A Mission to Preserve and Grow

The UK’s National Collection of Type Cultures (NCTC) is a unique collection of more than 5,000 expertly preserved and authenticated bacterial cultures, many of historical significance. Founded in 1920, NCTC is the longest established collection of its type anywhere in the world, with a history of its own that has reflected – and contributed to – the evolution of microbiology for more than 100 years.

But the NCTC is far from stuck in the past. In fact, it has been at the forefront of advances in the field, continually implementing the latest technologies in order to provide the best, most comprehensive resources to support microbiology research institutes worldwide charged with a broad range of missions.

In the 1930s, this meant the introduction of freeze-drying strains to ensure longevity and to streamline storage and shipment. In 1949, the NCTC began a 10-year effort to characterize every organism in the collection, generating records of colony morphology, biochemical test results, and freeze-drying status. In 1965, the bench-top staple Cowan and Steel’s Manual for the Identification of Bacteria, co-authored by NCTC curator Samuel Cowan, was published, and the NCTC began to develop computer-based bacterial identification methods.

With the advent of the genomics era, NCTC embraced the use of DNA-based methods for species identification. During the 1980s, Curator L. R. Hill introduced GC content analysis, DNA-DNA hybridization and restriction fragment length polymorphism (RFLP) analyses for more accurate species identification. 16S ribosomal RNA gene sequence analysis was introduced in the 1990s and NCTC strains continue to be routinely identified and authenticated using this method.

Currently, new strains are characterized for morphology, nutritional requirements, enzyme activity and subjected to serotyping, mass spectrometry, 16S sequencing, and, most recently, whole genome sequencing (WGS).

NCTC 3000 Project:
A Resource for Scientists Worldwide

In 2014, NCTC launched an ambitious five-year project, together with the Wellcome Sanger Institute, to generate high-quality reference genomes for 3,000 bacterial strains. The method of choice? Single Molecule, Real-Time (SMRT®) Sequencing. According to Professor Julian Parkhill of the Wellcome Sanger Institute, PacBio® technology was selected because it produces long-read genome assemblies with the highest consensus accuracy and uniform coverage.

“Whole genome sequencing is set to revolutionize medical microbiology. Genomic characterization of expertly-authenticated type and reference culture collection strains will contribute significantly to a wide range of clinical and research applications as we progress further into this new genomic era,” Prof Parkhill said. “SMRT Sequencing delivers long read lengths, generating the most comprehensive de novo assemblies.”

Julie E. Russell, Head of Culture Collections at Public Health England, added: “If NCTC is to continue to supply relevant authentic bacteria for use in scientific studies, then the quality of our own characterization and authentication data must be outstanding. Combining sequences, strain metadata and links to other resources in the public domain will ensure that this resource provides a unique comprehensive source of data to underpin microbial research and improve the provision of diagnostics and public health interventions for medically important bacteria and viruses.”

“If you’re trying to generate reference genomes that are going to be valuable to as many people as possible, with as much information in them as possible, then Pacific Biosciences has the edge in terms of generating more complete data.”

– Professor Julian Parkhill
Wellcome Sanger Institute

Julian Parkhill is a Professor at the Wellcome Sanger Institute

www.pacb.com/microbe
A Brief History of Microbiology

Laying the Foundation

1886: Robert Koch identifies a bacterium which he called "Bacterium coli communis," which would later be called *Escherichia coli*. A strain he isolated in 1886 is added to NCTC at its founding.

1890: German scientist Robert Koch provides proof of cause and effect between medical symptoms and culturable representative specimens of bacteria, through the isolation of pathogenic *Shigella"*, *Salmonella"*, and *Yersinia"* infections in volunteers.

1897: Robert Koch invents the agar-impregnated dish for cultivating bacteria, furthering early attempts at cultivating involved potato slices and gelatin.

1900: Alfred Wegener introduces *NCTC 160*, a strain of *Salmonella typhimurium* as a classical example of a vaccine strain during the First World War. It was one of the earliest vaccines to be used in human and veterinary medicine.

1890s: German scientist Robert Koch provides proof of cause and effect between medical symptoms and culturable representative specimens of bacteria, through the isolation of pathogenic *Shigella"*, *Salmonella"*, and *Yersinia"* infections in volunteers.

1900: Alfred Wegener introduces *NCTC 160*, a strain of *Salmonella typhimurium* as a classical example of a vaccine strain during the First World War. It was one of the earliest vaccines to be used in human and veterinary medicine.

1915: The isolation of the very first bacterial strain registered in the NCTC database, the *S. flexneri"* strain, is discovered by the bacteriologist Frederick William Twort in 1915.

1920s: Selman Waksman and Albert Schatz lead a pioneering effort to find new bacteria and antibiotics. Waksman discovers streptomycin (NCTC 4323) from which they isolated streptomycin.

1928: Alexander Fleming accidentally discovers penicillin. He returns from vacation and notices that a culture plate left lying out in his laboratory had become overlaid with fungi, except around a certain area. Fleming explores further after his former assistant Merie Price reminds him, “That’s how you discovered penicillin.” Over the next 20 years, Fleming deposits 5,000 penicillin strains, including a sample of *Herpetobacter"* isolated from his own nose in November 1935.

1930s: NCTC introduces freeze-drying of strains to ensure longevity and prevent storage and spoilage.

1941: Howard Florey and Ernst Chain begin mass-producing penicillin from soil cultures from the US and British governments after the bombing of Pearl Harbor. By D-Day in 1944, enough penicillin had been produced to treat all wounded Allied Forces.

1941: Howard Florey and Ernst Chain determine the antibacterial action of penicillin and begin mass-producing it for military use.

1947: Edward Tatum and Joshua Lederberg publish the first gene map of *E. coli*. A map of the *E. coli"* genome is published in 1947 by Tatum and Lederberg.


1950s: Waksman and Albert Schatz are awarded the Nobel Prize in Physiology or Medicine for their work on antibiotics.

1953: Howard Florey and Ernst Chain receive the Nobel Prize in Physiology or Medicine for their work on antibiotics.

1954: The biggest recorded outbreak of foodborne illness is documented, with over 200,000 cases of food poisoning. She eventually deposits over 200,000 *NCTC"* strains of bacteria associated with foodborne illness.

1960s: The European Collection of Microorganisms for Biotechnology (ECM) is established.

1977: CDC researchers Joseph McDade and Charles D. Shepard isolate *Legionella pneumophila"*, the bacterial pathogen behind the outbreak of Legionnaires’ disease in Philadelphia.

1977: Gilbert and Sanger independently develop methods to determine the exact sequence of DNA molecules.

1977: CDC researchers Joseph McDade and Charles D. Shepard isolate *Legionella pneumophila"*, the bacterial pathogen behind the outbreak of Legionnaires’ disease in Philadelphia.

1977: CDC researchers Joseph McDade and Charles D. Shepard isolate *Legionella pneumophila"*, the bacterial pathogen behind the outbreak of Legionnaires’ disease in Philadelphia.

1980s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

1990s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

2000s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

2010s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

The Genomics Era

2011: Pacific Biosciences announces the PacBio sequencing system, which is the first complete system for sequencing DNA, allowing scientists to sequence entire genomes in under a week.

2013: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

2015: NCTC scientists publish the genome of *NCTC 1", generated with SMRT Sequencing, and compare it to other *S. pneumoniae"* isolates collected in 1954, 1984, and 2002.

2015: NCTC and Wellcome Sanger Institute have announced a five-year project to sequence a 3,000 bacterial strains from the human gut.

The Advent of Antibiotics

1947: Edward Tatum and Joshua Lederberg publish the first gene map of *E. coli*. A map of the *E. coli"* genome is published in 1947 by Tatum and Lederberg.


1953: Waksman and Albert Schatz are awarded the Nobel Prize in Physiology or Medicine for their work on antibiotics.

1950s: Waksman and Albert Schatz are awarded the Nobel Prize in Physiology or Medicine for their work on antibiotics.

1954: The biggest recorded outbreak of foodborne illness is documented, with over 200,000 cases of food poisoning. She eventually deposits over 200,000 *NCTC"* strains of bacteria associated with foodborne illness.

1960s: The European Collection of Microorganisms for Biotechnology (ECM) is established.

1977: Gilbert and Sanger independently develop methods to determine the exact sequence of DNA molecules.

1977: CDC researchers Joseph McDade and Charles D. Shepard isolate *Legionella pneumophila"*, the bacterial pathogen behind the outbreak of Legionnaires’ disease in Philadelphia.

1977: CDC researchers Joseph McDade and Charles D. Shepard isolate *Legionella pneumophila"*, the bacterial pathogen behind the outbreak of Legionnaires’ disease in Philadelphia.

1980s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

1990s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

2000s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

2010s: The National Collection of Type Cultures (NCTC) begins a 10-year effort to characterize every organism in the collection.

Our understanding of microbiology has evolved enormously over the last 150 years. Few institutions have witnessed our collective progress move more closely than the National Collection of Type Cultures (NCTC). In fact, the collection itself is a record of the many milestones microbiologists have crossed, building on the discoveries of those who came before.

To date, 60% of NCTC’s historic collection now has a closed, finished reference genome, thanks to PacBio Single Molecule, Real-Time (SMRT) Sequencing. We are excited to be their partner in crossing this latest milestone on their quest to improve human and animal health by understanding the microscopic world.
Trustworthy Biological Resources for the Future

Historical collections are not only important in preserving the past— they are vital in helping us understand current pathogens, and in the development of future medical advances.

“Knowing very accurately what bacteria looked like before and during the introduction of antibiotics and vaccines and comparing them to current strains from the same collection shows us how they have responded to these treatments,” Prof Parkhill said. “This, in turn, helps us develop new antibiotics and vaccines.”

Plague, cholera, streptomycetes, and 250 strains of \textit{E. coli}, were among the reference genomes created and released in June 2018, as well as several of the most important known drug-resistant bacteria, such as tuberculosis and gonorrhea. The genome sequences of these highly valuable strains are fundamental for developing methods to identify specific infections in people, including tests that can be used in the field to rapidly identify the source of an outbreak and help contain infections. Applications could also include detection of bioterrorism agents, such as anthrax.

“Having these reference collections would allow more accurate assessment of the source of any eventual biological threat,” Prof Parkhill said.

The sequences will also be of great scientific significance. They include the ‘type strains’ of many bacteria in the collection—the first strains that describe the species and are used to classify them; 852 were bacterial species associated with human infection, and at least 298 of those type strains had no WGS data available in any public databases. Furthermore, the NCTC 3000 data set may reveal gaps in the range of the current collection, allowing NCTC to better curate the collection by identifying where strains are missing from clinically important lineages.

“This provides great potential for using the data in phylogeny and populations genetics studies,” Russell said.

The genomic data has been made publicly available through the European Bioinformatics Institute. An additional electronic portal that will bring together all the metadata associated the NCTC strains, including the raw and assembled genomic data is being created. The WGS information also benefits proteomic scientists who access the data to help to interpret the profiles generated by mass spectrometry.

“If NCTC is to remain scientifically important, it is essential that we embrace WGS technology and provide accurate reference genome sequences.”

— Julie E. Russell
Head of Culture Collections
Public Health England

“Our collection was established by scientists with incredible foresight in recognizing the need for trustworthy biological resources, and we are committed to ensuring that it remains scientifically relevant for the emerging challenges of the 21st century,” Russell said.

PacBio CSO Jonas Korlach said the project demonstrates the value of having complete genomes and praised the NCTC and Sanger Institute for making the most of its “living fossil” collection to bridge past, present and future.

“Rather than trying to understand and replicate an organism from the past by relying on a few bones, the NCTC has been able to get a much more complete picture from the living fossils in its collection, which will be of great benefit to all of us,” Korlach added.

Julie E. Russell is the Head of Culture Collections at Public Health England