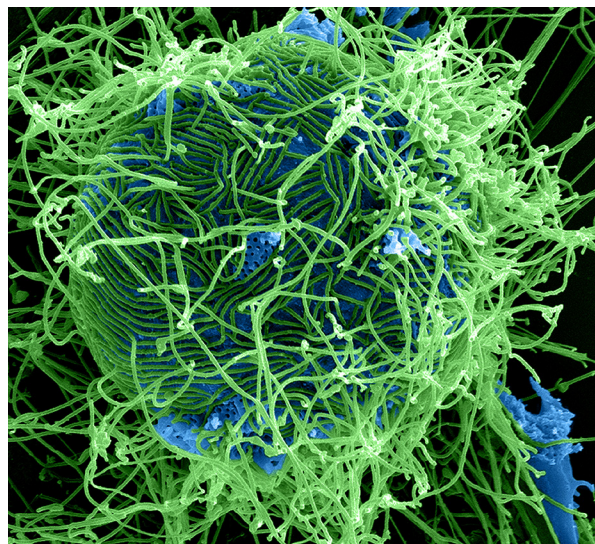
With improved high-throughput sequencing technologies, researchers can examine the genomes of deadly viruses and bacteria with unprecedented ease. Institute member Pardis Sabeti and colleagues at the Broad Institute study the Lassa and Ebola viral hemorrhagic fevers. They are examining the viral genome sequences in depth in order to understand how the viruses are changing over time, as well as which regions are mutating most rapidly. This information helps explain and predict transmission patterns and informs drug and vaccine development by providing information about which parts of the genome are most stable and represent the best potential targets. Genomic information is also critical in helping coordinate an effective ground response during an outbreak, as described in the following Ebola case study.

WATCHING AN OUTBREAK IN REAL TIME: THE EBOLA CASE FOR GENOMIC SURVEILLANCE

The 2014 Ebola outbreak has been unprecedented in the number of victims as well as its emergence in less remote and more populated areas in West Africa than had been affected by previous outbreaks. The virus is spread by contact with contaminated bodily fluids and affects great apes as well as humans with a 60 to 90 percent lethality rate. Current treatment for Ebola patients is largely restricted to supportive care. The viral genome encodes only seven gene products. The viral glycoprotein, a protein necessary for the virus to infect a cell, is being pursued as an interesting drug and vaccine target, but neither option is commercially available today.

Pardis Sabeti and her team worked with the Genomics Platform and Genome Sequencing and Analysis Program at the Broad to do almost real-time "genomic surveillance" of the outbreak, analyzing Ebola virus genome samples from patients in Sierra Leone and Guinea.

In this study, researchers deep-sequenced the Ebola virus genomes from 78 patients. (Deep sequencing is



Colorized scanning electron micrograph of filamentous Ebola virus particles (green) attached to and budding from a chronically infected VERO E6 cell (blue) (25,000x magnification). Credit: NIAID

sequencing done enough times to generate high confidence in the results.) These analyses made it possible to learn about the origin, lineage, and transmission of the virus. These findings are critical for setting up an effective ground response to such an outbreak because they could have important implications for field diagnostics, for example, if a sequence variation appears where primers (starting points for DNA synthesis) used in PCR-based diagnostic tests are located. It can also contribute to drug and vaccine development by pointing out areas of the viral genome that tend to evolve rapidly and might not be the most stable drug target. Sabeti explained, "Our sequencing data has provided a snapshot of how the virus is changing over the course of this outbreak. We see many new mutations emerging, a number of which have already affected diagnostic target sites, and a number of which affect the virus's proteins important for vaccines and therapies. We will want to continue to examine changes in the outbreak for implications on further surveillance and therapies." The approach shows that genomic surveillance of an outbreak, an approach applicable across diseases, can positively impact the ground response in the short term and inform therapies in the long term.

THE NEXT GENERATION OF INFECTIOUS DISEASE RESEARCH

The compound-by-compound, gene-by-gene approaches that preceded the Broad's current capabilities gave researchers important insights into the pathogens that cause many of the world's deadliest diseases, as well as introduced first lines of treatment that saved or improved countless lives. "The discovery of antibiotics was perhaps the most transforming event of modern medicine; however, the rise in antibiotic resistance is dramatically eroding into that effectiveness," said Deborah Hung. "We really need a renewed effort and focus on infectious diseases to find new drugs, develop more rapid diagnostics, and create more effective vaccines." Genomic and

computational analysis, chemical screening, and other powerful, high-throughput approaches are now ready to help bridge the gaps in understanding that cost millions of lives each year. These approaches also help the scientific community monitor and model pathogen evolution, run epidemiological risk analysis, and create integrated databases that will allow infectious disease researchers to optimally share information and take advantage of potential synergies.

FURTHER READING

Zumla A, Nahid P, Cole ST. [Advances in the development of new tuberculosis drugs and treatment regimens](#). *Nat Rev Drug Discov*. 2013 May;12(5). DOI: 10.1038/nrd4001.

Stanley SA, et al. [Identification of host-targeted small molecules that restrict intracellular *Mycobacterium tuberculosis* growth](#). *PLoS Pathog*. 2014 Feb 20;10(2): e1003946. DOI: 10.1371/journal.ppat.1003946. eCollection 2014.

Gire S, et al. [Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak](#). *Science*. 2014 Sep 12; 345(6202). DOI: 10.1126/science.1259657.

Barczak AK, et al. [RNA signatures allow rapid identification of pathogens and antibiotic susceptibilities](#). *Proc Natl Acad Sci U S A*. 2012 Apr 17;109(16). DOI: 10.1073/pnas.1119540109.

Palmer KL, Gilmore MS. [Multidrug-resistant enterococci lack CRISPR-cas](#). *MBio*. 2010 Oct 12; 1(4). DOI: 10.1128/mBio.00227-10.

Gevers D, et al. [The Human Microbiome Project: a community resource for the healthy human microbiome](#). *PLoS Biol*. 2012 Aug 14. DOI: 10.1371/journal.pbio.1001377.

Last updated December 2014